VER
Vascular & Endovascular Review

Developing Predictive Models for Carcinogen-Induced Mutation Pathways in Aquatic Species Using The Mutagenesispredictnet Algorithm

Dr. Gaurav Tamrakar, Shailesh Singh Thakur

Assistant Professor, Department of Mechanical, Kalinga University, Raipur, India. ku.gauravtamrakar@kalingauniversity.ac.in
Assistant Professor, Department of Mechanical, Kalinga University, Raipur, India. ku.shaileshsinghthakur@kalingauniversity.ac.in

ABSTRACT

Mutations in aquatic species induced by carcinogens represent many difficulties, not limited to complex environmental exposures and a variety of biological responses disrupting ecosystem health and biodiversity. Understanding mutation pathways due to carcinogenic exposure (e.g., polycyclic aromatic hydrocarbons (PAHs), heavy metals, industrial pollutants) for ecological risk assessments and measurements for food safety inspections requires myriad assessments [5] [7]. Individual studies are often limited by the scales of duration for the experiment and/or ethical issues for a particular aquatic species. Predictive computational modelling is a distinct form of assessment because it can fill gaps in existing data by simulating mutation pathways or identifying carcinogenic endpoints or even potential outcomes. This study introduces MutagenesisPredictNet, an algorithm to predict carcinogen-induced mutation pathways in aquatic organisms using environmental pollutant data along with molecular data (e.g., genetics). The algorithms use artificial intelligence to use machine learning and data analysis to model the interdependent nature between carcinogens and the DNA repair system, to predict where something will mutate (the hot spots), and any environmental condition that may affect the pathway. The findings of the study demonstrate that MutagenesisPredictNet can predict species mutation profiles specific to the environmental exposure using products of the pollutant, and give us information and understanding of mechanisms for carcinogenesis for future studies. The predictions can support environmental health surveillance, regulatory and policy efforts, and conservation actions. MutagenesisPredictNet advances aquatic toxicology by developing a modular approach enabling it to analyze multiple species, an approach that is more environmentally friendly than conducting new animal studies, as well as providing a platform to assess environmental exposure and degradation scenarios, as it reduces the potential carcinogenesis risk in aquatic ecosystems and important human food sources.

KEYWORDS: Carcinogenesis, Mutation Pathways, Aquatic Species, Predictive Modeling, MutagenesisPredictNet, Computational Toxicology, DNA Damage, Environmental Carcinogens, Machine Learning, Ecological Risk Assessment

How to Cite: Dr. Gaurav Tamrakar, Shailesh Singh Thakur, (2025) Developing Predictive Models for Carcinogen-Induced Mutation Pathways in Aquatic Species Using The Mutagenesispredictnet Algorithm, Vascular and Endovascular Review, Vol.8, No.1s, 27-30

INTRODUCTION

Aquatic ecosystems are increasingly threatened by carcinogenic pollutants from polycyclic aromatic hydrocarbons (PAHs), heavy metals, and industrial chemicals that bioaccumulate through water and sediment contamination. The exposure of aquatic wildlife to these carcinogens produces mutations that alter cellular functions, resulting in a greater frequency of cancer and genetic defects, and a better understanding of the mutational pathways induced by these carcinogens is needed to better delineate ecological impacts and defend protective biodiversity [1]. Unfortunately, these mutational pathways can be complicated and consist of several molecular interactions unique to each species.

Conventional experimentation to study carcinogen-induced mutations is expensive, time-consuming, and is also limited in its ability to account for environmental complexity and long-term duration. Current computational modelling efforts often do not incorporate the varied biological and environmental data at the same time, dramatically decreasing predictive accuracy and ecological usefulness [8].

Here, we present MutagenesisPredictNet, an innovative predictive modelling algorithm designed to capture the complexity of carcinogen-induced mutation pathways in aquatic species [2]. Our model leverages machine learning with multi-source datasets to provide improved estimation of mutation pathways while accounting for changes in exposure scenarios [1][12].

The research goals are to build a powerful computational framework for predicting carcinogenesis, validate its efficacy among aquatic species, and show the ecological relevance of this model in order to inform the assessment of environmental risk [3].

BIOLOGICAL BASIS OF CARCINOGEN-INDUCED MUTAGENESIS IN AQUATIC SPECIES

Mutagenesis in aquatic organisms due to carcinogens involves molecular interactions (or reactions) of environmental contaminants, as they interact with DNA within cells. Important mechanisms of carcinogenesis include the formation of DNA adducts, oxidative damage, and strand breaks from reactive intermediates of the carcinogen. Although aquatic organisms are equipped with DNA repair systems, nucleotide excision repair, and base excision repair systems to preserve genomic material, these systems can be overwhelmed, and the error rate can be high during periods of high contaminant exposure, leading to persistent mutations [6].

