






Article

Machine Learning Analysis of Essential Oils from Cuban Plants: Potential Activity against Protozoa Parasites

Renata Priscila Barros de Menezes ¹, Luciana Scotti ¹, Marcus Tullius Scotti ^{1,*}, Jesús García ², Rosalia González ³, Lianet Monzote ^{4,*} and William N. Setzer ^{5,6,*}

- ¹ Post-Graduate Program in Natural and Synthetic Bioactive Products, Federal University of Paraíba, João Pessoa 58051-900, Brazil; renatabarros@ltf.ufpb.br (R.P.B.d.M.); luciana.scotti@gmail.com (L.S.)
- ² Pharmacy Department, Faculty of Natural and Exact Sciences, University of Oriente, Santiago de Cuba 90500, Cuba; jgadi1990@gmail.com
- ³ Toxicology and Biomedicine Centre (TOXIMED), University of Medical Science, Santiago de Cuba 90500, Cuba; rosaliagonzalez.9110@gmail.com
- ⁴ Parasitology Department, Institute of Tropical Medicine “Pedro Kouri”, Havana 10400, Cuba
- ⁵ Aromatic Plant Research Center, 230 N 1200 E, Suite 100, Lehi, UT 84043, USA
- ⁶ Department of Chemistry, University of Alabama in Huntsville, Huntsville, AL 35899, USA
- * Correspondence: mtscotti@ccae.ufpb.br (M.T.S.); monzote@ipk.sld.cu (L.M.); wsetzer@chemistry.uah.edu or wsetzer@aromaticplant.org (W.N.S.)

Abstract: Essential oils (EOs) are a mixture of chemical compounds with a long history of use in food, cosmetics, perfumes, agricultural and pharmaceuticals industries. The main object of this study was to find chemical patterns between 45 EOs and antiprotozoal activity (antiplasmodial, antileishmanial and antitrypanosomal), using different machine learning algorithms. In the analyses, 45 samples of EOs were included, using unsupervised Self-Organizing Maps (SOM) and supervised Random Forest (RF) methodologies. In the generated map, the hit rate was higher than 70% and the results demonstrate that it is possible find chemical patterns using a supervised and unsupervised machine learning approach. A total of 20 compounds were identified (19 are terpenes and one sulfur-containing compound), which was compared with literature reports. These models can be used to investigate and screen for bioactivity of EOs that have antiprotozoal activity more effectively and with less time and financial cost.

Keywords: essential oil; Cuban plants; machine learning analysis; antiprotozoal activity



Citation: Barros de Menezes, R.P.; Scotti, L.; Scotti, M.T.; García, J.; González, R.; Monzote, L.; Setzer, W.N. Machine Learning Analysis of Essential Oils from Cuban Plants: Potential Activity against Protozoa Parasites. *Molecules* **2022**, *27*, 1366. <https://doi.org/10.3390/molecules27041366>

Academic Editors:

Francesca Mancianti and Valentina Virginia Ebani

Received: 16 December 2021

Accepted: 15 February 2022

Published: 17 February 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

An essential oil (EO) is a concentrated plant secondary metabolite composed of a mixture of volatile chemical compounds [1], with a long history of use in food, cosmetics, perfumes, agricultural and pharmaceuticals industries [2]. In the last decade, almost 5000 articles related to uses of EOs have been published, with a positive increment of more than 7% per year [3]. In this scenario, the scientific, economic, and biological importance of EOs is growing as alternatives to synthetic compounds commonly used in industry [4–6].

In particular, numerous studies demonstrated the wide pharmacological spectrum of EOs, including: antimicrobial [7,8], antifungal [9], antiparasitic [10], antiviral [10,11], insecticidal [12], anticarcinogenic [13–16], immunomodulatory [17], anti-inflammatory and antioxidant [14]. Nevertheless, the chemical composition of obtained EOs can unfortunately be different depending on the chosen method, geographical origins, the season, the type of soil and the agricultural conditions in which plants have grown. Thus, the same plant could produce different EO chemical composition profiles and therefore display different biological effects [3,8,14]. In this sense, some approaches have been used, such as the ‘chemotype concept’ with the aim to discern the bioactivity of EOs based on chemical profile. However, more complex questions have emerged due to the high complexity in the

chemical composition of EOs and interaction among constituents. Then, the role of each single constituent and synergistic/antagonist effects among components remain unclear in many potential EOs. During the last decade, computational studies using EOs have been reported in the continuous search for new therapeutic drugs or lead compounds. In particular, machine learning analysis has been used to identify new structures from EOs with potential antibacterial [18] and antiviral [3] activities.

Recently, potential EOs from Cuban plants was reviewed, which a variable number of components (terpenes, aliphatic derivatives, sulfur-containing compound, phenylpropanoids, alkaloids and amine-type compounds) and different biological activities (antiprotozoal, antibacterial, antifungal, anticancer, anthelmintic, larvicidal and insecticidal) were identified. However, correlation of potentialities of these EOs with chemical entities could not be linked due to scarce number of pure compounds that were tested and experiments related to synergistic/antagonistic effects were not found [19]. Then, in line with those previous studies (which common compounds were identified from different EOs with multiple biological effects), herein an extensive study of those EOs from Cuban plants is reported, with the main goal of finding chemical patterns between 45 samples and antiprotozoal activity, using unsupervised and supervised machine learning, which are Self-Organizing Maps (SOM) and Random Forest (RF), respectively.

Kohonen's self-organizing maps [20] are neural networks of the unsupervised type, based on the functioning principle of the central nervous system of animals [21–24]. SOM has competitive learning, with only one output neuron or local group of neurons that provides the final response to a current input signal. The data presented in the input neuron are mapped to the defined space of neurons with iterations and weight adjustments in a typically two-dimensional array (Kohonen map). The most common way to demonstrate the similarity in SOM is through the Euclidean distance between the vectors and the input data vector [21–24].

The U-matrix, created by Vesanto and collaborators [24], is used to visualize the SOM. This matrix allows the visualization and discrimination of the groups generated in the SOM, from the Euclidean distances. This degree of similarity is plotted in the third dimension generating a 3-D relief surface, in this way the clusters are represented in the form of "depressions", "valleys", and "peaks". The "depressions" and "valleys" of the relief surface of the U-matrix represent neurons belonging to the same cluster, while neurons that have a great distance from the adjacent neuron are represented by "peaks", they are cluster-discriminating neurons.

RF is an algorithm that will create several decision trees at random, thus obtaining a forest where each tree will be used in the result. It is a robust and complex algorithm, which can lead to a higher computational cost compared to others. A decision tree establishes rules for decision making, that is, the algorithm will generate a structure like a flowchart with "nodes" where a condition will be checked and if met, the flow follows one branch, otherwise, it follows another, always leading to the nearest "node" where further decision-making will take place, until the end of the tree. Thus, given a training set, the algorithm will analyze the data and look for the best conditions and where to insert each data into the flow [25–28].

In the literature, previous studies using these methods have been applied to EOs, which have been useful to select antiviral and low toxic samples [3], antibiofilm formation by *Staphylococcus aureus* [16,29], *S. epidermidis* [16] and *Pseudomonas aeruginosa* [18,30], as different biological activities such as antiviral, anthelmintic, anti-inflammatory, anticancer, antioxidant, antimicrobial, antifungal and cytotoxic activity [31]. However, the application on EOs with antiprotozoal activity has been scarcely documented. The study performed herein, demonstrated how multidisciplinary applications involving machine learning could represent a valuable tool in predicting the bioactive component in complex mixtures.

2. Results and Discussion

As previously reported, the analyses of 45 samples of EOs were included, which were obtained from 16 families, 33 species, and 408 different identified compounds [19]. Figure 1 represents the major compound class, as the components identified as main compounds in the different studied EOs previously reported [17].

