Journal of
Medicinal Chemistry

Article

Deep Learning Enhancing Kinome-Wide Polypharmacology Profiling: Model Construction and Experiment Validation

Xutong Li, Zhaojun Li, Xiaolong Wu, Zhaoping Xiong, Tianbiao Yang, Zunyun Fu, Xiaohong Liu, Xiaoqin Tan, Feisheng Zhong, Xiaozhe Wan, Dingyan Wang, Xiaoyu Ding, Ruirui Yang, hui hou, Chunpu Li, Hong Liu, Kaixian Chen, Hualiang Jiang, and Mingyue Zheng

J. Med. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.9b00855 • Publication Date (Web): 31 Jul 2019

Just Accepted

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

	Sciences, Liu, Hong; Shanghai Institute of Materia Medica Chinese Academy of Sciences, Center for Drug Discovery and Design Chen, Kaixian; Shanghai Institute of Materia Medica Chinese Academy of Sciences, Drug Descovery and Design Center Jiang, Hualiang; Shanghai Institute of Materia Medica Chinese Academy of Sciences, Drug Discovery and Design Center, State Key Laboratory of Drug Research; ShanghaiTech University, School of Life Science and Technology Zheng, Mingyue; Shanghai Institute of Materia Medica Chinese Academy of Sciences, Drug Descovery and Design Center

Deep Learning Enhancing Kinome-Wide Polypharmacology Profiling: Model Construction and Experiment Validation

Xutong Li,[†,#] Zhaojun Li,[†,∥] Xiaolong Wu, [†,§] Zhaoping Xiong, [†,‡] Tianbiao Yang,[†,#] Zunyun Fu,[†,#] Xiaohong Liu,[†,‡] Xiaoqin Tan,[†,#] Feisheng Zhong,[†,#] Xiaozhe Wan,[†,#] Dingyan Wang,[†,#] Xiaoyu Ding,[†,#] Ruirui Yang,[†,‡] Hui Hou, [†,⊥] Chunpu Li,[†] Hong Liu,[†] Kaixian Chen,[†,‡] Hualiang Jiang,[, †,‡] and Mingyue Zheng [*,†]*

[†]Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China.

[#]University of Chinese Academy of Sciences, No.19A Yuquan Road, Beijing 100049, China.

‡School of Life Science and Technology, ShanghaiTech University, 393 Huaxiazhong Road., Shanghai 200031, China.

§School of Pharmacy, East China University of Science and Technology, 130 Meilong Road, Shanghai 200237, China.

∥School of Information Management, Dezhou University, 566 West University Road, Dezhou 253023, China.

⊥School of Pharmacy, Fudan University, 220 Handan Road, Shanghai 201203, China

ABSTRACT

The kinome-wide virtual profiling of small molecules with high-dimensional structure-activity data is a challenging task in drug discovery. Here, we present a virtual profiling model against a panel of 391 kinases based on large-scale bioactivity data and the multitask deep neural network algorithm. The obtained model yields excellent internal prediction capability with an auROC of 0.90, and consistently outperforms conventional single-task models on external tests, especially for kinases with insufficient activity data.

Moreover, more rigorous experimental validations including 1,410 kinase-compound pairs showed a high-quality average auROC of 0.75 and confirmed many novel predicted "off-target" activities. Given the verified generalizability, the model was further applied to various scenarios for depicting the kinome-wide selectivity, and the association with certain diseases. Overall, the computational model enables us to create a comprehensive kinome interaction network for designing novel chemical modulators or drug repositioning and is of practical value for exploring previously less studied kinases.

Introduction

The protein kinase family is one of the largest enzyme families. The human kinome comprises more than 500 kinases, constituting approximately 1.7% of all human genes[1]. Dysregulation of protein kinases plays causal roles in numerous human diseases, including cancers, inflammatory diseases, central nervous system disorders, cardiovascular diseases, and complications of diabetes[2]. By December 2018, 48 small kinase inhibitors were approved by the U.S. FDA, approximately half of which were approved in the last five years. However, despite the success of the kinase inhibitor drugs,

some deficiencies cannot be ignored. On the one hand, only a small number of human kinases (approximately 80) have been successfully targeted by these drugs, and many kinase inhibitor drugs are used against the same targets in oncology[3]. More than 100 kinases (~25%) have completely unknown functions, and approximately 50% are largely uncharacterized, with little indication of how these targets influence the major signalling pathways[4]. Therefore, there are still many "untargeted" kinases and relevant diseases that require further investigation. On the other hand, many kinase inhibitor drugs have promiscuous profiles due to the conservation of the ATP binding site in the human kinome. As demonstrated by a recent study, even for kinase drugs in clinical use, the target space or bioactivities are still surprisingly poorly characterized[5]. Therefore, thorough characterization of the kinase spectrum of inhibitor drugs is important to explain undesired adverse effects and to enable drug repurposing, and the discovery of potent inhibitors with reasonable selectivity and polypharmacological profiles has become a very promising but challenging direction in novel drug development process[6].

A step towards accelerating the kinase drug discovery process is to quickly identify whether a compound interacts with a kinase. Recent advances in high-throughput

screening technologies have enabled the bioactivity profiling of hundreds of compounds against a panel of protein kinases[7], but the high cost and laborious process of chemical synthesis and biological characterization hinder its application in extensive chemical spaces. Many in silico modelling approaches have been developed to predict kinase inhibitory activity for large-scale compound libraries. Compared to traditional drug design methods (e.g., docking and standard QSAR), machine learning-based models such as naïve Bayesian (NB)[8], k-nearest neighbours (KNN) [9-10], random forest (RF)[11-13], support vector machine (SVM)[8-10, 14] and deep neural network (DNN)[15] have been established to predict a wider range of biological activities for a compound by employing high-dimensional datasets. Generally, these models have been separately trained with individual data sets relating to different tasks and often are more flexible and have high variance. For example, given a learning task with insufficient data, these models always show unsatisfactory predictive power and a tendency towards overfitting. In fact, standard machine learning algorithms are capable of learning meaningful chemical information from hundreds of compounds and become even more effective given additional data[16].

However, even a set of one hundred compounds is often an unavailable resource for building a machine learning model for some less studied or uncharacterized kinases.

In this study, we established a multitask deep neural network (MTDNN) classification model to predict the bioactivities of small molecules at a kinome-wide level, including kinases with both large and small amounts of available data are included. Previous studies of MTDNN applied to drug discovery suggested that MTDNN can obtain significantly better predictive accuracies than single-task methods for problems with multiple related tasks[17-21] due to the generalization ability and transfer learning effect of DNNs. These features make MTDNN an appropriate solution for bioactivity prediction against a spectrum of kinases with high homology to each other. Rodríguez-Perez et al. have showed that MTDNN resulted in a performance boost for other single-task and multi-task machine learning models in the prediction of highly potent inhibitors of 103 human kinases[22]. These results are illuminating, but a wider kinase panel and more rigorous experimental validations are in high demand. Here, the MTDNN model has been trained with over 140,000 bioactivity data points for 391 kinases, where one task corresponds to bioactivity prediction against a specific type of kinase. Extensive computational (on four

external datasets with significant sizes) and experimental validations (on 5 compounds against 282 kinases involving 1410 kinase-compound pairs) have been performed, and the MTDNN model consistently outperforms conventional single-task models in terms of auROC. In particular, it facilitates network information sharing across different kinases and compensates for the limited bioactivity data associated with any specific kinase, which enables reliable and comprehensive profiling of kinome-wide activity and reasonable estimations of overall and group-specific selectivity. We envisage that the MTDNN model could be used to quickly establish a comprehensive kinome interaction network for designing novel chemical modulators and for exploring previously less studied or untargeted kinases.

RESULTS AND DISCUSSION

Optimized MTDNN model

The predictive performance on the testing dataset of all the MTDNN models is summarized in Figure 1. The following points can be noted: (1) The influence of batch size in our search space is slight. (2) Models with very simple architecture show moderate performance, whereas those with complex structures are very sensitive to different

learning rates. In general, a more complex architecture means a more heavily parametrized model with a higher capacity, and such a model operates in a space with a larger dimension and has a more complex error surface than a thinner one. Thus, in such a complex error regime, the convergence process may either easily deviate from meaningful locations with a large learning rate or become trapped in local minima with a smaller learning rate. (3) Models with less favourable learning rate values always yield weak predictive power. A large learning rate might cause drastic updates leading to divergent behaviours, while a model with a small learning rate might require too many updates to achieve convergence. Here, models with moderate depth and fine-tuned learning rates can show satisfactory performance on the testing dataset, with auROCs higher than 0.89. With the layer size of [1024,391], the learning rate of 5×10^{-5} and the batch size of 128, on our internal test set, the model with the best performance has an auROC over 0.90, an F1-score of 0.74 and a BA of 0.83. This model was chosen for further external evaluations.

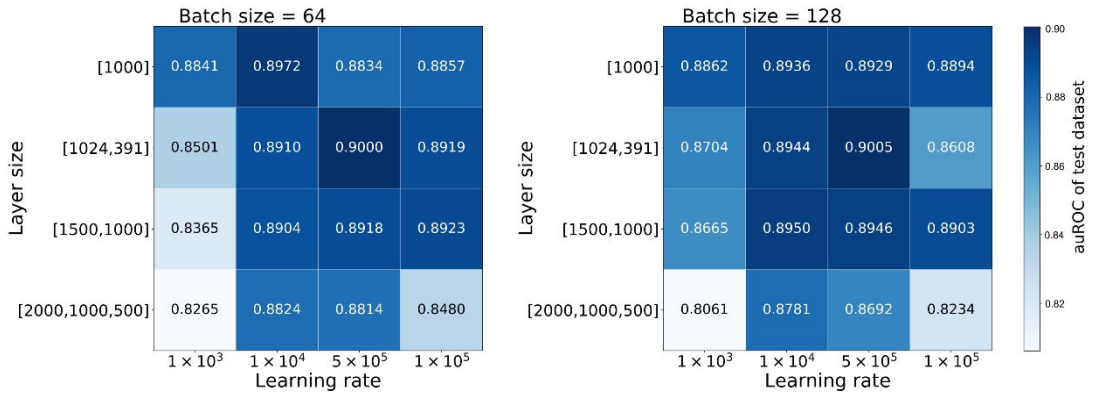


Figure 1. Model performances with different hyperparameters. The heat maps are coloured according to the auROC of the testing dataset, which is directly marked in each block. A darker colour corresponds to better performance.

Evaluation on external datasets

Despite the satisfactory predictive power of the internal testing dataset, it is valuable to investigate model performance on datasets that have different distributions from that of the modelled dataset to evaluate the generalization capability of the neural network. Thus, we prepared four different external datasets with diverse experimental methods and value types to evaluate the model.

