AmesNet: A Deep Learning Model Enhancing Generalization in Ames Mutagenicity Prediction

Authors
Tyler Umansky[1], Virgil Woods[1], Sean M. Russell[1], Daniel Haders[1]**

Affiliations
[1]Model Medicines

**Corresponding author

Abstract
For a drug candidate to make it past the preclinical trial stage, mutagenicity must be ruled out. In the preclinical pipeline, drug mutagenicity is assessed using the Salmonella/microsome mutagenicity assay, colloquially called the Ames test. In nearly all instances, the FDA necessitates the Ames test for IND submission and drug approval. Difficult to scale up, the Ames test has given way to artificial intelligence (AI)-driven computational alternatives, which predict strain-specific Ames mutations in quantitative structure-activity relationship (QSAR) experiments. Reported in a recent study, a multitask Ames QSAR algorithm, which predicts eight strain-specific mutations, matched state-of-the-art performance as measured by balanced accuracy. But, like most other neural network applications used to predict Ames genotoxicity, these multitask algorithms struggle with sensitivity and fail to generalize efficiently to out-of-domain data. We recently developed ChemPrint, a molecular geometric convolutional neural network, which generalizes well to out-of-domain data. Validated across multiple protein targets in our drug discovery programs, ChemPrint achieved a 46% (19/41) zero-shot hit rate in our AXL and BRD4 oncology program, a 100% (12/12) one-shot hit rate in our RdRp Thumb-1 antiviral program, and enabled off-target binding reductions ranging from 800-fold to over 15,000-fold. We report here AmesNet, an adaptation of ChemPrint, which predicts strain-specific Ames mutations. AmesNet learns from the concatenation of ChemPrint's molecular embeddings with Ames mutagenicity features. When benchmarked against leading Ames QSAR algorithms, AmesNet achieves state-of-the-art balanced accuracy and sensitivity across both in-domain and out-of-domain data. AmesNet attained a sensitivity of 0.75, marking a 30% improvement over the best-performing ML method on the test set. AmesNet is embedded in GALILEO, our AI-based drug discovery platform. The GALILEO ecosystem, which also includes bioinformatics and generative chemistry tools, will continue to grow. ChemPrint and AmesNet are the beginning in a series of AI tools that will greatly accelerate and derisk the drug discovery pipeline, by predicting every aspect of the drug safety landscape assessed in the preclinical setting—absorption, distribution, metabolism, excretion, and toxicity (ADMET).

Keywords
Ames test, multitask QSAR, artificial intelligence, ChemPrint, AmesNet, molecular geometric convolutional neural network, adaptive molecular embeddings, t-SNE, and ADMET.

Introduction

Many potential drugs with promising targeted activity do not make it into clinical studies because of unfavorable *in vivo* interactions. Measurements such as absorption, distribution, metabolism, excretion, toxicity (ADMET), and chemical mutagenicity are a bottleneck in the drug discovery pipeline. Mutagenicity is typically assessed experimentally only after a significant investment of time and money in optimizing on-target activity. Mutagenicity is screened for in the preclinical stage using the Salmonella/microsome mutagenicity test, colloquially known as the Ames test[1]. The Ames test is required for IND submission and drug approval by the FDA in nearly all cases[2]. The study must follow GLP guidelines and be conducted using five different bacterial strains, each tested with and without S9 metabolic activation, for a total of 10 tests[2]. This mutagenicity assay comprises a battery of histidine auxotroph *Salmonella* strains, which have a reversible mutation in the *his* gene that regulates histidine biosynthesis[1,3]. Compounds are tested in the presence or absence of a liver fraction (S9), which mimics metabolic activation in the human body. In each *Salmonella* strain, the mutation is either a single substitution or a frameshift with a unique sensitivity to mutagens[1,3]. If mutagenic, a drug candidate will revert specific mutation types. Having recovered histidine synthesis capabilities, the revertant *Salmonella* strains grow in histidine-free media (**Figure 1A**)[1,3].

