

DRUG-GENE INTERACTIONS IN ORO-DIGESTIVE CANCER VIA HISTONE DEMETHYLASE INHIBITORS

Swarnambiga Ayyachamy¹; Pradeep Kumar Yadalam² ;

Carlos M. Ardila^{2,3} 

1. Bioinformatics Lab, Department of Biomedical Engineering, Saveetha Engineering College, India, Department of Periodontist, Saveetha Dental College and Hospitals, SIMATS, India.
2. Department of Periodontics, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai - 600 077. Tamilnadu, India.
3. Basic Sciences Department, Faculty of Dentistry, Biomedical Stomatology Research Group, Universidad de Antioquia U de A, Medellín, 050010, Colombia.

EMAIL: martin.ardila@udea.edu.co

ABSTRACT

Introduction: Histone demethylase inhibitors (HDMIs) target histone demethylase enzymes, which play a key role in regulating gene expression. Quantum machine learning, particularly Variational Quantum Classifiers (VQCs), offers potential advantages by enabling efficient analysis of large-scale genomic data. **Objectives:** This study investigates the use of VQCs to predict and classify protein targets in drug-gene interactions involving HDMIs. VQCs can reveal

subtle patterns, manage the complexity of genomic datasets, and deliver highly accurate predictions, surpassing the capabilities of traditional machine learning approaches. **Methods:** Drug-gene interactions within the HDMI signaling pathways were compiled from a dataset of 28,280 drug-gene pairs. A subset of 4,000 protein interactions was selected for quantum machine learning analysis. The Variational Quantum Classifier was employed, integrating classical and quantum components to transform the data into a higher-dimensional space, with iterative optimization of the quantum circuit. **Results:** The VQC achieved a classification accuracy of 97.88% in predicting drug-gene interactions. However, the model's ranking performance was moderate, possibly due to the complexity of the task or the chosen feature mapping approach. On a test dataset, the model demonstrated an accuracy of 97.68%, with 100% recall and high precision. Despite these results, a relatively low F1 score indicates potential performance discrepancies across different classes. **Conclusions:** The Variational Quantum Classifier shows strong potential in predicting protein targets involved in drug-gene interactions for histone demethylase inhibitors, contributing to advancements in precision medicine for Oro-digestive cancers. Nevertheless, further refinement and optimization of the model are needed to enhance its overall performance and robustness.

KEYWORDS: Oro-Digestive cancer; drug-gene interactions; histone demethylase inhibitors; cancer therapies; gene expression.

INTERACCIONES FÁRMACO-GÉN EN CÁNCER ORO-DIGESTIVO MEDIANTE INHIBIDORES DE HISTONA DEMETILASAS

RESUMEN

Introducción: Los inhibidores de histona demetilases (HDMI) actúan sobre enzimas que regulan la expresión génica. El aprendizaje automático cuántico, en particular los Clasificadores Cuánticos Variacionales (VQC), ofrece ventajas potenciales al permitir el análisis eficiente de datos genómicos a gran escala. **Objetivos:** Este estudio investiga el uso de VQC para predecir y clasificar dianas proteicas en interacciones fármaco-gen con HDMI. Los VQC pueden detectar patrones sutiles, manejar la complejidad de conjuntos genómicos y proporcionar predicciones altamente precisas, superando los enfoques tradicionales de aprendizaje automático. **Métodos:** Se recopilaron interacciones fármaco-gen en rutas de señalización de HDMI a partir de un conjunto de 28.280 pares. Un subconjunto de 4.000 interacciones proteicas se seleccionó para análisis con aprendizaje cuántico. Se empleó un Clasificador Cuántico Variacional, integrando componentes clásicos y cuánticos para transformar los datos a un espacio de mayor dimensionalidad, con optimización iterativa del circuito cuántico. **Resultados:** El VQC alcanzó una precisión de clasificación del 97,88% en la predicción de interacciones. Sin embargo, su rendimiento en ranking fue moderado, posiblemente por la complejidad de la tarea o el mapeo de características utilizado. En un

conjunto de prueba, el modelo mostró un 97,68% de precisión, con 100% de recall y alta precisión. No obstante, un F1-score relativamente bajo sugiere discrepancias en el rendimiento entre clases. **Conclusiones:** El Clasificador Cuántico Variacional demuestra gran potencial para predecir dianas proteicas en interacciones fármaco-gen con HDMI, contribuyendo a avances en medicina de precisión para cánceres oro-digestivos. Sin embargo, se requieren mejoras en el modelo para optimizar su robustez y rendimiento global.

