Check for updates

Computational Methods for Structure-to-Function Analysis of Diet-Derived Catechins-Mediated Targeting of In Vitro Vasculogenic Mimicry

Cancer Informatics
Volume 20: 1–8
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11769351211009229
SAGE

Abicumaran Uthamacumaran[1], Narjara Gonzalez Suarez[2,3], Abdoulaye Baniré Diallo[3,4] and Borhane Annabi[2,3]

[1]Concordia University, Department of Physics, Montreal, QC, Canada. [2]Laboratoire d'Oncologie Moléculaire, Département de Chimie, Université du Québec à Montréal, Montreal, QC, Canada. [3]Centre de recherche CERMO-FC, Université du Québec à Montréal, Montreal, QC, Canada. [4]Laboratoire d'Algèbre Combinatoire et d'Informatique, Département d'Informatique, Université du Québec à Montréal, Montreal, QC, Canada.

ABSTRACT

BACKGROUND: Vasculogenic mimicry (VM) is an adaptive biological phenomenon wherein cancer cells spontaneously self-organize into 3-dimensional (3D) branching network structures. This emergent behavior is considered central in promoting an invasive, metastatic, and therapy resistance molecular signature to cancer cells. The quantitative analysis of such complex phenotypic systems could require the use of computational approaches including machine learning algorithms originating from complexity science.

PROCEDURES: *In vitro* 3D VM was performed with SKOV3 and ES2 ovarian cancer cells cultured on Matrigel. Diet-derived catechins disruption of VM was monitored at 24 hours with pictures taken with an inverted microscope. Three computational algorithms for complex feature extraction relevant for 3D VM, including 2D wavelet analysis, fractal dimension, and percolation clustering scores were assessed coupled with machine learning classifiers.

RESULTS: These algorithms demonstrated the structure-to-function galloyl moiety impact on VM for each of the gallated catechin tested, and shown applicable in quantifying the drug-mediated structural changes in VM processes.

CONCLUSIONS: Our study provides evidence of how appropriate 3D VM compression and feature extractors coupled with classification/regression methods could be efficient to study *in vitro* drug-induced perturbation of complex processes. Such approaches could be exploited in the development and characterization of drugs targeting VM.

KEYWORDS: Computational analysis, machine learning, vasculogenic mimicry, ovarian cancer, epithelial-mesenchymal transition, green tea catechins

RECEIVED: December 9, 2020. **ACCEPTED:** March 15, 2021.

TYPE: Methodology

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to BA.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: BA holds an Institutional Research Chair in Cancer Prevention and Treatment.

CORRESPONDING AUTHOR: Borhane Annabi, Laboratoire d'Oncologie Moléculaire, Département de Chimie, Université du Québec à Montréal, C.P. 8888, Succursale Centre-Ville, Montréal, QC H3C 3P8, Canada. Email: annabi.borhane@uqam.ca

Introduction

Vasculogenic mimicry (VM) is an adaptive cell behavior which defines the formation of fluid-containing channel-like microvascular structures by certain metastatic cancer cells.[1,2] VM is a complex system consisting of many interacting molecular processes which exhibit characteristics such as emergence, self-organization, criticality, and nonlinear dynamics.[3] As such, the equations modeling complex systems often do not have analytical solutions and require computational approximations.[4] In simple terms, a complex system denotes a system in which the collective whole cannot be defined by the sum of its interacting parts.[5] Classic statistical methods may therefore be insufficient in describing and quantifying the complex biology involved in such emergent structure formation processes. Whereas drug targeting of VM is currently arising in the therapeutic oncology field, new computational algorithms allowing to model such complex systems and extract relevant features to perform classification or regression methods must therefore be developed to optimize drug design strategies that would lead to the development of efficient drug-mediated targeting of VM.[6]

In a previous study, the relationship of green tea-derived catechins against transforming growth factor (TGF)-β-mediated epithelial-mesenchymal transition (EMT) in an ovarian clear cell carcinoma model was investigated, wherein the inhibitory effects of catechins carrying a galloyl moiety were demonstrated.[7] Specifically, (-)-epigallocatechin-3-gallate (EGCG), epicatechin gallate (ECG), catechin gallate (CG), and gallocatechin gallate (GCG) inhibited the TGF-β receptor-mediated downstream signaling cascades including phosphorylation of Smad3 and P38. These signaling pathways are involved in the induction of EMT and, thereby, the selection of more invasive phenotypes with increased cell chemotactic migration and VM.[7] Interestingly, EMT and *in vitro* VM have been shown to share a common molecular signature in part responsible for chemoresistance associated to poor cancer patients' survival.[8,9] One could try to extract these signatures

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

directly from the captured images of the EMT and *in vitro* VM. Several methods have been shown to be accurate on image compression and embedding, allowing extracting relevant features that can be used to classify these images or to identify relevant patterns. The success of this process relies on the capacity of modeling the complex system embedded in the images and transforming such networks in series of numerical indicators.

Here, we explored three approaches to assess in vitro capillary-like structure formation which mimics VM in 2 ovarian cancer cell models, the SKOV3 and ES2 cells. Such a procedure consists of 2 steps: first, the image processing steps to embed the 3D VM, transform it into a structure which allows the extraction of numerical features that represent best the capillary-like structure; second, the use of classification methods trained to recognize such structures into the embedded 3D VM.