The Ames test is expensive and difficult to scale, costing more than $10,000 per compound under GLP conditions[4]. This makes it impractical for routine screening of large compound libraries in drug discovery campaigns[5,6]. An alternative to the Ames test is quantitative structure-activity relationship (QSAR) algorithms, which can leverage deep learning networks to predict strain-specific Ames mutagenicity[7]. Most Ames QSAR algorithms are limited in scope as they restrict their predictions to one strain-specific mutation type (**Figure 1B**)[8]. Published recently, a multitask Ames QSAR algorithm outperforms its single-task predecessors in balanced accuracy and sensitivity, by predicting the eight strain-specific mutation types commonly assessed in the Ames test (**Figure 1C, D**)[9]. Despite these advances, current Ames QSAR algorithms demonstrate low sensitivity and generalize poorly to out-of-domain data[10,11,12]. Ames QSAR algorithms with low sensitivity are particularly concerning, as they produce more false negatives. A false negative in this context is the most serious type of error, as it allows a genotoxic compound to advance further in the drug discovery pipeline undetected. This results in wasted resources and costly follow-up studies. More critically, if such a compound also escapes detection in later Ames testing, it may advance to clinical trials, where it could lead to patient harm. Additionally, Ames QSAR algorithms show reduced performance when predicting genotoxicity for novel compound classes that were absent from the training set, a challenge known as out-of-domain generalization. This drop in performance on novel chemical structures greatly reduces the model's practical utility and trustworthiness, as it fails to provide reliable predictions for the majority of untested compounds encountered in real-world drug discovery. Without the ability to generalize beyond its training data, the model offers little value in screening efforts, where identifying previously unrecognized genotoxic liabilities is critical. These deficiencies were evident in the first and second Ames/QSAR International Challenge Projects, where widely used models such as Sarah Nexus and Derek Nexus, as well as newer models like the FDA's DeepAmes, demonstrated limited success[10,11,12].

Suboptimal deep learning algorithms in drug discovery lead to costly and misguided decisions, producing inaccurate predictions that fail to detect critical molecular properties, as seen in the current state of Ames QSAR algorithms. To address this problem in our drug discovery programs, we developed GALILEO, a computational drug development platform with AI-based generative chemistry capabilities[15]. Embedded in the GALILEO ecosystem, ChemPrint is a molecular geometric convolutional neural network that uniquely generalizes out-of-domain data effectively[15]. ChemPrint has been extensively validated across multiple protein targets within our drug discovery programs. In out-of-domain settings, it achieved a 46% (19/41) zero-shot hit rate in our AXL and BRD4 oncology program and a 100% (12/12) one-shot hit rate in our RdRp Thumb-1 antiviral program. Additionally, ChemPrint enabled the elimination of off-target binding, achieving selectivity improvements ranging from 800-fold to over 15,000-fold[15,16,17,18]. ChemPrint's out-of-domain generalization was achieved through adaptive molecular embeddings and a hyperparameter optimization strategy on t-distributed Stochastic Neighbor Embedding (t-SNE) data splits. ChemPrint is a moldable algorithm and it can perform different tasks when trained accordingly, without losing predictive power[15,16,17,18].

Here, we report AmesNet, an innovative version of ChemPrint, which predicts strain-specific Ames mutations. AmesNet learns to predict compound mutagenicity from the concatenation of ChemPrint's adaptive molecular embeddings to Ames-specific features (**Figure 2**). To assess AmesNet's performance, we compared it to a set of single-task and multitask Ames QSAR algorithms reported recently (**Figure 1B-D**)[9]. Our results showed that when predicting features from in-domain or out-of-domain data, AmesNet outperforms the comparator Ames QSAR algorithms in balanced accuracy and sensitivity. Like ChemPrint, which we have previously shown can generalize well, we expect AmesNet to behave in a similar way. ChemPrint and AmesNet are the first in a series of deep learning algorithms we plan to populate the GALILEO ecosystem with. Once finished, GALILEO will make drug discovery projects safer, faster, and cheaper.