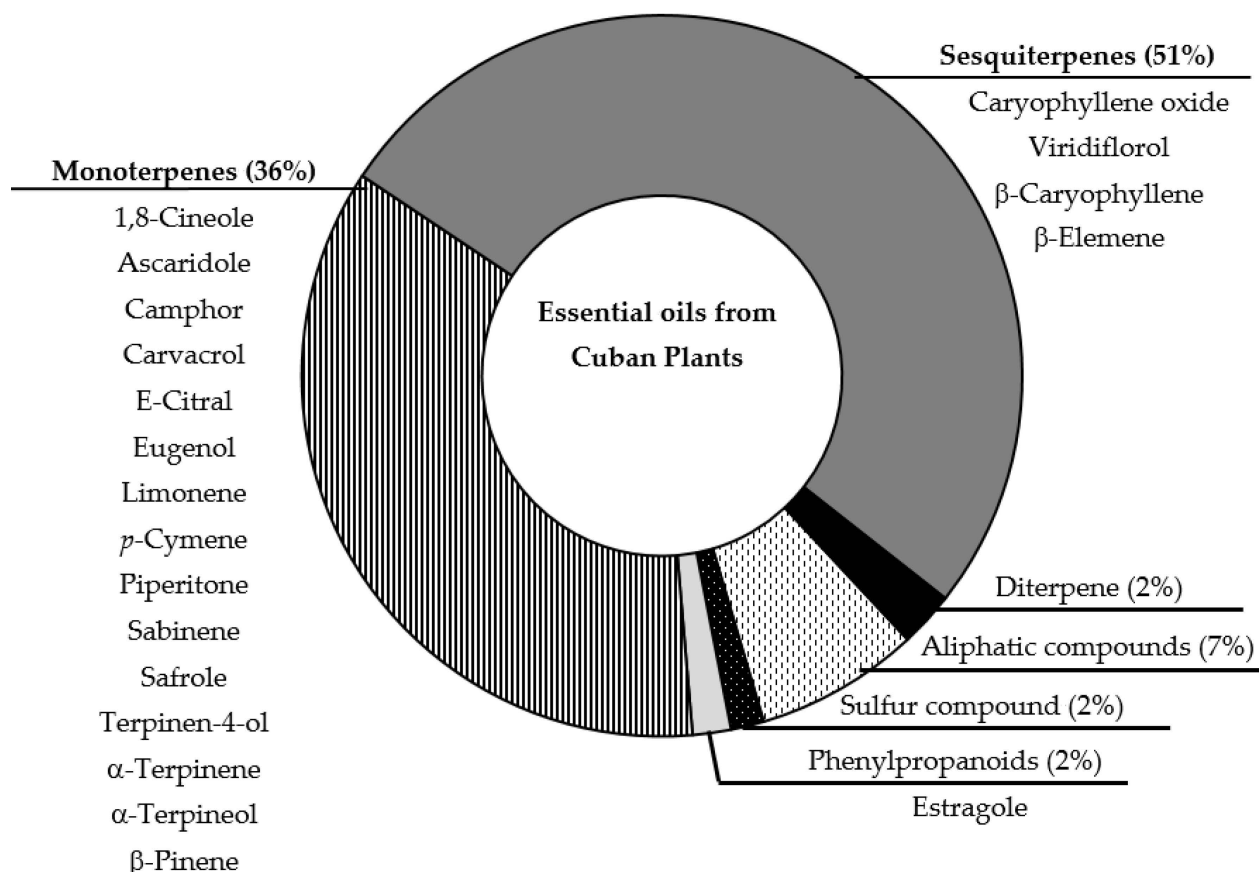


Figure 1. Schematic representation of chemical composition of analyzed EOs. Graphic represent the distribution of total components described for EOs; while list of compounds corresponds to main compounds identified in the EOs.

The dataset of 45 EOs was analyzed to find a chemical pattern between the EOs and antiprotozoal activity, including three activities: antiplasmodial, antileishmanial and antitrypanosomal. The analysis started with the use of the SOM, where the chemical composition of the 45 EOs was used as information to find patterns with the antiparasitic activity. Among them, 21 had some of the analyzed activities (antiplasmodial, antileishmanial and antitrypanosomal), with median inhibitory concentrations (IC_{50}) in in-vitro cultures $< 100 \mu\text{g/mL}$. The remaining EOs (24) had no antiprotozoal activity or activity had not been reported.

In the generated map, the hit rate was higher than 84%. The SOM validation was then performed using the 5-fold external cross-validation technique [32,33]; this means that the entire dataset is partitioned five times into a modeling set (training set) including 80% of the compounds in the set, and the external cross-validation data set, comprising the remaining 20% of the compounds in the data set. After this, only the modeling set is used to build the models and then the models are validated with the external cross-validation technique. In this sense, the dataset was subdivided into five training groups and five test groups, always keeping the ratio between active and not reported EOs. The validation results are described in Table 1.

Table 1. Accuracy statistics of the training and tests groups of the 5-fold external cross-validation of the Self-Organizing map (SOM).

Classification of EOs	Training					Average
	1	2	3	4	5	
True positive rate	0.90	0.95	0.99	0.90	0.80	0.91
True negative rate	0.70	0.65	0.68	0.70	0.75	0.71
Accuracy	0.81	0.80	0.83	0.81	0.78	0.81
Classification of EOs	Test					Average
	1	2	3	4	5	
True positive rate	0.60	0.60	0.95	0.60	0.99	0.75
True negative rate	0.75	0.75	0.70	0.95	0.80	0.80
Accuracy	0.67	0.67	0.83	0.77	0.90	0.78

Analyzing Table 1, we see that the hit rate for true positive rate (EOs active) and true negative rate (EOs that did not display antiprotozoal activity or had not been reported) both in training sets and in test sets were higher than 0.7, showing that the SOM model is robust. Model accuracy assessment gives information about the overall performance of the model, indicating the overall hit rate. The hit rate is the rate that evaluates how well the model correctly classified the EOs. Accuracy values vary between 0 and 1. Models with accuracy rate closer to 1 represents the higher model's hit rate; while an accuracy rate equal to or greater than 0.7 is considered models of optimal performance [23,27].

The SOM managed to find a chemical pattern between the chemical composition of EOs and antiprotozoal activity. In parallel, we chose to check if this chemical pattern is also found by using a supervised algorithm, known as RF. The RF model was generated using the 5-fold external cross-validation technique [32,33]; this means that the entire data set is partitioned five times into a modeling set (training set) including 80% of the compounds the set, and the external cross validation data set, comprising the remaining 20% of the compounds the data set. After this, only the modeling set is used to build the models and then the models are validated with the external cross validation technique. Its performance was evaluated through the statistics such as specificity, sensitivity, which obtained satisfactory values that corroborate the accuracy of the superior model, at 70%. The performances can be observed in Table 2, these parameters are an average between the five models. During the creation of the model, we also observed the domain of applicability to ensure that the samples tested were within the chemical space of each model.

Table 2. Summary of the statistics parameters of the RF model (average between the five models).

Model	Specificity	Sensitivity	Accuracy	PPV	NPV
RF	0.83	0.65	0.71	0.75	0.70

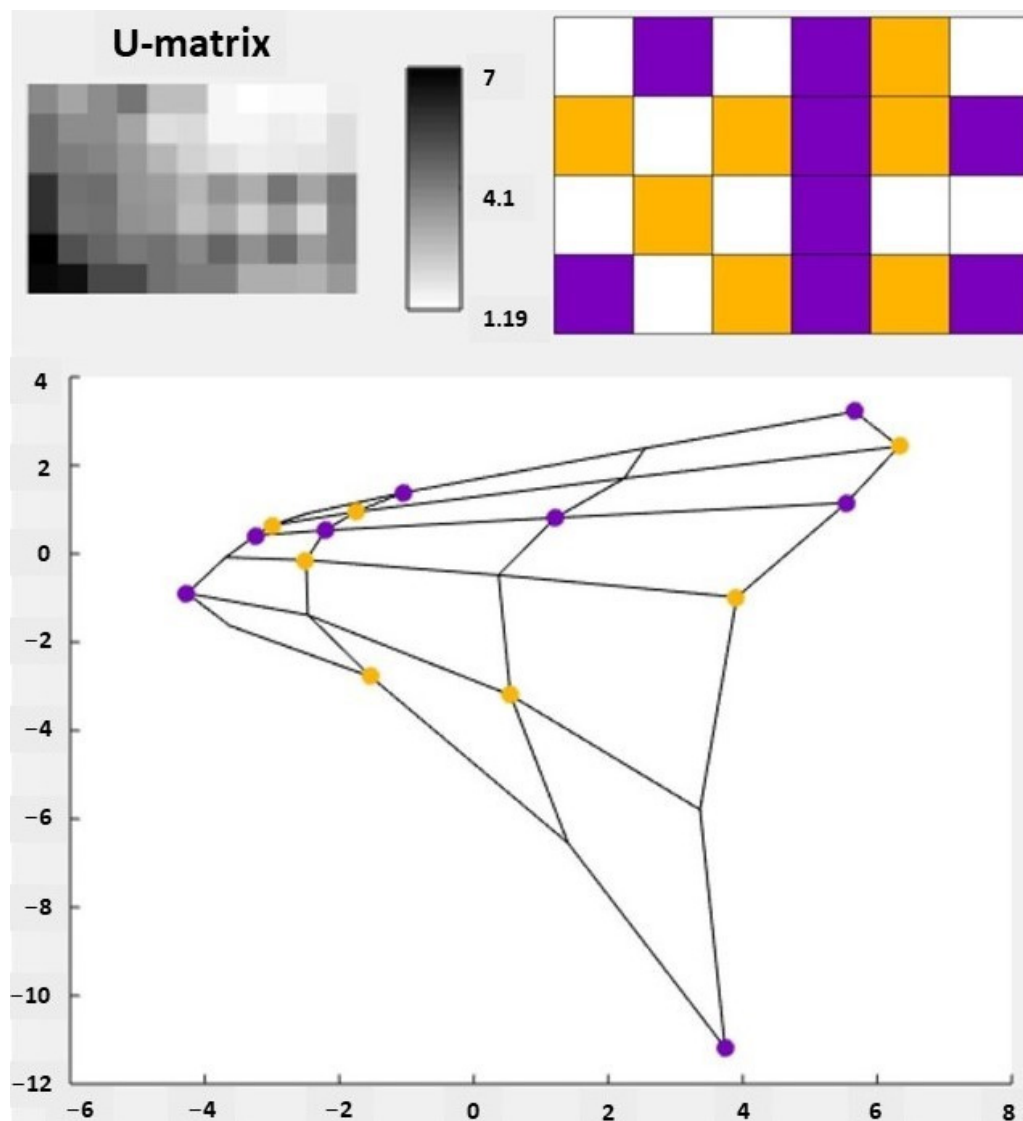
RF is random forest, PPV is positive predictive value, and NPV is negative predictive value.

In Table 3, we can see the accuracy and the global hit of both models and show that the chemical same pattern could be obtained use an unsupervised (SOM) and supervised (RF) machine learning. In both analyses, an accuracy rates higher than 0.7 were appreciated.