The first dataset was published by Davis et al.[23], who tested 72 known inhibitors against a panel of 442 kinase assays, resulting in a total of 9,424 K_d values for all potential protein-

ligand pairs. The second dataset was provided by Anastassiadis et al.[24] and reported the inhibitory percentage inhibition values of 178 compounds tested against 300 kinases at 0.5 μM. The other two datasets are PKIS1[25] and PKIS2[26], also referred to as Published Kinase Inhibitor Set 1 and 2, respectively. PKIS1 is a collection of 367 kinase inhibitors representing 31 diverse chemotypes, and their inhibitory percentage values evaluated against 232 kinases at 1 μM were collected. PKIS2 is composed of 645 small-molecule inhibitors representing 86 diverse chemotypes, which were all profiled at a concentration of 1 μM against a broader panel of 392 kinases. For the dataset of Anastassiadis et al., the activity threshold is set by converting the single point activity with the equation previously defined[27]. For datasets PKIS1 and PKIS2, an inhibition rate over 50% at 1 μM was defined as signifying activity[28]. Note that all compound-kinase pairs included in our training dataset were removed from these external datasets before further evaluation.

As summarized in Table 1, the classification performances on these external datasets are lower than those from the testing dataset. One of the potential reasons for this performance degradation could be the chemical distribution deviations. We first investigated the performance for the compounds dissimilar to the training samples by a

5-fold cluster cross validation[29] (Table 1). Clearly, the predictive performance appreciably decreased. To provide an overview of the different distributions between the training dataset and four external datasets, we used t-SNE[30] to visualize high-dimensional feature vectors by projecting each sample into three-dimensional space. As shown in Figure 2, the compounds in the training set were compared with the samples in each external dataset. The Davis set is relatively small, and all of the samples are distributed within the chemical space of the training dataset. For the Anastassiadis set and PKIS1, the chemical space of most of the samples overlaps with that of the training dataset. In contrast, for PKIS2, a number of samples are scattered outside the training set space, suggesting a distinct data distribution. As a result, MTDNN achieves the worst performance on PKIS2 of which the chemical space shows apparent deviation from the training set.

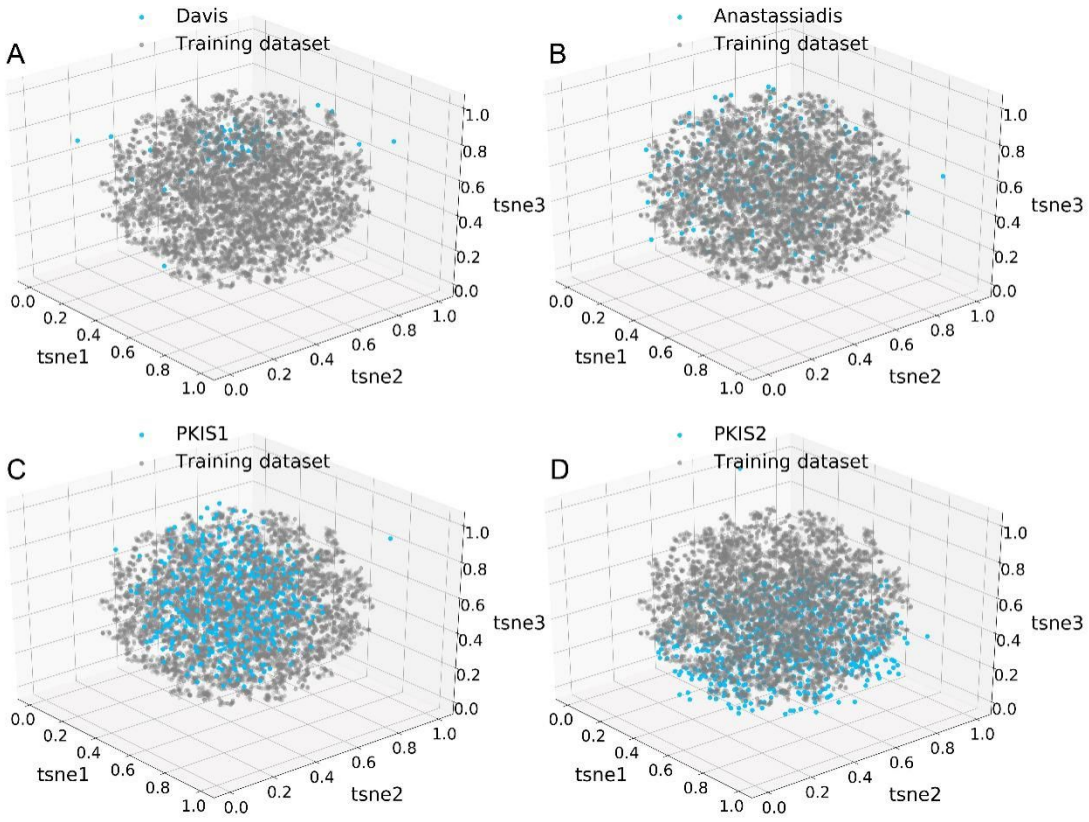


Figure 2. Visualizing datasets using t-SNE. The ECFP4 values of the compounds were used as input. The distribution of compounds from the training dataset (grey scatters) is compared to that from (A) the Davis dataset, (B) the Anastassiadis dataset, (C) PKIS1 and (D) PKIS2 (blue scatters).

One the other hand, the concordance of data from different sources could also limits the predictive performance. As highlighted by Sutherland et al., there is only modest concordance between assay panels from different sources, including the Davis set, the Anastassiadis set and part of our training set, i.e., the Metz set[27]. The discrepancies

between assays will certainly affect the predictive performance of the model on these external datasets, and it is therefore of interest to quantitatively analyse the data consistency between our training set and different external data sources. Towards this aim, the removed overlapping data between the training and external sources was used to perform diagnostic tests with auROCs, where a higher auROC value suggests a more consistent result between the different data sources. Specifically, given a molecule with multiple experimental data points relating to the same target, a TP is counted when both values in the training set and external sources are “active”. The auROCs of the Davis set, the Anastassiadis set, PKIS1 and PKIS2, with respect to our training set are 0.91, 0.82, 0.84 and 0.86, respectively. These results suggest that, despite the uncertainty in different sources and forms of activity measurements, significant consistency can be observed between our training and external data. Among them, the Davis set with the equilibrium dissociation constant data shares the highest concordance with our training data. The Anastassiadis set has the lowest concordance with the training set with an auROC of 0.82, and the MTDNN yielded a relatively low auROC of 0.72 even though it shows a similar distribution of chemical space to that of the training set.

Overall, the MTDNN still shows impressive performance on these external datasets. We supposed that by expanding the scale and increasing the diversity of the chemotypes of the training dataset, the generalization capability of the model could be further improved. To verify this point, with a series of increasing scales of modelling data, we built two series of models using random split and 5-fold cluster cross validation, respectively. The size of the training dataset was increased from 500 compounds to 25,000 compounds in increments of 2,500 compounds. As shown in Figure 3, for the random split validation, there is a continuous uptrend of the auROC curve due to the rising probability of finding similar compounds in the training set. While for the cluster cross validation, auROC curve stops to increase after modelling compounds reach 5,500. Nonetheless, the performances of cluster cross validation still demonstrated that the model is of decent predictive power for compounds with unseen structures. Moreover, we may find that the auROCs for external datasets of both series of models rise steadily with the increasing size of training dataset, which supports the argument that the expansion of the chemical space of training set will improve model generalization capability.

Table 1. Model performance on internal testing, cluster-cross-validation, and external validations.

	Testing dataset	Cluster-cross-	Davis	Anastassiadis	PKIS1	PKIS2
auROC	0.90	0.82±0.02	0.78	0.72	0.79	0.69
Recall	0.78	0.63±0.05	0.58	0.44	0.62	0.43
Precision	0.70	0.66±0.09	0.39	0.38	0.15	0.18
F1-score	0.74	0.64±0.07	0.47	0.41	0.24	0.26
BA	0.83	0.75±0.02	0.71	0.66	0.72	0.64

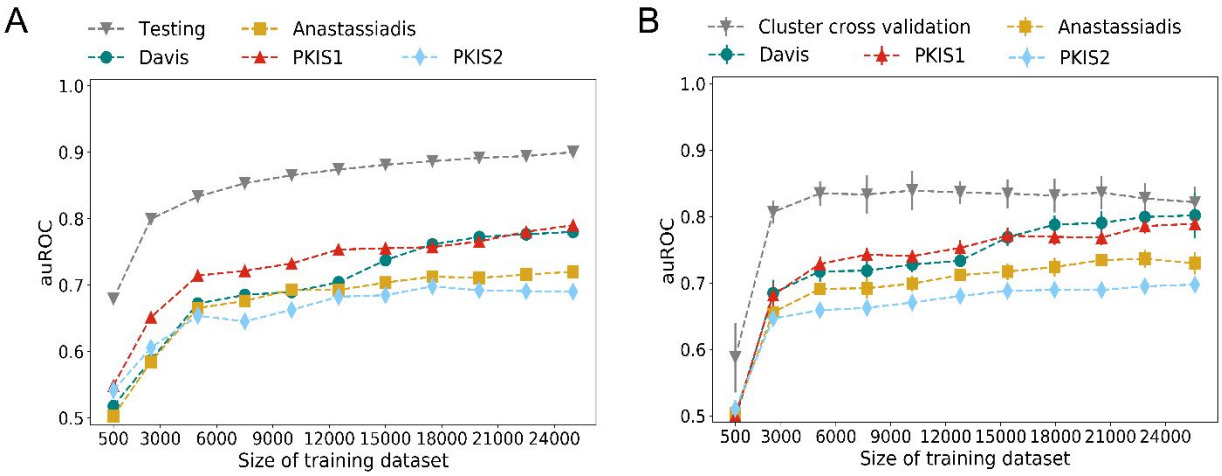


Figure 3. The auROC on internal validation and external testing datasets of models with increasing numbers of compounds in the training dataset. (**A**) Random split validation. (**B**) Five-fold cluster-cross-validation.

Comparison with previously reported models

In recent years, several models have been proposed for predicting the kinome-wide polypharmacological effect of small molecules [8-13, 31-32]. Niijima et al. developed a deconvolution approach to dissecting the Kinase SARfari database by which kinase−inhibitor pairs are represented by residues and fragments and built kinome-wide activity classification models with dual-component naïve Bayes (DCNB) and dual-component SVMs (DCSVMs)[8]. DCSVM achieved good performance on an internal validation set, showing an unbalanced accuracy over 0.85, but it did not achieve consistently good performance on external sets. Merget et al. reported ligand-based activity classification models for over 280 kinases by RF methods, based on an extensive dataset combining both proprietary and open bioactivity data[12]. The models yielded good prediction results with an average auROC of 0.76, and high quality (auROC > 0.7) was achieved for over 200 kinases. Janssen et al. presented Drug Discovery Maps (DDM), a t-SNE-based model that can map chemical and target space and predict the activities of novel kinase inhibitors across the kinome[32]. Using DDM, they discovered new inhibitors for FMS-like tyrosine kinase 3 (FLT3).