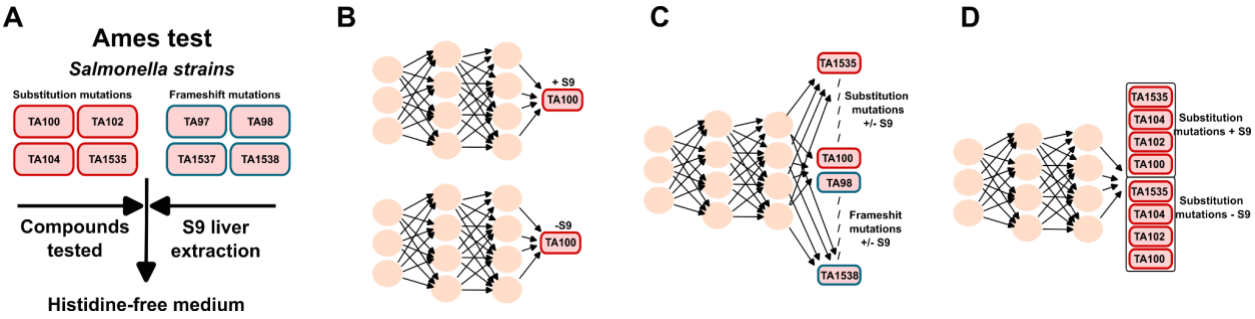


Figure 1. Ames test and predictive algorithms. A) Diagram depicting the Ames test, where the rectangles represent eight strains commonly used. Compounds are tested in the presence or absence of a liver fraction (S9), which mimics metabolic activation in the human body. **B)** Single-task Ames QSAR algorithm (ST) trained to predict a strain-specific feature (in this case, strain TA100), with or without metabolic activation (S9). **C)** Ungrouped multitask Ames QSAR algorithm (uMT), trained to predict eight strain-specific features, with or without metabolic activation (S9). **D)** Grouped multitasked Ames QSAR algorithm, trained to predict only substitution strains-specific features with or without metabolic activation (S9).

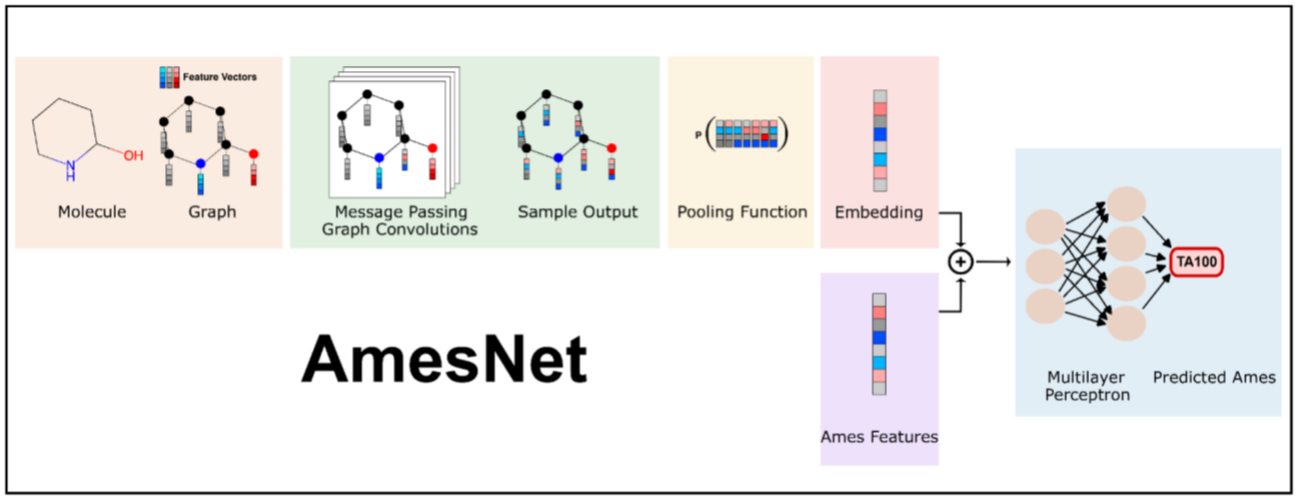


Figure 2. The AmesNet Architecture. AmesNet shares the same graph convolutional neural network as ChemPrint and also contains the addition of Ames specific features.

Results and Discussion

The goal of this study was to benchmark AmesNet against the three Ames QSAR algorithms reported recently and described above in **Figures 1B-D**: single task (ST), ungrouped multitask (uMT), and grouped multitask (gMT)[9]. All data used in this study were sourced directly from the reference publication[9]. The test datasets, both in-domain and out-of-domain, remained unchanged across all experiments to ensure consistency in evaluation. We trained and optimized AmesNet models using two different approaches. In the first approach, we used the same training and validation datasets as those reported in the reference publication[9]. In the second approach, we partitioned the training and validation datasets using our unique t-SNE-based methodology. See the "*Methods*" section below for a detailed explanation of the computational workflow. Four metrics were used in these comparisons: balanced accuracy, sensitivity, specificity, and Area Under the Receiver Operating Characteristic Curve (AUROC).

When using the in-domain datasets from the previous publication, AmesNet outperformed the three Ames QSAR models used as comparators (ST, uMT, and gMT). AmesNet performed even better when using the t-SNE approach to partition in-domain data **(Figure 3 and Table 1)**. These results mean that in the case of in-domain data, AmesNet tuned on the reference paper's data split is enough to outperform current QSAR algorithms in balanced accuracy and sensitivity. The t-SNE data split for hyperparameter optimization offers AmesNet an extra advantage over the Ames QSAR comparators by increasing balanced accuracy and sensitivity rate. Like ChemPrint, AmesNet was designed for maintaining predictive power on out-of-domain data. When compared against the three Ames QSAR comparators on out-of-domain data, AmesNet was superior, with and without using our t-SNE hyperparameter optimization approach (**Figure 4 and Table 2**). AmesNet attained a balanced accuracy of 0.82 and a sensitivity of 0.75, representing improvements of approximately 10% and 30%, respectively, over the best-performing ML method on the test set. This ability to generalize on unseen data with a high balanced accuracy and sensitivity is what makes AmesNet a unique algorithm. Sensitivity is the most critical metric in Ames mutagenicity prediction, as false negatives, or missed detections of genotoxic compounds,

