Brazilian Journal of Biology®

ISSN 1519-6984 (Print)
ISSN 1678-4375 (Online)

Original Article

The use of Artificial Intelligence in predicting Respiratory Syncytial Virus-inhibiting flavonoids

O uso da Inteligência Artificial na predição de flavonoides inibidores do Vírus Sincicial Respiratório

B. R. P. Lopes[a], T. T. Albertini[a,b], M. F. Costa[a,c], A. S. Ferreira[a,b], K. A. Toledo[a,c*] and J. C. Rocha[a,b]

[a]Universidade Estadual Paulista – UNESP, Faculdade de Ciências e Letras, Departamento de Ciências Biológicas, Assis, SP, Brasil

[b]Universidade Estadual Paulista – UNESP, Faculdade de Ciências e Letras, Laboratório de Matemática Aplicada, Assis, SP, Brasil

[c]Universidade Estadual Paulista – UNESP, Instituto de Biociências Letras e Ciências Exatas, Pós-graduação em Microbiologia, São José do Rio Preto, SP, Brasil

Abstract

Human Respiratory Syncytial Virus (hRSV) infection results in death and hospitalization of thousands of people worldwide each year. Unfortunately, there are no vaccines or specific treatments for hRSV infections. Screening hundreds or even thousands of promising molecules is a challenge for science. We integrated biological, structural, and physicochemical properties to train and to apply the concept of artificial intelligence (AI) able to predict flavonoids with potential anti-hRSV activity. During the training and simulation steps, the AI produced results with hit rates of more than 83%. The better AIs were able to predict active or inactive flavonoids against hRSV. In the future, in vitro and/or in vivo evaluations of these flavonoids may accelerate trials for new anti-RSV drugs, reduce hospitalizations, deaths, and morbidity caused by this infection worldwide, and be used as input in these networks to determine which parameter is more important for their decision.

Keywords: artificial intelligence, respiratory syncytial virus, flavonoids, antiviral.

Resumo

A infecção pelo Vírus Sincicial Respiratório Humano (hRSV) resulta na morte e hospitalização de milhares de pessoas em todo o mundo a cada ano. Infelizmente, não existem vacinas ou tratamentos específicos para tais infecções. A testagem de centenas, ou mesmo milhares, de moléculas promissoras é um desafio para a ciência. Neste trabalho, nós integramos propriedades biológicas, estruturais e físico-químicas para treinar e aplicar o conceito de inteligência artificial (IA) capaz de prever flavonoides com potencial atividade anti-hRSV. Durante as etapas de treinamento e simulação, a IA produziu resultados com taxas de acerto superiores a 83%, sendo capaz de prever flavonoides ativos ou inativos contra o hRSV. No futuro, avaliações *in vitro* e/ou *in vivo* desses flavonoides poderão acelerar os testes de novas drogas anti-RSV, reduzir hospitalizações, mortes e morbidade causadas por essa infecção. Além disso, a validação futura destes dados poderá determinar qual parâmetro tem maior peso na decisão da inteligência.

Palavras-chave: inteligência artificial, vírus sincicial respiratório, flavonoides, antiviral.

1. Introduction

Human respiratory syncytial virus (hRSV) or human orthopneumovirus is a member of the Pneumoviridae family of negative-sense single-stranded RNA viruses. Human infection with hRSV causes Acute Lower Respiratory Tract Infections (ALRTI) in newborns and children and is considered a public health problem worldwide, due to high mortality and hospitalization rates and the high treatment costs. hRSV infections affect approximately 70% of newborns in their first year of life and 95% of children up to 2 years of age, resulting in more than 3 million hospitalizations and approximately 200,000 deaths per year (Noor and Krilov 2018).

Currently, there is no vaccine available that can prevent hRSV infection. Only two drugs for the treatment of hRSV have been approved for human use: Palivizumab (a neutralizing monoclonal antibody) and Ribavirin (a broad-spectrum antiviral). Treatment costs limit Palivizumab therapy in high-risk patients. However, the low cost of Ribavirin does not outweigh its limited efficacy and risk of severe side effects (Anderson et al., 1990). Recently, two small-molecule anti-RSV therapeutics have been investigated in phase II clinical trials: GS-5806 or Presatovir, an allosteric entry inhibitor and ALS-8176 or Lumicitabine, a ribonucleoside analog (DeVincenzo et al. 2014; Wang et al. 2015).

*e-mail: karina.toledo@unesp.br
Received: December 29, 2022 – Accepted: April 8, 2023

Therefore, new therapeutic options against hRSV disease are urgently needed to address this unmet clinical need. In this context, flavonoids could be an interesting option. Flavonoids are natural, heterogeneous, and numerous (more than 6,000) compounds synthesized in plants in response to stress conditions and play an important role in the defense of plant cells against pathogens and insects (Nabavi et al. 2020). In vitro and in vivo studies have shown that flavonoids have low toxicity and have a synergistic effect with other drugs.

Chemically, flavonoids are hydroxylated phenolic molecules that have shown positive results in studies against variety of DNA and RNA viruses (Lalani and Poh 2020). Given the thousands of flavonoids that have been described, a small number of these compounds have been evaluated for their anti-hRSV activity. Some of them have shown positive results (Lopes et al. 2020; Ma et al. 2002; Wang et al. 2012; Chung et al. 2013). Investigating the anti-hRSV activity of this large number of molecules requires a large investment of time and money (Wouters et al., 2020). To address these obstacles, we propose the use of Artificial Intelligence (AI) techniques by combining artificial neural networks (ANN's) and genetic algorithms (GAs) techniques.

Currently, AI techniques are widely used in complex problems of classification, clustering, pattern recognition, and prediction. The applications are diverse, including face and voice recognition, medicine for disease diagnosis, weather prediction, and classification of reproductive data (Abiodun et al. 2018; Fernandez et al. 2020). They are also used to increase the efficiency and speed of therapeutic drug discovery (Mandlik et al., 2016; Sutariya et al., 2014).

In this study, various artificial intelligences were fed with experimental information from anti-hRSV assays and biological, structural, and physicochemical parameters of flavonoids known to be active or inactive for anti-hRSV activity. Nine artificial intelligences were selected to evaluate 489 untested flavonoids from the existing literature for their anti-hRSV activity. In this blind test, the AIs were able to classify the compounds as active and inactive for future testing *in vitro* and/or *in vivo*.

2. Materials and Methods

2.1. Definitions for the AI variables

Considering the correlation between molecular structure, biological activity, and physicochemical properties, we selected variables related to these characteristics, to be used as input data for AI. The variables are described as follows:

2.2. Biological variables from in vitro assays

The PubMed online platform was searched for articles testing the antiviral activity of flavonoids in vitro against hRSV using HEp-2 cells (Gamble, 2017). Various information was extracted from the articles, such as the viral strain tested (Long, A or B), the PFU (plaque forming unit), the concentration (µg/mL) of flavonoids, their effectiveness as anti-hRSV agents, and the type of test conducted (screening, virucidal, pre- and post-treatment). If necessary, the multiplicity of infection (MOI) was converted to PFU. Experimental conditions in which the flavonoid inhibited ≥ 50% of the viral infection were classified as effective (active flavonoid). Values below 50% were considered ineffective (inactive flavonoids). Experimental data was referred to as "Empirical data" Experimental data obtained by linear regression from the results were referred to as "Theoretical data". When the two data sources were combined, the data was referred to as "Total data" When it was not possible to extract all necessary data, the article was excluded from our research. All flavonoids included in this study are listed in Table 1.

2.3. Physicochemical flavonoids variables

The physicochemical properties of each flavonoid was determined and calculated using two public tools: PubChem and Open Babel software (O'Boyle et al., 2011). From PubChem, we extracted CID identification (or compound identifier number) and SDF files containing the three-dimensional coordinates of the atoms that make up each flavonoid (atom_block and bond_block). From Open Babel, we obtained standard molecular properties such as molecular weight, number of atoms, bonds and rings, molar refractive power, octanol-water position, and topological polar surface area. In addition, we calculated the energetic properties of each flavonoid (bond stretching, angular bending, stretch bending, torsional energy, out-of-plane bending, Van der Waals energy, electrostatic energy, and total energy). All data is listed in Table 2.

2.4. Input dataset

The biological and physicochemical variables were divided into three distinct sets: Empirical, Theoretical, and Total data. Each set was further divided subdivided into two sub-datasets: one referring to training, validation, and testing, the other referring to simulation. In the first sub-dataset (training), the Empirical, Theoretical, and Total data presented was 1106, 3649 and 4755 variables from the input dataset, respectively. In the second sub-dataset (validation), the Empirical, Theoretical, and Total data presented was 200, 650 and 850 variables of the input dataset, respectively, representing 15% of the total of each database.

2.5. Artificial Intelligence

The application of the AI technique consists of an association between artificial neural network techniques and genetic algorithms. The ANN technique was responsible for modeling the ANN architectures and the GA technique for ANN optimization. Artificial neural networks are inspired by early models of sensory processing in the brain. Using algorithms that mimic the processes of real neurons, it is possible to have the network learn to solve many types of problems. Here, we used a Multilayer Perceptron ANN, that has one or more hidden layers, in addition to the input and output layers (Pagel and Kirshtein, 2017). Each layer consists of nodes (artificial neurons), and for each node, a weighted sum, an input "value x corresponding weights", and an activation function is performed to produce an output (Figure 1, Artificial Neuron). The resulting value of this equation will or will not, activate the node (Fernandez et al., 2020).