Figure 2 shows the U-matrix of the SOM, i.e., the visual analysis of the SOM. The U-matrix is constructed by measuring the Euclidean distance in the vector space between adjacent neurons [21,24,34]. It is possible to normalize the distances to be represented by colors or in shades of gray [21,24]. What is represented in the U-Matrix are the clusters mapped by the SOM and not the individual samples.

Table 3. Summary of test averages corresponding to 5-fold cross-validation using the different machine learning algorithms, self-organizing maps (SOM) and random forest (RF).

Classification of EOs	Average	
	SOM (Unsupervised)	RF (Supervised)
Active	0.85	0.75
Not reported	0.83	0.70
Accuracy	0.84	0.71

**Figure 2.** Visualization of the self-organizing map (SOM) of essential oil (EO) data. In the upper corner we have the U-matrix. The left U-matrix does not identify the activities of the EOs while the right U-matrix identifies those activities by color: active is violet and yellow represents samples with no antiprotozoal activity or that have not been reported. The values shown on the scale between the two U-matrices represent the values of the percent of molecules present in the EOs, varying between 1.19 and 7. These values were used to group the EOs by activities. At the bottom, we have the principal component analysis (PCA) projection of the SOM measured by its two eigenvectors with higher eigenvalues. The activities were plotted using the same identification colors as the U-matrix.

Forty-five EOs were used for the SOM analysis. After mapping the SOM, the 45 EOs were correctly grouped into active and inactive (EOs that do not display antiprotozoal activity or have not been reported). There was also the separation of groups of greater similarity and difference between them, taking into account the chemical composition of the EOs, which were approximated or distanced in the SOM. Thus, in the U-matrix, each square represents a group of EOs that are organized both by activity and chemical similarity, with the purple ones relating to active EOs and the yellow ones to inactive EOs.

It is also worth noting that the U-matrix is a visual representation of the topological mapping of the SOM, in this way, the white squares are valleys that separate the clusters that were generated.

It is also possible to observe in Figure 2 the principal component analysis (PCA) graph, which was generated from the correlation matrix of the EOs dataset used in the generation of the SOM. PCA is used to reduce the dimensionality of the data and allow a better visualization of the clusters, since it allows representing the input data as linear combinations of their projections [23]. The PCA performed in this study has an explained variance of 25.47%, that is, using only two variables it is possible to explain a quarter of the entire variance.

While in the U-matrix we have the white squares representing valleys that distance the clusters, in the PCA graph neighboring cartographic units are connected by lines to make the map view clearer and more defined [23].

After the general analysis with the 45 EOs, the SOM was constructed considering the chemical patterns of each sample. The most significant molecules for the chemical pattern separation of active and not reported EOs obtained with SOM Toolbox tool are shown in Figure 3. In this sense, 20 compounds present in the EOs were associated with at least one of these three biological activities, of which 19 are terpenes (10 monoterpenes and 9 sesquiterpenes) and one sulfur-containing compound. As is evident, a high predominance of terpene-type compounds was observed. Previously, the role of terpene compounds has been reviewed, suggesting the promising therapeutic value against protozoa parasites [35–37].

The identification of the most significant molecules is made by observing the region in the U-matrix of active EOs. Once the region was identified, we observed the most expressive molecules in that region. For example, when analyzing the U-matrix, we observe that in the lower right corner, there is a region in purple color, indicating a region of active EOs. Following the analysis, we will observe which molecules are most representative of that region; thus, we have the molecules (*E*)- β -ocimene, (*Z*)- β -ocimene and β -phellandrene. Note, in Figure 3, that the individual matrices of these molecules indicate their greater presence in the lower right region, the region of active EOs.

In a general comparison of listed components between Figures 1 and 3, note that only three compounds match as major component of EOs and as significant molecules generated by SOM strategy: camphor, piperitone and safrole. In general, pharmacological studies of EOs suggest that major identified components could be responsible for the biological activity. However, some studies did not correlate the main compound with the antiprotozoal effect [38–40]. Thus, using the present model, we selected molecules present in EOs that can influence in the antiprotozoal activity of studied EOs, and could suggest other EOs based in the complete chemical composition and not only in the major components. In addition, it is interesting to specify that in the used data, camphor was identified in 5 samples with concentrations between 0.1 to 17.1%, piperitone was present in 7 samples ranging from 0.1 to 23.7%, and safrole was documented in 3 samples from 1.6 to 71.8% [17]. In regard to antiprotozoal activity, analysis of the samples with these compounds with concentrations higher than 5%, we note that, for example, camphor was reported in the EOs from *Piper aduncum* L. and *Piper aduncum* var. *ossanum* (C.DC.) Trel. that showed antiplasmodial, antileishmanial and antitrypanosomal activity, as well as piperitone. Safrole, in contrast, was identified in *Piper auritum* Kunt as major compound and displayed antileishmanial activity [17]. These examples could corroborate the observed

results from the SOM analysis and probably could highlight *Piper* as a promising genus to study antiprotozoal properties, related with the main compounds or synergism resulting from the presence of these components in this genus. In fact, antileishmanial potentialities of the *Piper* genus was recently reviewed [41].

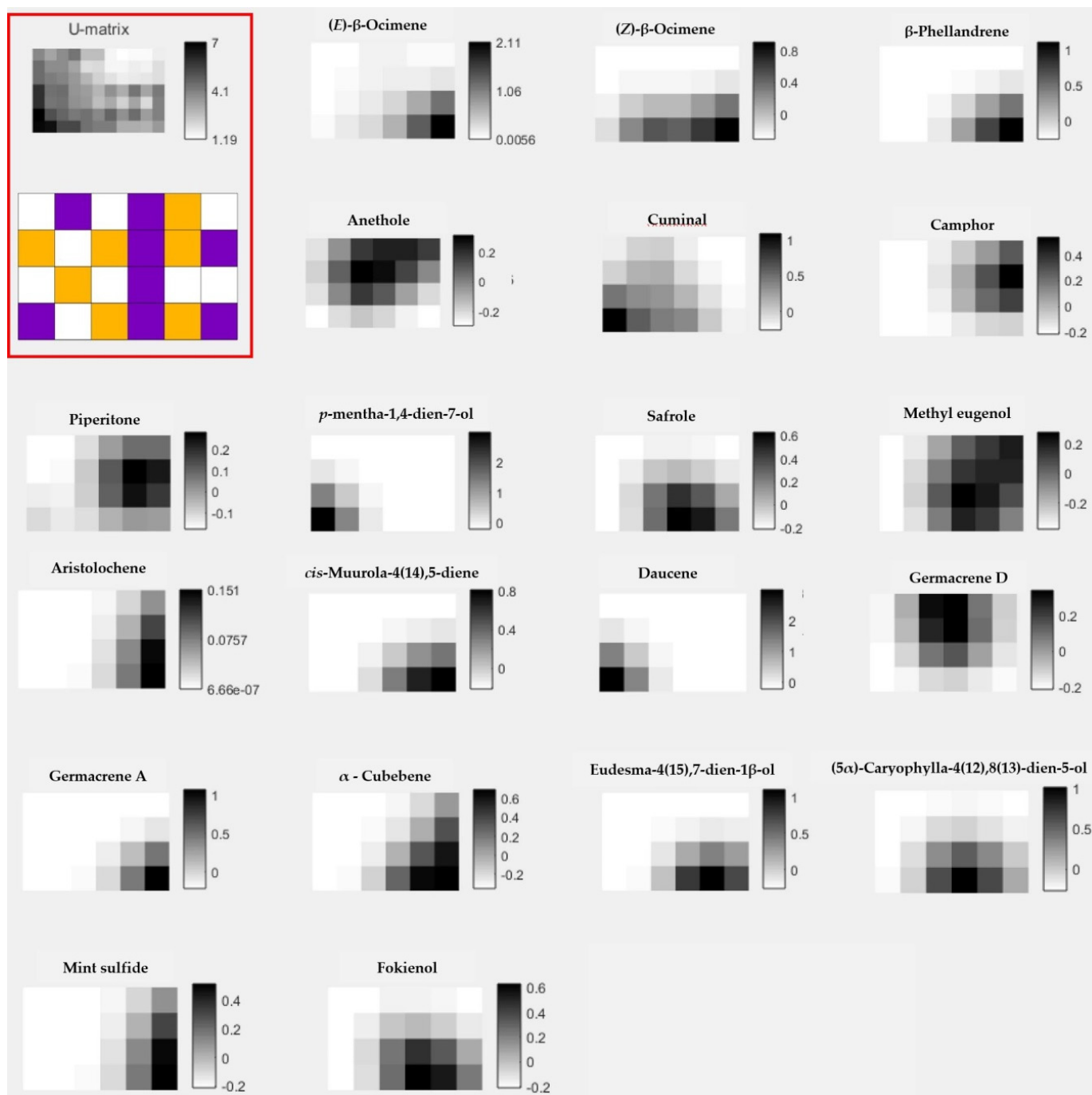


Figure 3. The most significant molecules for the EOs by activity. In the upper left we have the U-matrix of the self-organizing map generated in the study, with the upper U-matrix not identifying the tribes and the lower U-matrix identifying the tribes by color; active is violet and yellow represents samples with no antiprotozoal activity or that have not been reported.

A quick literature search in Pubmed Electronic Database was carried out with 20 identified compounds. The most important results confirm that: β -ocimene and safrole have shown activity against *Trypanosoma brucei* with IC_{50} values of 1.1 $\mu\text{g}/\text{mL}$ [42] and

18.4 µg/mL [43], respectively; while methyleugenol had an IC₅₀ of 5.7 µg/mL against *Plasmodium falciparum* [44]. However, it was noted that several of the identified compounds were not evaluated against these protozoa parasites, which could be addressed in further screening assays.