pose the greatest risk. A model with low sensitivity fails to identify hazardous compounds early, allowing them to advance through the drug discovery pipeline undetected. This leads to wasted resources, costly follow-up studies, and an increased risk of genotoxic compounds progressing into clinical trials, where late-stage failures are significantly more consequential. Again, the t-SNE data split for hyperparameter optimization provides AmesNet with an extra generalization advantage, which is more pronounced when presented with out-of-domain data. These results put AmesNet ahead of the comparator QSARs, even those that competed in this year's Ames/QSAR International Challenge Projects[9,10,11,12].

Analysis of compound structures in the out-of-domain test data reveals a distinct advantage of AmesNet over the gMTL model: its ability to detect genotoxic polyaromatic compounds. Polyaromatic compounds are a well-known source of false negatives in Ames QSAR mutagenicity prediction and have proven difficult for models to classify accurately[13,14]. AmesNet correctly identified all polyaromatic compounds in **Figure 5** as genotoxic. In comparison, the gMTL model misclassified all compounds in **Figure 5** as non-genotoxic, failing to detect their mutagenic potential. AmesNet's performance on the polyaromatic compound class demonstrates the impact of its improved sensitivity. These results help validate AmesNet's ability to address limitations in existing models and improve predictive accuracy for high-risk chemical classes.

AmesNet is a variation of ChemPrint, which we described and benchmarked recently. In previous studies we showed that ChemPrint, which leverages our t-SNE splitting approach for data training and hyperparameter optimization, exhibits remarkable superiority over existing algorithms. Its ability to generate highly selective ligands with potent on-target activity while minimizing off-target interactions highlights its versatility in drug discovery. ChemPrint has demonstrated strong predictive performance across multiple programs, achieving a 46% (19/41) zero-shot hit rate in our AXL and BRD4 oncology program and a 100% (12/12) one-shot hit rate in our RdRp Thumb-1 antiviral program. Additionally, it significantly improved selectivity, reducing off-target binding by factors ranging from 800-fold to over 15,000-fold[15,16,17,18]. These results showed that ChemPrint can be trained to perform different functions (identify ligands, and thereafter remove those with polyaffinity features) within the same drug discovery pipeline. Leveraging ChemPrint's unique t-SNE splitting approach was instrumental in this case. Without the t-SNE splitting approach, ChemPrint would have failed to generalize beyond known chemical entities.

Mutagenicity is one of many toxicity features that must be assessed in the preclinical setting. Having successfully developed the AmesNet algorithm, we will soon introduce a suite of tools capable of filtering out additional toxicity and ADME features commonly identified in preclinical studies. These tools will allow for the optimization and refinement of molecular candidates by eliminating non-specific activity. This capability can be extended to any known systemically toxic drug-protein interaction, allowing for precise optimization of therapeutic selectivity. These algorithms, once incorporated into the GALILEO toolkit, will expedite the drug discovery process and minimize both time and financial expenditures.

Table 1. Averaged In-Domain Test Set Prediction Values by AmesNet versus Multitask QSAR algorithms. Values are presented with 95% confidence intervals in parentheses.

DL Architecture	Balanced Accuracy	Sensitivity	Specificity	ROC AUC
STL (reference)	0.610 (0.536–0.689)	0.273 (0.128–0.423)	0.948 (0.911–0.979)	0.745 (0.635–0.842)
uMTL (reference)	0.754 (0.662–0.842)	0.571 (0.391–0.738)	0.937 (0.895–0.972)	**0.894 (0.831–0.944)**
gMTL (reference)	0.771 (0.676–0.857)	0.603 (0.425–0.767)	0.940 (0.898–0.975)	0.875 (0.800–0.934)
AmesNet - base HPO	0.789 (0.772–0.806)	0.620 (0.586–0.653)	**0.959 (0.953–0.965)**	0.880 (0.864–0.894)
AmesNet - t-SNE HPO	**0.796 (0.780–0.812)**	**0.668 (0.635–0.700)**	0.924 (0.916–0.932)	0.878 (0.863–0.893)

The best learning architecture for each strain task, whether STL, uMTL, gMTL, or AmesNet.