Table 1. Flavonoids previously investigated in the literature on their anti-hRSV activity.

Flavones**		Flavans**	
CID*	**Name**	**CID***	**Name**
14005	Quercetin pentaacetate[1]	932	Naringenin[3]
64982	Baicalin[2]	9064	Cianidanol[3]
72344	Nobiletin[2]	10621	Hesperidin[3]
114776	Isoorientin[2]	65084	Gallocatechin[3]
124034	Swertisin[2]	68071	Pinocembrin[3]
155692	Spinosin[2]	161557	Ampelopsin[3]
162350	Isovitexin[2]	165506	4'-Hydroxyflavanone[3]
188323	Cirsimaritin[2]	442428	Naringin[3]
261859	Gardenin A[2]	503737	Liquiritin[4]
442659	Swertiajaponin[3]	11669392	7-O-Galloyltricetiflavan[5]
3084961	Oroxindin[3]	11999968	7,4'-Di-O-galloyltricetiflavan[5]
5271991	Ganhuangenin[3]	12309904	Isookanin[6]
5280343	Quercetin[1]		
5280443	Apigenin[3]		
5280445	Luteolin[3]	**Isoflavonoids****	
5280862	Isokaempferide[3]	CID*	Name
5281600	Amentoflavone[3]	124052	Glabridin[7]
5281605	Baicalein[3]	187808	Glycitin[7]
5281607	Chrysin[3]	480783	6,8-Diprenylgenistein[7]
5281616	Galangin[3]	5280373	Biochanin[8]
5281628	Hispidulin[3]	5280448	Calycosin[8]
5281672	Myricetin[3]	5280961	Genistein[8]
5281697	Scutellarein[3]	5281704	Afrormosin[5]
5281703	Wogonin[3]	5281708	Daidzein[5]
5281954	Tectochrysin[3]	5281812	Texasin[9]
5318997	Icariin[3]	5282074	Hydroxygenistein[9]
5320315	Oroxylin A[3]	5481948	Semilicoisoflavone B[9]
5320438	Pectolinarigenin[3]		
5481982	Baohuoside II[3]		
10005544	10-O-Methylquercetin tetraacetate[3]		
11726019	Isothymonin[3]	**Aurone****	
16681753	10',5',6,7-tetramethoxyflavone[3]	CID*	Name
44144321	Quercetin-10-rham(4)-gluc[3]	5281295	Sulfuretin[4]

*CID = compound identification from PubChem site. **classification flavonoids classes. References: [1] Lopes et al. (2020) [2] Wang et al. (2012) [3] Chung et al. (2013) [4] Ma et al. (2001) [5] Li et al. (2006) [6] Song et al. (2016) [7] Kaul et al. (1985) [8] Ma et al. (2002) [9] Shi et al. (2016)

Table 2. Biological and Physicochemical variables used to develop and train the ANN.

Source	Abbr.	Definition	Type	Unit	ANN Structure
DOI	CID	Identification in PubChem	String	None	Not Input
SDF file	AB	Cartesian 3D Position	(R) Number	Å	Input
SDF file	BB	Connections Between Atoms	(N) Number	Covalent Bond	Input
Open Babel	MW	Molecular Mass	(R) Number	g/mol	Input
Open Babel	NA	Total N. Atoms	(N) Number	Atoms	Input
Open Babel	NB	Total N. Bonds	(N) Number	Covalent Bond	Input
Open Babel	NR	N. of Chemical Rings	(W) Number	None	Input
Open Babel	LP	Log(10) of Partition Coefficient	(R) Number	None	Input
Open Babel	PSA	Polar Surface Area	(R) Number	Å^2	Input
Open Babel	BS	Energy of Bonds Stretching	(R) Number	Kcal/mol	Input
Open Babel	AB	Energy in Angle Between Atoms Bonded	(R) Number	Kcal/mol	Input
Open Babel	SB	Energy of Stretch Bending	(R) Number	Kcal/mol	Input
Open Babel	TE	Energy of Torsional Rotation Between Atoms	(R) Number	Kcal/mol	Input
Open Babel	OPB	Energy of Bending Outside of the Plan	(R) Number	Kcal/mol	Input
Open Babel	VDWE	Van der Waals Energy	(R) Number	Kcal/mol	Input
Open Babel	EE	Molecular Surface Charge	(R) Number	Kcal/mol	Input
Open Babel	TE	Sum of all OBprop Energy in F.F.	(R) Number	Kcal/mol	Input
DOI	S	hRSV Strain	(N) Number	None	Input
DOI	PFU	Particle Forming Units	(R) Number	N. particles/mL	Input
DOI	T	Type of Assay	(N) Number	None	Input
DOI	C	Concentration of Assay	(R) Number	µg/ml	Input
ARN	E	Have Effect	(W) Number	None	Output
ARN	NE	Don't Have Effect	(W) Number	None	Output

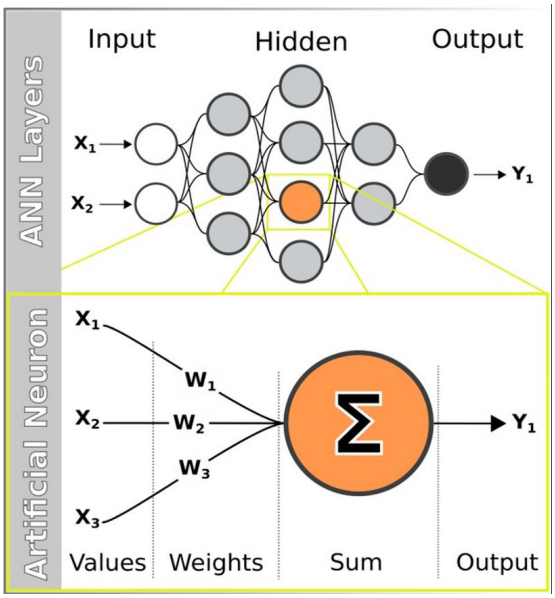

Figure 1. Multilayer Perception ANN structure and Artificial neuron structure.

The backpropagation algorithm was used. Thus, the results were compared with the real results of antiviral activity, and the weights were changed to obtain fewer errors. Each architecture consists of input, output, and hidden layers (between 1 and 3) with a random number of neurons ranging between (10-600), transfer functions (logsig, purelin, tansig, hardlim, tribas, radbas, and satlin), and training functions (trainrp, trainscg, traincgb, traincgf, traincgp, traingdm, and traingd) (Beale et al., 2017). The data processed by ANN was divided into three sets: training (70%), validation (15%), and testing (15%). Each set was randomly generated for each training session (Benzer and Benzer, 2015). At the end of the process, the percentage of success and failure of the network during learning was determined. To obtain the best ANN for determining the antiviral activity of flavonoids, the genetic algorithm technique was added to select the best network architectures (Linden et al., 2012).

2.6. Genetic Algorithm

The Genetic Algorithm (GA) technique was inspired by C. Darwin's theory. According to this theory, the principle of selection privileges the stronger individuals in terms of longevity and reproduction. Individuals who have more offspring have more chances to pass on their genetic codes to the next generation. These genetic codes represent the identity of each individual and are represented by chromosomes (Linden et al., 2012). In GAs, a chromosome undergoes an evolution process consisting of training, selection, recombination, and mutation. At the end of several evolutionary cycles, stronger individuals must be contained (Rosa and Luz, 2009). The aim of applying of the GA technique was to obtain ANN which represents the fewest errors in the flavonoids classification into "active" or "inactive" for the viral activity of the hRSV virus. Thus, the first step was to randomly create an initial population of different ANN's, ranging from 100 to 1000 individuals. For each ANN, the parameters that determine the architecture were defined. These parameters define the genes of the chromosomes that make up the GA populations. These parameters consisted of the number of neurons in the first, second, and third hidden layers, the transfer function for the first, second, and third hidden layers, the transfer function for the output layer, the training function used, and the number of hidden layers to be used. The data were ordered and classified according to the accuracy of the ANN. To select 20% of the best individuals for the next generation, a selection procedure known as elitism was used (Rosa and Luz, 2009). Moreover, 60% of the new generation is generated from the previous generation by crossing-over and mutation (this last generation is not more than 5%). The remaining 20% were provided by migration, that is, they were randomly generated according to the parameters needed in the initial population. After these stages, a new population was created using the best ANN. This cycle is called a generation. The maximum number of generations is 300 so, at the end of the generations the software shows the best ANN architecture to solve the problems of this study, the activity or inactivity of the virus.