Nevertheless, several studies in the literature have already confirmed the antiparasitic action of EOs with identified compounds obtained from plants in other geographical locations, which is summarized in Table 4 together with results of EOs from Cuban plants (supplementary material). The higher number of reports from Cuban and other EOs was found for camphor. For example, antiprotozoal activity was evaluated for EOs from Cuban plants against *Plasmodium falciparum*, *Leishmania* spp. and *Trypanosoma* spp. from *Alpinia zerumbet* (Pers.) B. L. Burtt & R. M. Smith [45], *Piper aduncum* L. [46] and *Piper ossanum* (C.DC.) Trel [47]; while the rest of EOs displayed activity only against kinetoplastid parasites from *Alpinia speciosa* K. Schum. [39], *Artemisia absinthium* L. [42], *Piper cubeba* L. [48], and *Thymus hirtus* sp. *algeriensis* Boiss. et Reut [30], which camphor proved to be one of the major substances in all included samples.

However, although in the literature, piperitone was only found in an EO from Benin with antitrypanosomal and antiplasmodial activity [49], in Cuban samples, it was found in higher concentrations of EOs (19 to 24%) that showed a broad spectrum of antiprotozoal effects mainly from *Piper* species [46,47]. In contrast, a diverse number of studies from worldwide plants, EOs with germacrene D and with antikinetoplastid activity correlated with antiplasmodial activity shown by Cuban EOs with this compound [50].

Table 4. In vitro antiprotozoal profile of Cuban and according literature review of EOs that present identified compounds in this study (previous shown in Figure 3).

Compound	Country	Plant	Compound %	Targeted Protozoa (Result)	Ref.
(E)-β-Ocimene	Brazil	<i>Annona vepretorum</i> Mart.	6.8%	<i>Trypanosoma cruzi</i> (IC ₅₀ = 32 µg/mL)	[51]
		<i>Syzygium cumini</i> (L.) Skeels.	11.7%	<i>Leishmania amazonensis</i> (IC ₅₀ = 60 mg/L)	[52]
		<i>Xylopia frutescens</i> Aubl.	6.8%	<i>Trypanosoma cruzi</i> (IC ₅₀ = 15 to 30 µg/mL)	[53]
	Cuba *	<i>Bursera graveolens</i> Triana & Planch	13%	<i>Leishmania amazonensis</i> (IC ₅₀ = 36.7 mg/L)	[54]
		<i>Piper auritum</i> Kunt	0.49%	<i>Leishmania major</i> (IC ₅₀ = 29.1 µg/mL) <i>Leishmania mexicana</i> (IC ₅₀ = 63.3 µg/mL) <i>Leishmania braziliensis</i> (IC ₅₀ = 52.1 µg/mL) <i>Leishmania donovani</i> (IC ₅₀ = 12.8 µg/mL)	[55]
(Z)-β-Ocimene	Brazil	<i>Syzygium cumini</i> (L.) Skeels.	29%	<i>Leishmania amazonensis</i> (60 mg/L)	[53]
	Cuba *	<i>Bursera graveolens</i> Triana & Planch	0.9%	<i>Leishmania amazonensis</i> (IC ₅₀ = 36.7 mg/L)	[54]
		<i>Piper ossanum</i> (C.DC.) Trel	0.14%	<i>Plasmodium falciparum</i> (IC ₅₀ = 1.5 µg/mL) <i>Trypanosoma brucei</i> (IC ₅₀ = 8.1 µg/mL) <i>Trypanosoma cruzi</i> (IC ₅₀ = 8.0 µg/mL) <i>Leishmania amazonensis</i> (IC ₅₀ = 19.3 µg/mL)	[47]
β-Phellandrene	Cameroon	<i>Ocimum gratissimum</i> L.	21.1%	<i>Plasmodium berghei</i> (at 200, 300 and 500 mg/kg caused a suppression of parasitaemia of 55.0%, 75.2% and 77.8%, respectively)	[56]
	Cuba *	<i>Bixa orellana</i> L.	0.2%	<i>Leishmania amazonensis</i> (8.5 mg/L)	[57]
		<i>Piper ossanum</i> (C.DC.) Trel	2.1%	<i>Plasmodium falciparum</i> (IC ₅₀ = 2.8 µg/mL) <i>Trypanosoma brucei</i> (IC ₅₀ = 8.4 µg/mL) <i>Trypanosoma cruzi</i> (IC ₅₀ = 8.6 µg/mL)	[47]

Table 4. Cont.

Compound	Country	Plant	Compound %	Targeted Protozoa (Result)	Ref.	
Camphor	Brazil	<i>Alpinia speciosa</i> K. Schum	17.1%	<i>Trypanosoma cruzi</i> (IC ₅₀ = 92 µg/mL) <i>Leishmania brasiliensis</i> (IC ₅₀ = 67 µg/mL)	[58]	
		<i>Ocotea odorifera</i> (Vell) Rohwer	6.5%	<i>Leishmania amazonensis</i> (IC ₅₀ = 11.7 µg/mL)	[59]	
		<i>Piper cubeba</i> L	5.6%	<i>Trypanosoma cruzi</i> (IC ₅₀ = 87.9 µg/mL)	[40]	
	Cuba *	<i>Alpinia zerumbet</i> (Pers.) B.L.Burt & R.M.Smith	0.1%	<i>Plasmodium falciparum</i> (IC ₅₀ = 66.2 µg/mL)	[45]	
		<i>Piper aduncum</i> L.	17.1%	<i>Leishmania amazonensis</i> (IC ₅₀ = 23.8 µg/mL) <i>Leishmania donovani</i> (IC ₅₀ = 7.7 µg/mL) <i>Leishmania infantum</i> (IC ₅₀ = 8.1 µg/mL)	[46]	
		<i>Piper ossanum</i> (C.DC.) Trel	13.8 and 9.4%	<i>Plasmodium falciparum</i> (IC ₅₀ = 1.5 and 2.8 µg/mL) <i>Trypanosoma brucei</i> (IC ₅₀ = 8.1 and 8.4 µg/mL) <i>Trypanosoma cruzi</i> (IC ₅₀ = 8.0 and 8.6 µg/mL)	[47]	
	Ethiopia	<i>Artemisia absinthium</i> L.	27.4%	<i>Leishmania aethiopica</i> (IC ₅₀ = 8 µg/mL) <i>Leishmania donovani</i> (IC ₅₀ = 42 µg/mL)	[60]	
	Morocco	<i>Rosmarinus officinalis</i> L.	18.7%	<i>Leishmania major</i> (IC ₅₀ = 1.2 µg/mL)	[61]	
	Spain	<i>Artemisia absinthium</i> L.	4.5%	<i>Trypanosoma cruzi</i> (84% of inhibition at 200 µg/mL)	[62]	
		<i>Artemisia pedemontana</i> subsp. <i>assoana</i> (Willk.) Rivas Mart.	7.7%	<i>Trypanosoma cruzi</i> (20 to 70% of inhibition at 200 µg/mL)	[62]	
	Turkey	<i>Salvia recognita</i> Fisch. & Meyer	42%	<i>Plasmodium falciparum</i> (IC ₅₀ = 17 to 12 µg/mL)	[63]	
	Tunisia	<i>Thymus hirtus</i> sp. <i>algeriensis</i> Boiss. et Reut	13.8%	<i>Leishmania major</i> (IC ₅₀ = 0.43 µg/mL) <i>Leishmania infantum</i> (IC ₅₀ = 0.25 µg/mL)	[39]	
Germacrene D	Brazil	<i>Casearia sylvestris</i> Sw.	19.6%	<i>Leishmania amazonensis</i> (IC ₅₀ = 24.2 µg/mL)	[64]	
		<i>Eugenia gracillima</i> Kiaersk.	16.1%	<i>Leishmania braziliensis</i> (IC ₅₀ = 74.6 µg/mL) <i>Leishmania infantum</i> (IC ₅₀ = 80.4 µg/mL)	[65]	
		<i>Guatteria australis</i> A. St.-Hil.	22.2%	<i>Leishmania infantum</i> (IC ₅₀ = 30.7 µg/mL)	[66]	
		<i>Lantana camara</i> L.	11.7%	<i>Trypanosoma cruzi</i> (IC ₅₀ = 201.94 µg/mL) <i>Leishmania braziliensis</i> (IC ₅₀ = 72.31 µg/mL)	[67]	
		<i>Melampodium divaricatum</i> (Rich. ex Rich.) DC.	12.7%	<i>Leishmania amazonensis</i> (IC ₅₀ = 10.7 µg/mL)	[64]	
		<i>Piper cernuum</i> Vell.	12.7%	<i>Leishmania amazonensis</i> (Infection index of 115 at 10 µg/mL)	[68]	
		<i>Piper duckei</i> C. DC.	14.7%	<i>Leishmania amazonensis</i> (IC ₅₀ = 42–46 µg/mL) <i>Leishmania guyanensis</i> (IC ₅₀ = 15.2 µg/mL)	[69]	
		<i>Vernonia polyanthes</i> (Spreng.) Vega & M. Dematteis	4.3%	<i>Leishmania infantum</i> (IC ₅₀ = 19.4 µg/mL)	[70]	
			<i>Xylopi frutescens</i> Aubl.	17.8%	<i>Trypanosoma cruzi</i> (IC ₅₀ = 15 to 30 µg/mL)	[53]
		Cuba *	<i>Tagetes lucida</i> Cav.	0.3%	<i>Plasmodium berghei</i> (IC ₅₀ = 72 µg/mL)	[50]

Table 4. Cont.