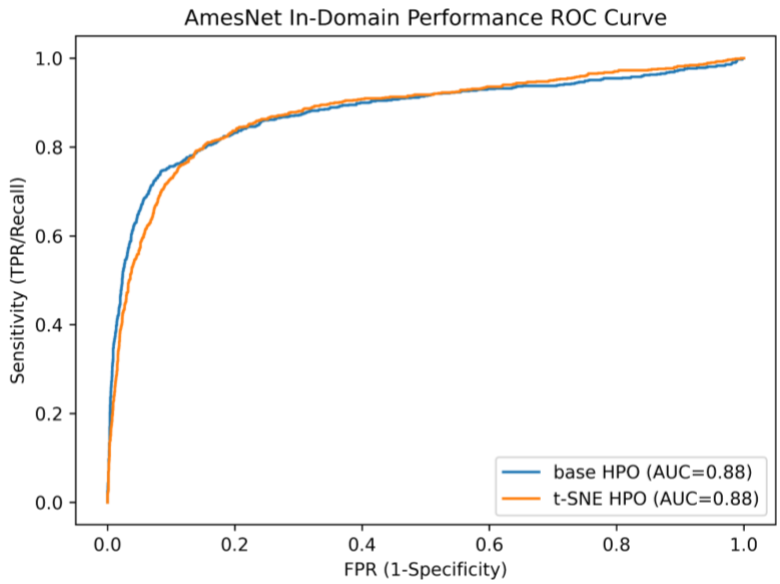


Figure 3. Area Under the Receiver Operating Characteristic Curve (AUROC) for AmesNet trained on in-domain data. AmesNet model 1, which used random splitting with chemical diversity of in-domain data, is in blue (base HPO), while AmesNet model 2, which used t-SNE splitting with chemical diversity of in-domain data, is in orange (tSNE-HPO).

Table 2. Averaged Out-of-Domain Test Set Prediction Values by AmesNet versus Multitask QSAR algorithms. Values are presented with 95% confidence intervals in parentheses.

DL Architecture	Balanced Accuracy	Sensitivity	Specificity	ROC AUC
STL (reference)	0.700 (0.592–0.813)	0.448 (0.238–0.675)	**0.951 (0.913–0.979)**	0.799 (0.664–0.915)
uMTL (reference)	0.716 (0.602–0.828)	0.542 (0.317–0.749)	0.890 (0.834–0.943)	0.823 (0.713–0.913)
gMTL (reference)	0.745 (0.628–0.858)	0.575 (0.347–0.788)	0.915 (0.863–0.957)	0.832 (0.721–0.923)
AmesNet - base HPO	0.788 (0.771–0.808)	0.669 (0.636–0.708)	0.907 (0.898–0.916)	0.861 (0.843–0.879)
AmesNet - t-SNE HPO	**0.818 (0.800–0.835)**	**0.749 (0.718–0.783)**	0.886 (0.877–0.896)	**0.882 (0.865–0.900)**

The best learning architecture for each strain task, whether STL, uMTL, gMTL, or AmesNet.