2.7. Blind-test

As the three best AIs were determined for each dataset (Empirical, Theoretical, and Total) 489 flavonoids that had not yet been tested for anti-hRSV activity in the literature were analyzed using the nine better AIs (Supplementary Material Table 1). For this purpose, we set the following experimental conditions: the main three-dimensional flavonoid conformation available on the PubChem website, a hRSV strain, an inoculum of 100 PFU and four types of experimental treatment (screening, virucidal, pre- and post-treatment) in the presence of 16 µg/ml of the flavonoid. The values returned by AI were classified as "active" or "inactive" compounds. Agreement and disagreement between AIs were reviewed. A list of "most promising" compounds and "least promising" compounds was generated.

2.8. Further statistical analysis

Statistical analysis of the results was performed using the Receiver Operating Characteristic (ROC) and the corresponding Area Under the Curve (AUC). The AUC is defined as the area under the ROC curve and is calculated based on the sensitivity and specificity of the results obtained, which allows., an independent analysis of the results presented (Fawcett, 2006).

3. Results

Following the methodology described above, after the networks were trained, validated and tested, three better ANN architectures were selected using the GA technique. The parameters of these ANN's are listed in Table 3 for each dataset (E1-E3 for the Empirical dataset; T1-T3 for the Theoretical dataset and TT1-TT3 Total dataset). The Table 3 shows the number of neurons (N) in each layer and the functions that were used by the software to generate the outputs and to update the weight values in the backpropagation algorithm (Transfer Function; TF).

The performances of the AIs created with the best ANN are shown in Table 4. It can be observed that the best AIs in Total databases have success rates for training and simulation (>83%). AIs trained with Empirical, Theoretical, and Total data can recognize active/inactive flavonoids at 91–87, 97–99 and 83–92%, respectively. Details of these three performances can be observed by analyzing the confusion matrices for the training and AI simulation (Figure 2), and for the other AIs can be found in the Supplementary Material, S1. For example, Figure 2 shows the confusion matrices for training and simulation from the entire dataset (TT1). When considering the known active flavonoids (TP + FN), AI had an accuracy of 81.7% and an error of 18.3% (c values in Figure 2). When considering known inactive flavonoids (FP + TN), AI had an accuracy of 93.1% and an error of 6.9% (d values). When we mixed known active and inactive flavonoids, AI was 83% (a value) and 92.5% (b value) to predict active and inactive flavonoids, respectively. Overall, when trained with 4755 different entries, the AI was 89.8% correct and 10.2% incorrect (e values) in predicting the anti-hRSV activity of flavonoids.

During simulation, the AI was 71.6% (h) and 90.3% (i) for the known active and inactive flavonoids, respectively. When the active and inactive conditions were mixed, the AI was 77.6% (f) and 87.2% (g), respectively. In the simulation phase, 194 data points were correctly classified as active and 523 as inactive, giving an accuracy of 84.4% from a total of 850 data points. The same analysis can be performed if we consider the confusion matrix for the data of E1 and T1 (Figure 2). The confusion matrices for the other datasets (TT2, TT3, E2, E3, T2, and T3) are shown in Figure S1.

To evaluate the results of training and simulation, the receiver operator characteristic (ROC) curve and area under the ROC curve were used in this study (Figure 3 and Supplementary Material, S1). The ROC is a probability curve and has two evaluation parameters: the false positive rate (represented by the x-axis) and the true positive rate (represented by the y-axis). Therefore, the ROC curves were calculated to predict antiviral activity (active data or inactive data). In the three cases (Total, Empirical, and Theoretical data), the positioning of the curves in the northwestern region showed that the AI did not randomly select the exits, which would be the case if the curve of the results were close to the diagonal between the x-axis and y-axis, represented by the gray line in the ROC curve graph. Based on this result, it can be concluded that a learning process has taken place and the AI has shown satisfactory performance in predicting the viral effect of the flavonoid. Complementarily, the AUC (measures the two-dimensional area under the ROC curve), which summarizes the ROC curve into a unique value, showed the following results for the training Total (TT1), Empirical (E1), and Theoretical (T1) datasets, considering the antiviral activity or inactivity, 0.97 - 0.969, 0.938 - 0.936, and 0.998- 0.999, respectively (Figure 3).

Table 3. Parameters of the three best ANN architecture from each dataset.

Data set	ANN	Hidden layer						Training	Output
		1		2		3			
		N	T.F.	N	T.F.	N	T.F.		
Total	TT1	189	purelin	371	radbas	78	satlin	trainscg	satlin
	TT2	173	purelin	246	radbas	73	hardlim	trainscg	purelin
	TT3	228	logsig	117	radbas	67	tansig	trainrp	logsig
Empirical	E1	47	radbas	86	tansig	46	tribas	trainrp	tansig
	E2	122	logsig	79	logsig	81	tribas	trainrp	logsig
	E3	241	tansig	54	logsig	23	tansig	trainrp	tansig
Theoretical	T1	139	logsig	79	tribas	56	logsig	trainrp	purelin
	T2	389	purelin	166	radbas	45	satlin	trainscg	tansig
	T3	332	purelin	173	radbas	38	satlin	trainscg	satlin

Where N = number of neurons; T.F. = transfer function.

Table 4. Anti-hRSV activity predictions (%) with best parameters from each dataset.

Data set	AIs	Training		Simulation	
		Active	Inactive	Active	Inactive
Total	TT1	83.0	92.5	77.6	87.2
	TT2	78.6	93.0	72.0	87.8
	TT3	85.8	89.2	84.4	84.7
Empirical	E1	91.4	87.6	84.4	84.5
	E2	82.8	91.6	75.6	89.7
	E3	79.3	94.0	71.1	92.9
Theoretical	T1	97.2	99.3	99.0	97.6
	T2	96.5	99.2	97.5	96.9
	T3	97.7	96.6	95.5	93.6

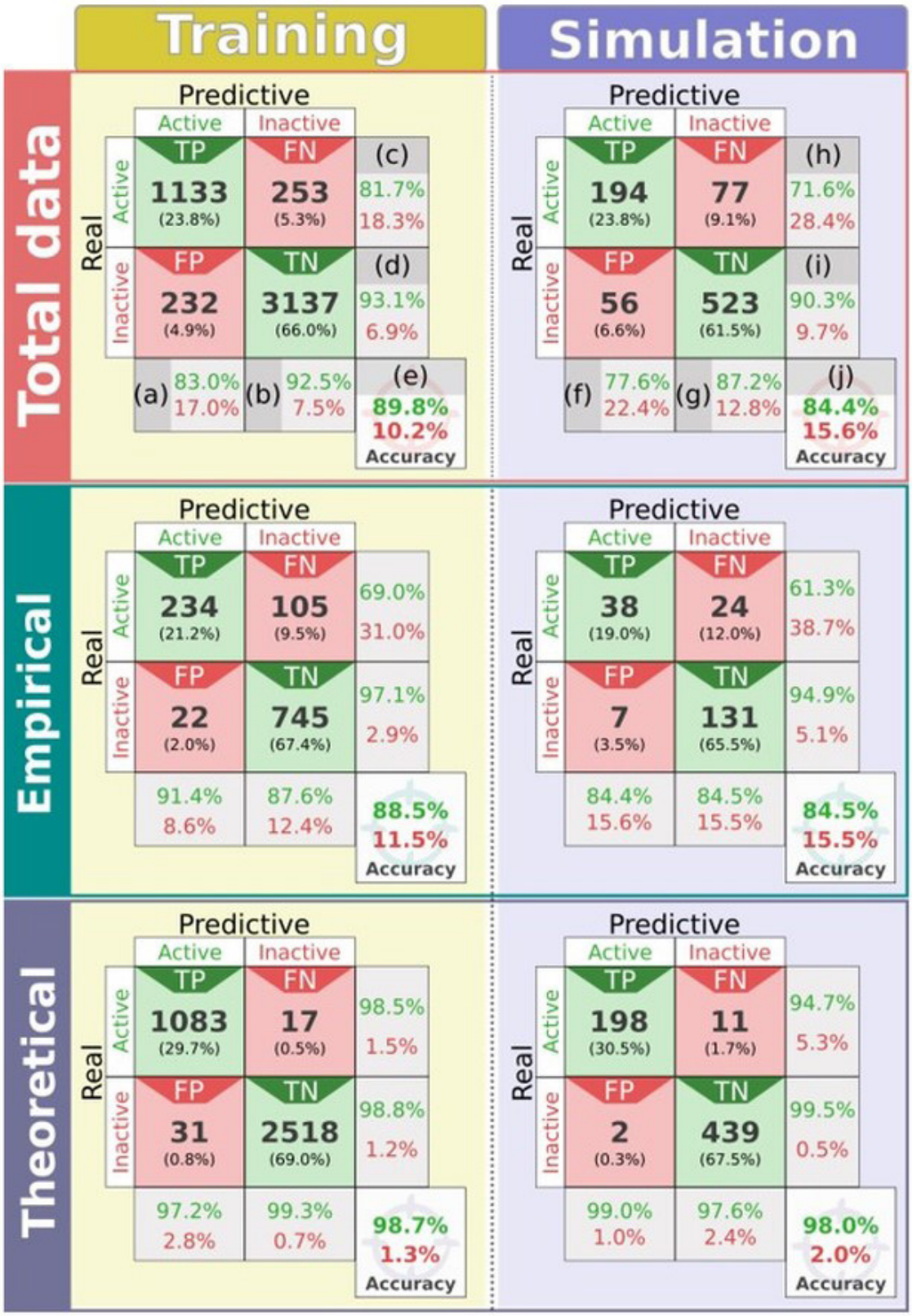

Figure 2. Confusion Matrices in training and simulation from the Total (TT1), Empirical (E1) and Theoretical (T1) datasets. The diagonal cells (green cells) correspond to the correctly classified data and the cells outside the diagonal (red cells) correspond to misclassified data. TP: True Positive; TN: True Negative; FN: False Negative; FP: False Positive; a–j: percentage of correct and error.