PALABRAS CLAVE: cáncer oro-digestivo; interacciones entre fármacos y genes; inhibidores de histona demetilasas; tratamientos contra el cáncer; regulación de la expresión génica.

INTRODUCTION

Histone demethylase inhibitors (HDIMs) are a class of drugs that inhibit the activity of histone demethylase enzymes, which play a crucial role in regulating gene expression by removing methyl groups from histone proteins (1). By inhibiting these enzymes, HDIMs can modify the methylation patterns of histones, leading to changes in gene expression. A specific group of HDIMs,

including lysine-specific demethylase 1 (LSD1) inhibitors, as well as inhibitors of Jumonji domain-containing protein 2 (JMJD2) and Jumonji domain-containing protein 3 (JMJD3), has been studied in the context of Oro-Digestive cancers. LSD1, a histone demethylase, specifically removes methyl groups from lysine 4 of histone H3, and its inhibition has been shown to suppress the growth and metastasis of

esophageal squamous cell carcinoma and gastric cancer cells in preclinical models (2).

In drug discovery and development, understanding the complex interactions between drugs and genes is critical for designing effective targeted therapies. Single protein targets refer to specific proteins that drugs bind to, modulating their activity, which in turn affects gene expression and determines therapeutic responses. This interaction between drugs and genes has transformed the field of precision medicine, enabling the development of tailored treatment strategies based on an individual patient's genetic profile.

Epigenetic enzyme-targeted therapies hold promise in cancer treatment by targeting enzymes such as histone deacetylases, DNA

methyltransferases, and lysine demethylases, including KDM4s. Researchers are focusing on designing potent KDM4 inhibitors as potential anticancer agents, with particular emphasis on selective small molecule inhibitors, scaffold candidates, and cyclic skeletons (3,4).

Quantum machine learning has the potential to significantly enhance the analysis of drug-gene interactions by efficiently processing and analyzing large-scale genomic datasets. This technology can detect subtle patterns and relationships in drug-gene interactions, providing deeper insights into underlying biological mechanisms. Quantum machine learning is particularly adept at handling the complexity and high dimensionality of

genomics data, resulting in more accurate predictions (5).

Variational Quantum Classifiers (VQCs) take advantage of the intrinsic parallelism of quantum computers, enabling them to represent data in higher-dimensional spaces compared to classical models. This capability is especially useful for analyzing complex datasets with intricate relationships between features. VQCs utilize quantum gradients, which can potentially lead to faster training times compared to classical gradient descent algorithms (6). This is particularly relevant for problems where the classical gradient landscape is challenging and prone to being trapped in local minima. VQCs are universally expressive, meaning they can represent any classification function, giving them the

ability to address problems that are intractable for classical algorithms. The parallel processing capabilities of quantum computing enable VQCs to analyze data simultaneously, potentially resulting in faster training times compared to classical algorithms (7,8).

Previous studies have demonstrated the power of variational autoencoders in drug discovery. For instance, a tool based on RNA sequencing data predicted the tissue of origin in cancer patients with 96% accuracy, guiding 73% of first-line therapy decisions, which led to positive treatment responses by matching diagnostic predictions to gene fusions. Another recent study utilized a deep neural network-based variational autoencoder (VAE) combined with gradient-boosted decision trees to predict

chemotherapy response in various cancers. This study showed that VAE-encoded tumor transcriptome features outperformed original gene expression profiles, providing better classification performance than PCA or ICA components (9,10).

VQCs are an advanced form of variational classifiers, leveraging quantum superposition and entanglement for superior feature extraction and discrimination. Their ability to perform parallel computations makes them faster and more effective than classical machine learning algorithms. While few studies have explored drug-gene interactions using quantum machine learning, this research aims to investigate the application of quantum classifiers in predicting and classifying single protein targets within

drug-gene interactions involving histone demethylase inhibitors.