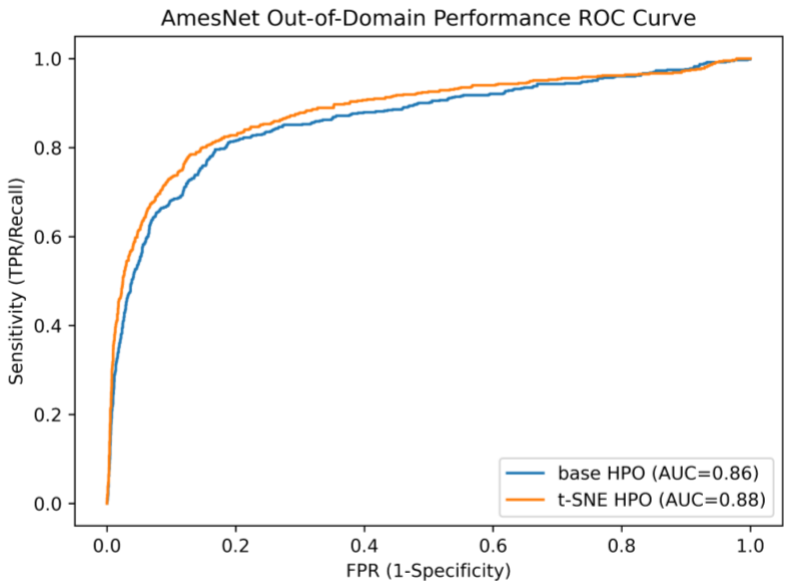
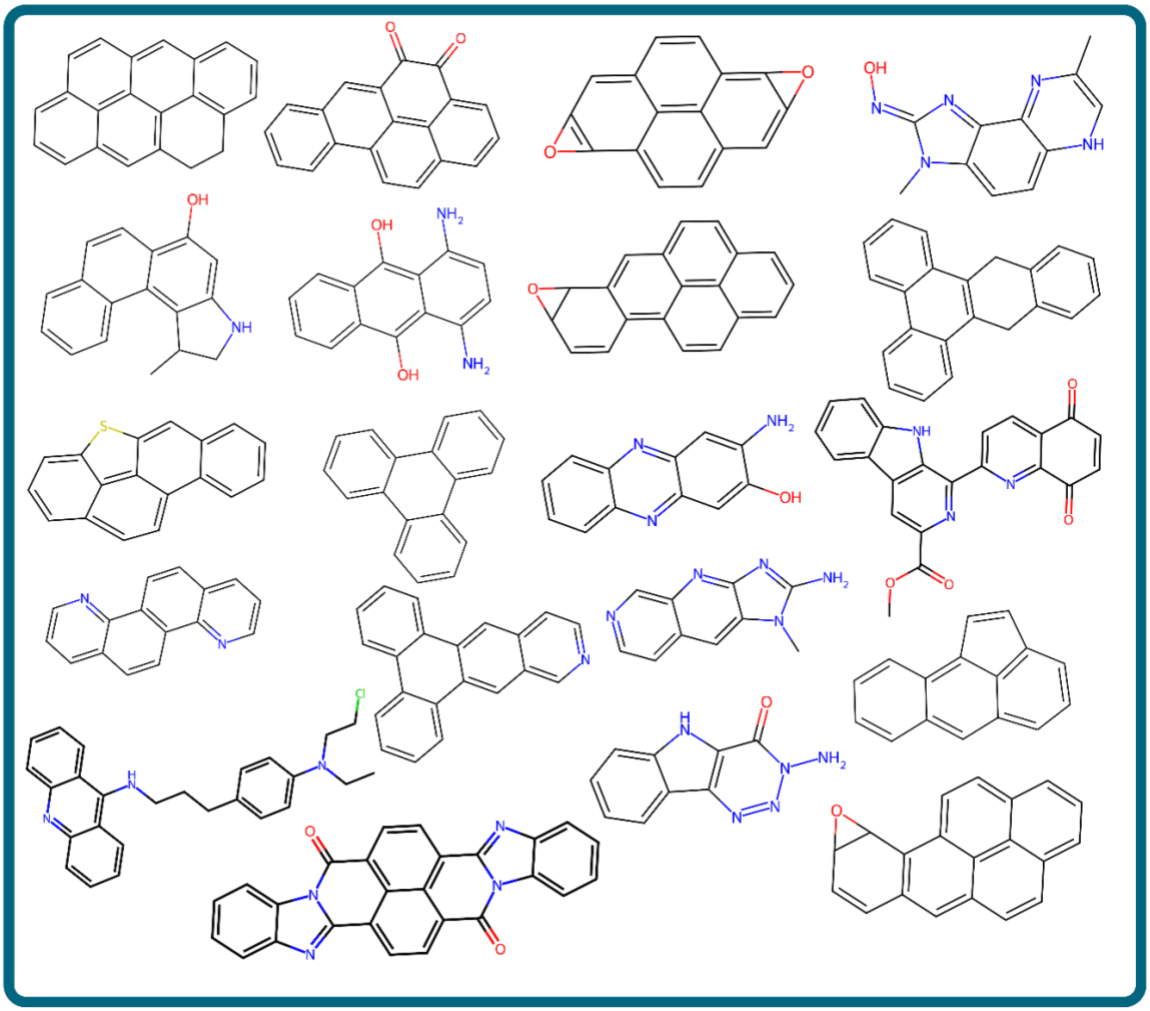


Figure 4. Area Under the Receiver Operating Characteristic Curve (AUROC) for AmesNet trained on out-of-domain data. AmesNet model 3, which used random splitting with chemical diversity of out-of-domain data, is in blue (base HPO), while AmesNet model 4, which used t-SNE splitting with chemical diversity of out-of-domain data, is in orange (t-SNE HPO).