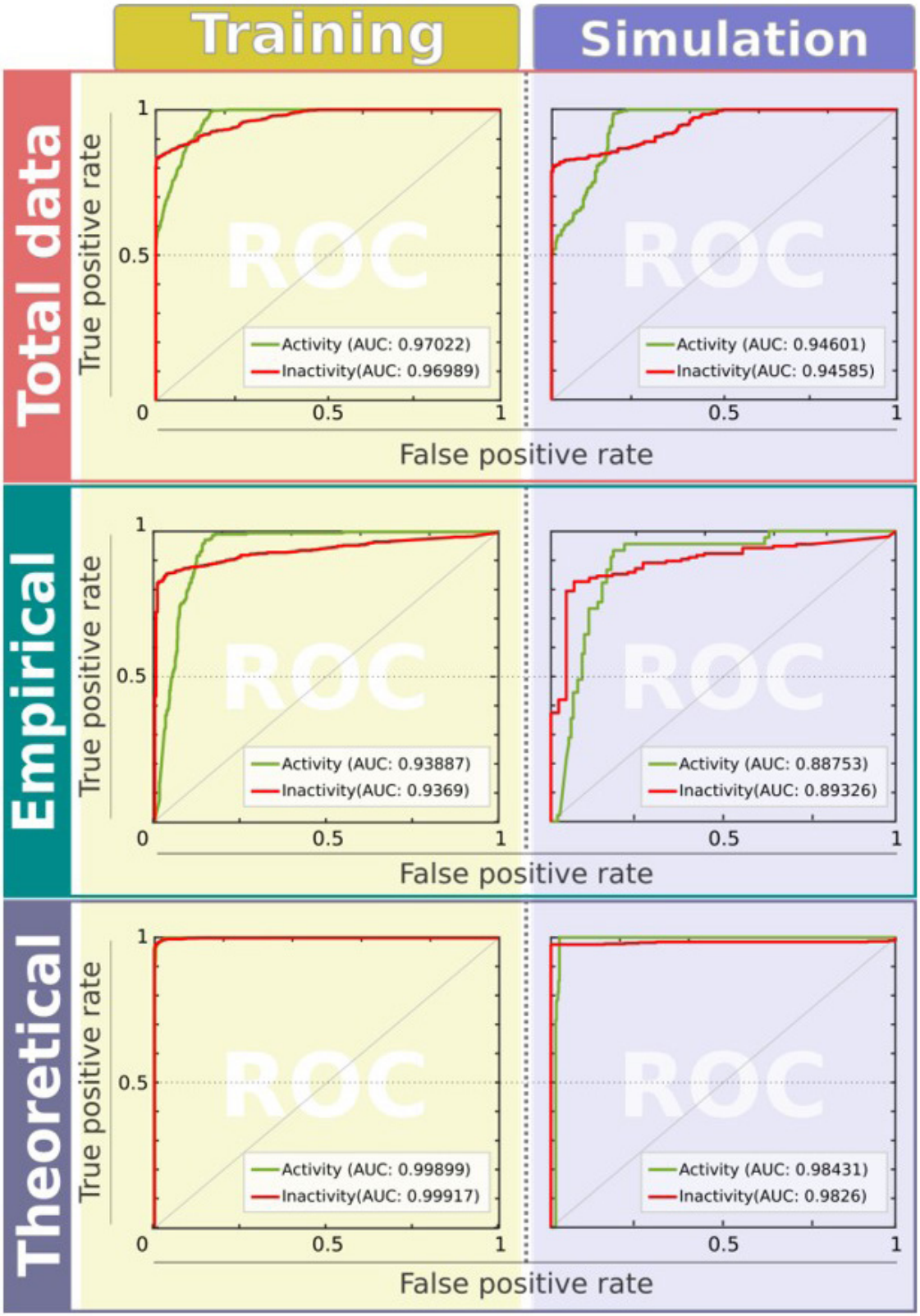

Figure 3. ROC curve and AUC values for training and simulation steps. The graphs show the ROC curves and the AUC values for the training and simulation steps of the Total (TT1), Empirical (E1) and Theoretical (T1) datasets. The AUC values provide information on the assertiveness of the AI's assertiveness in predicting flavonoids activity or inactivity of against hRSV.

Also, for the simulation data, we have the following results for the AUC: 0.946 - 0.945, 0.887 - 0.893 and 0.984 - 0.982, for the Total (TT1), Empirical (E1), and Theoretical (T1) datasets, respectively (Figure 3). In Figure 4, we show the performance of the AIs for Total (TT1), Empirical (E1), and Theoretical (T1) data, where we observe Mean Squared Errors (MSE) of the order of 10^{-2} for all ANN's. The ROC curves, the AUC, and the performance values from the other AIs (TT2 - 3, E2 - 3, and T2 - 3) presented similar results, which are presented in Figure S1.

Since the data described so far indicates a high hit rate of AIs (> 83%), we submitted 489 flavonoids with unknown anti-hRSV activity to AI to predict them. For this purpose, a dataset was created that included physicochemical and biological variables of these flavonoids under different experimental conditions.

The following experimental conditions were set for this blind test: inoculum of 100 PFU of strain A, four types of treatment (screening, virucidal, pre and post-treatment) in the presence of 16 µg/ml of the flavonoid. The combination of the biological and physicochemical variables with the experimental conditions of the 489 flavonoids resulted in 1956 input datasets that were submitted for AI analysis. For this purpose, the nine best AI previous described previously (E1-E3, T1-T3 and TT1-TT3) were used. At the end of the blind test, the results provided by the different AIs indicate the probability that the flavonoid is experimentally active or inactive. These results were compared to determine the agreement between the AIs (Table 5). Table 5 shows that of the 1956 input datasets, the three (E1-E3) AIs generated from the Empirical data concordantly classified 500 input datasets as inactive and another 348 as active. For the AIs generated from the Theoretical data (T1-T3), they agree in classifying 471 input data as inactive and 132 as active. Finally, the AIs generated from the Total data (TT1-TT3) they agreed to classify 713 input data as inactive and another 31 as active.

Lastly, it was possible to list the compounds that were considered with more or less potential against hRSV by the AIs depending on the experimental conditions and the biological and physicochemical variables provided. For this purpose, we selected the 20 compounds for which a larger number of AIs, specifically nine, agreed to classify them as active or inactive. At least seven AIs agreed to classify 10 compounds as active. These compounds are identified by CIDs: 71307295, 5379096, 10449654, 485522, 629964, 5318869, 5320399, 5458461, 5746354 and 5321398 and belong to the flavan, flavone, and isoflavone class of flavonoids. The AIs were also evaluated for the type of antiviral assay in which each of these compounds would show promising results against hRSV. The most frequently reported type of antiviral assay was post-treatment, followed by screening, virucidal, and pre-treatment.

For the compounds classified as inactive, 100% of AIs agreed in 61 cases. To compile the final list of compounds, we selected only those among the 61 where the platform ZINC15 indicated at least five suppliers, resulting in in the selection of 10 flavonoids, as shown in Table 6.

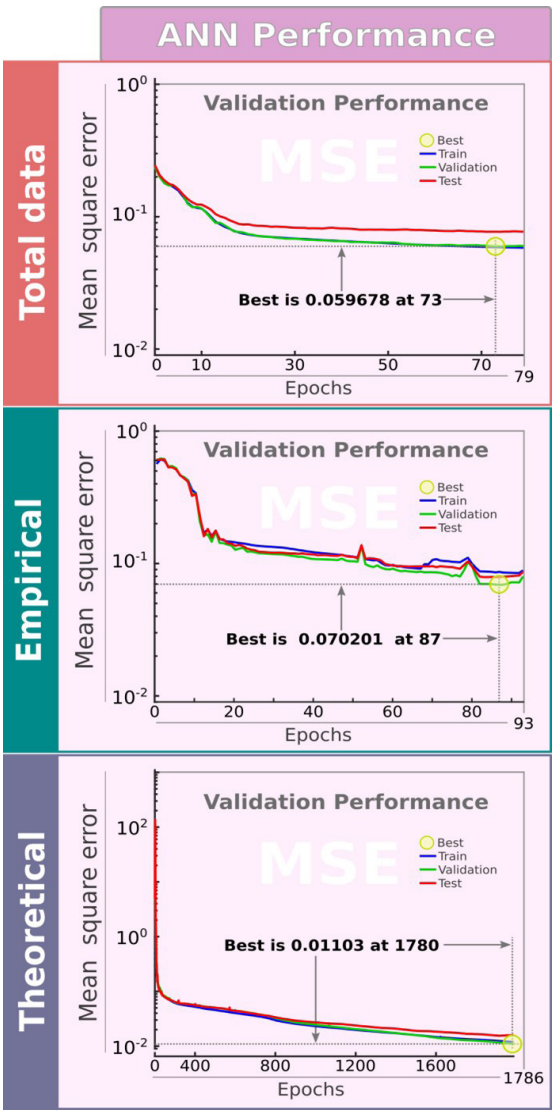

Figure 4. Total performance of ANN for each dataset. The performance graphs relate the Mean Square Error (MSE) to the epochs of the AIs for the Total (TT1), Empirical (E1) and Theoretical (T1) datasets. The curves represent the performance of the different stages of AI tests (train, validation and test) and the green dot indicates the better number of epochs.