Materials and Methods

Dataset Preparation

Drug-gene interaction data related to HDMI signaling were obtained from the Probes & Drugs database (<https://www.probes-drugs.org/>) (11). The dataset, comprising 28,280 drug-gene interactions, was annotated and pre-processed to remove outliers. From this dataset, 4,000 single protein interactions were selected for further analysis. The data were split into training and test sets, with 80% allocated for training and 20% for testing. Five-fold cross-validation was

used to ensure robustness. Only single protein interactions were considered as targets for quantum machine learning, and other protein classes were excluded from the analysis.

Variational Quantum Classifier (VQC)

The VQC is a hybrid classification method that integrates both classical and quantum components. It maps classical data into a higher-dimensional quantum space using a feature map, then constructs a variational circuit composed of ansatz gates and variational parameters. The architecture of the VQC includes several key elements: a feature map, ansatz gates, variational parameters, optimization, Pauli-Z encoding,

quantum operations, and expectation value calculations.

The feature map encodes classical data into quantum states, capturing essential information for classification.

Ansatz gates allow the circuit to explore various quantum states, with the gate structure tailored to the specific problem being addressed.

Variational parameters control the quantum states generated by the circuit and are adjusted during the optimization process. The circuit is iteratively optimized to maximize the likelihood of correctly predicting class labels.

To classify new data, the VQC utilizes the optimized variational parameters and quantum operations. The classification is determined by calculating the expectation value of the Pauli-Z operator on qubit 0, which yields the predicted class label.

Results

The quantum classifier achieved a high accuracy of 97.88% in classifying and predicting drug-gene interactions. However, its ranking performance, reflected by a moderate ROC AUC of 0.6670, was less impressive. This could be attributed to factors such as the complexity of the task or the selection of the feature map and ansatz

gates in the VQC architecture. Overall, the classifier's performance is commendable. When applied to the test dataset, the model achieved 97.68% accuracy, 97.68% precision, and 100% recall. Despite these strong metrics, the macro-averaged F1 score of 0.494, which balances precision and recall, indicates potential variability in performance across different classes. While the quantum classifier performed well overall on both the training and test datasets, the lower macro-averaged F1 score suggests uneven performance among the different classes in the dataset.

Figure 1 presents the ROC curve for 5-fold cross-validation, illustrating the quantum machine learning model's performance across five different data subsets. Each line corresponds to a different fold, while the

blue dashed line represents the mean ROC curve. The x-axis shows the false positive rate, and the y-axis shows the true positive

rate. The AUC for each fold and the mean ROC are included in the legend.