In the first step, we used the 3 following main approaches: Fractal dimension, wavelet analysis, and percolation. (1) Fractal dimension allows the extraction of box-counting statistics based on the fractal analysis measure which recursively divides a complex structure on the image into smaller and smaller boxes to measure the self-similarity between them. The higher the fractal dimension, the greater the self-similarity and roughness of the structure in higher dimensions of space.[10,11] (2) Wavelet analysis represents the cumulative structures distribution of the characteristic coefficients of the image using a histogram that shows the frequencies and their location for a given image.[12,13] Haar wavelets are conceptually the simplest wavelet basis. Using continuous wavelet analysis, one can then study the way spectral features evolve over time, identify common time-varying patterns in between signals, and perform time-localized filtering. To transform a signal into wavelets, at each resolution the signal is convolved using a series of wavelet coefficients. The continuous wavelet transform could also be used for multifractal analysis if 3D imaging was made available (not performed); and (3) Percolation scores are the critical exponents in a site-lattice percolation model computed based on the Newman-Ziff algorithm.[14] It allows to study the phase transitions in a network when a certain number of nodes or links are removed.[15,16] In principle, percolation algorithms can characterize the connectivity of clusters (branching networks) and predict their emergence during VM. Critical scaling exponents quantitatively characterize the geometrical phase transition (ie, emergence of VM structures).

In the second step, the extracted features from the image processing algorithms can be applied to train classification and regression learners to distinguish between the different drug-treated groups. There exists several types of machine learning approaches that could be applied to either classify images into discrete classes (eg, assess whether 3D capillary-like structure is present or not) or perform regression analyses (allowing to project the emergence of patterns during time). Methods such as support vector machines (SVM) are shown to perform well in such tasks.

In this study, we highlight that the structural complexity of the collective system regulating in vitro 3D capillary-like structure formation can be modeled using fractal dimension, wavelet analysis, and percolation, allowing the quantification of the catechins-mediated structural changes in in vitro VM. This could further be envisioned to be exploited in the future development and characterization of drugs targeting VM.

Materials and Methods

Materials and cell culture

Matrigel and all catechins were purchased from Sigma-Aldrich Canada (Oakville, ON). Human SKOV3 ovarian adenocarcinoma cells as well as human ES2 ovarian clear cell carcinoma cells were purchased from the American Type Culture Collection (ATCC), and monolayers cultured as previously described with McCoy's 5a Modified Medium for ES2 cells (Wisent, 317-010-CL), and DMEM medium (Wisent, 319-005-CL) for SKOV3 cells both containing 10% fetal bovine serum (Life Technologies, 12483-020), 100 U/mL penicillin, and 100 mg/mL streptomycin (Wisent, 450-202-EL).[7] This study was exempted from institutional review board (IRB) as no human participants or cells derived from them were used.

In vitro vasculogenic mimicry assay

In vitro vasculogenic mimicry (VM) of ES2 and SKOV3 ovarian cancer cells was assessed by Matrigel tube formation as previously described.[17] In brief, each well of a 96-well plate was precoated with 50 μL of Matrigel. After gel solidification, cell suspension in culture media (1.8×10^4 cells/200 μL) was seeded into the wells. Various gallated and ungallated catechins (30 μM) were added to the cell culture media and incubated at 37°C in a CO_2 incubator. Optimal capillary-like structures were formed upon 24 hours, and pictures were taken using a digital camera attached to a phase-contrast inverted microscope.

Data sets

A duplicate of 48 images were collected from gallated- and ungallated catechins-treated cells, where each image in duplicate sets was divided into 4 quadrants to increase the training set sample size.

Features extraction

Features are extracted from images using the 3 approaches implemented in MATLAB. For fractal dimension analysis, box-counting statistics are computed using the *Boxcount toolbox* (add-on feature in MATLAB).[18] For the wavelet approach, the maximal 2D wavelet coefficients are computed with the Haar basis using the *Wavelet Analyzer feature*. Finally for the Percolation approach, the percolation scores are computed from percolation clusters using a modified Newman-Ziff algorithm. Additional details of this algorithm can be found in Kruk et al.[19]