Table 5. Agreement among all AI's.

Data set	AIs	Agreement	
		inactivity	activity
Total	TT1–3	713	31
Theoretical	T1–3	471	132
Empirical	E1–3	500	348

Table 6. Selection of flavonoids identified by AIs as active or inactive.

Prediction	CID	Class	Treatment*	AI agreement (%)	ZINC ID	Antiviral activity in literature
ACTIVE	71307295	Flavan	3	88.8	14642741	
	5379096	Flavone	2	88.8	14779854	Papilomavirus[1]
	10449654	Flavone	3	88.8	14762971	
	485522	Flavone	4	77.7	6483423	
	629964	Flavone	4	77.7	2585767	
	5318869	Flavone	3	77.7	5732364	Sars-Cov-2[2], Picornavirus[2]
	5320399	Flavone	1	77.7	2392262	
	5458461	Flavone	4	77.7	1763468	Influenza[3], Herpes simplex[4]
	5746354	Isoflavone	3	88.8	14762971	
	5321398	Isoflavone	2	77.7	1792763	Influenza[5], Sars-Cov[6]
INACTIVE	11483087	Flavanone	2	100	14728393	HIV[7]
	440735	Flavanone	2	100	58116	Sars-Cov-2[8]
	73202	Flavanone	4	100	14806381	
	5280681	Flavone	4	100	5998596	Hepatitis B[9], Poliovirus[10]
	5318761	Flavone	4	100	49823026	
	71437113	Flavone	2	100	85867362	
	15382687	Flavone	4	100	15271731	
	5281617	Flavone	2	100	5732375	Sars-Cov-2[11], African swine fever virus[12]
	56776173	Flavone	2	100	67903379	
	5281801	Isoflavone	2	100	6092209	HIV[13], HCV[14], Rhinovirus[15]

*The numbers for antiviral treatment indicate the kind of antiviral assay was predicted by the AI (1 for pre-treatment; 2 for virucidal; 3 for post-treatment and 4 for screening). Abbreviations: HIV, human immunodeficiency virus; HCV, hepatitis C virus. References: [1] Kumar et al. (2015) [2] Leal et al. (2021) [3] Walther et al. (2016) [4] Likhitwitayawuid et al. (2006) [5] Liu et al. (2008) [6] Yu et al. (2012) [7] Esposito et al. (2013) [8] Deshpande et al. (2020) [9] Lin et al. (2005) [10] Robin et al. (2001) [11] Nouadi et al. (2021) [12] Hakobyan et al. (2019) [13] Tewtrakul et al. (2007) [14] Lee et al. (2018) [15] Choi et al. (2010)

This parameter was included, considering that a better chance may exist in the future to test this compound in vitro and/or in vivo. The complete list of the 40 flavonoids unanimously ranked as inactive can be found in the Supplementary Material (Table S2). The CID of the compounds classified as inactive were: 11483087, 440735, 73202, 5280681, 5318761, 71437113, 15382687, 5281617, 56776173 and 5281801, belonging to the classes of flavanone, flavone, and isoflavone. The AIs further indicated that these flavonoids did not inhibit hRSV when tested in virucidal and screening assays.

4. Discussion

The hRSV is one of the most important causative etiological agents of infantile and senile bronchiolitis. Although hRSV infections result in thousands of cases worldwide each year, there are no vaccines. Treatments focus on supportive or preventive measures, such as the expensive monoclonal antibodies. The current scenario is forcing the scientific community to look for efficient and cost-effective drugs (Battles et al., 2016).

Externally, the hRSV virion has three structural proteins anchored in its membrane: F, G, and SH. Protein F has been the target of pharmacological strategies in the search for anti-hRSV compounds, as it has been shown to be important in the adhesion and internalization stages of the viral cycle (Battles et al., 2016). During these steps, the F protein trimer (F0) adopts an elongated structural conformation (F1) that facilitates the entry of the virion into the target cell (Krarup et al., 2015). Studies have shown that small molecules have great potential to inhibit the hRSV virus as they are able to interact in the central cavity of the F0 protein, preventing the conformational change of F1. Furthermore, hRSV-infected cells express the F protein in their plasma membrane, allowing the formation of large syncytia by cell fusion. Syncytia facilitates the spread of the virus in tissues (Battles et al., 2016).

Flavonoids are plant metabolites with a wide chemical and biological diversity, including antiviral activity. Since flavonoids are small molecules found in our daily diet, readily available, and relatively inexpensive, they have potential to be used as drugs to combat hRSV. Considering the hundreds of flavonoids that have been described, a small number of these compounds have been evaluated for their anti-hRSV activity. Several of them have been shown to have antiviral activity and others have yet to be analyzed (Lopes et al., 2020; Wang et al., 2012; Chung et al., 2013; Ma et al., 2001). These studies have also shown that the anti-hRSV activity of flavonoids is related to the inhibition of the first phases of the viral infection cycle, adhesion, and internalization, as well as to their chemical structural compounds (Li et al., 2006; Song et al., 2016; Kaul et al., 1985; Shi et al. 2016). Lopes et al., 2020 proposed that the anti-hRSV mechanism of action of flavonoids is related to the interaction of these compounds in the central cavity of the F0 protein, preventing its transition to the F1 conformation (Lopes et al., 2020).

Thus, the search for other flavonoids that may exhibit anti-hRSV activity is warranted. There are hundreds of them deposited in databases such as PubChem. When one considers the possibility of chemically altering these compounds through synthetic modifications, the number of compounds to be tested could number in the thousands.

In a recent article, the authors put the average cost of developing a new drug between $314 million to $2.8 billion over a decade or more (Wouters et al., 2020). Several research groups incorporated mathematical and/or computational approaches into their new drug studies to reduce the costs and time for drug development (Liu et al., 2008; Anusuya and Gromiha, 2017; Ganesan et al., 2017; Costa et al., 2016; Guimarães et al., 2018; Perilla et al., 2015; Scotti et al., 2015; Teixeira et al., 2017; Uriarte-Pueyo and Calvo, 2010).

Studies addressing the anti-hRSV activity of flavonoids have mainly involved in vitro or in vivo analyzes before mathematical and/or computational approaches were used (Cichero et al., 2017; Hao et al., 2011; Jiménez-Somarribas et al., 2017; Xia et al., 2016). The need to perform in-house tests beforehand does not lead to an optimal reduction in research time and cost. Other studies have evaluated the chemical or structural properties of flavonoids (González-Díaz et al., 2005; Ji et al., 2015). They have not considered the biological and experimental properties in their analysis.

Here, we propose the use of an AI technique, by feeding the artificial neural networks associated with genetic algorithms, with real input parameters from the literature. The data were organized into matrices, containing biological and physicochemical properties of flavonoids and experimental antiviral parameters of hRSV. This strategy is expected to provide a more robust, complete analysis in less time and at a lower cost in predicting the anti-hRSV activity of thousands of flavonoids that have not yet been tested in the literature.

In short, the artificial neural network is an artificial intelligence technique whose architecture mimics the knowledge acquisition and organizational capabilities of the human brain (Goh, 1995). The unique features of this computational model include robust performance in dealing with noisy or incomplete input patterns, high fault tolerance, and the ability to generalize from training data (Bertolaccini et al., 2017). These properties are responsible for various applications of ANN's, such as image processing, pattern recognition and prediction (Cristea, 2009; Choi et al., 2009). Once trained, the network becomes extremely fast, which is attractive for solving complex problems that require real-time processing. Combining this technique with genetic algorithms, a powerful optimization method based on the principles of genetics and natural selection, increases the efficiency of the network (Cristea, 2009).

The results generated by AI in this study can be divided into two parts. In the first part, nine AIs with a high percentage of reliable predictions were generated through training and simulation based on experimental parameters and biological and physicochemical data from the literature (> 83%). Next, the architecture of these AIs was used to evaluate flavonoids that have not yet been studied in the literature for their anti-hRSV activity. The input data included the structural and physicochemical parameters of 489 flavonoids under four different treatment conditions (pre, post, virucidal, and screening), always performed with 100 PFU of A strain and 16 µg/mL of flavonoid. Under these conditions, the AIs returned their active or inactive status. Evaluation of maximum agreement among AIs showed that at least seven AIs agreed in predicting 10 compounds as active.