To verify the generalizability and transfer learning effect of our MTDNN model, a parallel comparison with previously reported models was carried out. Here, we chose the method of Merget et al. as a reference because it requires the same form of input and because its source code for deriving prediction models is available. Based on the same sets of training data used for MTDNN, we built a total of 391 single-task random forest models (RFs) with their provided code. First, we evaluated the RFs on our internal testing dataset, among which 226 kinases had auROCs higher than 0.7 and 134 kinases had auROCs higher than 0.8. These results agree well with the high-quality performance (i.e., auROC> 0.7 was achieved for ~200 kinases, of which 118 prediction models had an auROC ≥ 0.8) reported by Merget et al., confirming that these RFs are correctly established. In comparison with these high-quality RFs, we may notice that MTDNN further improved the predictive capability. As shown in Figure 4A, the MTDNN showed consistently higher average auROC values than those of the RFs on all four external datasets and achieved more high-quality (auROC>0.7) predictions than RFs.

For all individual tasks, we further investigated the relation between the size of the training datasets and their auROCs on the external datasets (Table S2). As shown in

Figure 4B, MTDNN demonstrated decent performance even when the number of training bioactivity data points for the kinases was quite small. In contrast, for these small-data tasks, RF can give the results only barely better than random prediction (auROC of ~0.5). Given a kinase with fewer than 100 training data points, the performance of MTDNN significantly exceeded that of RF. As the amount of training data increases, the gap between MTDNN and RF gradually narrows but remains. These results corroborate the transfer learning effect of MTDNN, which can take advantage of tasks with larger amounts of training data to improve the predictive performance of tasks with smaller training data. Obviously, the predictive capability for kinases without sufficient training data is an essential feature of our model, resulting in its practical value for developing a selective inhibitor for the previously less studied or untargeted kinases.

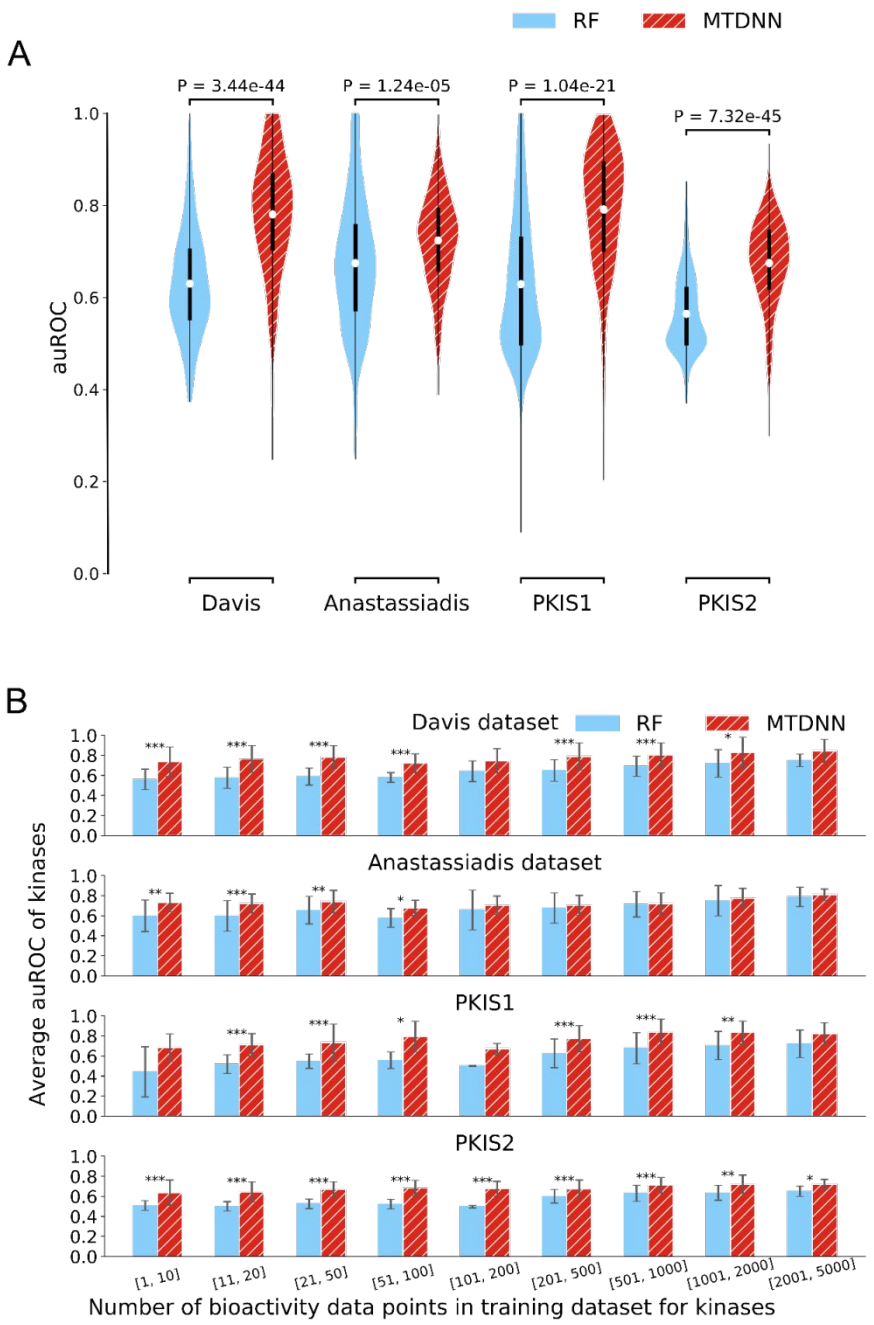


Figure 4. Prediction performance comparison between MTDNN and RF of Merget et al. (**A**) Violin plots of auROCs for kinases, grouped by underlying external datasets and coloured by method (RF in blue and MTDNN in red). The white dots are the average

auROCs. The upper and lower endpoints of the black segments are the first and the third quartile, respectively. The p-values of the t-tests are shown above the violin. (**B**) Bar plots of model performance on tasks with different amounts of bioactivity data in external datasets. A bar indicates the average auROC of tasks with the number of bioactivity data points within the underlying range. '*': 0.01<p<0.05; '**': 0.001<p<0.01; '***': p<0.001.

Experimental validation

In addition to retrospective model validation with previously reported data, we also performed more rigorous prospective experimental validation to investigate the practical benefits of the model for drug discovery or repurposing applications. Five kinase inhibitors with unknown activity profiling data and diverse structures were used for this validation, including the clinically approved drugs BAY73-4506 and K-115 and the investigational inhibitors NVP-BHG712, E31, and DC381016 (Table 2). BAY73-4506, K-115 and NVP-BHG712 were purchased from Selleck, E31 and DC381016 were provided by our co-workers[33-34]. These five kinase inhibitors have different primary targets under investigation, while their complete kinase spectra have not been reported by the time we

performed the experimental validation. All commercially available compounds were used without any further purification. Reversed-phase HPLC analyses were performed on an Agilent 1100 HPLC system with a DAD detector (area normalization). Purity of all compounds for biological evaluation was confirmed by HPLC to be >95%.

Figure 5 also summarizes the predicted kinase activity profile of the five molecules. To validate the prediction, the commercial KinaseProfiler Service (Eurofins Scientific, Inc.) was utilized to evaluate the experimental activity of each molecule against a panel of 405 kinases, among which 282 fall in our MTDNN capacity panel (391 kinases). Filter-binding radiometric kinase activity assays were performed at a concentration of 1 μM, with the active threshold defined as 50% (Table S3). It is gratifying to note that, for the 1,410 kinase-compound pairs in our MTDNN capacity panel, the model produced high-quality predictions that were generally in agreement with the experimental data with an average auROC of 0.75 (Table 2). Moreover, in addition to the activities reported previously, many novel “off-target” activities have been successfully predicted by the model and confirmed by subsequent experiments at lower compound concentrations.

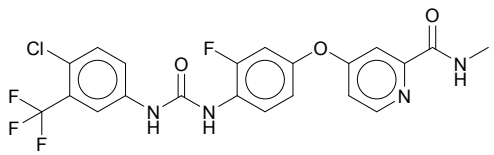
BAY73-4506 (Regorafenib) is an orally available antineoplastic agent launched for the treatment of colorectal cancer[35], gastrointestinal stromal tumour (GIST)[36] and hepatocellular carcinoma[37]. On the one hand, many previously published inhibitory activities, e.g., its known potent inhibitory effects against VEGFR1, VEGFR2, VEGFR3, PDGFRβ, Kit, RET and c-Raf[38], have been accurately identified by our model. On the other hand, there are inhibitory activities for several other kinases that have not been reported (Table 3). Examples include the following: (1) Targeting p38α, ZAK, and TAO1, which are involved in p38 MAPK signal transduction pathways. As a member of the p38 MAPK family that plays a critical role in cancer cell biology, p38α is an investigational target for the treatment of several types of cancer, including colorectal cancer and cardiovascular disorders. ZAK and TAO1 have been shown to trigger p38 MAPK activation[39]. (2) Targeting TrkB and TrkC, which are important therapeutic targets for neuroblastoma, non-small cell lung cancer (NSCLC) and colorectal cancer, for patients harbouring alterations in TRK expression or activity[40]. (3) Targeting DDR1, which has recently been identified as a new therapeutic target for colorectal cancer[41] and could also be a biomarker of epithelial ovarian cancer and a prognostic marker for NSCLC patients[42].

(4) Targeting LOK, which is a previously untargeted kinase with little indication of its function[25]. Overall, the successful prediction of kinase "off-targets" of BAY73-4506 provides additional explanations for its anticancer mechanisms of action and informs the search for new clinical applications. Moreover, identifying BAY73-4506 as a potent inhibitor of LOK may help to accelerate the related functional study of this untargeted kinase.

K-115 is a Rho-kinase inhibitor approved and launched in Japan as an ophthalmic solution for the treatment of glaucoma and ocular hypertension in 2014[43]. Several "off-target" inhibitory activities against AGC kinases, including PKG and PKC, that could contribute to clinical cancer therapy were revealed by the MTDNN model[44-46]. NVP-BHG712 is a specific EphB4 receptor inhibitor that blocks vascular endothelial growth factor-mediated angiogenesis in vivo[47]. In addition to c-Src, NVP-BHG712 was predicted to inhibit other Src family kinases (SFKs), including Lyn, Hck, Lck and Yes. Thus, NVP-BHG712 could potentially inhibit the proliferation and survival of cancer cells, especially pancreatic cancer and non-small cell lung cancer[48-49]. E31[34] was predicted to inhibit additional SFKs (Src, Yes, Blk, Lyn, Fgr, Lck), but the subsequent experiment excluded

the activity against Src. This false positive prediction may indicate molecular structural similarity with known Src inhibitors, which suggested that a relatively minor structural modification could improve the inhibitory activity of E31 against Src and thus increase its therapeutic potential. For DC381016, a compound designed and synthesized for c-Met inhibition[33], the MTDNN model also revealed inhibitory effects against several important or novel therapeutic kinases, including Abl, Trk, DDR1 and ACK1[50]. In summary, based on the impressive experimental verification results obtained here, in addition to its high throughput, we envisage that the MTDNN model can be used as a practical tool for large-scale target identification and drug redirection.