A few examples of carcinogenic contaminants within aquatic environments are the class of compounds known as polycyclic aromatic hydrocarbons (PAHs) and heavy metals, such as mercury and cadmium, and persistent organic pollutants. PAHs are metabolically activated to reactive intermediates that bind DNA, thereby initiating mutagenesis. Heavy metals induce oxidative stress and inhibit DNA repair enzymes, thereby compounding genetic injury.

The exposure of susceptible aquatic organisms to contaminants activates a cascade from DNA injury to mutagenesis and potentially to carcinogenesis due to the

disruption of intracellular regulation of apoptotic pathways and proliferation. Fragile cells in bioaccumulating species will carry higher contaminant burdens, which ultimately lead to a greater susceptibility, such as the case with filter-feeding bivalves [10].
While DNA damage due to chemicals can have serious effects on aquatic organisms, understanding these biological processes, as well as species-specific vulnerabilities, is important for predicting carcinogenic effects of contaminants within aquatic ecosystems and for establishing effective regulations to protect the environment [11].

DESIGN AND IMPLEMENTATION OF MUTAGENESISPREDICTNET ALGORITHM

The MutagenesisPredictNet algorithm is designed as a multi-layer predictive model using multiple biological and environmental data to model carcinogen-induced mutation pathways in aquatic organisms. The primary data sources are the pollutant profiles (chemical dose, duration of exposure), the genomic and transcriptomic data from the target species, and the mutation databases. The biological datasets, including genomic and transcriptomic data and the mutation databases, must be pre-processed via normalization, feature extraction, and noise suppression to ensure underlying data quality and consistency [9] [14].
If your model predicts mutation probability P_m for a pathway i in species s under carcinogen c, you could express it as:

$$P_m(s,c,i) = \sigma\left(\sum_{j=1}^{n} \omega_j \cdot f_j\left(X_{s,c}\right) + b\right)$$

Where:

- $X_{s,c}$ Feature vector combining pollutant concentration, exposure time, and genomic markers for species s under carcinogen ccc
- $f_j(\cdot)$ Feature transformation functions (e.g., embeddings from neural network layers)
- ω_j Learned weights for each transformed feature
- b Model bias term
- $\sigma(\cdot)$ Sigmoid activation function mapping scores to probabilities between 0 and 1
- n Number of features after preprocessing

The prediction model itself is built on machine learning frameworks primarily (deep neural networks) and also includes network analysis to capture biologically-affine linkages in the interactions of carcinogens within chemical pathways (and DNA repair networks) [15]. Each layer of processing consists of embedded features to represent chemical within-genomic features and also predicts the likelihood of mutation and pathway disruption [13].

The first important output of the model is that it generates probabilistic mutation maps (i.e., risk scores) by chemical and species; the second output is the overall pathways impact score by chemical species; and thirdly is the species-specific risk profiles across chemical exposure scenarios. Performance of the model is compared against curated datasets of known mutations and environmental exposure cases; the performance measures include accuracy, precision, recall, and area under the receiver operating characteristic curve (AUC-ROC). Sensitivity analyses were undertaken to test robustness across various species and chemical-pollutant scenarios.

This design helps NPA and other assessment partners' offices generate predictive outcomes that are scalable, interpretable, and can assist in future decision-making about ecological risk assessments and targeted experiments.

PREDICTIVE INSIGHTS AND CASE STUDIES

MutagenesisPredictNet was utilized to model important aquatic organisms, the zebrafish (Danio rerio) and blue mussel (Mytilus edulis), both having important ecological relevance and characterized genomes. The prediction model was able to identify mutation pathways from common aquatic carcinogens, including benzo[a]pyrene (a PAH) and cadmium, and identified specific genes involved in DNA repair, apoptosis, and cell cycle regulation for the mutant targets.

For the benzo[a]pyrene model, predicted works identified increased probabilities for mutation pathways related to cytochrome P450 metabolism and nucleotide excision repair, which were consistent with experimental observations. Cadmium predictions modelled mutation pathways, which resulted in disruption of oxidative stress response pathways and metal binding proteins, which is consistent with known mechanisms of metal-induced genotoxicity.

Model outputs reflect biologically plausible scenario-specific susceptibility and bioavailability of pollutants in different environmental scenarios, presenting information that would allow prioritizing risk and monitoring of sensitive species and habitats.