Compound	Country	Plant	Compound %	Targeted Protozoa (Result)	Ref.
Methyl eugenol	Brazil	<i>Aniba canelilla</i> (H.B.K.) Mez	14.8%	<i>Trypanosoma evansi</i> (Growth inhibition at 0.5%, 1.0% and 2.0%)	[71]
		<i>Hyponia salzmannii</i> (Benth.) Harley.	5.6%	<i>Trypanosoma cruzi</i> (IC ₅₀ = 35–42 µg/mL)	[72]
	Cuba *	<i>Piper auritum</i> Kunt	0.6%	<i>Leishmania major</i> (IC ₅₀ = 29.1 µg/mL) <i>Leishmania mexicana</i> (IC ₅₀ = 63.3 µg/mL) <i>Leishmania braziliensis</i> (IC ₅₀ = 52.1 µg/mL) <i>Leishmania donovani</i> (IC ₅₀ = 12.8 µg/mL)	[55]
	Bhutan	<i>Pleurospermum amabile</i> W.W.Smith,	3.8%	<i>Plasmodium falciparum</i> (IC ₅₀ = 79 µg/mL)	[73]
Piperitone	Benin	<i>Cymbopogon schoenanthus</i> Spreng.	60.3%	<i>Trypanosoma brucei</i> (IC ₅₀ = 2.1 µg/mL) <i>Plasmodium falciparum</i> (IC ₅₀ = 43.1 µg/mL)	[49]
	Cuba *	<i>Alpinia zerumbet</i> (Pers.)B.L.Burtt&R.M.Smith	0.1%	<i>Plasmodium falciparum</i> (IC ₅₀ = 71.4 µg/mL)	[45]
		<i>Bursera graveolens</i> Triana & Planch	0.1%	<i>Leishmania amazonensis</i> (IC ₅₀ = 36.7 mg/L)	[54]
		<i>Piper aduncum</i> L.	23.7%	<i>Leishmania amazonensis</i> (IC ₅₀ = 23.8 µg/mL) <i>Leishmania donovani</i> (IC ₅₀ = 7.7 µg/mL) <i>Leishmania infantum</i> (IC ₅₀ = 8.1 µg/mL)	[46]
		<i>Piper ossanum</i> (C.DC.) Trel	20.1 and 19.0%	<i>Plasmodium falciparum</i> (IC ₅₀ = 1.5 and 2.8 µg/mL) <i>Trypanosoma brucei</i> (IC ₅₀ = 8.1 and 8.4 µg/mL) <i>Trypanosoma cruzi</i> (IC ₅₀ = 8.0 and 8.6 µg/mL)	[47]
Safrole	Brazil	<i>Myroxylon peruiferum</i> L.f.	8.3%	<i>Leishmania amazonensis</i> (IC ₅₀ = 54–162 µg/mL)	[74]
		<i>Ocotea odorifera</i> (Vell) Rohwer	6.5%	<i>Leishmania amazonensis</i> (IC ₅₀ = 11.7 µg/mL)	[59]
	Cuba *	<i>Piper auritum</i> Kunt	71.8%	<i>Leishmania major</i> (IC ₅₀ = 29.1 µg/mL) <i>Leishmania mexicana</i> (IC ₅₀ = 63.3 µg/mL) <i>Leishmania braziliensis</i> (IC ₅₀ = 52.1 µg/mL) <i>Leishmania donovani</i> (IC ₅₀ = 12.8 µg/mL)	[75]

* Data of EOs from Cuba used in this study.

3. Materials and Methods

3.1. Essential Oils Database

The 31 articles previously selected and analyzed by Monzote et al. [19] were used. A database with identified compounds with a concentration > 0.1% were performed and stored in Excel spreadsheet, which traces were not included. In parallel, the described pharmacological properties to each EO were assigned.

3.2. Machine Learning Analysis

3.2.1. Self-Organizing Maps (SOMs)

The database contained information on 45 essential oils, with chemical composition and biological activity. For the realization of the neural maps, the information of the composition of the each EO from the dataset was used like descriptors. The chemical components were analyzed with SOMs in Matlab 6.5 and SOM Toolbox 2.0 [24]. The SOM Toolbox tool is a set of Matlab functions that can be used for the elaboration and implementation of neural networks, since it contains functions for the creation, visualization, and analysis of SOMs.

The data set was presented to the network before any adjustments were made. Subsequently, the data group was partitioned according to the regions of the weight vectors

of the map, in each training stage. Then, the correct prediction of these sets and the total correct predictions of the compounds were evaluated. In the most relevant models, the set was divided into training and test sets to assess the forecasting capacity. Training and test performance were assessed by calculating the proportion of the number of samples correctly classified by SOM. For each map, 5 cross-validations were performed, being partitioned into 80% training and 20% testing. In the SOM, sites containing molecules for each descriptor were identified to highlight existing chemical patterns. The SOM was generated with a 4×6 rectangular GRID.

3.2.2. Principal Component Analysis (PCA)

PCA analysis was calculated using the SOM toolbox 2.0 [24]. The utilization of PCA for dimension reduction lies in the fact that the PCs are generated so that they explain maximal amounts of variance [27].

The PCA was calculated using the database contained information on 45 essential oils.

3.2.3. Random Forest Model

Knime 4.4.1 software (KNIME 4.4.1 the Konstanz Information Miner Copyright, 2021, www.knime.org, last accessed on 14 February 2022) [76,77] was used to perform the analyses and to generate the model, in silico. The EOs dataset were divided using a “Partitioning” tool, with the “Stratified sample” option, separated between training and testing datasets, which represented 80% and 20% of all compounds, respectively. Molecules in the training and testing datasets were randomly selected, but the same proportions of active and not reported substances were maintained for both databases. The information of the composition of the EOs was used like descriptors.

The model utilized a “5-fold external cross-validation” procedure and the Random Forest (RF) algorithm. The RF parameters selected for all models generated 100 total forests to be built, and $-5,440,374,124,525,988,069$ static random seeds (get reproducible results) were generated using random numbers for the model.

The external performances of the selected models were analyzed for sensitivity (true positive rate, which represents the active rate), specificity (true negative rate, which represents the inactive rate), and accuracy (general predictability).

The Applicability domain (APD) corresponds to the chemical space that surrounds the descriptors of the molecules used in the construction of the model. In this way, the applicability domain will provide information about the similarity between what is being tested and what was used to build the model [78–80].

The APD was used to assess whether predictions for the compounds in each dataset were reliable. The APD is based on Euclidean distances, and measures of similarity between the training set descriptors are used to define the APD. Therefore, if a compound in the test set has distances and similarities beyond the APD limit, its prediction will not be reliable. APD can be calculated using the following formula:

$$APD = d + Z\sigma \quad (1)$$

where d and σ are the Euclidean distances for the mean and standard deviation of the compounds in the training set, respectively. Z is an empirical cutoff value, which was set to 0.5 in this study [81].

4. Conclusions

Scientific studies corroborate the results found in this study. Thus, this study of EO analysis establishes a way to find chemical pattern between EOs and antiparasitic activity (antileishmanial, antitrypanosomal and antimalarial). This finding makes it possible to direct studies and biological tests for EOs that have antiparasitic activity more effectively and with less time and financial cost. In particular, we strongly suggest further antiprotozoal studies with EOs from species of the *Piper* genus and the pure compound camphor taking into account data from Cuban EOs. Nevertheless, machine learning analysis studies will be

interesting for EOs from different geographical locations to predict bioactive components with potential antiplasmodial, antileishmanial, and antitrypanosomal activity.

Supplementary Materials: The following are available online, Table S1: Database of essential oils of Cuban plants.

Author Contributions: Conceptualization, M.T.S. and W.N.S.; methodology, R.P.B.d.M., J.G. and R.G.; formal analysis, M.T.S. and L.M.; investigation, all authors R.P.B.d.M., L.S., M.T.S., J.G., R.G., L.M., W.N.S.; resources, J.G. and R.G.; data curation, R.P.B.d.M. and M.T.S.; writing—original draft preparation, R.P.B.d.M., M.T.S. and L.M.; writing—review and editing, all authors R.P.B.d.M., L.S., M.T.S., J.G., R.G., L.M., W.N.S.; project administration, W.N.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are available in the article and the supplementary material.

Acknowledgments: The present work is a collaboration of members of the Research Network Natural Products against Neglected Diseases (ResNetNPND, www.resnetnpnd.org/mobile, latest accessed on 14 February 2022) and a contribution of the Aromatic Plant Research Center (APRC, www.aromaticplant.org, latest accessed on 14 February 2022).

Conflicts of Interest: The authors declare no conflict of interest.

Sample Availability: There are no samples available.