Table 2. Details for experimental compounds.

Compounds	Reported primary targets	Model performance				
		auROC	Recall	Precision	F1-score	BA
 BAY73-4506	VEGFR1, VEGFR2, VEGFR3, PDGFRβ, Kit, RET, c-Raf[38]	0.74	0.39	0.76	0.52	0.67

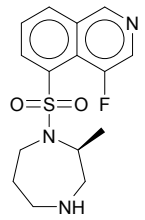
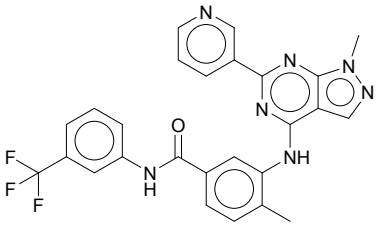
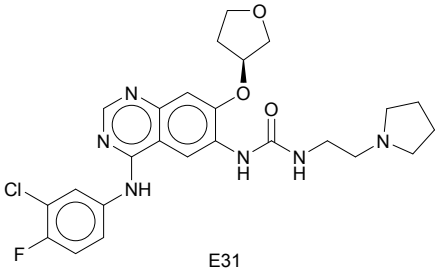
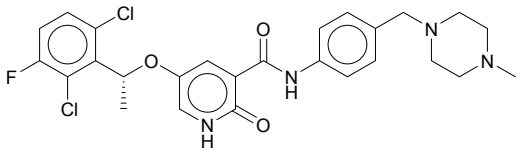
 K-115	ROCK1, ROCK2[51-52]	0.76	0.52	0.32	0.40	0.68
 NVP-BHG712	EphB4, c-Raf, c-Src, c-Abl[53]	0.86	0.69	0.41	0.51	0.80
 E31	RET, EGFR, ERBB2, ERBB4[34]	0.71	0.42	0.31	0.36	0.65
 DC381016	MET[33]	0.69	0.45	0.50	0.80	0.66
Average		**0.75**	0.49	0.46	0.45	0.69

Table 3. The predicted active probability and experimental %activity of BAY73-4506 against kinases.

UniProt ID	Kinase	Family	Predicted probability	%activity at 1 μM[a]	at 0.1 μM[b]	at 0.01 μM[c]
P10721	**Kit**[d]	TK	0.7795	-1		
P35916	**VEGFR3**	TK	0.9825	0	1	38
P07949	**Ret**	TK	0.9928	0	6	50
P17948	**VEGFR1**	TK	0.9885	3	3	43
P04049	**Raf1**	TKL	0.8489	8	57	93
P35968	**VEGFR2**	TK	0.9950	17	19	68
P09619	**PDGFRβ**	TK	0.9713	55		
Q16539	p38α	CMGC	0.5766	-6	52	100
P00520	Abl(m)	TK	0.9867	-1	-7	22
Q9NYL2	ZAK	TKL	0.9506	-1	2	27
Q8NE63	HIPK4	CMGC	0.9961	-1	3	31
Q9Y4K4	MAP4K5	STE	0.5848	0	37	104
P16277	Blk(m)	TK	0.5559	1	45	52
Q16620	TrkB	TK	0.9862	1	53	100
Q16288	TrkC	TK	0.9487	1	17	54
P07948	Lyn	TK	0.9373	1	21	32
P42685	PTK5	TK	0.9961	3	18	64
P29322	EphA8	TK	0.9835	5	44	85
P29317	EphA2	TK	0.9902	5	58	90
P15056	B-Raf	TKL	0.8673	8	66	101
P00519	Abl	TK	0.9352	8	49	79
P04049	c-RAF	TKL	0.8489	8	57	93
Q4JIM5	Arg(m)	TK	0.5631	9	29	97
P07333	Fms	TK	0.9437	10	36	73
Q08345	DDR1	TK	0.9612	11	43	92
Q9H422	HIPK3	CMGC	0.9742	12	68	103
Q9H2X6	HIPK2	CMGC	0.7037	13	61	96
O15146	MuSK	TK	0.9955	21	92	89
P36888	Flt3	TK	0.9801	31		

O94804	LOK	STE	0.9604	34
Q9HBH9	Mnk2	CAMK	0.8779	34
P16234	PDGFRα	TK	0.9929	35
P23443	p70S6K	AGC	0.6446	36
Q7L7X3	TAO1	STE	0.9578	42
P53667	LIMK1	TKL	0.8753	46
Q92772	CDKL2	CMGC	0.9881	47

[a, b, c] The %activities of a kinase at BAY73-4506 concentrations of 1 μM, 0.1 μM and 0.01 μM. A null value indicates that the activity was not tested. [d] Kinases in bold type have been previously reported as targets of BAY73-4506.

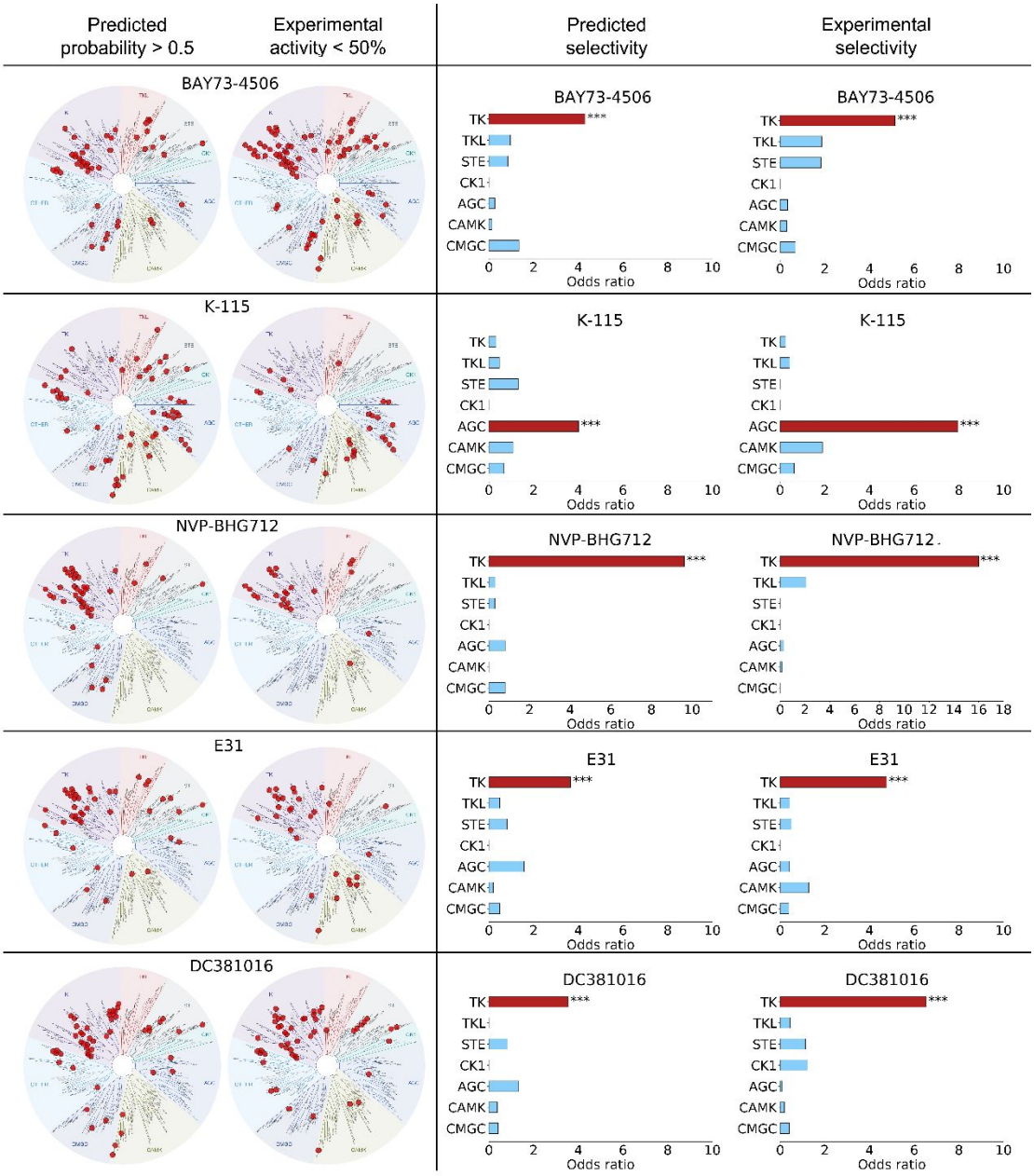


Figure 5. Kinome map of the predicted and experimental bioactivities. Kinases coloured in red are experimental %activity<50% at 1 μM or predicted active probability>0.5 (left). Bar plots of predicted and experimental group selectivity. When a group has an odds ratio significantly greater than 1, the bar is coloured red (right). '*': 0.01<p<0.05; '**': 0.001<p<0.01; '***': p<0.001.

Selectivity

Tracking the selectivity profile for the inhibitors is a particularly important step before large-scale biochemical assays for kinase drug discovery. It is of great interest to investigate whether the MTDNN model can provide reliable predictions of the selectivity. Here, we adopted a standard score[54] to measure the model's ability to rank compounds in terms of their overall selectivity, which is calculated as the number of kinase hits above or below a threshold value divided by the number of kinases tested. As an accepted quantitative measurement of selectivity, the standard score is simple but effective. The selectivity scores of the Davis dataset were calculated based on reported and predicted inhibitory activities, respectively. We set K_d=1 μM and predicted probability=50% as the activity threshold for reported and predicted inhibitory activities, respectively. The results show that the predicted selectivity scores have a significant rank correlation with the reported experimental kinase selectivity, with a Spearman correlation of 0.747 (Figure 6).