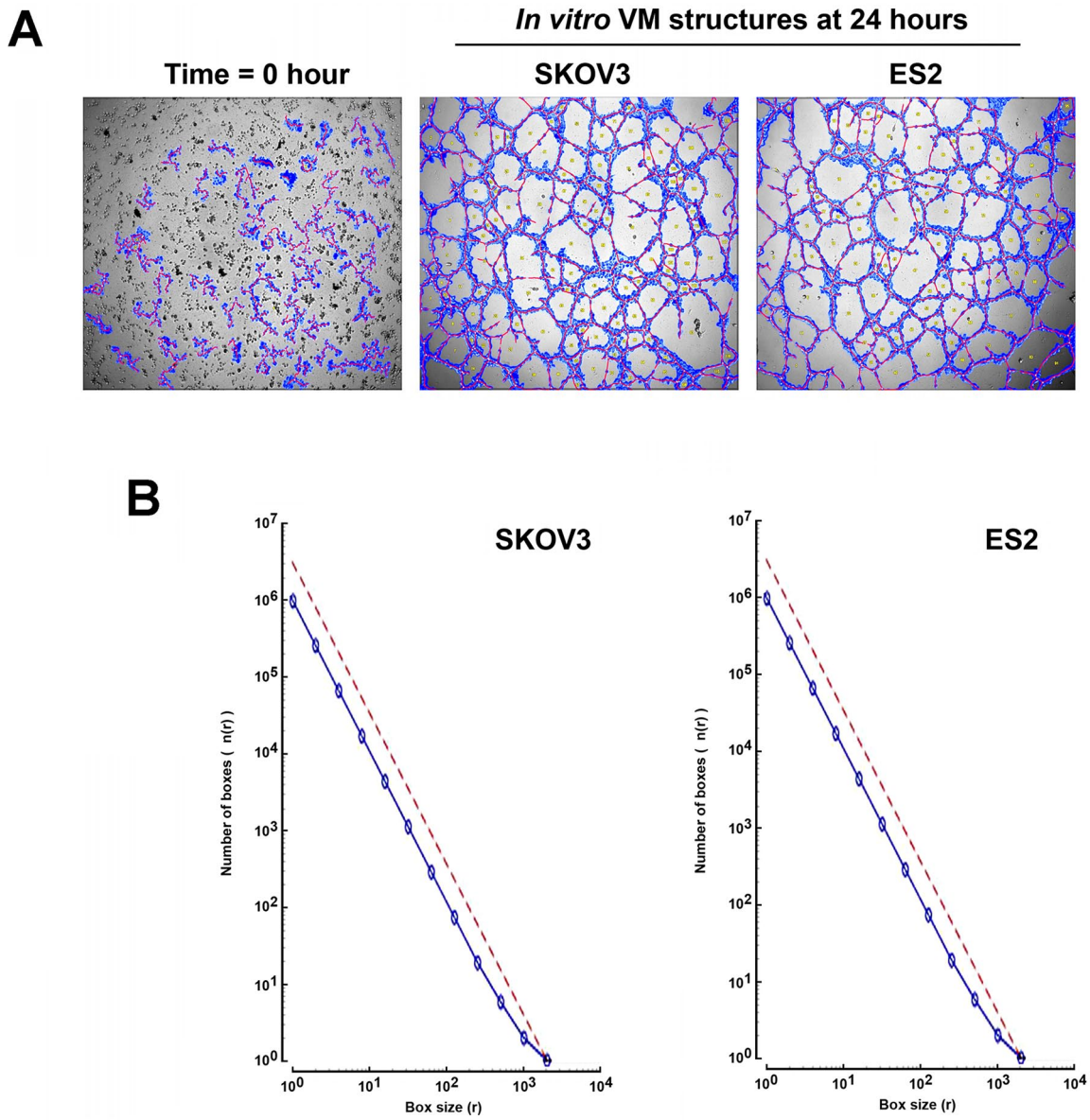

Figure 1. (A) *In vitro* VM in 2 ovarian cancer cell models. Representative capillary-like structure formation at time 0 hour and time 24 hours from ES2 and SKOV3 ovarian cancer cells. (B) Fractal dimension analysis on SKOV3 and ES2 *in vitro* VM. Fractal dimension, Df = 2 ± 8.0825e-16 was computed for both models. The box-count plots are shown, where the space-filling (red dashed line) denotes the fractal curve filling a 2D unit box to map the recognized VM structures. The actual box-count (blue line) adjusts the space-filling curve to the measured local scaling exponents computed by the algorithm. Here, n(r) is the number of boxes and r is the box size of the structures in the box-counting algorithm. VM indicates vasculogenic mimicry.

Regression/classification analyses

For each of the extracted features, a SVM regression model was trained with the wavelet coefficients from the ungallated catechins-treated cells, where each image in duplicate sets was divided into 4 quadrants to increase the training set sample size. During the training, we used an 80% training set, a 20% testing set model, and a 5-fold cross-validation. Two SVM kernels have been tested in this study: a linear SVM kernel and Gaussian SVM that had a Gaussian kernel with different scales presented in results section. To assess the accuracy of the regression of the response curve, R-square and the root mean square error (RMSE) were computed between the true and the regressed values. A small RMSE corresponds to high prediction accuracy of the algorithm.

Statistics

Statistical significance testing and graphical analyses were performed using Origin Pro version 8.5, Graphpad Prism v. 8.4.3, XLSTAT add-in for Microsoft Excel and MATLAB.

Results and Discussion

In vitro vasculogenic mimicry (VM) was triggered as described in the Methods section by ES2 and SKOV3 cells cultured on top of Matrigel, and representative pictures of capillary-like structures taken (Figure 1A). Figure 1B shows the number of boxes as a function of the box sizes used by the algorithm to compute the fractal dimension. The fractal dimension (Df) would correspond to the slope of the log-log plot of function, as indicated by the red dashed arrow. In the case of both

SKOV3 and ES2 cell models, the fractal dimension was Df = 2 ± 8.0825e-16 and showed no difference between the two ovarian cancer models, indicating identical self-similarity in the VM molecular processes involved across these models. However, there is a limit imposed on the fractal dimension algorithm, since 2D images are used to study a dynamic 3D structure formation process. That is, in general, the box-counting algorithm imposes a Df of 2.00 as its maximum limit of stratification when processing 2-dimensional (2D) images. When the algorithm converges to a Df of 2.00 with an uncertainty, it implies that the fractal dimension can be higher than computed. Multiple fractal dimensions may also be needed to characterize the complexity of the VM structures. When more than one fractal dimension exists, a set of statistical approaches known as multifractal analysis is required. As such, our findings indicate that multifractal analysis algorithms may be exploited upstream in the experimental design and in quantifying texture features with 3D imaging techniques.

Structure-to-function impact of 8 diet-derived catechins was next explored experimentally to describe their pharmacological effect on VM. Given the observed similar response in forming capillary-like structures between the 2 cell models, we decided to focus on the ES2 cells only to further apply the fractal dimension algorithm to compute images of the various catechin-treated cells. We found that all molecules bearing the galloyl moiety, CG, ECG, GCG, and EGCG demonstrated a general decline in fractal dimension indicating these drug molecules reduced the fractal characteristics of the structures by lowering branching complexity and self-similarity (Table 1). The impact of the galloyl moiety on the tested catechins was analyzed statistically and the findings confirmed to be significant, with a 2-tailed P value $<.0001$ from a 2-sample t test.