Alternatively, at least nine AIs agreed in predicting 40 compounds as inactive, of which we selected 10 to highlight. The selection considered the commercial availability of the compounds which will facilitate future in vitro and in vivo tests. Among the active compounds, seven are flavones, two isoflavones, and one flavan, of which jaceosidin, kumatakenin, sophoricoside, and artocarpin have already been studied for some viruses (Table 6). Among the 10 inactive compounds, six are flavones, three flavanones, and one isoflavone, including 3-O-methylquercetin, orobol, and eriodictyol, which have already been studied for their antiviral activity against some viruses (Table 6).

In the future, these flavonoids could be tested in vitro and/or in vivo against hRSV. The results of these tests could indicate promising compounds to combat the hRSV virus, increase our knowledge about the activity of flavonoids against this virus and provide new data that could be incorporated into AI. Further studies may also indicate which input parameters are the most influential in AI's decision making, and thus improve its predictive ability.

5. Conclusion

Here, we have developed an artificial intelligence capable of predicting the active or inactive status of antiviral flavonoid activity against the hRSV virus. This important tool could accelerate the studies to find new anti-hRSV drugs, and thus reduce the number of hospitalizations, deaths, and illnesses caused by this viral infection worldwide.

Acknowledgements

We thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES; doctor's scholarship) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; Grant 2020/08588-0) for financial assistance and UNESP for structure support.

References

ABIODUN, O.I., JANTAN, A., OMOLARA, A.E., DADA, K.V., MOHAMED, N.A.E. and ARSHAD, H., 2018. State-of-the-art in artificial neural network applications: a survey. *Heliyon*, vol. 4, no. 11, pp. e00938. http://dx.doi.org/10.1016/j.heliyon.2018.e00938. PMid:30519653.

ANDERSON, L.J., ROBERT, A.P. and RAYMOND, L.S., 1990. Association between respiratory syncytial virus outbreaks and lower respiratory tract deaths of infants and young children. *The Journal of Infectious Diseases*, vol. 161, no. 4, pp. 640-646. http://dx.doi.org/10.1093/infdis/161.4.640. PMid:2319164.

ANUSUYA, S. and GROMIHA, M.M., 2017. Quercetin derivatives as non-nucleoside inhibitors for dengue polymerase: molecular docking, molecular dynamics simulation, and binding free energy calculation. *Journal of Biomolecular Structure & Dynamics*, vol. 35, no. 13, pp. 2895-2909. http://dx.doi.org/10.1080/07391102.2016.1234416. PMid:27608509.

BATTLES, M.B., LANGEDIJK, J.P., FURMANOVA-HOLLENSTEIN, P., CHAIWATPONGSAKORN, S., COSTELLO, H.M., KWANTEN, L., VRANCKX, L., VINK, P., JAENSCH, S., JONCKERS, T.H., KOUL, A., ARNOULT, E., PEEPLES, M.E., ROYMANS, D. and MCLELLAN, J.S., 2016. Molecular mechanism of respiratory syncytial virus fusion inhibitors. *Nature Chemical Biology*, vol. 12, no. 2, pp. 87-93. http://dx.doi.org/10.1038/nchembio.1982. PMid:26641933.

BEALE, M.H., HAGAN, M.T. and DEMUTH, H.B., 2017. *Neural Network Toolbox™: user's guide.* Natick: MathWorks.

BENZER, R. and BENZER, S., 2015. Appication of artificial neural network into the freshwater fish cought in Turkey. *International Journal of Fisheries and Aquatic Studies*, vol. 4, no. 06

BERTOLACCINI, L., SOLLI, P., PARDOLESI, A. and PASINI, A., 2017. An overview of the use of artificial neural networks in lung cancer research. *Journal of Thoracic Disease*, vol. 9, no. 4, pp. 924-931. http://dx.doi.org/10.21037/jtd.2017.03.157. PMid:28523139.

CHOI, H.J., BAE, E.Y., SONG, J.H., BAEK, S.H. and KWON, D.H., 2010. Inhibitory effects of orobol 7-O-d-Glucoside from Banaba (Lagerstroemia Speciosa L.) on human rhinoviruses replication. *Letters in Applied Microbiology*, vol. 51, no. 1, pp. 1-5. http://dx.doi.org/10.1111/j.1472-765X.2010.02845.x. PMid:20497313.

CHOI, H.J., KIM, J.H., LEE, C.H., AHN, Y.J., SONG, J.H., BAEK, S.H. and KWON, D.H., 2009. Antiviral activity of quercetin 7-rhamnoside against porcine epidemic diarrhea virus. *Antiviral Research*, vol. 81, no. 1, pp. 77-81. http://dx.doi.org/10.1016/j.antiviral.2008.10.002. PMid:18992773.

CHUNG, D.H., MOORE, B.P., MATHARU, D.S., GOLDEN, J.E., MADDOX, C., RASMUSSEN, L., SOSA, M.I., ANANTHAN, S., WHITE, E.L., JIA, F., JONSSON, C.B. and SEVERSON, W.E., 2013. A cell based high-throughput screening approach for the discovery of new inhibitors of respiratory syncytial virus. *Virology Journal*, vol. 10, no. 1, pp. 19. http://dx.doi.org/10.1186/1743-422X-10-19. PMid:23302182.

CICHERO, E., TONELLI, M., NOVELLI, F., TASSO, B., DELOGU, I., LODDO, R., BRUNO, O. and FOSSA, P., 2017. Benzimidazole-based derivatives as privileged scaffold developed for the treatment of the rsv infection: a computational study exploring the potency and cytotoxicity profiles. *Journal of Enzyme Inhibition and Medicinal Chemistry*, vol. 32, no. 1, pp. 375-402. http://dx.doi.org/10.1080/14756366.2016.1256881. PMid:28276287.

COSTA, M.F., JESUS, T.I., LOPES, B.R.P., ANGOLINI, C.F.F., MONTAGNOLLI, A., GOMES, L.P., PEREIRA, G.S., RUIZ, A.L., CARVALHO, J.E., EBERLIN, M.N., SANTOS, C. and TOLEDO, K.A., 2016. Eugenia aurata and eugenia punicifolia HBK inhibit inflammatory response by reducing neutrophil adhesion, degranulation and NET release. *BMC Complementary and Alternative Medicine*, vol. 16, no. 1, pp. 403. http://dx.doi.org/10.1186/s12906-016-1375-7. PMid:27770779.

CRISTEA, P.D. 2009. Application of neural networks in image processing and visualization. In: R.D. AMICIS, R. STOJANOVIC, G. CONTI, eds. *GeoSpatial visual analytics*. Dordrecht: Springer. http://dx.doi.org/10.1007/978-90-481-2899-0_5.

DESHPANDE, R.R., TIWARI, A.P., NYAYANIT, N. and MODAK, M., 2020. In silico molecular docking analysis for repurposing therapeutics against multiple proteins from SARS-CoV-2. *European Journal of Pharmacology*, vol. 886, pp. 173430. http://dx.doi.org/10.1016/j.ejphar.2020.173430. PMid:32758569.

DEVINCENZO, J.P., WHITLEY, R.J., MACKMAN, R.L., SCAGLIONI-WEINLICH, C., HARRISON, L., FARRELL, E., MCBRIDE, S., LAMBKIN-WILLIAMS, R., JORDAN, R., XIN, Y., RAMANATHAN, S., O'RIORDAN, T., LEWIS, S.A., LI, X., TOBACK, S.L., LIN, S.L. and CHIEN, J.W., 2014. Oral GS-5806 activity in a respiratory syncytial virus challenge study. *The New England Journal of Medicine*, vol. 371, no. 8, pp. 711-722. http://dx.doi.org/10.1056/NEJMoa1401184. PMid:25140957.

ESPOSITO, F., SANNA, C., DEL VECCHIO, C.D., CANNAS, V., VENDITTI, A., CORONA, A., BIANCO, A., SERRILLI, A.M., GUARCINI, L., PAROLIN, C., BALLERO, M. and TRAMONTANO, E., 2013. Hypericum Hircinum L. components as new single-molecule inhibitors of both HIV-1 reverse transcriptase-associated DNA polymerase and ribonuclease H activities. *Pathogens and Disease*, vol. 68, no. 3, pp. 116-124. http://dx.doi.org/10.1111/2049-632X.12051. PMid:23821410.

FAWCETT, T., 2006. An introduction to ROC analysis. *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874. http://dx.doi.org/10.1016/j.patrec.2005.10.010.

FERNANDEZ, E.I., FERREIRA, A.S., CECÍLIO, M.H.M., CHÉLES, D.S., SOUZA, R.C.M., NOGUEIRA, M.F.G. and ROCHA, J.C., 2020. Artificial intelligence in the IVF laboratory: overview through the application of different types of algorithms for the classification of reproductive data. *Journal of Assisted Reproduction and Genetics*, vol. 37, no. 10, pp. 2359-2376. http://dx.doi.org/10.1007/s10815-020-01881-9. PMid:32654105.

GAMBLE, A., 2017. PubMed Central (PMC). *The Charleston Advisor*, vol. 19, no. 2, pp. 48-54. http://dx.doi.org/10.5260/chara.19.2.48.