References

1. Yang, S.-K.; Tan, N.-P.; Chong, C.-W.; Abushelaibi, A.; Lim, S.-H.; Lai, K.-S. The missing piece: Recent approaches investigating the antimicrobial mode of action of essential oils. *Evol. Bioinform.* **2021**, *17*, 1176934320938391. [[CrossRef](#)] [[PubMed](#)]
2. Galvan, D.; Efftig, L.; Neto, L.T.; Conte-Junior, C.A. An overview of research of essential oils by self-organizing maps: A novel approach for meta-analysis study. *Compr. Rev. Food Sci. Food Saf.* **2021**, *20*, 3136–3163. [[CrossRef](#)]
3. Sabatino, M.; Fabiani, M.; Božović, M.; Garzoli, S.; Antonini, L.; Marcocci, M.E.; Palamara, A.T.; De Chiara, G.; Ragno, R. Experimental data based machine learning classification models with predictive ability to select in vitro active antiviral and non-toxic essential oils. *Molecules* **2020**, *25*, 2452. [[CrossRef](#)] [[PubMed](#)]
4. Bakkali, F.; Averbeck, S.; Averbeck, D.; Idaomar, M. Biological effects of essential oils—A review. *Food Chem. Toxicol.* **2008**, *46*, 446–475. [[CrossRef](#)] [[PubMed](#)]
5. Falleh, H.; Ben Jemaa, M.; Saada, M.; Ksouri, R. Essential oils: A promising eco-friendly food preservative. *Food Chem.* **2020**, *330*, 127268. [[CrossRef](#)] [[PubMed](#)]
6. Lucia, A.; Guzmán, E. Emulsions containing essential oils, their components or volatile semiochemicals as promising tools for insect pest and pathogen management. *Adv. Colloid Interface Sci.* **2020**, *287*, 102330. [[CrossRef](#)] [[PubMed](#)]
7. Agreles, M.A.A.; Cavalcanti, I.D.L.; Cavalcanti, I.M.F. The role of essential oils in the inhibition of efflux pumps and reversion of bacterial resistance to antimicrobials. *Curr. Microbiol.* **2021**, *78*, 3609–3619. [[CrossRef](#)]
8. Wińska, K.; Mączka, W.; Łyczko, J.; Grabarczyk, M.; Czubaszek, A.; Szumny, A. Essential oils as antimicrobial agents—myth or real alternative? *Molecules* **2019**, *24*, 2130. [[CrossRef](#)]
9. Cai, J.; Yan, R.; Shi, J.; Chen, J.; Long, M.; Wu, W.; Kuca, K. Antifungal and mycotoxin detoxification ability of essential oils: A review. *Phytother. Res.* **2021**, *36*, 62–72. [[CrossRef](#)]
10. Luna, E.C.; Luna, I.S.; Scotti, L.; Monteiro, A.F.M.; Scotti, M.T.; De Moura, R.O.; De Araújo, R.S.A.; Monteiro, K.L.C.; De Aquino, T.M.; Ribeiro, F.F.; et al. Active essential oils and their components in use against neglected diseases and arboviruses. *Oxidative Med. Cell. Longev.* **2019**, *45*, 6587150. [[CrossRef](#)]
11. Reichling, J. Antiviral and virucidal properties of essential oils and isolated compounds – A scientific approach. *Planta Med.* **2021**. [[CrossRef](#)] [[PubMed](#)]
12. Jankowska, M.; Rogalska, J.; Wyszowska, J.; Stankiewicz, M. Molecular targets for components of essential oils in the insect nervous system—A review. *Molecules* **2018**, *23*, 34. [[CrossRef](#)] [[PubMed](#)]
13. Andrade, M.; Braga, M.A.; Cesar, P.H.S.; Trento, M.V.C.; Espósito, M.A.; Silva, L.F.; Marcussi, S. Anticancer properties of essential oils: An overview. *Curr. Cancer Drug Targets* **2018**, *18*, 957–966. [[CrossRef](#)] [[PubMed](#)]

14. Spisni, E.; Petrocelli, G.; Imbesi, V.; Spigarelli, R.; Azzinnari, D.; Sarti, M.D.; Campieri, M.; Valerii, M.C. Antioxidant, anti-inflammatory, and microbial-modulating activities of essential oils: Implications in colonic pathophysiology. *Int. J. Mol. Sci.* **2020**, *21*, 4152. [[CrossRef](#)]
15. Poma, P.; Labbozzetta, M.; Ramarosandratana, A.; Rosselli, S.; Tutone, M.; Sajeva, M.; Notarbartolo, M. In vitro modulation of P-glycoprotein activity by *Euphorbia intisy* essential oil on acute myeloid leukemia cell line HL-60R. *Pharmaceuticals* **2021**, *14*, 111. [[CrossRef](#)] [[PubMed](#)]
16. Costa, R.G.A.; da Anunciação, T.A.; Araujo, M.D.S.; Souza, C.A.; Dias, R.B.; Sales, C.B.S.; Rocha, C.A.G.; Soares, M.B.P.; da Silva, F.M.A.; Koolen, H.H.F.; et al. In vitro and in vivo growth inhibition of human acute promyelocytic leukemia HL-60 cells by *Guatteria megalophylla* Diels (Annonaceae) leaf essential oil. *Biomed. Pharmacother.* **2020**, *122*, 109713. [[CrossRef](#)]
17. Sandner, G.; Heckmann, M.; Weghuber, J. Immunomodulatory activities of selected essential oils. *Biomolecules* **2020**, *10*, 1139. [[CrossRef](#)]
18. Patsilnakos, A.; Artini, M.; Papa, R.; Sabatino, M.; Božović, M.; Garzoli, S.; Vrenna, G.; Buzzi, R.; Manfredini, S.; Selan, L.; et al. Machine learning analyses on data including essential oil chemical composition and in vitro experimental antibiofilm activities against *Staphylococcus* species. *Molecules* **2019**, *24*, 890. [[CrossRef](#)]
19. Monzote, L.; García, J.; González, R.; Scotti, M.T.; Setzer, W.N. Bioactive essential oils from Cuban plants: An inspiration to drug development. *Plants* **2021**, *10*, 2515. [[CrossRef](#)]
20. Kohonen, T. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* **1982**, *43*, 59–69. [[CrossRef](#)]
21. Kohonen, T.; Kaski, S.; Lagus, K.; Salojärvi, J.; Honkela, J.; Paatero, V.; Saarela, A. Self organization of a massive document collection. *IEEE Trans. Neural Netw.* **2000**, *11*, 574–585. [[CrossRef](#)] [[PubMed](#)]
22. Livingstone, D. *Data Analysis for Chemists*; Oxford University Press, Ed.; Oxford Science Publications: Oxford, UK, 1995; ISBN 0-19-855728-0.
23. Menezes, R.; Sessions, Z.; Muratov, E.; Scotti, L.; Scotti, M. Secondary metabolites extracted from Annonaceae and chemotaxonomy study of terpenoids. *J. Braz. Chem. Soc.* **2021**, *32*, 2061–2070. [[CrossRef](#)]
24. Vesanto, J.; Himberg, J.; Alhoniemi, E.; Parhankangas, J. Self-organizing map in matlab: The SOM toolbox. *Proc. Matlab DSP Conf.* **1999**, *99*, 35–40.
25. Salzberg, S.L. Book Review: C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Mach. Learn.* **1994**, *16*, 235–240. [[CrossRef](#)]
26. Mao, J.; Akhtar, J.; Zhang, X.; Sun, L.; Guan, S.; Li, X.; Chen, G.; Liu, J.; Jeon, H.-N.; Kim, M.S.; et al. Comprehensive strategies of machine-learning-based quantitative structure-activity relationship models. *Iscience* **2021**, *24*, 103052. [[CrossRef](#)]
27. Barros, R.P.C.; Da Cunha, E.V.L.; Catão, R.M.R.; Scotti, L.; Souza, M.S.R.; Brás, A.A.Q.; Scotti, M.T. Virtual screening of secondary metabolites of the genus *Solanum* with potential antimicrobial activity. *Rev. Bras. Farmacogn.* **2018**, *28*, 686–691. [[CrossRef](#)]
28. Kovdienko, N.A.; Polishchuk, P.G.; Muratov, E.N.; Artemenko, A.G.; Kuz'Min, V.E.; Gorb, L.; Hill, F.; Leszczynski, J. Application of random forest and multiple linear regression techniques to QSPR prediction of an aqueous solubility for military compounds. *Mol. Inform.* **2010**, *29*, 394–406. [[CrossRef](#)]
29. Papa, R.; Garzoli, S.; Vrenna, G.; Sabatino, M.; Sapienza, F.; Relucenti, M.; Donfrancesco, O.; Fiscarelli, E.V.; Artini, M.; Selan, L.; et al. Essential oils biofilm modulation activity, chemical and machine learning analysis. Application on *Staphylococcus aureus* isolates from cystic fibrosis patients. *Int. J. Mol. Sci.* **2020**, *21*, 9258. [[CrossRef](#)]
30. Artini, M.; Papa, R.; Vrenna, G.; Lauro, C.; Ricciardelli, A.; Casillo, A.; Corsaro, M.M.; Tutino, M.L.; Parrilli, E.; Selan, L. Cold-adapted bacterial extracts as a source of anti-infective and antimicrobial compounds against *Staphylococcus aureus*. *Future Microbiol.* **2019**, *14*, 1369–1382. [[CrossRef](#)]
31. El-Attar, N.E.; Hassan, M.K.; Alghamdi, O.A.; Awad, W.A. Deep learning model for classification and bioactivity prediction of essential oil-producing plants from Egypt. *Sci. Rep.* **2020**, *10*, 21349. [[CrossRef](#)]
32. Fourches, D.; Pu, D.; Tassa, C.; Weissleder, R.; Shaw, S.Y.; Mumper, R.J.; Tropsha, A. Quantitative nanostructure–activity relationship modeling. *ACS Nano* **2010**, *4*, 5703–5712. [[CrossRef](#)]
33. Cherkasov, A.; Muratov, E.N.; Fourches, D.; Varnek, A.; Baskin, I.I.; Cronin, M.; Dearden, J.C.; Gramatica, P.; Martin, Y.C.; Todeschini, R.; et al. QSAR modeling: Where have you been? Where are you going? *J. Med. Chem.* **2014**, *57*, 4977–5010. [[CrossRef](#)]
34. Arcoverde, G.F.B.; Almeida, C.M.; Ximenes, A.D.C.; Maeda, E.E.; Araújo, L.S. Identification of priority areas for forest restoration using self-organizing maps neural network. *Bol. Ciênc. Geod.* **2011**, *1*, 379–400. [[CrossRef](#)]
35. Isah, M.; Ibrahim, M.A.; Mohammed, A.; Aliyu, A.B.; Masola, B.; Coetzer, T. A systematic review of pentacyclic triterpenes and their derivatives as chemotherapeutic agents against tropical parasitic diseases. *Parasitology* **2016**, *143*, 1219–1231. [[CrossRef](#)] [[PubMed](#)]
36. Scotti, M.T.; Scotti, L.; Ishiki, H.; Ribeiro, F.F.; Da Cruz, R.M.D.; De Oliveira, M.P.; Mendonça, F.J.B. Natural products as a source for antileishmanial and antitrypanosomal agents. *Comb. Chem. High Throughput Screen.* **2016**, *19*, 537–553. [[CrossRef](#)] [[PubMed](#)]
37. Nweze, J.A.; Mbaoji, F.N.; Li, Y.-M.; Yang, L.-Y.; Huang, S.-S.; Chigor, V.N.; Eze, E.A.; Pan, L.-X.; Zhang, T.; Yang, D.-F. Potentials of marine natural products against malaria, leishmaniasis, and trypanosomiasis parasites: A review of recent articles. *Infect. Dis. Poverty* **2021**, *10*, 1–19. [[CrossRef](#)] [[PubMed](#)]
38. Pandey, A.K.; Singh, P. The genus *Artemisia*: A 2012–2017 literature review on chemical composition, antimicrobial, insecticidal and antioxidant activities of essential oils. *Medicines* **2017**, *4*, 68. [[CrossRef](#)] [[PubMed](#)]