We next went on with 2D wavelet analysis of in vitro VM using sample images comparing, for a proof-of-concept purpose only, C- and CG-treated cells (Figure 2A, upper panels). The original images are contrasted with the wavelet decomposition using the Haar basis at level 3 (Figure 2A, lower panels). In wavelet analysis, the signal is decomposed into low and high frequencies which, at each stepwise, integer increase in decomposition level, half the frequencies of the previous level are discarded. For analytical purposes specific to the studied model, level 3 was shown sufficient for the convolution and appears to globally well represent and preserve the statistical features of the branching structures. The latter mean branching clusters were identified by the 2D wavelet analysis on the original images (all of identical size input: 1280 × 1024 pixels; Table 2). Comparison between the ungallated- and gallated catechins-treated cells confirmed that the mean cluster sizes were lower in the gallated catechins-treated groups (Figure 2B). Statistical analysis with 2-sample t test shows a t value of 9.395 (degrees of freedom = 30) in comparing the gallated groups and non-gallated groups, and a P value $<.0001$.

Table 1. Fractal dimension analysis on ES2 model.

MOLECULES	FRACTAL DIMENSION (DF)
C1	>2.00
C2	>2.00
CG1	1.93
CG2	1.93
EC1	>2.00
EC2	>2.00
ECG1	1.91
ECG2	1.87
EGC1	>2.00
EGC2	>2.00
EGCG1	1.86
EGCG2	1.80
GC1	>2.00
GC2	>2.00
GCG1	1.77
GCG2	1.81

The fractal dimension (Df) computed using the Box-counting algorithm on ES2 images from all 8 catechin groups. A fractional value indicates the shape of the structures is best defined as pertaining to a dimension in between 1D (a line) and 2D (a flat surface). As explained in the Discussion section, Df > 2.00 indicates the structures are best defined using 3D multifractal analysis (not performed).

The globally largest 2D, horizontal level 3 wavelet coefficients for the various tested molecules were next explored. Vertical and diagonal coefficients are not shown since the horizontal component was dominant which resulted in a global coefficient many-fold larger in absolute value for the ungallated molecules in comparison to the gallated ones (Figure 3). This indicates that the frequency modes corresponding to the observed branching from capillary-like structures have been specifically reduced in the gallated catechins groups. Accordingly, the 2-sample t test had a 2-tailed P value $<.0001$, implying the observed structural differences between the gallated and ungallated samples are statistically significant.

Next, percolation clustering analysis was performed on the VM images from the various tested molecules. The 2D percolation scores show an incline in the gallated catechins-treated groups (Table 2; critical exponents). The lower the percolation score, the more complex the network structures are. Thus, the percolation invasion becomes less complex when cells are treated with molecules bearing the galloyl moiety. The interpretation for the higher critical exponents in the gallated molecules-treated groups is 2-fold: (1) the critical threshold for

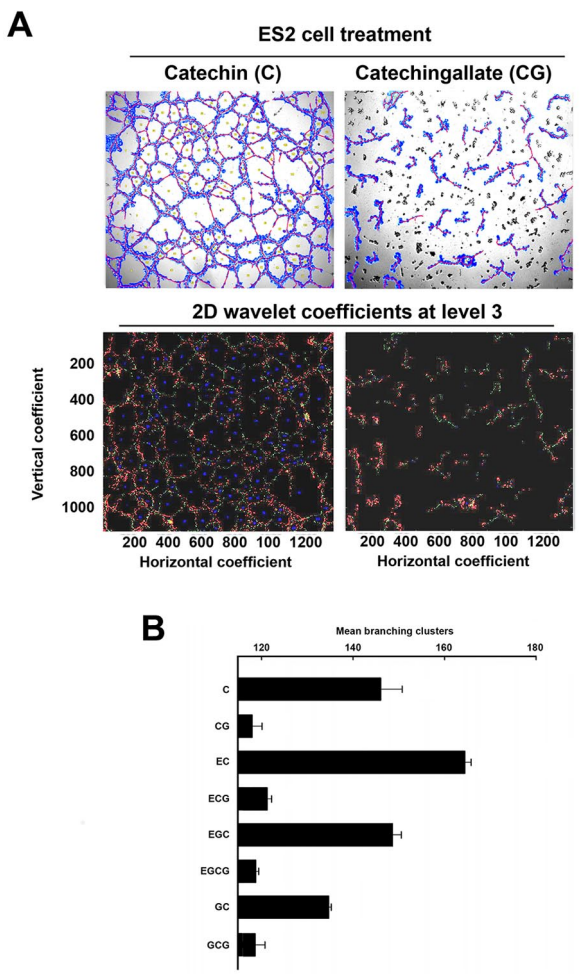

Figure 2. (A) Sample images of 2D wavelet analysis of ES2 cells. The wavelet analyses on the C and CG VM images (upper panels) are shown as decomposed by the MATLAB wavelet analyzer application (lower panels) from the original images, where the axes correspond to the pixel sizes. The wavelet decomposition images were analyzed with the Haar wavelet basis at level 3 where the axes correspond to the wavelet coefficients (diagonal coefficient not shown). (B) Mean branching clusters assessed by 2D wavelet analysis of ES2 VM. Histogram statistics from 2D wavelet analysis in ES2 model showing the mean percolation clusters observed with all 8 catechins tested ($P < .0001$, $r^2 = 0.999$ when comparing gallated vs ungallated groups). VM, vasculogenic mimicry; C, catechin; CG, catechin gallate; EC, epicatechin; ECG, epicatechin gallate; EGC, epigallocatechin; EGCG, epigallocatechin-3-gallate; GC, gallocatechin; GCG, gallocatechin gallate.

randomly distributed cells to phase transition to orderly network structures is higher, and (2) the cells tend to form random clusters rather than self-organizing into network structures (ie, emergence). The Mann-Whitney 2-tailed test between ungallated and gallated molecules-treated groups percolation scores computed a P value of .012, where $P < .05$ indicates there is a significant difference between the 2 groups of molecules.