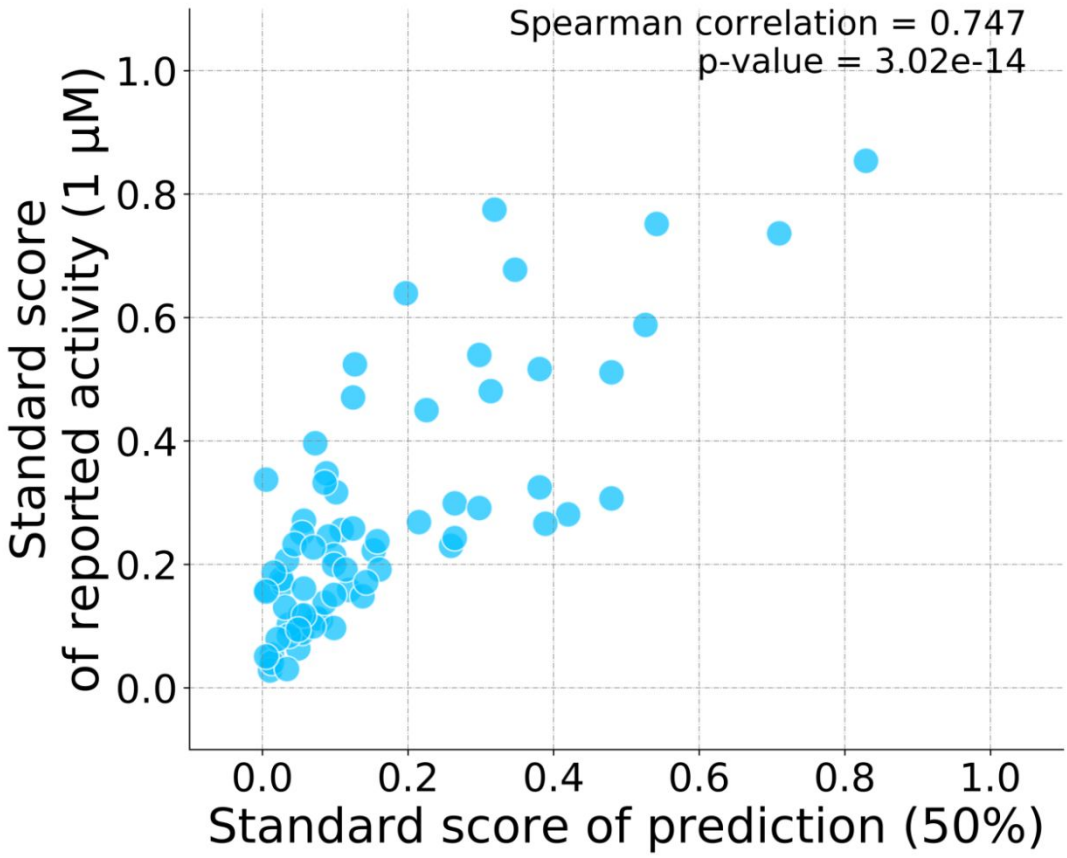


Figure 6. Scatter plot of correlation between the experimental and predicted overall selectivity. The x-axis and y-axis for a point are the standard selectivity scores calculated from predicted active probabilities (>50% as active) and reported K_d values (>1 μM as active), respectively.

Moreover, we defined the odds ratio (OR)[55] as a measure of group-specific selectivity, which reveals whether an inhibitor preferentially targets a specific kinase subfamily or a group of kinases of particular interest. For example, to calculate the strength of the

association between an inhibitor for the group of TK, the OR statistic can be calculated as:

$$OR = (N_{TP}/N_{OP})/(N_{TN}/N_{ON})$$

where N_{TP} refers to the number of positive interactions of the inhibitor within the TK family and N_{OP} refers to the number of its positive interactions with kinases other than TK. Similarly, N_{TN} and N_{ON} are defined as the number of negative interactions for TK and other kinases, respectively. If the OR is significantly greater than 1.0, the kinases inhibited by this inhibitor can be considered to be enriched in the TK group, indicating that the inhibitor is TK-selective. Figure 5 compares the ORs of the abovementioned 5 inhibitors for each of the major groups described by the kinome phylogenetic tree based on the activity spectrum predicted by the MTDNN. Clearly, the predicted group-specific selectivity revealed a high correlation with the experimental results. Even for molecules with relatively low predictive precision results, e.g., K-115, the ORs still provide a pronounced effect separating the kinase subfamily most associated with their primary studied targets.

In addition to subfamily selectivity, predicting disease-specific selectivity is a compelling application of the model in a clinical setting. Taking NSCLC as an example, which is

known to be associated with a variety of oncogenes encoding protein kinases, we defined a group of NSCLC-related kinases, including EGFR, ALK, ROS1, B-raf, Ret, ErbB2, Met, FGFR1, IGF1R and Src[49, 56-60]. A total of 48 FDA-approved small-molecule kinase inhibitors were collected for analysis of whether they can be used in NSCLC treatment or exhibit repositioning potential for such treatment based on the predicted group selectivity. As shown in Figure 7, some well-known NSCLC drugs, such as gefitinib, crizotinib, and erlotinib, are ranked high among all those drugs. Interestingly, the top ranked three drugs are not initially approved for NSCLC treatment, but all of them have been reported to exhibit monotherapy activities for NSCLC. Examples include the following: (1) Lapatinib is an ErbB2 and EGFR kinase inhibitor launched for the treatment of advanced or metastatic HER2 (ErbB2)-positive breast cancer[61]. A randomized phase II study has shown that although lapatinib monotherapy does not induce a significant number of tumour regressions in NSCLC, it showed equivalent progression-free survival as first-line chemotherapy in the 1,500 mg once daily group[62]. In addition, another study has demonstrated that lapatinib single treatment may be an effective option for the therapy of KRAS-mutated NSCLC that is resistant to erlotinib and gefitinib. (2) Ponatinib is a potent,

oral multi-targeted kinase inhibitor of Abl, PDGFRα, VEGFR2, FGFR1 and Src and has been approved for the treatment of patients with resistant or intolerant chronic myeloid leukaemia (CML)[63]. Recently, Ren et al. reported that ponatinib can be used as a treatment for established NSCLC cell lines with FGFR1 overexpression, resulting in marked cell growth inhibition[64]. (3) Bosutinib is a dual Src/Abl inhibitor for the treatment of CML. A phase I study of bosutinib has shown good tolerance and efficiency in NSCLC patients[65]. Also worth noting are vandetanib and ibrutinib, which have shown efficacy in patients with NSCLC harbouring RET rearrangement[66] and NSCLC cell lines carrying EGFR mutations[67], respectively. These results highlighted the practicality of our model in finding novel medical indications for pioneering drugs. With a clearer definition of therapeutic targets related to a certain disease, the capability of the MTDNN model to guide the development of drugs with precise multi-targeting selectivity can be envisaged.

In summary, the above analyses illustrated how MTDNN can be used to better understand the intra-family and inter-species selectivity of kinase inhibitors, which highlighted the significant value of the model in repositioning inhibitors for new kinase targets and exploring their unknown therapeutic potential. Moreover, MTDNN can also be

used to prioritize compounds on the basis of their overall selectivity or selectivity towards certain kinase subfamilies or diseases.

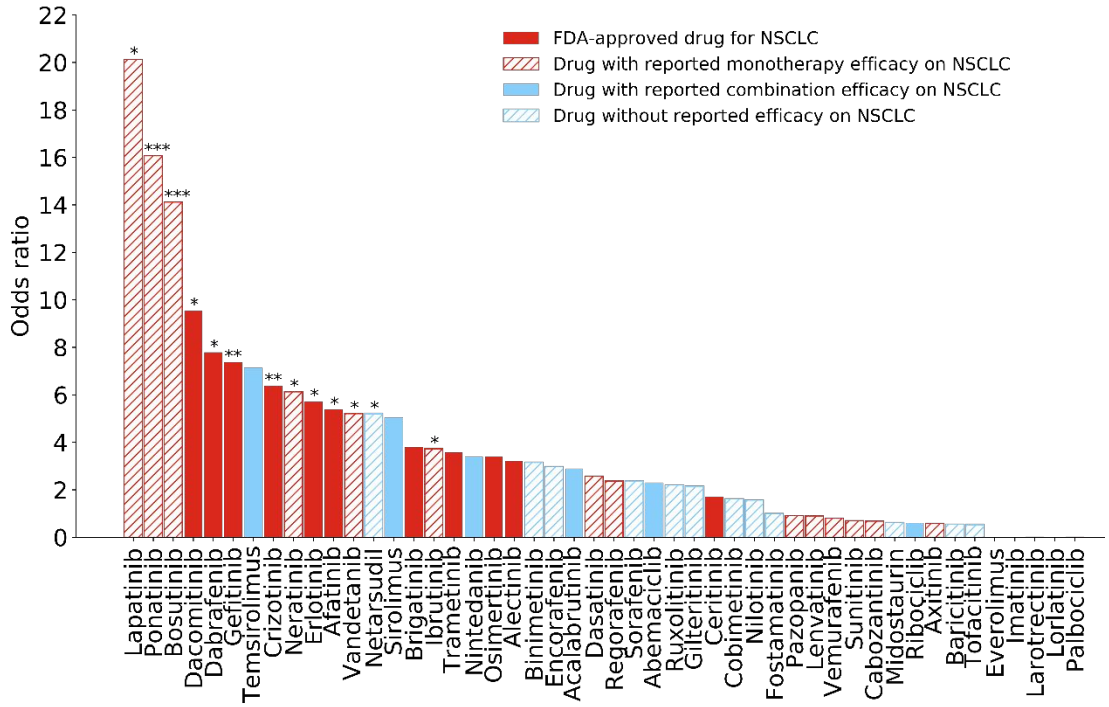


Figure 7. The odds ratio of 48 FDA-approved small-molecule kinase inhibitors for the group of NSCLC-related kinases. Red bars indicate the FDA-approved drugs for NSCLC. The bars with red stripes indicate that there have been reported evidence showing monotherapy efficacy against NSCLC for the drugs. Blue bars indicate that there have been reported evidence showing efficacy against NSCLC when combined with other drugs. The bars filled with blue stripes indicate drugs that, to the best of our knowledge,

have not been reported for treating NSCLC. '*': 0.01<p<0.05; '**': 0.001<p<0.01; '***': p<0.001.

CONCLUSION

In this study, a virtual kinase chemogenomics model was developed for predicting the interaction profiles of kinase inhibitors against a panel of 391 kinases based on large-scale bioactivity data and the MTDNN algorithm. As a result of the high relatedness among tasks (meaning widespread cross-reactivity of kinase inhibitors) and the transfer learning effect of MTDNN, the obtained model yields excellent prediction ability with an auROC of 0.90 on an internal testing dataset. On external datasets, the MTDNN model also shows impressive high-quality prediction results, despite the apparent deviation of chemical diversity distribution and the uncertainty in different data sources. The analysis revealed that the prediction results of the model could be further improved by expanding the scale and increasing the diversity of chemotypes in the training dataset. Compared with conventional single-task RF models, the model consistently shows higher auROCs on external datasets, especially for kinases with insufficient activity data. Moreover,

rigorous experimental validations were performed using compounds with diverse structures and unknown kinase activity profiles. The predicted spectrum shows significant agreement with experimental data, with a high-quality average auROC of 0.75, and many novel and therapeutic “off-target” activities have been successfully predicted by the model and confirmed by subsequent experiments. Based on the predicted kinase profile, the MTDNN model can also be used to depict the overall selectivity, the selectivity towards a subfamily of kinases, and the strength of association with certain diseases such as NSCLC. Overall, MTDNN enables us to create a comprehensive kinome interaction network for designing novel chemical modulators or drug repositioning and is of practical value for exploring previously less studied kinases.