39. Fampa, P.; Florencio, M.; Santana, R.C.; Rosa, D.; Soares, D.C.; Guedes, H.L.D.M.; da Silva, A.C.; Chaves, D.S.A.; Pinto-Da-Silva, L.H. Anti-*Leishmania* effects of volatile oils and their isolates. *Rev. Bras. Farmacogn.* **2021**, 1–18. [[CrossRef](#)]
40. de Moraes, M.C.; de Souza, J.V.; da Silva Maia Bezerra Filho, C.; Dolabella, S.S.; de Sousa, D.P. Trypanocidal essential oils: A review. *Molecules* **2020**, *25*, 4568. [[CrossRef](#)]
41. Peixoto, J.F.; Ramos, Y.J.; Moreira, D.D.L.; Alves, C.R.; Gonçalves-Oliveira, L.F. Potential of *Piper* spp. as a source of new compounds for the leishmaniasis treatment. *Parasitol. Res.* **2021**, *120*, 2731–2747. [[CrossRef](#)]
42. Kamte, S.L.N.; Ranjbarian, F.; Cianfaglione, K.; Sut, S.; Dall'Acqua, S.; Bruno, M.; Afshar, F.H.; Iannarelli, R.; Benelli, G.; Cappellacci, L.; et al. Identification of highly effective antitrypanosomal compounds in essential oils from the Apiaceae family. *Ecotoxicol. Environ. Saf.* **2018**, *156*, 154–165. [[CrossRef](#)] [[PubMed](#)]
43. Nibret, E.; Wink, M. Trypanocidal and antileukaemic effects of the essential oils of *Hagenia abyssinica*, *Leonotis ocyimifolia*, *Moringa stenopetala*, and their main individual constituents. *Phytomedicine* **2010**, *17*, 911–920. [[CrossRef](#)] [[PubMed](#)]
44. Greve, H.L.; Kaiser, M.; Brun, R.; Schmidt, T.J. Terpenoids from the oleo-gum-resin of *Boswellia serrata* and their antiplasmodial effects in vitro. *Planta Med.* **2017**, *83*, 1214–1226. [[CrossRef](#)] [[PubMed](#)]
45. Mendiola, J.; Pino, J.A.; Fernández-Calienes, A.; Mendoza, D.; Herrera, P. Chemical composition and in vitro antiplasmodial activity of essential oils of leaves and flowers of *Alpinia zerumbet* grown in Cuba. *PharmacologyOnline* **2015**, *2*, 1–5.
46. Monzote, L.; Scull, R.; Cos, P.; Setzer, W.N. Essential oil from *Piper aduncum*: Chemical analysis, antimicrobial assessment, and literature review. *Medicines* **2017**, *4*, 49. [[CrossRef](#)]
47. Gutiérrez, Y.; Montes, R.; Scull, R.; Sánchez, A.; Cos, P.; Monzote, L.; Setzer, W.N. Chemodiversity associated with cytotoxicity and antimicrobial activity of *Piper aduncum* var. *ossanum*. *Chem. Biodivers.* **2016**, *13*, 1715–1719. [[CrossRef](#)]
48. Esperandim, V.R.; Ferreira, D.D.S.; Rezende, K.C.S.; Magalhães, L.G.; Souza, J.M.; Pauletti, P.M.; Januário, A.H.; Laurentz, R.D.S.D.; Bastos, J.K.; Símara, G.V.; et al. In vitro antiparasitic activity and chemical composition of the essential oil obtained from the fruits of *Piper cubeba*. *Planta Med.* **2013**, *79*, 1653–1655. [[CrossRef](#)]
49. Kpoviessi, S.; Bero, J.; Agbani, P.; Gbaguidi, F.; Kpadonou-Kpoviessi, B.; Sinsin, B.; Accrombessi, G.; Frédérick, M.; Moudachirou, M.; Quetin-Leclercq, J. Chemical composition, cytotoxicity and in vitro antitrypanosomal and antiplasmodial activity of the essential oils of four *Cymbopogon* species from Benin. *J. Ethnopharmacol.* **2014**, *151*, 652–659. [[CrossRef](#)]
50. Regalado, E.L.; Fernández, M.D.; Pino, J.A.; Mendiola, J.; Echemendia, O.A. Chemical composition and biological properties of the leaf essential oil of *Tagetes lucida* Cav. from Cuba. *J. Essent. Oil Res.* **2011**, *23*, 63–67. [[CrossRef](#)]
51. Costa, E.V.; Dutra, L.; Nogueira, P.C.D.L.; Moraes, V.R.D.S.; Salvador, M.; Ribeiro, L.H.G.; Gadelha, F.R. Essential oil from the leaves of *Annona vepretorum*: Chemical composition and bioactivity. *Nat. Prod. Commun.* **2012**, *7*, 265–266. [[CrossRef](#)]
52. Dias, C.N.; Rodrigues, K.A.F.; Carvalho, F.A.A.; Carneiro, S.M.P.; Maia, J.G.S.; Andrade, E.H.A.; Moraes, D.F.C. Molluscicidal and leishmanicidal activity of the leaf essential oil of *Syzygium cumini* (L.) SKEELS from Brazil. *Chem. Biodivers.* **2013**, *10*, 1133–1141. [[CrossRef](#)] [[PubMed](#)]
53. Da Silva, T.B.; Menezes, L.; Sampaio, M.F.C.; Meira, C.S.; Guimarães, E.T.; Soares, M.; Prata, A.P.D.N.; Nogueira, P.C.D.L.; Costa, E. Chemical composition and anti-*Trypanosoma cruzi* activity of essential oils obtained from leaves of *Xylopia frutescens* and *X. laevigata* (Annonaceae). *Nat. Prod. Commun.* **2013**, *8*, 403–406. [[CrossRef](#)] [[PubMed](#)]
54. Monzote, L.; Hill, G.M.; Cuellar, A.; Scull, R.; Setzer, W.N. Chemical composition and anti-proliferative properties of *Bursera graveolens* essential oil. *Nat. Prod. Commun.* **2012**, *7*, 1531–1534. [[CrossRef](#)] [[PubMed](#)]
55. Rodríguez, E.J.; Saucedo-Hernández, Y.; Heyden, Y.V.; Simó-Alfonso, E.F.; Ramis-Ramos, G.; Lerma-García, M.J.; Monteagudo, U.; Bravo, L.; Medinilla, M.; De Armas, Y.; et al. Chemical analysis and antioxidant activity of the essential oils of three Piperaceae species growing in the central region of Cuba. *Nat. Prod. Commun.* **2013**, *8*, 1325–1328. [[CrossRef](#)] [[PubMed](#)]
56. Tchoumboungang, F.; Amvam Zollo, P.H.; Dagne, E.; Mekonnen, Y. In vivo antimalarial activity of essential oils from *Cymbopogon citratus* and *Ocimum gratissimum* on mice infected with *Plasmodium berghei*. *Planta Med.* **2005**, *71*, 20–23. [[CrossRef](#)]
57. Monzote, L.; García, M.; Scull, R.; Cuellar, A.; Setzer, W.N. Antileishmanial activity of the essential oil from *Bixa orellana*. *Phytother. Res.* **2013**, *28*, 753–758. [[CrossRef](#)]
58. Pereira, P.S.; Maia, A.J.; Duarte, A.E.; Oliveira-Tintino, C.D.M.; Tintino, S.R.; Barros, L.M.; Vega-Gomez, M.C.; Rolón, M.; Coronel, C.; Coutinho, H.D.; et al. Cytotoxic and anti-kinetoplastid potential of the essential oil of *Alpinia speciosa* K. Schum. *Food Chem. Toxicol.* **2018**, *119*, 387–391. [[CrossRef](#)]
59. Alcoba, A.E.T.; De Melo, D.C.; De Andrade, P.M.; Dias, H.J.; Pagotti, M.C.; Magalhães, L.G.; Júnior, W.G.F.; Crotti, A.; Miranda, M.L.D. Chemical composition and in vitro antileishmanial and cytotoxic activities of the essential oils of *Ocotea dispersa* (Nees) Mez and *Ocotea odorifera* (Vell) Rohwer (Lauraceae). *Nat. Prod. Res.* **2017**, *32*, 2865–2868. [[CrossRef](#)]
60. Tariku, Y.; Hymete, A.; Hailu, A.; Rohloff, J. In vitro evaluation of antileishmanial activity and toxicity of essential oils of *Artemisia absinthium* and *Echinops kebericho*. *Chem. Biodivers.* **2011**, *8*, 614–623. [[CrossRef](#)]
61. Bouyahya, A.; Et-Touys, A.; Bakri, Y.; Talbau, A.; Fellah, H.; Abrini, J.; Dakka, N. Chemical composition of *Mentha pulegium* and *Rosmarinus officinalis* essential oils and their antileishmanial, antibacterial and antioxidant activities. *Microb. Pathog.* **2017**, *111*, 41–49. [[CrossRef](#)]
62. Sainz, P.; Andrés, M.F.; Díaz, M.; Bailén, M.; Rocha, J.N.; Martínez-Díaz, R.A.; González-Coloma, A. Chemical composition and biological activities of *Artemisia pedemontana* subsp. *assoana* essential oils and hydrolate. *Biomolecules* **2019**, *9*, 558. [[CrossRef](#)] [[PubMed](#)]