To assess the predictive power of the extracted features in combination with a classifier, we performed several regression and classification analyses. First, a linear SVM regression model was trained with the wavelet coefficients from the ungallated catechins-treated cells, where each image in duplicate sets was divided into 4 quadrants to increase the training set sample size

Table 2. Percolation clustering statistics in ES2 model.

MOLECULES	SCORE (X)	SCORE (Y)
C1	0.4554	0.5102
C2	0.4405	0.5244
CG1	0.4761	0.5438
CG2	0.4748	0.5426
EC1	0.4403	0.5102
EC2	0.4349	0.5238
ECG1	0.4785	0.5625
ECG2	0.5031	0.5426
EGC1	0.4432	0.5222
EGC2	0.4502	0.5104
EGCG1	0.5178	0.5597
EGCG2	0.4785	0.571
GC1	0.4534	0.5291
GC2	0.4502	0.5288
GCG1	0.4784	0.5485
GCG2	0.4642	0.5459

The 2D critical exponents for phase transition in ES2 model from random cell distributions to branching clusters.

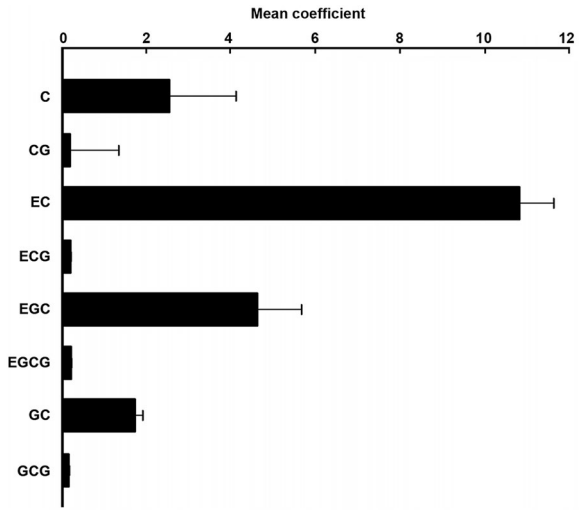

Figure 3. Horizontal wavelet coefficients. Higher coefficients generally imply greater levels of cluster connectivity. The globally largest level 3 horizontal mean coefficients of each catechin-treated group observed in the 2D wavelet analysis of ES2 cells. The coefficients are a global measure of the level of convolution at which the branching clusters identified in Figure 2B can be reconstructed using the Haar wavelet basis. C, catechin; CG, catechin gallate; EC, epicatechin; ECG, epicatechin gallate; EGC, epigallocatechin, EGCG, epigallocatechin-3-gallate; GC, gallocatechin; GCG, gallocatechin gallate.

(n = 48 images). The linear SVM is a supervised learning algorithm available in both regression and classification tasks. It attempts to cluster the data into 2 separate groups, ungallated and