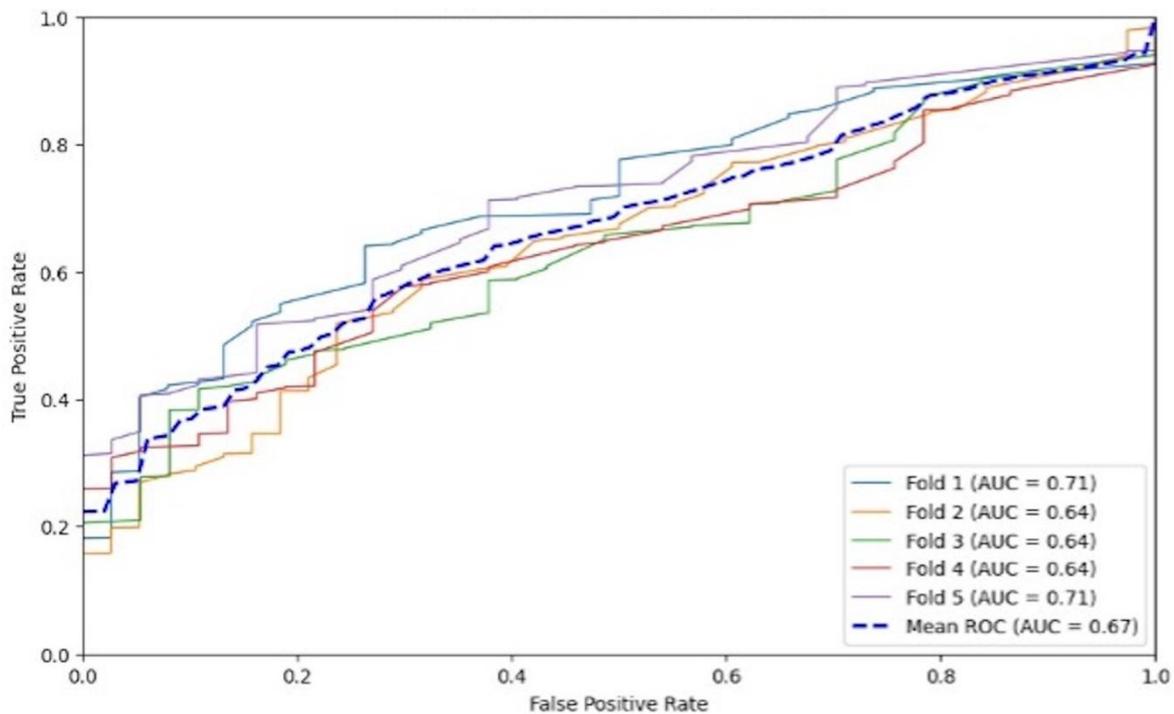


Figure 1. ROC Curve shows the five-fold cross-validation of a single protein family.

The trained quantum classifier exhibited strong performance on both training and test datasets, with a consistently high accuracy of 97.91% during training. As the model trained, the cost function decreased,

indicating effective learning. After training, the model maintained an accuracy of 97.68% on the test dataset, demonstrating good generalization to unseen data. With a recall of 100%, the model showed excellent

capability in identifying positive classes. However, the macro-averaged F1 score of approximately 0.494 on the test dataset suggests some performance challenges in classifying the minority classes. The mean accuracy was 97.88%, and the mean ROC AUC of 0.6670 indicates moderate class separation capability.

The performance metrics for the "single protein" class were particularly high, with a precision of 0.98, recall of 1.00, and an F1-score of 0.99. Although the overall dataset showed moderate results, with macro-averages of approximately 0.49 for precision, recall, and F1-score, the weighted averages for these metrics were much higher, at around 0.96, indicating strong overall performance. However, the model needs improvement in classifying the

minority classes, as shown by lower precision, recall, and F1-scores for these groups.

Table 1 highlights that the model's accuracy in predicting the "single protein" class was 97%. In this classification task, the precision and recall values were particularly high, with 98% of predicted instances being correct. The recall was perfect at 1.00, indicating that all instances of the "single protein" class were correctly identified. The F1-score, the harmonic mean of precision and recall, was 0.99, indicating strong overall performance. The dataset's support of 1,723 reflects the number of instances in the "single protein" class, further demonstrating the model's high accuracy in predicting this class.

Table 1. Prediction on the single-protein family.

Sequence Number	Class-single protein	Precision	Recall	F1 -score	Support
1	Single protein	0.98	1.00	0.99	1723

Discussion

The complex molecular mechanisms underlying Oro-Digestive cancers pose significant challenges in developing effective biomarkers and therapeutic strategies. Histone lysine-specific demethylase 1 (LSD1), an oncogene implicated in various malignancies, plays a crucial role in cellular processes associated with tumor progression (2,3,12). Previous studies have highlighted the oncogenic functions of LSD1 in tumorigenesis and its potential as a therapeutic target. Elevated

levels of LSD1 in oral squamous cell carcinoma (OSCC) have been linked to more aggressive tumor phenotypes and poorer overall survival. Inhibition or depletion of LSD1 has been shown to reduce tumor cell proliferation, migration, invasion, tumorsphere formation, and xenograft growth, underscoring its relevance as a therapeutic target (4,13,14). Exploring drug-gene interactions, particularly involving LSD1, can help address the biological complexity of these cancers.

Machine learning has proven invaluable for predicting drug-gene interactions, which are key to designing targeted therapies for Oro-Digestive cancers (15–17). In this study, we leveraged quantum machine learning (QML) to enhance the prediction of drug-gene interactions for targeted therapy. QML offers superior feature extraction capabilities compared to traditional machine learning approaches, potentially improving the accuracy and efficacy of predictive models.