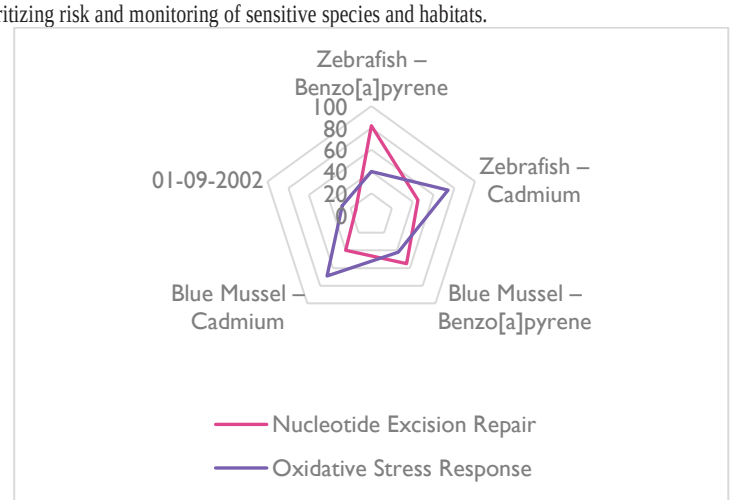

Figure 1: Comparative Mutation Probability Profiles for Key Pathways in Zebrafish and Blue Mussel Under Benzo[a]pyrene and Cadmium Exposure

Figure 1 illustrates the expected mutation probabilities in four major biological pathways for two aquatic species under two significant carcinogens. Benzo[a]pyrene predominantly affects nucleotide excision repair and cytochrome P450 metabolism, and cadmium predominantly alters oxidative stress and metal-binding protein production. The pathway profiles demonstrate species-specific susceptibility and pollution-specific mutation frequencies, which enhances the confidence in the accuracy of the MutagenesisPredictNet predictions.

Table 1: Predicted Mutation Probabilities for Key Pathways in Zebrafish and Blue Mussel under Benzo[a]pyrene and Cadmium Exposure Using MutagenesisPredictNet

Species	Carcinogen	Pathway Affected	Mutation Probability (%)
Zebrafish	Benzo[a]pyrene	Nucleotide Excision Repair	82
Zebrafish	Benzo[a]pyrene	Cytochrome P450 Metabolism	78
Zebrafish	Cadmium	Oxidative Stress Response	74
Zebrafish	Cadmium	Metal-Binding Protein Synthesis	70
Blue Mussel	Benzo[a]pyrene	Cytochrome P450 Metabolism	79
Blue Mussel	Benzo[a]pyrene	Nucleotide Excision Repair	55
Blue Mussel	Cadmium	Oxidative Stress Response	69
Blue Mussel	Cadmium	Metal-Binding Protein Synthesis	68

Table 1, which shows the predicted mutation probabilities from MutagenesisPredictNet in four particular biological pathways for zebrafish and blue mussel exposed to two main aquatic carcinogens. The predictions elucidate that benzo[a]pyrene was predicted to impact the DNA repair and metabolic pathways very strongly, whereas cadmium was predicted to affect the oxidative stress and metal detoxification pathways. These values provide quantitative evidence of species-specific and pollution-specific susceptibility traits.

There are limitations on model development derived from existing genomic databases for species (toxicogenomic) and pollutant (chemical pollutants) presence-absence, because they do not impede emerging contaminants or atypical (rare) mutations. Future refinements may include modelling incorporating molecular docking simulations, which improve the model's characterization of chemical-DNA interaction, and integrating toxicogenomic datasets for refining pathway-level mutation modelling and increased precision and ecological relevance.

CONCLUSION

MutagenesisPredictNet represents a significant advancement in the predictive Modeling of mutations from carcinogen-induced pathways in aquatic organisms. The predictive algorithm draws together data on pollution exposure with information on the genome, using sophisticated machine-learning and network analysis, to give predictions about species-specific mutation risk and disruptions in pathways. This represents a further determination of understanding the mechanisms of carcinogenesis in aquatic systems and is expanding how we can develop quantitative tools for ecological risk assessment. The predictive ability of MutagenesisPredictNet holds great significance for ecological monitoring and regulatory policy, potentially creating interventions to support the health of vulnerable species and ecosystems against exposures to carcinogens, to inform policies on the type and level of contaminants anticipated, the risk of shifts in species-community structures, and where mutation hotspot predictive capability might emerge in the environment. Making determinations on whether contaminant-based mutations may occur with presently expected exposures is challenging. MutagenesisPredictNet enables enhanced decision-making capability in assuring the health of the environment while addressing public health opportunities. Some limitations of MutagenesisPredictNet relate to relying on the existing genomic and pollution level datasets to parametrize some of the species ranges and the type of contaminants modelled. There are also challenges in fully dynamic interactions in the environment, as multi-omics (different DNA, RNA, and protein structural interactions and connections) will eventually need to be considered. This also includes the acknowledgement that the ultimate and cumulative effects of contaminant exposures that are imposed hierarchically within ecological interactions cannot yet be fully expressed here. Future work will focus on further developing the MutagenesisPredictNet framework to include more diverse aquatic species and multi-omic data types (like transcriptomics, proteomics). By design, the final framework might converge through experimental processes to provide the package with increasingly useful and robust information to support modelling adaptations the health of aquatic populations, using MutagenesisPredictNet, together with model-derived simulation and experimental testing programs that ideally integrate scales of exposure (global to regional – 'micro' and 'macro'), become increasingly impactful in expanding the biological relevance of predictive capabilities over time, and these scientific achievements will continue to strengthen the probability of further deterrent, and protective, developments.