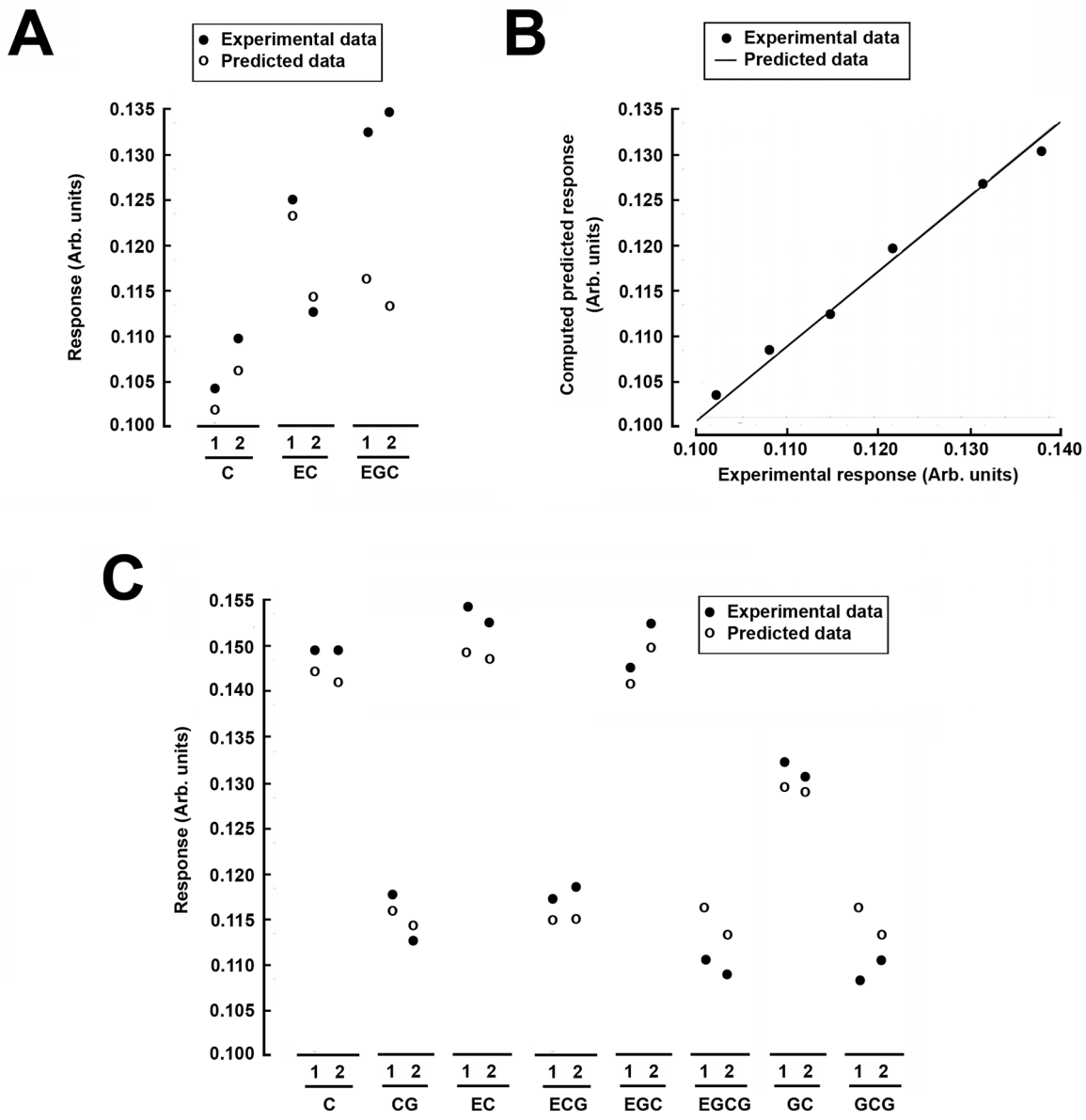

Figure 4. (A) Linear SVM regression predictor on VM by images from ungallated catechin-treated cells. Response curve for prediction of a control data set using linear SVM trained on 2D wavelet coefficients (from Figure 3) of treatment groups. Only ungallated catechins were used to train the SVM. (B) Linear SVM performance curve. Prediction curve of treatment groups' Wavelet coefficients—trained linear SVM on 6 of the 8 catechin-treated groups. (C) Gaussian SVM model on wavelet clusters. Fine Gaussian SVM performance on wavelet analysis mean cluster scores in all 8 catechin-treated groups. Analysis was performed in 2 independent data sets for each catechin tested in (A) and (C). C, catechin; CG, catechin gallate; EC, epicatechin; ECG, epicatechin gallate; EGC, epigallocatechin; EGCG, epigallocatechin-3-gallate; GC, gallocatechin; GCG, gallocatechin gallate; SVM, support vector machine; VM, vasculogenic mimicry.

gallated, by drawing a hyperplane of maximum-margin which divides them. When the treatment group-trained SVM was presented with ungallated molecules coefficients, poor prediction performance was observed with the greatest deviation caused by the EGC1/2 groups (Figure 4A). The resultant response curve had an R-squared value of 0.12, indicating the SVM predictions did not have a linear correlation (ie, poor estimation).

However, when gallated catechins-treated group coefficients were presented from a new set of 24 images, the linear SVM had a high prediction performance with an R-squared value of 0.99 and a RMSE of 0.0013489 ± 0.1209e-5 (n = 24; $P < .01$; Figure 4B). Interestingly, the prediction speed increased from 79 observations/second (Figure 4A) to 920 observations/second (Figure 4B). As seen in Figure 4A and B, the data mining analysis for the wavelet coefficients were restricted to 6 of the 8 catechin groups only. The GCG catechin had globally largest wavelet coefficients very similar to the EGCG and ECG catechin (Figure 3). Similar values resulted in poor training of the linear classifier on wavelet coefficients since it could not find a distinction between these catechins. However, the mean branching clusters show significant difference in these catechin groups (Figure 2B), and this observation is in good agreement with the chemical structures of these 4 gallated catechins. When the mean branching cluster statistics from 2D wavelet analysis of all 8 catechins (n = 32) were used to train a Gaussian SVM regression (Figure 4 C), there was good