EXPERIMENTAL SECTION

Datasets

The SARfari dataset and Metz dataset were merged into one set to generate the classification model. (1) The SARfari dataset refers to Kinase SARfari database (accessed Nov. 2017), an integrated chemogenomics workbench focused on kinases, which is composed of 54,189 compounds, 989 different kinase domains and 532,155 data

points in the form of IC_{50}, K_i, K_d and other values. (2) The second dataset, the Metz dataset[68], contains 1498 compounds with known structures, 173 human kinases and 107,791 pKi data points. The inhibition activity in the merged dataset was converted to two classes: active (pKi/pKd/pIC50≥6) and inactive (pKi/pKd/pIC50＜6). After the deletion of mutant kinases and kinases without both active and inactive data points, the final dataset contains over 170,000 bioactivity data points composed of 391 kinases (Table S1) and ~32,000 compounds. The dataset is divided into 80% training and 20% testing datasets by random selection of compounds.

Cluster-cross-validation

In contrast to conventional k-fold cross-validation, which distributes the compounds randomly cross the folds, cluster-cross-validation[29] identifies clusters of compounds in the modelling dataset and distributes them to folds to guarantee that compounds of the same cluster are present either only in the training or only in the test set.

We performed a 5-fold cluster-cross-validation using the hierarchical clustering single linkage algorithm. Single linkage represents a way to measure the dissimilarity between

groups of samples, in which the dissimilarity between group G and H is the smallest dissimilarity between two points in opposite groups:

$$\mathrm{D}(G, H) = \min_{i \in G,\, j \in H} d_{ij}.$$

With this property, single linkage cluster guarantees a minimum distance between compounds of all folds. The distances d_{ij} between compounds are measured by Jaccard distances (1 - Tanimoto similarity) on binarized ECFP4 compound representations. The minimum distance was set as 0.3. As a result, all compounds were clustered into 7,861 clusters with 1,640 compounds for the largest cluster, which were then merged to 5 different folds with about 6,400 compounds each.

Data balancing methods

For many datasets, the numbers of data points that belong to different classes are significantly different. The direct development of machine learning models using imbalanced datasets will fail to properly represent the distributive characteristics of the data and thus provide unfavourable accuracies across the categories. Instead of creating balanced data distributions through different sampling strategies, a cost-sensitive method

algorithm[69] was applied in our model, in which the cross entropy is weighted based on the ratio of the numbers of different categories. To address the issue that there are far fewer active data points than inactive data points for many kinases, a penalty for misclassifying the active data points coefficient is introduced into the cost matrix. Accordingly, the cross entropy cost matrix of a data point for each kinase is converted to:

$$H(\mathbf{y},\mathbf{y}') = \begin{cases} -N_i/N_a \sum_m \mathbf{y}_m' \log \mathbf{y}_m, & \text{active data point} \\ \sum_m \mathbf{y}_m' \log \mathbf{y}_m, & \text{inactive data point} \end{cases}$$

where N_i and N_a refer to the number of inactive data points and the number of active data points for a kinase, respectively; and $\mathbf{y}_m'$ and $\mathbf{y}_m$ are the m-th class of true labels and predicted outputs, respectively.

Multitask deep neural networks (MTDNNs)

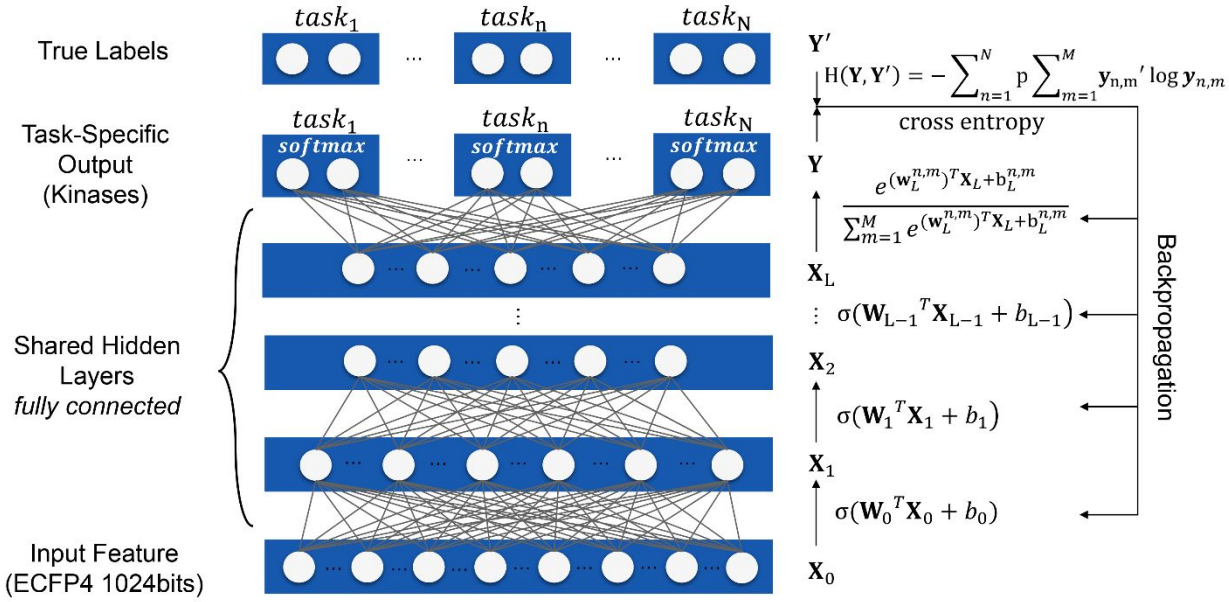


Figure 8. The architecture of a multitask deep neural network.

MTDNNs implement a straightforward solution to problems with multiple related labels. The shared hidden layers among all tasks can help the model learn a shared representation and thereby obtain relatively strong abstracting capabilities. Xu et al.[17] have demonstrated that an MTDNN can outperform single-task DNNs by utilizing the activities of molecules from the other tasks if the molecules from these tasks share similar structures and correlated activities (either positively or negatively correlated). Ma et al.[70] also proposed that an MTDNN can take advantage of the tasks with a larger training dataset to improve the predictive performance for the tasks with a smaller training set,

and the regularization effect produced by various task can help the MTDNN avoid potential overfitting. For kinases, as most inhibitors interact with the hinge motif in the highly conserved catalytic domain, the structures and activities of inhibitors of different kinases are substantially correlated. Thus, MTDNN provides an appropriate solution to kinome spectrum prediction, and the cross-reaction of inhibitors can boost the predictive performance for related kinases, especially for kinases with insufficient interaction data. In this study, we constructed a multitask classification DNN architecture with shared hidden layers among all tasks (Figure 8), where each task represents the activity to be predicted against a specific kinase. Extended connectivity fingerprints[71] with a radius of 2 continuous bonds (ECPF4) and a length of 1024 bits were adopted to featurize each molecule and fed into the input layer, which means that only two-dimensional chemical structural information is needed for making predictions based on the resulting model. The input was fed into one or more (L) fully connected hidden layers, and the rectified linear unit (ReLU)[72] was chosen to perform the nonlinear transformation as follows between any two adjacent layers:

$$\mathbf{X}_{l+1} = \sigma\left(\mathbf{W}_l^{T}\mathbf{X}_l + b_l\right)$$

where $\mathbf{X}_l$, $\mathbf{W}^l$ and b_l represent the input, weight matrix and bias for the l-th layer, respectively, and σ refers to the ReLU activation function. After L transformations, the L-th layer is fed into a task-specific output layer, which consists of N softmax classifiers corresponding to N tasks with M labels (M=2 for binary classification) for predicting whether the input molecule is active or inactive towards a panel of N kinases. For the n-th task, the probability that the input has a label of m is:

$$y_{n,m} = \frac{e^{(\mathbf{w}_L^{n,m})^T \mathbf{X}_L + b_L^{n,m}}}{\sum_{m=1}^{M} e^{(\mathbf{w}_L^{n,m})^T \mathbf{X}_L + b_L^{n,m}}}$$

The cross entropy between true labels and predicted labels is calculated as the loss function, and thus, the weight matrix and bias vectors can be updated using the backpropagation algorithm:

$$H(\mathbf{Y},\mathbf{Y}') = -\sum_{n=1}^{N} p \sum_{m=1}^{M} \mathbf{y}_{n,m}' \log \mathbf{y}_{n,m}$$

where p refers to N_i/N_a or 1 according to the class of true labels as mentioned in the balancing method. The process of network training will repeat multiple epochs until the loss or other evaluation metrics converge.

Optimization of hyperparameters

In general, DNN is sensitive to the choice of hyper parameters, and so is MTDNN. To optimize the generalizability of the model, we explored a number of different hyperparameters through a grid search[73], including network architecture (layer size), learning rate, and batch size. An early stopping[74] scheme was introduced during the process of training, where the auROC on the validation set was monitored every five epochs. The training is stopped once the auROC decreases four times continuously.

The layer size was varied from shallow (one hidden layer with 1000 nodes, i.e., [1000]) to moderate (two hidden layers of 1500 and 1000 nodes, i.e., [1500,1000]) and deep ([2000,1000,500]). Additionally, as there are 1024 nodes in the input layer and 391 tasks in the output layer, we also tried another two-hidden-layer architecture [1024,391]. Since network architectures with different levels of complexity require different learning rates to achieve stable and improved performance, the Adaptive Moment Estimation (Adam) method[75] was used for backpropagation with initial learning rates varying among 1×10^{-3}, 1×10^{-4}, 5×10^{-5} and 1×10^{-5}. Two sizes of mini batch (64 and 128) were also tested. To regularize our network, we used a dropout of 0.5 as well as a moderate L2 weight decay

of 0.002 for each hidden layer in every network. These two strategies have been proven to work in concert to avoid overfitting[76].

MTDNN models are developed in Tensorflow (Version 1.6.0) and DeepChem (Version 2.1.0). All trainings are performed on standard NVIDIA GPUs. The code is developed in Python 3.6.

Evaluation metrics

For model quality assessment, the auROC (area under the ROC curve), recall, precision, F1-score and BA (balanced accuracy) were evaluated (Table 4).

Table 4. Description of the evaluation metrics.

Evaluation metric	Equation[a]
Recall	$\mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$
Precision	$\mathrm{TP}/(\mathrm{TP}+\mathrm{FP})$
F1-score	$2\times\frac{\mathrm{Precision}\times\mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}}$
BA	$(\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}+\frac{\mathrm{TN}}{\mathrm{FP}+\mathrm{TN}})/2$

[a]TP is the number of correctly predicted actives (true positives), TN is the number of correctly predicted inactives (true negatives), FP is the number of incorrectly identified actives (false positives) and FN is the number of incorrectly identified inactives (false negatives).