GANESAN, A., COOTE, M.L. and BARAKAT, K., 2017. Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discovery Today*, vol. 22, no. 2, pp. 249-269. http://dx.doi.org/10.1016/j.drudis.2016.11.001. PMid:27890821.

GOH, A.T.C., 1995. Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, vol. 9, no. 3, pp. 143-151. http://dx.doi.org/10.1016/0954-1810(94)00011-S.

GONZÁLEZ-DÍAZ, H., CRUZ-MONTEAGUDO, M., VIÑA, D., SANTANA, L., URIARTE, E. and DE CLERCQ, E., 2005. QSAR for anti-RNA-virus activity, synthesis, and assay of anti-rsv carbonucleosides given a unified representation of spectral moments, quadratic, and topologic indices. *Bioorganic & Medicinal Chemistry Letters*, vol. 15, no. 6, pp. 1651-1657. http://dx.doi.org/10.1016/j.bmcl.2005.01.047. PMid:15745816.

GUIMARÃES, G.C., PIVA, H.R.M., ARAÚJO, G.C., LIMA, C.S., REGASINI, L.O., MELO, F.A., FOSSEY, M.A., CARUSO, I.P. and SOUZA, F.P., 2018. Binding investigation between M2-1protein from HRSV and acetylated quercetin derivatives: 1H NMR, fluorescence spectroscopy, and molecular docking. *International Journal of Biological Macromolecules*, vol. 111, pp. 33-38. http://dx.doi.org/10.1016/j.ijbiomac.2017.12.141. PMid:29292149.

HAKOBYAN, A., ARABYAN, E., KOTSINYAN, A., KARALYAN, A., SAHAKYAN, H., ARAKELOV, V., NAZARYAN, K., FERREIRA, F. and ZAKARYAN, H., 2019. Inhibition of african swine fever virus infection by Genkwanin. *Antiviral Research*, vol. 167, pp. 78-82. http://dx.doi.org/10.1016/j.antiviral.2019.04.008. PMid:30991087.

HAO, M., LI, Y., WANG, Y. and ZHANG, S., 2011. A classification study of Respiratory Syncytial Virus (RSV) inhibitors by variable selection with random forest. *International Journal of Molecular Sciences*, vol. 12, no. 2, pp. 1259-1280. http://dx.doi.org/10.3390/ijms12021259. PMid:21541057.

JI, D., YE, E. and CHEN, H.F., 2015. Revealing the binding mode between respiratory syncytial virus fusion protein and benzimidazole-based inhibitors. *Molecular BioSystems*, vol. 11, no. 7, pp. 1857-1866. http://dx.doi.org/10.1039/C5MB00036J. PMid:25872614.

JIMÉNEZ-SOMARRIBAS, A., MAO, S., YOON, J.J., WEISSHAAR, M., COX, R.M., MARENGO, J.R., MITCHELL, D.G., MOREHOUSE, Z.P., YAN, D., SOLIS, I., LIOTTA, D.C., NATCHUS, M.G. and PLEMPER, R.K., 2017. Identification of non-nucleoside inhibitors of the respiratory syncytial virus polymerase complex. *Journal of Medicinal Chemistry*, vol. 60, no. 6, pp. 2305-2325. http://dx.doi.org/10.1021/acs.jmedchem.6b01568. PMid:28245119.

KAUL, T.N., MIDDLETON JUNIOR, E. and OGRA, P.L., 1985. Antiviral effect of flavonoids on human viruses. *Journal of Medical Virology*, vol. 15, no. 1, pp. 71-79. http://dx.doi.org/10.1002/jmv.1890150110. PMid:2981979.

KRARUP, A., TRUAN, D., FURMANOVA-HOLLENSTEIN, P., BOGAERT, L., BOUCHIER, P., BISSCHOP, I.J.M., WIDJOJOATMODJO, M.N., ZAHN, R., SCHUITEMAKER, H., MCLELLAN, J.S. and LANGEDIJK, J.P.M., 2015. A Highly stable prefusion RSV F vaccine derived from structural analysis of the fusion mechanism. *Nature Communications*, vol. 6, no. 1, pp. 8143. http://dx.doi.org/10.1038/ncomms9143. PMid:26333350.

KUMAR, S., JENA, L., MOHOD, K., DAF, S. and VARMA, A.K., 2015. Virtual screening for potential inhibitors of high-risk human papillomavirus 16 E6 protein. *Interdisciplinary Sciences, Computational Life Sciences*, vol. 7, no. 2, pp. 136-142. http://dx.doi.org/10.1007/s12539-015-0008-z. PMid:26199214.

LALANI, S. and POH, C.L., 2020. Flavonoids as antiviral agents for enterovirus A71 (EV-A71). *Viruses*, vol. 12, no. 2, pp. 184. http://dx.doi.org/10.3390/v12020184. PMid:32041232.

LEAL, C.M., LEITÃO, S.G., SAUSSET, R., MENDONÇA, S.C., NASCIMENTO, P.H.A., CAIO, C.F., ESTEVES, M.E.A., LEAL DA SILVA, M., GONDIM, T.S., MONTEIRO, M.E.S., TUCCI, A.R., FINTELMAN-RODRIGUES, N., SIQUEIRA, M.M., MIRANDA, M.D., COSTA, F.N., SIMAS, R.C. and LEITÃO, G.G., 2021. Flavonoids from siparuna cristata as potential inhibitors of SARS-CoV-2 replication. *Revista Brasileira de Farmacognosia*, vol. 31, no. 5, pp. 658-666. http://dx.doi.org/10.1007/s43450-021-00162-5. PMid:34305198.

LEE, W.P., LAN, K.L., LIAO, S.X., HUANG, Y.H., HOU, M.C. and LAN, K.H., 2018. Inhibitory effects of amentoflavone and orobol on daclatasvir-induced resistance-associated variants of hepatitis C virus. *The American Journal of Chinese Medicine*, vol. 46, no. 4, pp. 835-852. http://dx.doi.org/10.1142/S0192415X18500441. PMid:29737209.

LI, Y., LEUNG, K.T., YAO, F., OOI, L.S.M. and OOI, V.E.C., 2006. Antiviral flavans from the leaves of pithecellobium clypearia. *Journal of Natural Products*, vol. 69, no. 5, pp. 833-835. http://dx.doi.org/10.1021/np050498o. PMid:16724853.

LIKHITWITAYAWUID, K., CHAIWIRIYA, S., SRITULARAK, B. and LIPIPUN, V., 2006. Antiherpetic flavones from the heartwood of artocarpus gomezianus. *Chemistry & Biodiversity*, vol. 3, no. 10, pp. 1138-1143. http://dx.doi.org/10.1002/cbdv.200690115. PMid:17193228.

LIN, Y.L., SHEN, C.C., HUANG, Y.J. and CHANG, Y.Y., 2005. Homoflavonoids from ophioglossum petiolatum. *Journal of Natural Products*, vol. 68, no. 3, pp. 381-384. http://dx.doi.org/10.1021/np0401819. PMid:15787440.

LINDEN, D.E.J., HABES, I., JOHNSTON, S.J., LINDEN, S., TATINENI, R., SUBRAMANIAN, L., SORGER, B., HEALY, D. and GOEBEL, R., 2012. Real-time self-regulation of emotion networks in patients with depression. *PLoS One*, vol. 7, no. 6, pp. e38115. http://dx.doi.org/10.1371/journal.pone.0038115. PMid:22675513.

LIU, A.L., WANG, H.D., LEE, S.M.Y., WANG, Y.T. and DU, G.H., 2008. Structure-activity relationship of flavonoids as influenza virus neuraminidase inhibitors and their in vitro anti-viral activities. *Bioorganic & Medicinal Chemistry*, vol. 16, no. 15, pp. 7141-7147. http://dx.doi.org/10.1016/j.bmc.2008.06.049. PMid:18640042.

LOPES, B.R.O., COSTA, M.F., RIBEIRO, A.G., LIMA, C.S., CARUSO, I.P., ARAÚJO, G.C., KUBO, L.H., IACOVELLI, F., FALCONI, M., DESIDERI, A., OLIVEIRA, J., REGASINI, L.O., de SOUZA, F.P. and TOLEDO, K.A., 2020. Quercetin pentaacetate inhibits in vitro human respiratory syncytial virus adhesion. *Virus Research*, vol. 276, pp. 197805. http://dx.doi.org/10.1016/j.virusres.2019.197805. PMid:31712123.

MA, S.C., BUT, P.P.H., OOI, V.E.C., HE, Y.H., LEE, S.H.S., LEE, S.F. and LIN, R.C., 2001. Antiviral amentoflavone from selaginella sinensis. *Biological & Pharmaceutical Bulletin*, vol. 24, no. 3, pp. 311-312. http://dx.doi.org/10.1248/bpb.24.311. PMid:11256492.

MA, S.C., DU, J., BUT, P.P.H., DENG, S.L., ZHANG, Y.W., OOI, V.E.C., XU, H.X., LEE, S.H.S. and LEE, S.F., 2002. Antiviral chinese medicinal herbs against respiratory syncytial virus. *Journal of Ethnopharmacology*, vol. 79, no. 2, pp. 205-211. http://dx.doi.org/10.1016/S0378-8741(01)00389-0. PMid:11801383.