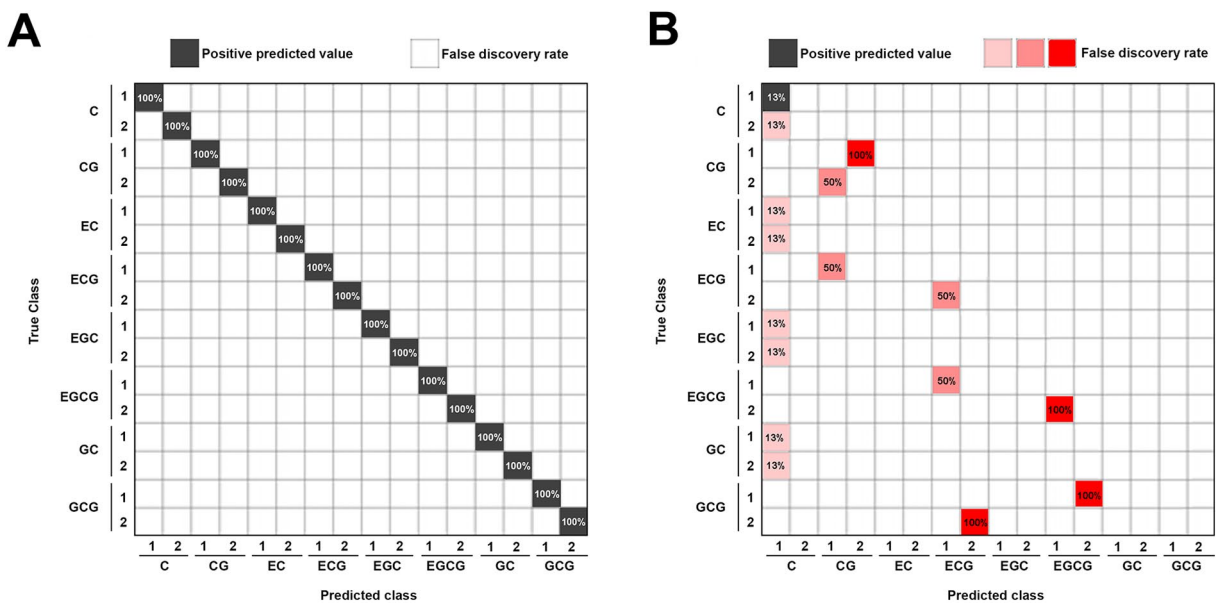

Figure 5. (A) Classification predictive performance chart on Percolation scores. To assess whether the data fed to the classifiers are correctly spotted (predicted value) with the actual data label (true value), SVM classification on percolation statistics was done as described in the Methods section. Positive predictive values for the Fine Gaussian SVM classifier on percolation scores with 5-fold cross-validation was performed in 2 independent data sets for each catechin tested. (B) Linear SVM's poor performance on fractal dimension scores. This demonstrates the fractal dimension is not enough to capture the predictive power of VM structures. C, catechin; CG, catechin gallate; EC, epicatechin; ECG, epicatechin gallate; EGC, epigallocatechin; EGCG, epigallocatechin-3-gallate; GC, gallocatechin; GCG, gallocatechin gallate; SVM, support vector machine; VM, vasculogenic mimicry.

prediction performance (RMSE of 2.469 ± 0.0078 and r^2 of 0.98). Linear SVMs showed poorer performance in comparison to the Gaussian SVM kernel. These findings highlight that the wavelet clustering scores can be used to train data mining algorithms in VM pattern recognition.

Finally, the percolation scores of the ungallated and gallated catechin-treated groups were subjected to a fine Gaussian SVM classifier with identical training conditions as defined for the linear SVM. The fine Gaussian SVM had a Gaussian kernel scale of 0.35, with a one-vs-one multiclass method. When presented with a new set of images (n = 32; ie, 1 image for each of the 8 catechin groups split into 4 quadrants), a perfect prediction was found each time with prediction speed of 200 observations/second (Figure 5A). However, a linear SVM was found to be a poor classifier of the data with a 56.9% accuracy. Following, the SVM classifier was assessed on the fractal dimension analyses. As shown, the SVM classifier's performance on the fractal analyses failed due to the ambiguity of the FD scores for the gallated groups as indicated by the greater than sign > 2.00 (Figure 5B). The FD scores were too close to each other amid the different groups leading to a poor classification. When presented with a new set of images (n = 32), the Gaussian SVM had 0% classification accuracy, whereas the linear SVM had 6.3% accuracy (Figure 5B). These findings demonstrate the Box-counting algorithm's lack of predictive power in assessing drug-mediated changes in VM structures. Multifractal analysis techniques such as the calculation of Hölder exponents or Hurst indices (with time-series analysis) are suggested as a direction required to assess the performance of classifiers or other machine learning algorithms on this fractal dimension feature extractor. Alternately, other types of fractal dimension estimators such as the sandbox method can be attempted.

One can notice that, it is important to choose the appropriate kernel according to whether the data are linearly separable or not. Although the performance is high, the quality of the data sets and its design do not reflect the overall landscape of 3D VM embeddings. The results for image classification of gallated and ungallated catechins could deteriorate with the addition of more data sets. Further studies could be performed with large and more heterogeneous data sets. This deterioration of SVM-based approaches could be compensated by classification algorithms with greater predictive outcomes such as artificial neural networks based on random forests.[20,21] Neural networks are a set of algorithms capable of pattern recognition from complex data sets.[22,23]

Conclusion

The presented findings collectively highlights, in 2 independent ovarian cancer cell models, that the galloyl moiety is a specific structural feature which can be quantified and which associated features can be extracted through computational techniques. These features can provide high predictive power, in part, highlighting the chemopreventive properties of diet-derived catechins by characterizing their ability to inhibit in vitro branching networks involved in VM processes. Most importantly, our study suggests that the prior discussed chemopreventive activities of galloylated catechins can be computationally corroborated using the following 3 image processors

and feature extractors: fractal dimension analysis, wavelet analysis, and percolation clustering. Fractal dimension analysis, at least using a 2D box-counting algorithm, was shown to be an ineffective method in assessing drug-mediated changes in VM structures. Wavelet analysis and percolation clustering outperformed the predictive power of the fractal dimension as VM feature extractors. These findings have been verified with simple classification and regression algorithms. The problem of overfitting by the machine learning algorithms has been challenged by the lack of the fractal dimension's performance in classifying the treatment groups in comparison to the other 2 feature extractors.

Furthermore, the assessment of VM targeting highlights the critical and specific structural features dictating efficient inhibition of other gallated catechins numerous intracellular targets, including HuR, MT1-MMP, and the kinase activity of the TGF-βR.[7,24-26] A reduction in the fractal dimension and 2D wavelet coefficients in the gallated catechins-treated ES2 cells strongly support these conclusions. Furthermore, a higher set of critical exponents in the percolation statistics of the treated groups also confirmed it was more difficult for the ungallated catechins to alter *in vitro* VM structures. Distinctive patterns (differences) amid the ungallated- and gallated-catechins-treated groups that can be captured with the extracted features show how they can be predictive assets on future drug design strategies. Collectively, our findings suggest that the metastatic and chemoresistant phenotypes responsible for the emergence of complex systems such as those observed in VM processes can be suppressed by the gallated catechins. Furthermore, our study demonstrates image-based classification methods could be powerful tools in the emerging field of computational oncology.