ASSOCIATED CONTENT

Supporting Information.

Table S1: Kinase information (XLSX)

Table S2: Prediction performance comparisons between MTDNN and RF (XLSX)

Table S3: Experimental data (XLSX)

Molecular formula strings and the associated biochemical and biological data (CSV)

Molecular information (DOCX)

AUTHOR INFORMATION

Corresponding Author

*Phone: +86-21-50806600-1303. E-mail: hljiang@simm.ac.cn (Hualiang Jiang)

*Phone: +86-21-50806600-1308. E-mail: myzheng@simm.ac.cn (Mingyue Zheng).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

We gratefully acknowledge financial support from the National Natural Science Foundation of China (81773634 to M.Z. and 81430084 to K.C.), National Science & Technology Major Project “Key New Drug Creation and Manufacturing Program”, China (Number: 2018ZX09711002 to H.J.), and “Personalized Medicines—Molecular Signature-based Drug Discovery and Development”, Strategic Priority Research Program of the Chinese Academy of Sciences (XDA12050201 to M.Z.), and the open fund of state key laboratory of Pharmaceutical Biotechnology, Nanjing University, China (KF-GN-201706 to H.J.). Image generated using TREEspot™ Software Tool and reprinted with permission from KINOMEscan®, a division of DiscoveRx Corporation, © DISCOVERX CORPORATION 2010.

ABBREVIATIONS USED

NB, naïve Bayesian; KNN, k-nearest neighbours; RF, random forest; SVM, support vector machine; DNN, deep neural network; MTDNN, multitask deep neural network; auROC, area under receiver operating characteristic; BA, balanced accuracy; PKIS1, published kinase inhibitor set 1; PKIS2, published kinase inhibitor set 2; t-SNE, t-distributed stochastic neighbor embedding; ECFP4, extended connectivity fingerprint 4; DCNB, dual-component naïve Bayes; DCSVM, dual-component SVM; DDM, drug discovery maps; FLT3, FMS-like tyrosine kinase 3; DAD, Diode-array detectors; VEGFR, vascular endothelial growth factor receptor; PDGFRβ, platelet-derived growth factor receptor β; Kit, mast/stem cell growth factor receptor Kit; RET, proto-oncogene tyrosine-protein kinase receptor Ret; c-Raf, RAF proto-oncogene serine/threonine-protein kinase; p38α, mitogen-activated protein kinase 14; ZAK, sterile α motif and leucine zipper containing kinase AZK; TAO1, serine/threonine-protein kinase; MAPK, mitogen-activated protein kinase; TrkB, BDNF/NT-3 growth factors receptor; TrkC, NT-3 growth factor receptor; NSCLC, non-small cell lung cancer; DDR1, epithelial discoidin domain-containing receptor 1; LOK, serine/threonine-protein kinase 10; PKG, protein kinase G; PKC, protein kinase C; Src, proto-oncogene tyrosine-protein kinase Src; SFKs, Src family

kinases; Lyn, tyrosine-protein kinase Lyn; Hck, hematopoietic cell kinase; Lck, leukocyte c-terminal Src kinase; Yes, Tyrosine-protein kinase Yes; Blk, B lymphocyte kinase; Fgr, tyrosine-protein kinase Fgr; c-Met, hepatocyte growth factor receptor; Abl, abelson murine leukemia viral oncogene homologue 1; ACK1, activated CDC42 kinase 1; ROCK1, rho-associated protein kinase 1; ROCK2, rho-associated protein kinase 2; EGFR, epidermal growth factor receptor; ERBB2, receptor tyrosine-protein kinase erbB-2; ERBB4, receptor tyrosine-protein kinase erbB-4; OR, odds ratio; ALK, ALK tyrosine kinase receptor; ROS1, proto-oncogene tyrosine-protein kinase ROS; B-raf, serine/threonine-protein kinase B-raf; FGFR1, fibroblast growth factor receptor 1; IGF1R, insulin-like growth factor 1 receptor; KRAS, GTPase KRas; PDGFRα, platelet-derived growth factor receptor α; VEGFR2, vascular endothelial growth factor receptor 2; CML, chronic myeloid leukaemia; HER2, receptor tyrosine-protein kinase erbB-2; ReLU, rectified linear unit; GPU, graphics processing unit; TP, true positive; TN, true negative; FP, false positive; FN, false negative.

REFERENCES

(1) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002**, *298*, 1912-1934.

(2) Wu, P.; Nielsen, T. E.; Clausen, M. H. FDA-approved small-molecule kinase inhibitors. *Trends Pharmacol. Sci.* **2015**, *36*, 422-439.

(3) Fabbro, D.; Cowan‐Jacob, S. W.; Moebitz, H. Ten things you should know about protein kinases: IUPHAR Review 14. *Br. J. Pharmacol.* **2015**, *172*, 2675-2700.

(4) Fedorov, O.; Müller, S.; Knapp, S. The (un)targeted cancer kinome. *Nat. Chem. Biol.* **2010**, *6*, 166-169.

(5) Klaeger, S.; Heinzlmeir, S.; Wilhelm, M.; Polzer, H.; Vick, B.; Koenig, P.-A.; Reinecke, M.; Ruprecht, B.; Petzoldt, S.; Meng, C. The target landscape of clinical kinase drugs. *Science* **2017**, *358*, eaan4368.

(6) Knight, Z. A.; Lin, H.; Shokat, K. M. Targeting the cancer kinome through polypharmacology. *Nat. Rev. Cancer* **2010**, *10*, 130.

(7) Goldstein, D. M.; Gray, N. S.; Zarrinkar, P. P. High-throughput kinase profiling as a platform for drug discovery. *Nat. Rev. Drug Discovery* **2008**, *7*, 391.

(8) Niijima, S.; Shiraishi, A.; Okuno, Y. Dissecting kinase profiling data to predict activity and understand cross-reactivity of kinase inhibitors. *J. Chem. Inf. Model.* **2012**, *52*, 901-912.

(9) Lapins, M.; Wikberg, J. E. Kinome-wide interaction modelling using alignment-based and alignment-independent approaches for kinase description and linear and non-linear data analysis techniques. *BMC Bioinformatics* **2010**, *11*, 339-339.

(10) Stephan C, S.; Steven M, M. Kinome-wide activity modeling from diverse public high-quality data sets. *J. Chem. Inf. Model.* **2013**, *53*, 27-38.

(11) Cao, D. S.; Zhou, G. H.; Liu, S.; Zhang, L. X.; Xu, Q. S.; He, M.; Liang, Y. Z. Large-scale prediction of human kinase–inhibitor interactions using protein sequences and molecular topological structures. *Anal. Chim. Acta* **2013**, *792*, 10-18.

(12) Merget, B.; Turk, S.; Eid, S.; Rippmann, F.; Fulle, S. Profiling prediction of kinase inhibitors: toward the virtual assay. *J. Med. Chem.* **2016**, *60*, 474-485.

(13) Bora, A.; Avram, S.; Ciucanu, I.; Raica, M.; Avram, S. Predictive models for fast and effective profiling of kinase inhibitors. *J. Chem. Inf. Model.* **2016**, *56*, 895-905.

(14) Yabuuchi, H.; Niijima, S.; Takematsu, H.; Ida, T.; Hirokawa, T.; Hara, T.; Ogawa, T.; Minowa, Y.; Tsujimoto, G.; Okuno, Y. Analysis of multiple compound–protein interactions reveals novel bioactive molecules. *Mol. Syst. Biol.* **2011**, *7*, 472.

(15) Manallack, D. T.; Pitt, W. R.; Gancia, E.; Montana, J. G.; Livingstone, D. J.; Ford, M. G.; Whitley, D. C. Selecting screening candidates for kinase and G protein-coupled receptor targets using neural networks. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1256-1262.

(16) Subramanian, G.; Ramsundar, B.; Pande, V.; Denny, R. A. Computational modeling of β-secretase 1 (BACE-1) inhibitors using ligand based approaches. *J. Chem. Inf. Model.* **2016**, *56*, 1936-1949.

(17) Xu, Y.; Ma, J.; Liaw, A.; Sheridan, R. P.; Svetnik, V. Demystifying multitask deep neural networks for quantitative structure–activity relationships. *J. Chem. Inf. Model.* **2017**, *57*, 2490-2504.

(18) Ramsundar, B.; Liu, B.; Wu, Z.; Verras, A.; Tudor, M.; Sheridan, R. P.; Pande, V. Is multitask deep learning practical for pharma? *J. Chem. Inf. Model.* **2017**, *57*, 2068-2076.

(19) Ramsundar, B.; Kearnes, S.; Riley, P.; Webster, D.; Konerding, D.; Pande, V. Massively multitask networks for drug discovery. *Arxiv Preprint Arxiv:1502.02072* **2015**.

(20) Wu, K.; Wei, G.-W. Quantitative toxicity prediction using topology based multitask deep neural networks. *J. Chem. Inf. Model.* **2018**, *58*, 520-531.

(21) Lenselink, E. B.; Ten Dijke, N.; Bongers, B.; Papadatos, G.; Van Vlijmen, H. W.; Kowalczyk, W.; IJzerman, A. P.; Van Westen, G. J. Beyond the hype: deep neural networks outperform established methods using a ChEMBL bioactivity benchmark set. *J. Cheminform.* **2017**, *9*, 45.

(22) Rodríguez-Pérez, R.; Bajorath, J. r. Multitask Machine Learning for Classifying Highly and Weakly Potent Kinase Inhibitors. *ACS Omega* **2019**, *4*, 4367-4375.

(23) Davis, M. I.; Hunt, J. P.; Sanna, H.; Pietro, C.; Wodicka, L. M.; Gabriel, P.; Michael, H.; Treiber, D. K.; Zarrinkar, P. P. Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2011**, *29*, 1046-1051.

(24) Theonie, A.; Deacon, S. W.; Karthik, D.; Haiching, M.; Peterson, J. R. Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat. Biotechnol.* **2011**, *29*, 1039-1045.

(25) Elkins, J. M.; Fedele, V.; Szklarz, M.; Azeez, K. R. A.; Salah, E.; Mikolajczyk, J.; Romanov, S.; Sepetov, N.; Huang, X. P.; Roth, B. L. Comprehensive characterization of the Published Kinase Inhibitor Set. *Nat. Biotechnol.* **2015**, *34*, 95.

(26) Drewry, D. H.; Wells, C. I.; Andrews, D. M.; Angell, R.; Al-Ali, H.; Axtman, A. D.; Capuzzi, S. J.; Elkins, J. M.; Ettmayer, P.; Frederiksen, M. Progress towards a public chemogenomic set for protein kinases and a call for contributions. *PLoS One* **2017**, *12*, e0181585.