Figure 5. Polyaromatic Compounds Correctly Classified as Genotoxic by AmesNet but Misclassified by the gMTL Model



Conclusions

We show in this paper that AmesNet, an adaptation of ChemPrint for the prediction of compound mutagenicity, outperforms existing multitask Ames QSAR algorithms (**Figure 2, Table 1, Table 2**). Sensitivity is the most important metric in mutagenicity prediction, as low sensitivity increases the risk of false negatives, allowing genotoxic compounds to advance undetected through the drug discovery pipeline. AmesNet differentiates itself from existing Ames QSAR models by achieving superior sensitivity while maintaining high balanced accuracy, minimizing the likelihood of toxic compounds evading detection. AmesNet marks the beginning of more preclinical models to come, which will populate the GALILEO ecosystem, for the benefit of our AI-driven drug development pipeline. In fact, the simplicity with which we turned ChemPrint into a predictor of compound mutagenesis, puts us on course toward targeting other equally important safety features assessed in the preclinical stage. For example, graph convolutional neural network

QSAR algorithms already exist for the prediction of drug absorption, distribution, metabolism, and excretion (ADME) features[19]. These tools, however, lack AmesNet's predictive power when presented with out-of-domain data. Concatenating ADME-specific features to adaptive molecular embeddings will populate GALILEO with a battery of tools devoted to drug safety in preclinical AI-based drug assessment. These models will complement our already published discovery capabilities, where ChemPrint has demonstrated exceptional generalization to out-of-domain data, achieved high hit rates across oncology and antiviral targets, and enabled significant off-target binding reductions.

Methods

Data availability

We attempted to obtain the dataset from the Ames/QSAR International Challenge Project by reaching out to the authors multiple times but did not receive a response[11,12]. As a result, we used the best available publicly accessible dataset for this study[9].

Dataset source

For model training we used the reference dataset described in the publication we are comparing AmesNet against, and for which the authors provided a link in their GitHub repository[9]. https://github.com/luiraym/gmtames

Compiled from four individual data sources (ISSSTY, OASIS, EFSAP, and MHLW), downloaded from the OECD QSAR Toolbox, the dataset used in this study for model training, includes Ames annotations from eight Salmonella typhimurium strains, each indicative of a specific mutation type: TA100 (base-pair substitution type), TA102 (base-pair substitution type), TA104 and TA1535 (base-pair substitution type); TA1537 (frameshift type); and TA1538, TA97, and TA98 (frameshift type). Tested in the presence or absence of S9 fraction, which is an extract that mimics compound activation in the liver.

https://github.com/luiraym/gmtames/blob/main/gmtames/data/master_datasets/master_dataset_curation/gmtamesQSAR_structures.sdf

Dataset clean up approach

The small molecules in the dataset were converted into SMILES format and standardized using a custom methodology. High-quality SMILES strings were obtained having canonical tautomerism and consistent stereochemistry throughout the dataset used in this study. After SMILES standardization, the data was merged with the pre-partitioned in-domain and out-of-domain datasets using the compound identifier. The merged dataset was further cleaned by organizing strain and S9 condition information into separate columns. Both the in-domain and out-of-domain datasets, each containing 48,735 rows, were prepared for model training, validation, and testing according to the predefined data splits.

Dataset partition approach

For the data partition we used our t-SNE chemical diversity split. The splitting proportions were 45% training, 45% validation, and 10% test. The partition proportions used in the comparator publication were 70% training, 20% validation, and 10% test.

Code Availability: GitHub repository for the t-SNE splitting method is archived at https://github.com/Model-Medicines/tSNE-Chemical-Data-Splitting
In-domain Partitioning CSV File:
https://github.com/luiraym/gmtames/blob/main/gmtames/data/master_datasets/gmtamesQSAR_endpoints_maxmin.csv
Out-of-domain Partitioning CSV File:
https://github.com/luiraym/gmtames/blob/main/gmtames/data/master_datasets/gmtamesQSAR_endpoints_scaffold.csv

Hyperparameter optimization approach
Grid search was used for hyperparameter optimization, using the highest area under the receiver operating characteristic curve (AUROC), value as the target parameter.

Model Retraining
Having obtained optimized hyperparameters, the four AmesNet models were retrained with the recombination of the train and validation data. The probabilities used were averaged over 10 models, and thereafter bootstrapped.

Reference Model Comparison
AmesNet was benchmarked against three reference models: single-task (ST), ungrouped multitask (uMT), and grouped multitask (gMT)[9]. The "overall" model was excluded as it benefits from post hoc selection, making it an ensemble-like construct rather than a single, directly comparable model. Although the "overall" model's results are not presented in our figures, AmesNet still achieves higher Balanced Accuracy and Sensitivity.