In the future, these findings must be reproduced using larger data sets allowing to test heterogeneity and performance of state-of-the-art image classification approaches and pattern recognition including deep learning architectures.[23,27] These deep learning architectures are known to perform better than SVM, suggesting a predictive power improvement with larger image data sets or time-lapse imaging, of structural features in drug molecules that target VM processes. Whereas the studied system consists of dynamic 3D structures analyzed as 2D images, computational algorithms will next require to be extended to 3D computational analysis ideally using 3D time-lapse imaging. For instance, multifractal analysis and 3D wavelet analysis are pertinent algorithms for such prospective studies and would allow for more precise strategies in drug design and development. It will also be important to assess the predictive power of such computational algorithms on inputs from VM data sets from different experimental conditions.

Author Contributions

AU and BA conceptualized the study and wrote the original draft of the article. AU, NGS, and ABD conducted the formal analysis and investigation. All authors reviewed and edited the article.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID iD

Narjara Gonzalez Suarez https://orcid.org/0000-0003-1430-2264

REFERENCES

1. Folberg R, Maniotis AJ. Vasculogenic mimicry. *APMIS*. 2004;112:508-525.
2. Luo Q, Wang J, Zhao W, et al. Vasculogenic mimicry in carcinogenesis and clinical applications. *J Hematol Oncol*. 2020;13:19.
3. Nicolis G, Rouvas-Nicolis C. Complex systems. *Scholarpedia*. 2007;2:1473.
4. Gros C. *Complex and Adaptive Dynamical Systems: A Primer*. 2nd ed. Berlin, Germany: Springer-Verlag; 2011.
5. Hanel R, Thurner S, Klimek P. *Introduction to the Theory of Complex Systems*. Oxford, UK: Oxford University Press; 2018.
6. Yang XG, Zhu LC, Wang YJ, Li YY, Wang D. Current advance of therapeutic agents in clinical trials potentially targeting tumor plasticity. *Front Oncol*. 2019;9:887.
7. Sicard AA, Dao T, Suarez NG, et al. Diet-derived gallated catechins prevent TGF-β-mediated epithelial-mesenchymal transition, cell migration and vasculogenic mimicry in chemosensitive ES-2 ovarian cancer cells. *Nutr Cancer*. 2021;73:169-180.
8. Liu Q, Qiao L, Liang N, et al. The relationship between vasculogenic mimicry and epithelial-mesenchymal transitions. *J Cell Mol Med*. 2016;20:1761-1769.
9. Zhang J, Qiao L, Liang N, et al. Vasculogenic mimicry and tumor metastasis. *J BUON*. 2016;21:533-541.
10. Baish JW, Jain RK. Fractals and cancer. *Cancer Res*. 2000;60:3683-3688.
11. Esgiar AN, Naguib RN, Sharif BS, Bennett MK, Murray A. Fractal analysis in the detection of colonic cancer images. *IEEE Trans Inf Technol Biomed*. 2002;6:54-58.
12. Holschneider M. *Wavelets: An Analysis Tool* (Oxford Mathematical Monographs). Oxford, UK: Oxford University Press; 1999.
13. Roberts T, Newell M, Auffermann W, et al. Wavelet-based scaling indices for breast cancer diagnostics. *Stat Med*. 2017;36:1989-2000.
14. Newman ME, Ziff RM. Efficient Monte Carlo algorithm and high-precision results for percolation. *Phys Rev Lett*. 2000;85:4104-4107.
15. Gould H, Tobochnik J, Christian W. *An Introduction to Computer Simulation Methods: Applications to Physical Systems*. 3rd ed. London, England: Pearson; 2006.
16. Saberi AA. Recent advances in percolation theory and its applications. *Phys Rep*. 2015;578:1-32.
17. Annabi B, Lee YT, Turcotte S, et al. Hypoxia promotes murine bone-marrow-derived stromal cell migration and tube formation. *Stem Cells*. 2003;21:337-347.
18. Moisy F. Boxcount toolbox. https://www.mathworks.com/matlabcentral/fileexchange/13063-boxcount. Updated 2008.
19. Kruk M, Świderski B, Osowski S, Kurek J, Słowińska M, Walecka I. Melanoma recognition using extended set of descriptors and classifiers. *J Image Video Proc*. 2015;2015:43.
20. Makris GM, Pouliakis A, Siristatidis C, et al. Image analysis and multi-layer perceptron artificial neural networks for the discrimination between benign and malignant endometrial lesions. *Diagn Cytopathol*. 2017;45:202-211.
21. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer Series in Statistics). 2nd ed. New York, NY: Springer; 2017.
22. Bishop CM. *Neural Networks for Pattern Recognition*. Oxford, UK: Oxford University Press; 1995.
23. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press; 2016.
24. Annabi B, Currie JC, Moghrabi A, Béliveau R. Inhibition of HuR and MMP-9 expression in macrophage-differentiated HL-60 myeloid leukemia cells by green tea polyphenol EGCg. *Leuk Res*. 2007;31:1277-1284.
25. Djerir D, Iddir M, Bourgault S, Lamy S, Annabi B. Biophysical evidence for differential gallated green tea catechins binding to membrane type-1 matrix metalloproteinase and its interactors. *Biophys Chem*. 2018;234:34-41.
26. Sicard AA, Suarez NG, Cappadocia L, Annabi B. Functional targeting of the TGF-βR1 kinase domain and downstream signaling: a role for the galloyl moiety of green tea-derived catechins in ES-2 ovarian clear cell carcinoma. *J Nutr Biochem*. 2021;87:108518.
27. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436-444.