Quantum computing and QML are gaining traction in healthcare applications, particularly for classification tasks where they show potential to outperform classical models (5,6). Variational quantum classifiers (VQCs), designed to harness quantum computational speed-ups, hold

promise for future advancements, although they may face limitations with current quantum devices. Mutual Information Feature Selection (MIFS) has been explored to improve model interpretability, but classical computational methods often struggle with large-scale data, making QML algorithms more attractive. QML strategies can be classified into three broad categories: quantum machine learning, quantum-inspired machine learning, and hybrid classical-quantum approaches, all of which extend classical machine learning methods into the quantum domain.

Previous studies support the potential of QML in outperforming classical machine learning. For instance, a study comparing classical models with a quantum variational circuit for estimating brain age and

predicting gender using structural MRI data found that the quantum models delivered superior accuracy and lower error rates (20). Another study employed a quantum neural network (QNN) to differentiate between solitary large brain metastases (LBM) and high-grade gliomas (HGG) on contrast-enhanced T1-weighted brain MRI, achieving a test ROC-AUC of 0.76 and balanced accuracy (bACC) of 0.74 (21,22). These results are comparable to our study, where the VQC achieved an accuracy of 97% in predicting drug-gene interactions for a single protein target (Figure 1, Table 1).

Despite these promising results, QML faces several challenges, including limited quantum data, computational resource constraints, quantum noise, and interpretability issues. Quantum systems

are particularly susceptible to noise and errors, which can impact model accuracy and reliability. Moreover, the interpretability of quantum models remains a challenge, making it difficult to fully understand their decision-making processes.

Future improvements in QML will require several advancements. Generating more quantum data, developing robust quantum algorithms, and improving computational resources are essential. Additionally, enhancing the interpretability of quantum models and addressing concerns regarding data privacy and security will be crucial. As quantum computing technology advances, with innovations such as increased qubit counts and improved error correction techniques, more complex models will

become feasible. Developing hybrid classical-quantum approaches and methods specifically designed for QML will further enhance its application in precision medicine.

Conclusion

The application of a Variational Quantum Classifier for predicting and classifying single protein targets within drug-gene interactions involving histone demethylase inhibitors demonstrates significant promise, particularly in the context of precision medicine for Oro-Digestive cancers. The model's high accuracy and precision highlight its capability to identify complex relationships between drugs and genes. However, further refinement in feature selection and hyperparameter tuning may improve its performance in ranking drug-

gene interactions. Overall, this study emphasizes the potential of quantum machine learning to drive advancements in the development of targeted therapies and personalized treatment strategies.

Data availability statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author Contributions:

SA, PY, PN and CA designed the research study. SA, PY, PN and CA performed the research. SA, PY, PN and CA analyzed the data. SA, PY, PN and CA wrote the manuscript. All authors contributed to editorial changes in the manuscript. All

authors read and approved the final manuscript.

Ethics Approval and Consent to Participate:

Not required

Acknowledgment: We would like to thank the Center of Medical and Bioallied Health Sciences and Research, Ajman University, Ajman, UAE

Funding: This research received no external funding

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

REFERENCES

1. Tobin SJ, Chang H, Kent MS, Davies AE. JARID1-targeted histone H3 demethylase

inhibitors exhibit anti-proliferative activity and overcome cisplatin resistance in canine oral melanoma cell lines. *Vet Comp Oncol.* 2021;19(3):518–28.

2. Hodges S, Cooney J. Novel lysine-specific histone demethylase 1 inhibitor in acute myeloid leukaemia transformed from essential thrombocythaemia. *Cancer Rep (Hoboken).* 2022;5(8):e1588.

3. Long X, Zhang S, Wang Y, Chen J, Lu Y, Hou H, et al. Targeting JMJD1C to selectively disrupt tumor T(reg) cell fitness enhances antitumor immunity. *Nat Immunol.* 2024;25(3):525–36.

4. Wang Y, Zhu Y, Wang Q, Hu H, Li Z, Wang D, et al. The histone demethylase LSD1 is a novel

oncogene and therapeutic target in oral cancer.
Cancer Lett. 2016;374(1):12–21.