REFERENCES

1. Habeeb, A. A., & Kazaz, Q. N. N. (2023). Bayesian and Classical Semi-parametric Estimation of the Balanced Longitudinal Data Model. International Academic Journal of Social Sciences, 10(2), 25–38. https://doi.org/10.9756/IAJSS/V10I2/IAJSS1010
2. Saidova, K., Madraimov, A., Kodirova, M., Madraimov, A., Kodirova, K., Babarakhimov, T., Urinova, S., & Zokirov, K. (2024). Assessing the impact of invasive species on native aquatic ecosystems and developing management strategies. International Journal of Aquatic Research and Environmental Studies, 4(S1), 45-51. https://doi.org/10.70102/IJARES/V4S1/8
3. Akyol, O., Çoker, T., Toprak, H. B., & Capapé, C. (2023). Capture of a Rare Smoothback Angelshark Squatina oculata (Squatinidae) in Turkish 2 Waters, with Updated Records from the eastern Mediterranean Sea. Natural and Engineering Sciences, 8(1), 38-45. http://doi.org/10.28978/nesciences.1281611
4. Saidov, A., Yakhshieva, Z., Makhkamova, N., Gudalov, M., Djuraeva, N., Umirzaqov, o., Adilova, o., & Juraev, A. (2024). Examining environmental impact through geological interactions and earth's layers. Archives for Technical Sciences, 2(31), 230–239. https://doi.org/10.70102/afts.2024.1631.230
5. Whitmore, J., & Fontaine, I. (2024). Techniques for Creating, Extracting, Separating, and Purifying Food and Feed Using Microalgae. *Engineering Perspectives in Filtration and Separation, 2*(4), 28-33.
6. Ansari, H., & Parmar, J. (2024). Tracing Human Evolution through Ancient DNA: Insights from Paleogenomic Studies. *Progression Journal of Human Demography and Anthropology*, *2*(3), 13-16.

7. Nair, R., & Reddy, D. (2024). Sustainable Food Systems for Reducing Food Waste and Loss. *International Journal of SDG's Prospects and Breakthroughs*, *2*(4), 18-23.
8. Nabende, P. (2025). Design of a Bio-Inspired Legged Robot for Rough Terrain Locomotion. *Association Journal of Interdisciplinary Technics in Engineering Mechanics*, *3*(2), 14-17.
9. Sharma, N., & Rajput, A. (2024). Development of A Genomic-based Predictive Model for Warfarin Dosing. *Clinical Journal for Medicine, Health and Pharmacy*, *2*(2), 11-19.
10. Ziwei, M., & Han, L. L. (2023). Scientometric Review of Sustainable Land Use and Management Research. *Aquatic Ecosystems and Environmental Frontiers*, *1*(1), 21-24.
11. Panneerselvam, K., Al-Farouni, M. H., Ganesan, A., Harini, B., & Velanganni, R. (2025). AI-driven mobility prediction for efficient session continuity. Journal of Internet Services and Information Security, 15(2), 942–952. https://doi.org/10.58346/JISIS.2025.I2.062
12. Xue, M. (2024). Assessing the Recreational Fishers and their Catches based on Social Media Platforms: Privacy and Ethical Data Analysis Considerations. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 15(3), 521-542. https://doi.org/10.58346/JOWUA.2024.I3.033
13. Sulfath, K. K., Ramakrishnan, P. R., Shareef, P. M., & Shanmugam, H. (2025). Enhancing IT Service Management in Indian IT Organizations: A Technological Integration of ISO 20000 with AI, Blockchain, Predictive Analytics, and Zero Trust Security. Indian Journal of Information Sources and Services, 15(1), 267–273. https://doi.org/10.51983/ijiss-2025.IJISS.15.1.34
14. Klein, D., & Dech, S. (2024). The Role of Big Data Analytics in Enhancing Customer Relationship Management. International Academic Journal of Innovative Research, 11(3), 27–33. https://doi.org/10.71086/IAJIR/V11I3/IAJIR1121
15. Kumar, P., & Tiwari, R. (2024). AI-Based Predictive Maintenance for Industrial IOT Applications. International Academic Journal of Science and Engineering, 11(4), 10–15. https://doi.org/10.71086/IAJSE/V11I4/IAJSE1164