References

1. Ames, B. N., Lee, F. D. & Durston, W. E. An improved bacterial test system for the detection and classification of mutagens and carcinogens. *Proc Natl Acad Sci U S A* **70**, 782–6 (1973).
2. U.S. Food and Drug Administration. S2(R1) Genotoxicity Testing and Data Interpretation for Pharmaceuticals Intended for Human Use. FDA Guidance for Industry (2012).
3. Mortelmans, K. & Zeiger, E. The Ames Salmonella/microsome mutagenicity assay. *Mutat Res* **455**, 29–60 (2000).
4. Hsiao, Y., Su, B.-H. & Tseng, Y. J. Current development of integrated web servers for preclinical safety and pharmacokinetics assessments in drug development. *Brief Bioinform* **22**, 3 (2020).
5. Aljallal, M. A., Chaudhry, Q. & Price, N. R. Assessment of performance of the profilers provided in the OECD QSAR toolbox for category formation of chemicals. *Sci Rep* **14**, 18330 (2024).
6. Basile, A. O., Yahi, A. & Tatonetti, N. P. Artificial Intelligence for Drug Toxicity and Safety. *Trends Pharmacol Sci* **40**, 624–635 (2019).

7. Benigni, R. & Bossa, C. Data-based review of QSARs for predicting genotoxicity: the state of the art. *Mutagenesis* **34**, 17–23 (2019).
8. Varnek, A. *et al.* Inductive transfer of knowledge: application of multi-task learning and feature net approaches to model tissue-air partition coefficients. *J Chem Inf Model* **49**, 133–44 (2009).
9. Lui, R., Guan, D. & Matthews, S. Mechanistic Task Groupings Enhance Multitask Deep Learning of Strain-Specific Ames Mutagenicity. *Chem Res Toxicol* **36**, 1248–1254 (2023).
10. Li, T., Liu, Z., Thakkar, S., Roberts, R. A. & Tong, W. DeepAmes: A deep learning-powered Ames test predictive model with potential for regulatory application. *Regul. Toxicol. Pharmacol.* **142**, 105486 (2023).
11. Honma, M. *et al.* Improvement of quantitative structure-activity relationship (QSAR) tools for predicting Ames mutagenicity: outcomes of the Ames/QSAR International Challenge Project. *Mutagenesis* **34**, 3–16 (2019).
12. Furuhama, A. *et al.* Evaluation of QSAR models for predicting mutagenicity: outcome of the Second Ames/QSAR international challenge project. *SAR QSAR Environ Res* **34**, 983–1001 (2023).
13. Feeney, S. V., Lui, R., Guan, D. & Matthews, S. Multiple Instance Learning Improves Ames Mutagenicity Prediction for Problematic Molecular Species. *Chem Res Toxicol* **36**, 1227–1237 (2023).
14. Hillebrecht, A., Muster, W., Brigo, A., Kansy, M., Weiser, T. & Singer, T. Comparative Evaluation of in Silico Systems for Ames Test Mutagenicity Prediction: Scope and Limitations. *Chem Res Toxicol* **24**, 843–854 (2011).
15. Tyler Umansky, Navya Ramesh, Virgil Woods, Sean M. Russell, Davey Smith, Daniel Haders. ChemPrint: An AI-Driven Framework for Enhanced Drug Discovery. *https://www.biorxiv.org/content/10.1101/2024.03.22.586314v4* (2024).
16. Tyler Umansky, Virgil Woods, Sean M. Russell, David S. Garvey, Davey Smith, Daniel Haders. GALILEO Generatively Expands Chemical Space and Achieves One-Shot Identification of a Library of Novel, Specific, Next Generation Broad-Spectrum Antiviral Compounds at High Hit Rates. *https://www.biorxiv.org/content/10.1101/2025.01.17.633620v1* (2025).
17. Virgil Woods, Tyler Umansky, Sean M. Russell, David S. Garvey, Davey Smith, Daniel Haders. Discovery of RdRp Thumb-1 as a Novel Broad-Spectrum Antiviral Family of Targets and MDL-001 as a Potent Broad-Spectrum Inhibitor Thereof. https://www.biorxiv.org/content/10.1101/2024.03.29.587401v3 (2024).
18. Virgil Woods, Tyler Umansky, Sean M. Russell, Briana L. McGovern, Romel Rosales, M. Luis Rodriguez, Harm van Bakel, Emilia Mia Sordillo, Viviana Simon, Adolfo García-Sastre, Kris M. White, William F. Brubaker, Davey Smith, Daniel Haders. MDL-001: An Oral, Safe, and Well-Tolerated Broad-Spectrum Inhibitor of Viral Polymerases. https://www.biorxiv.org/content/10.1101/2025.01.13.632836v3 (2025).
19. Montanari, F., Kuhnke, L., Ter Laak, A. & Clevert, D.-A. Modeling Physico-Chemical ADMET Endpoints with Multitask Graph Convolutional Networks. *Molecules* **25**, (2019).