63. Tabanca, N.; Demirci, B.; Baser, K.H.C.; Aytac, Z.; Ekici, M.; Khan, S.I.; Jacob, M.R.; Wedge, D.E. Chemical composition and antifungal activity of *Salvia macrochlamys* and *Salvia recognita* essential oils. *J. Agric. Food Chem.* **2006**, *54*, 6593–6597. [[CrossRef](#)] [[PubMed](#)]
64. Moreira, R.R.D.; Dos Santos, A.G.; Carvalho, F.A.; Perego, C.H.; Crevelin, E.J.; Crotti, A.E.M.; Cogo, J.; Cardoso, M.L.C.; Nakamura, C.V. Antileishmanial activity of *Melampodium divaricatum* and *Casearia sylvestris* essential oils on *Leishmania amazonensis*. *Rev. Inst. Med. Trop.* **2019**, *61*, e33. [[CrossRef](#)] [[PubMed](#)]
65. Sampaio, M.G.V.; Dos Santos, C.R.B.; Vandesmet, L.C.S.; Dos Santos, B.S.; Santos, I.B.D.S.; Correia, M.T.D.S.; Martins, A.L.D.B.; Da Silva, L.C.N.; Menezes, I.R.D.A.; Gomez, M.C.V.; et al. Chemical composition, antioxidant and antiprotozoal activity of **Eugenia gracillima** Kiaersk. leaves essential oil. *Nat. Prod. Res.* **2019**, *35*, 1914–1918. [[CrossRef](#)]
66. Siqueira, C.A.T.; Serain, A.F.; Pascoal, A.C.R.F.; Andrezza, N.L.; De Lourenço, C.C.; Góis Ruiz, A.L.T.; De carvalho, J.E.; De Souza, A.C.O.; Tonini Mesquita, J.; Tempone, A.G.; et al. Bioactivity and chemical composition of the essential oil from the leaves of *Guatteria australis* A.St.-Hil. *Nat. Prod. Res.* **2015**, *29*, 1966–1969. [[CrossRef](#)]
67. Barros, L.M.; Duarte, A.E.; Morais-Braga, M.F.B.; Waczuk, E.P.; Vega, C.; Leite, N.F.; De Menezes, I.R.A.; Coutinho, H.D.M.; Rocha, J.B.T.; Kamdem, J.P. Chemical characterization and trypanocidal, leishmanicidal and cytotoxicity potential of *Lantana camara* L. (Verbenaceae) essential oil. *Molecules* **2016**, *21*, 209. [[CrossRef](#)]
68. Capello, T.M.; A Martins, E.G.; de Farias, C.F.; Figueiredo, C.R.; Matsuo, A.L.; Felipe Passero, L.D.; Oliveira-Silva, D.; Sartorelli, P.; Henrique Lago, J.G. Chemical composition and in vitro cytotoxic and antileishmanial activities of extract and essential oil from leaves of *Piper cernuum*. *Nat. Prod. Commun.* **2015**, *10*, 285–288. [[CrossRef](#)]
69. Carmo, D.F.M.D.; Amaral, A.C.F.; Machado, G.M.C.; Leon, L.; Silva, J.R.D.A. Chemical and biological analyses of the essential oils and main constituents of *Piper* species. *Molecules* **2012**, *17*, 1819–1829. [[CrossRef](#)]
70. Moreira, R.R.D.; Martins, G.Z.; Varandas, R.; Cogo, J.; Perego, C.H.; Roncoli, G.; Sousa, M.D.C.R.; Nakamura, C.V.; Salgueiro, L.; Cavaleiro, C. Composition and leishmanicidal activity of the essential oil of *Vernonia polyanthes* Less (Asteraceae). *Nat. Prod. Res.* **2017**, *31*, 2905–2908. [[CrossRef](#)]
71. Giongo, J.L.; Vaucher, R.A.; Da Silva, A.S.; Oliveira, C.B.; de Mattos, C.B.; Baldissera, M.D.; Sagrillo, M.R.; Monteiro, S.G.; Custódio, D.L.; de Matos, M.S.; et al. Trypanocidal activity of the compounds present in *Aniba canelilla* oil against *Trypanosoma evansi* and its effects on viability of lymphocytes. *Microb. Pathog.* **2017**, *103*, 13–18. [[CrossRef](#)]
72. De Souza, L.I.O.; Bezzera-Silva, P.C.; do Amaral Ferraz Navarro, D.M.; da Silva, A.G.; dos Santos Correia, M.T.; da Silva, M.V.; de Figueiredo, R.C.B.Q. The chemical composition and trypanocidal activity of volatile oils from Brazilian Caatinga plants. *Biomed. Pharmacother.* **2017**, *96*, 1055–1064. [[CrossRef](#)] [[PubMed](#)]
73. Wangchuk, P.; Keller, P.; Pyne, S.; Taweechotipatr, M.; Kamchonwongpaisan, S. GC/GC-MS analysis, isolation and identification of bioactive essential oil components from the Bhutanese medicinal plant, *Pleurospermum amabile*. *Nat. Prod. Commun.* **2013**, *8*, 1305–1308. [[CrossRef](#)] [[PubMed](#)]
74. Andrade, M.; Azevedo, C.D.S.; Motta, F.N.; dos Santos, M.L.; Silva, C.L.; De Santana, J.M.; Bastos, I.M.D. Essential oils: In vitro activity against *Leishmania amazonensis*, cytotoxicity and chemical composition. *BMC Complement. Altern. Med.* **2016**, *16*, 1–8. [[CrossRef](#)]
75. Monzote, L.; García, M.; Montalvo, A.M.; Scull, R.; Miranda, M. Chemistry, cytotoxicity and antileishmanial activity of the essential oil from *Piper auritum*. *Mem. Inst. Oswaldo Cruz.* **2010**, *105*, 168–173. [[CrossRef](#)] [[PubMed](#)]
76. Berthold, M.R.; Cebron, N.; Dill, F.; Di Fatta, G.; Gabriel, T.R.; Georg, F.; Meinel, T.; Ohl, P.; Sieb, C.; Wiswedel, B. KNIME: The Konstanz Information Miner. *ACM SIGKDD Explor. Newsl.* **2006**, *11*, 58–61.
77. Berthold, M.R.; Cebron, N.; Dill, F.; Gabriel, T.R.; Kötter, T.; Meinel, T.; Ohl, P.; Thiel, K.; Wiswedel, B. KNIME-the Konstanz information miner: Version 2.0 and beyond. *AcM SIGKDD Explor. Newsl.* **2009**, *11*, 26–31. [[CrossRef](#)]
78. Gramatica, P. Principles of QSAR Modeling. *Int. J. Quant. Struct. Relatsh.* **2020**, *5*, 61–97. [[CrossRef](#)]
79. Roy, K.; Kar, S.; Das, R.N. *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*, 1st ed.; Academic Press: San Diego, CA, USA, 2015; ISBN 9780128016336.
80. Kar, S.; Roy, K.; Leszczynski, J. Applicability domain: A step toward confident predictions and decidability for QSAR Modeling. *Methods Mol. Biol.* **2018**, *1800*, 141–169. [[CrossRef](#)]
81. Scotti, M.; Speck-Planche, A.; Tavares, J.; Sobral da Silva, M.; Cordeiro, N.D.; Scotti, L. Virtual screening of alkaloids from Apocynaceae with potential antitrypanosomal activity. *Curr. Bioinform.* **2015**, *10*, 509–519. [[CrossRef](#)]