5. Ciliberto C, Herbster M, Ialongo AD, Pontil M, Rocchetto A, Severini S, et al. Quantum machine learning: a classical perspective. Proc Math Phys Eng Sci. 2018;474(2209):20170551.

6. Zeguendry A, Jarir Z, Quafafou M. Quantum Machine Learning: A Review and Case Studies. Entropy (Basel). 2023;25(2).

7. Cerezo M, Verdon G, Huang HY, Cincio L, Coles PJ. Challenges and opportunities in quantum machine learning. Nat Comput Sci. 2022;2(9):567–76.

8. Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S. Quantum machine learning. Nature. 2017;549(7671):195–202.

9. Rahimi M, Asadi F. Oncological Applications of Quantum Machine Learning. Technol Cancer Res Treat. 2023;22:15330338231215214.

10. Felefly T, Roukoz C, Fares G, Achkar S, Yazbeck S, Meyer P, et al. An Explainable MRI-Radiomic Quantum Neural Network to Differentiate Between Large Brain Metastases and High-Grade Glioma Using Quantum Annealing for Feature Selection. J Digit Imaging. 2023;36(6):2335–46.

11. Skuta C, Popr M, Muller T, Jindrich J, Kahle M, Sedlak D, et al. Probes & Drugs portal: an

interactive, open data resource for chemical biology. *Nat Methods*. 2017;14(8):759–60.

12. Benedetti R, Dell'Aversana C, De Marchi T, Rotili D, Liu NQ, Novakovic B, et al. Inhibition of Histone Demethylases LSD1 and UTX Regulates ERα Signaling in Breast Cancer. *Cancers (Basel)*. 2019;11(12).

13. Yang CY, Lin CK, Tsao CH, Hsieh CC, Lin GJ, Ma KH, et al. Melatonin exerts anti-oral cancer effect via suppressing LSD1 in patient-derived tumor xenograft models. *Oncotarget*. 2017;8(20):33756–69.

14. Chen TM, Huang CM, Setiawan SA, Hsieh MS, Sheen CC, Yeh CT. KDM5D Histone Demethylase Identifies Platinum-Tolerant Head

and Neck Cancer Cells Vulnerable to Mitotic Catastrophe. *Int J Mol Sci*. 2023;24(6).

15. Yadalam PK, Natarajan PM, Mosaddad SA, Heboyan A. Graph neural networks-based prediction of drug gene association of P2X receptors in periodontal pain. *J Oral Biol Craniofac Res*. 2024;14(3):335–8.

16. Yadalam PK, Arumuganainar D, Ronsivalle V, Di Blasio M, Badnjevic A, Marrapodi MM, et al. Prediction of interactomic hub genes in PBMC cells in type 2 diabetes mellitus, dyslipidemia, and periodontitis. *BMC Oral Health*. 2024;24(1):385.

17. Ramalingam K, Yadalam PK, Ramani P, Krishna M, Hafedh S, Badnjevic A, et al. Light gradient boosting-based prediction of quality of

life among oral cancer-treated patients. BMC Oral Health. 2024;24(1):349.

18. Jang HY, Song J, Kim JH, Lee H, Kim IW, Moon B, et al. Machine learning-based quantitative prediction of drug exposure in drug-drug interactions using drug label information. NPJ Digit Med. 2022;5(1):88.

19. Liu M, Wu Y, Chen Y, Sun J, Zhao Z, Chen X wen, et al. Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. J Am Med Inform Assoc. 2012;19(e1):e28-35.

20. Amin J, Anjum MA, Sharif M, Jabeen S, Kadry S, Moreno Ger P. A New Model for Brain Tumor Detection Using Ensemble Transfer

Learning and Quantum Variational Classifier. Comput Intell Neurosci. 2022;2022:3236305.

21. Park S, Park DK, Rhee JKK. Variational quantum approximate support vector machine with inference transfer. Sci Rep. 2023;13(1):3288.

22. Jäger J, Krems R V. Universal expressiveness of variational quantum classifiers and quantum kernels for support vector machines. Nat Commun. 2023;14(1):576.