MANDLIK V., BEJUGAM, P.R. and SINGH, S., 2016. Application of artificial neural networks in modern drug discovery. In: M. PURI, ed. *Artificial neural network for drug design, delivery and disposition*. London: Academic Press.

NABAVI, S.M., ŠAMEC, D., TOMCZYK, M., MILELLA, L., RUSSO, D., HABTEMARIAM, S., SUNTAR, I., RASTRELLI, L., DAGLIA, M., XIAO, J., GIAMPIERI, F., BATTINO, M., SOBARZO-SANCHEZ, E., NABAVI, S.F., YOUSEFI, B., JEANDET, P., XU, S. and SHIROOIE, S., 2020. Flavonoid biosynthetic pathways in plants: versatile targets for metabolic engineering. *Biotechnology Advances*, vol. 38, pp. 107316. http://dx.doi.org/10.1016/j.biotechadv.2018.11.005. PMid:30458225.

NOOR, A. and KRILOV, L.R., 2018. Respiratory syncytial virus vaccine: where are we now and what comes next? *Expert Opinion on Biological Therapy*, vol. 18, no. 12, pp. 1247-1256. http://dx.doi.org/10.1080/14712598.2018.1544239. PMid:30426788.

NOUADI, B., EZAOUINE, A., MESSAL, M.E., BLAGHEN, M., BENNIS, F. and CHEGDANI, F., 2021. Prediction of anti-COVID 19 therapeutic power of medicinal moroccan plants using molecular docking. *Bioinformatics and Biology Insights*, vol. 15, pp. 11779322211009199. http://dx.doi.org/10.1177/11779322211009199. PMid:33888980.

O'BOYLE, N.M., BANCK, M., JAMES, C.A., MORLEY, C., VANDERMEERSCH, T. and HUTCHISON, G.R., 2011. Open Babel: an open chemical toolbox. *Journal of Cheminformatics*, vol. 3, no. 1, pp. 33. http://dx.doi.org/10.1186/1758-2946-3-33. PMid:21982300.

PAGEL J.F. and KIRSHTEIN, P., 2017. Neural networks: the hard and software logic. In: J.F. PAGEL and P. KRISHTEIN, eds. *Machine dreaming and consciousness*. London: Academic Press.

PERILLA, J.R., GOH, B.C., CASSIDY, C.K., LIU, B., BERNARDI, R.C., RUDACK, T., YU, H., WU, Z. and SCHULTEN, K., 2015. Molecular dynamics simulations of large macromolecular complexes. *Current Opinion in Structural Biology*, vol. 31, pp. 64-74. http://dx.doi.org/10.1016/j.sbi.2015.03.007. PMid:25845770.

ROBIN, V., IRURZUN, A., AMOROS, M., BOUSTIE, J. and CARRASCO, L., 2001. Antipoliovirus flavonoids from psiadia dentata. *Antiviral Chemistry & Chemotherapy*, vol. 12, no. 5, pp. 283-291. http://dx.doi.org/10.1177/095632020101200503. PMid:11900347.

ROSA, T.O. and LUZ, H., 2009. Conceitos básicos de algoritmos genéticos: teoria e prática. In: *XI Encontro de Estudantes de Informática do Tocantins* (pp. 27-37). Palmas: Centro Universitário Luterano de Palmas.

SCOTTI, L., MENDONÇA FILHO, F., ISHIKI, H., RIBEIRO, F., SINGLA, R., BARBOSA FILHO, J.M., SILVA, M. and SCOTTI, M., 2015. Docking studies for multi-target drugs. *Current Drug Targets*, vol. 18, no. 5, pp. 592-604. http://dx.doi.org/10.2174/1389450116666150825111818.

SHI, H., REN, K., LV, B., ZHANG, W., ZHAO, Y., TAN, R.X. and LI, E., 2016. Baicalin from scutellaria baicalensis blocks Respiratory Syncytial Virus (RSV) infection and reduces inflammatory cell infiltration and lung injury in mice. *Scientific Reports*, vol. 6, no. 1, pp. 35851. http://dx.doi.org/10.1038/srep35851. PMid:27767097.

SONG, M., GAO, M.H., HUANG, W.H., MAN-MEI, L., HUA, L., YAO-LAN, L., XIAO-QI, Z. and WEN-CAI, Y., 2016. Flavonoids from the seeds of hovenia acerba and their in vitro antiviral activity. *Journal of Pharmaceutical and Biomedical Sciences*, vol. 6, no. 6, pp. 401-409.

SUTARIYA, V., GROSHEV, A., SADANA, P., BHATIA, D. and PATHAK, Y., 2014. Artificial neural network in drug delivery and pharmaceutical research. *The Open Bioinformatics Journal*, vol. 7, no. 1, pp. 49-62. http://dx.doi.org/10.2174/1875036201307010049.

TEIXEIRA, T.S.P., CARUSO, I.P., LOPES, B.R.P., REGASINI, L.O., TOLEDO, K.A., FOSSEY, M.A. and SOUZA, F.P., 2017. Biophysical characterization of the interaction between M2-1 protein of HRSV and quercetin. *International Journal of Biological Macromolecules*, vol. 95, pp. 63-71. http://dx.doi.org/10.1016/j.ijbiomac.2016.11.033. PMid:27851930.

TEWTRAKUL, S., SUBHADHIRASAKUL, S., CHEENPRACHA, S. and KARALAI, C., 2007. HIV-1 protease and HIV-1 integrase inhibitory substances from eclipta prostrata. *Phytotherapy Research*, vol. 21, no. 11, pp. 1092-1095. http://dx.doi.org/10.1002/ptr.2252. PMid:17696192.

URIARTE-PUEYO, I. and CALVO, M.I., 2010. Structure-activity relationships of acetylated flavone glycosides from galeopsis Ladanum L. (Lamiaceae). *Food Chemistry*, vol. 120, no. 3, pp. 679-683. http://dx.doi.org/10.1016/j.foodchem.2009.10.060.

WALTHER, E., XU, Z., RICHTER, M., KIRCHMAIR, J., GRIENKE, U., ROLLINGER, J.M., KRUMBHOLZ, A., SALUZ, H.P., PFISTER, W., SAUERBREI, A. and SCHMIDTKE, M., 2016. Dual acting neuraminidase inhibitors open new opportunities to disrupt the lethal synergism between streptococcus pneumoniae and influenza virus. *Frontiers in Microbiology*, vol. 7, pp. 357. http://dx.doi.org/10.3389/fmicb.2016.00357. PMid:27047471.

WANG, G., DEVAL, J., HONG, J., DYATKINA, N., PRHAVC, M., TAYLOR, J., FUNG, A., JIN, Z., STEVENS, S.K., SEREBRYANY, V., LIU, J., ZHANG, Q., TAM, Y., CHANDA, S.M., SMITH, D.B., SYMONS, J.A., BLATT, L.M. and BEIGELMAN, L., 2015. Discovery of 4′-Chloromethyl-2′-Deoxy-3′,5′-Di- O -Isobutyryl-2′-Fluorocytidine (ALS-8176), a first-in-class RSV polymerase inhibitor for treatment of human respiratory syncytial virus infection. *Journal of Medicinal Chemistry*, vol. 58, no. 4, pp. 1862-1878. http://dx.doi.org/10.1021/jm5017279. PMid:25667954.

WANG, Y., CHEN, M., ZHANG, J., ZHANG, X.L., HUANG, X.J., WU, X., ZHANG, Q.W., LI, Y.L. and YE, W.C., 2012. Flavone C-Glycosides from the Leaves of Lophatherum Gracile and Their in Vitro Antiviral Activity. *Planta Medica*, vol. 78, no. 1, pp. 46-51. http://dx.doi.org/10.1055/s-0031-1280128. PMid:21870321.

WOUTERS, O.J., MCKEE, M. and LUYTEN, J., 2020. Estimated research and development investment needed to bring a new medicine to market, 2009-2018. *Journal of the American Medical Association*, vol. 323, no. 9, pp. 844-853. http://dx.doi.org/10.1001/jama.2020.1166. PMid:32125404.

XIA, C., LI, M.M., LI, Y.I. and WU, X., 2016. Study on anti-RSV activities and QSAR of natural caffeoylquinic acid derivates. *Zhong Yao Cai*, vol. 39, no. 2, pp. 383-388.

YU, M.S., LEE, J., LEE, J.M., KIM, Y., CHIN, Y.W., JEE, J.G., KEUM, Y.S. and JEONG, Y.J., 2012. Identification of myricetin and scutellarein as novel chemical inhibitors of the SARS coronavirus helicase, NsP13. *Bioorganic & Medicinal Chemistry Letters*, vol. 22, no. 12, pp. 4049-4054. http://dx.doi.org/10.1016/j.bmcl.2012.04.081. PMid:22578462.

Supplementary Material

Supplementary material accompanies this paper.

Table S1 – The 489 compounds used to train the AI.

Table S2 – Flavonoids unanimously classified as inactive by ANN.

Figure S1 – Confusion Matrices in training and simulation.

This material is available as part of the online article from https://doi.org/10.1590/1519-6984.270776