(27) Sutherland, J. J.; Gao, C.; Cahya, S.; Vieth, M. What general conclusions can we draw from kinase profiling data sets? *Biochim. Biophys. Acta* **2013**, *1834*, 1425-1433.

(28) Posy, S. L.; Hermsmeier, M. A.; Wayne, V.; Karl-Heinz, O.; Gordon, T.; Lippy, J. S.; Trainor, G. L.; Loughney, D. A.; Johnson, S. R. Trends in kinase selectivity: insights for target class-focused library screening. *J. Med. Chem.* **2011**, *54*, 54-66.

(29) Mayr, A.; Klambauer, G.; Unterthiner, T.; Steijaert, M.; Wegner, J. K.; Ceulemans, H.; Clevert, D.-A.; Hochreiter, S. Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. *Chem. Sci.* **2018**, *9*, 5441-5451.

(30) Maaten, L. v. d.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579-2605.

(31) Avram, S.; Bora, A.; Halip, L.; Curpăn, R. Modeling Kinase Inhibition Using Highly Confident Data Sets. *J. Chem. Inf. Model.* **2018**, *58*, 957-967.

(32) Janssen, A. P. A.; Grimm, S. H.; Wijdeven, R. H. M.; Lenselink, E. B.; Neefjes, J.; van Boeckel, C. A. A.; van Westen, G. J. P.; van der Stelt, M. Drug Discovery Maps, a

Machine Learning Model That Visualizes and Predicts Kinome–Inhibitor Interaction Landscapes. *J. Chem. Inf. Model.* **2019**, *59*, 1221-1229.

(33) Zhang, D.; Ai, J.; Liang, Z.; Zhu, W.; Peng, X.; Chen, X.; Ji, Y. C.; Jiang, H.; Luo, C.; Geng, M. Novel 5-(benzyloxy)pyridin-2(1 H)-one derivatives as potent c-Met inhibitors. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2408-2413.

(34) Zhang, X.; Peng, T.; Ji, X.; Li, J.; Tong, L.; Li, Z.; Yang, W.; Xu, Y.; Li, M.; Ding, J. Design, synthesis and biological evaluation of novel 4-anilinoquinazolines with C-6 urea-linked side chains as inhibitors of the epidermal growth factor receptor. *Bioorg. Med. Chem.* **2013**, *21*, 7988-7998.

(35) Grothey, A.; Cutsem, E. V.; Sobrero, A.; Siena, S.; Falcone, A.; Ychou, M.; Humblet, Y.; Bouché, O.; Mineur, L.; Barone, C. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet* **2013**, *381*, 303-312.

(36) Demetri, G. D.; Reichardt, P.; Kang, Y. K.; Blay, J. Y.; Rutkowski, P.; Gelderblom, H.; Hohenberger, P.; Leahy, M.; Mehren, M. V.; Joensuu, H. Efficacy and safety of

regorafenib for advanced gastrointestinal stromal tumours after failure of imatinib and sunitinib (GRID): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet* **2013**, *381*, 295-302.

(37) Bruix, J.; Qin, S.; Merle, P.; Granito, A.; Huang, Y. H.; Bodoky, G.; Pracht, M.; Yokosuka, O.; Rosmorduc, O.; Breder, V. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **2016**, *389*, 56-66.

(38) Wilhelm, S. M.; Dumas, J.; Adnane, L.; Lynch, M.; Carter, C. A.; Schütz, G.; Thierauch, K. H.; Zopf, D. Regorafenib (BAY 73-4506): a new oral multikinase inhibitor of angiogenic, stromal and oncogenic receptor tyrosine kinases with potent preclinical antitumor activity. *Int. J. Cancer* **2011**, *129*, 245-255.

(39) Koul, H. K.; Pal, M.; Koul, S. Role of p38 MAP Kinase Signal Transduction in Solid Tumors. *Genes Cancer* **2013**, *4*, 342-359.

(40) Lange, A.; Lo, H.-W. Inhibiting TRK proteins in clinical cancer therapy. *Cancers* **2018**, *10*, 105.

(41) Sirvent, A.; Lafitte, M.; Roche, S. DDR1 inhibition as a new therapeutic strategy for colorectal cancer. *Mol. Cell Oncol.* **2018**, *5*, e1465882.

(42) Yupeng, L.; Xiaoyun, L.; Xiaomei, R.; Ke, D. Small molecule discoidin domain receptor kinase inhibitors and potential medical applications. *J. Med. Chem.* **2015**, *58*, 3287-3301.

(43) Tanihara, H.; Inoue, T.; Yamamoto, T.; Kuwayama, Y.; Abe, H.; Araie, M. Phase 2 randomized clinical study of a Rho kinase inhibitor, K-115, in primary open-angle glaucoma and ocular hypertension. *Am. J. Ophthalmol.* **2013**, *156*, 731-736.

(44) Prêtre, V.; Wicki, A. Inhibition of Akt and other AGC kinases: A target for clinical cancer therapy? *Semin. Cancer Biol.* **2018**, *48*, 70-77.

(45) Browning, D. D.; Kwon, I.-K.; Wang, R. cGMP-dependent protein kinases as potential targets for colon cancer prevention and treatment. *Future Med. Chem.* **2010**, *2*, 65-80.

(46) Rocha, A. B., Da; Mans, D. R. A.; Regner, A., .; Schwartsmann, G., . Targeting protein kinase C: new therapeutic opportunities against high-grade malignant gliomas? *Oncologist* **2002**, *7*, 17-33.

(47) Kathawala, R. J.; Liuya, W.; Nagaraju, A.; Kang, C.; Atish, P.; Saeed, A.; Yun-Kai, Z.; Yi-Jun, W.; Kamlesh, S.; Amal, K. The small molecule tyrosine kinase inhibitor NVP-BHG712 antagonizes ABCC10-mediated paclitaxel resistance: a preclinical and pharmacokinetic study. *Oncotarget* **2015**, *6*, 510-521.

(48) Wook, J. D.; Young Moon, O.; Young Geon, J.; Yunkyung, C.; Hyeon, L. D. The inhibition of SRC family kinase suppresses pancreatic cancer cell proliferation, migration, and invasion. *Pancreas* **2014**, *43*, 768-776.

(49) Wislez, M. SRC-family kinases are activated in non-small cell lung cancer and promote the survival of epidermal growth factor receptor-dependent cell lines. *Am. J. Pathol.* **2007**, *170*, 366-376.

(50) Wu, X.; Zahari, M. S.; Renuse, S.; Kelkar, D. S.; Bharbuiya, M. A.; Rojas, P. L.; Stearns, V.; Gabrielson, E.; Malla, P.; Sukumar, S. The non-receptor tyrosine kinase

TNK2/ACK1 is a novel therapeutic target in triple negative breast cancer. *Oncotarget* **2016**, *8*, 2971.

(51) Isobe, T.; Mizuno, K.; Kaneko, Y.; Ohta, M.; Koide, T.; Tanabe, S. Effects of K-115, a Rho-Kinase Inhibitor, on Aqueous Humor Dynamics in Rabbits. *Curr. Eye Res.* **2014**, *39*, 813-822.

(52) Kotaro, Y.; Kazuichi, M.; Noriko, H.; Kazuko, O.; Yu, Y.; Yukihiro, S.; Ryu, M.; Toru, N. The novel Rho kinase (ROCK) inhibitor K-115: a new candidate drug for neuroprotective treatment in glaucoma. *Invest. Ophthalmol. Vis. Sci.* **2014**, *55*, 7126-7136.

(53) Martiny-Baron, G.; Holzer, P.; Billy, E.; Schnell, C.; Brueggen, J.; Ferretti, M.; Schmiedeberg, N.; Wood, J. M.; Furet, P.; Imbach, P. The small molecule specific EphB4 kinase inhibitor NVP-BHG712 inhibits VEGF driven angiogenesis. *Angiogenesis* **2010**, *13*, 259-267.

(54) Karaman, M. W.; Sanna, H.; Treiber, D. K.; Paul, G.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Pietro, C.; Davis, M. I.; Edeen, P. T. A quantitative analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* **2008**, *26*, 127-132.

(55) Bland, J. M.; Altman, D. G. Statistics notes. The odds ratio. *BMJ* **2000**, *320*, 1468-1468.

(56) Emma, S.; Thomas, H.; Simon, G. R.; Dennis, P. A.; Otterson, G. A.; Raphael, B.; Ravi, S. Molecular pathways and therapeutic targets in lung cancer. *Oncotarget* **2014**, *5*, 1392-1433.

(57) Janku, F.; Garrido-Laguna, I.; Petruzelka, L. B.; Stewart, D. J.; Kurzrock, R. Novel Therapeutic Targets in Non-small Cell Lung Cancer. *J. Thorac. Oncol.* **2011**, *6*, 1601-1612.

(58) Thanyanan, R.; Grace Kho, D. Targeted therapies in development for non-small cell lung cancer. *J. Carcinog.* **2013**, *12*, 22.

(59) Chan, B. A.; Hughes, B. G. M. Targeted therapy for non-small cell lung cancer: current standards and the promise of the future. *Transl. Lung Cancer Res.* **2015**, *4*, 36-54.

(60) Giaccone, G., .; Zucali, P. A. Src as a potential therapeutic target in non-small-cell lung cancer. *Ann. Oncol.* **2008**, *19*, 1219-1223.

(61) Rusnak, D. W.; Lackey, K.; Affleck, K.; Wood, E. R.; Alligood, K. J.; Rhodes, N.; Keith, B. R.; Murray, D. M.; Knight, W. B.; Mullin, R. J. The effects of the novel, reversible epidermal growth factor receptor/ErbB-2 tyrosine kinase inhibitor, GW2016, on the growth of human normal and tumor-derived cell lines in vitro and in vivo. *Mol. Cancer Ther.* **2001**, *1*, 85-94.

(62) Ross, H. J.; Blumenschein, G. R.; Aisner, J.; Damjanov, N.; Dowlati, A.; Garst, J.; Rigas, J. R.; Smylie, M.; Hassani, H.; Allen, K. E. Randomized phase II multicenter trial of two schedules of lapatinib as first-or second-line monotherapy in patients with advanced or metastatic non–small cell lung cancer. *Clin. Cancer Res.* **2010**, *16*, 1938-1949.

(63) O'Hare, T.; Shakespeare, W. C.; Zhu, X.; Eide, C. A.; Rivera, V. M.; Wang, F.; Adrian, L. T.; Zhou, T.; Huang, W.-S.; Xu, Q.; Metcalf, C. A., 3rd; Tyner, J. W.; Loriaux, M. M.; Corbin, A. S.; Wardwell, S.; Ning, Y.; Keats, J. A.; Wang, Y.; Sundaramoorthi, R.; Thomas, M.; Zhou, D.; Snodgrass, J.; Commodore, L.; Sawyer, T. K.; Dalgarno, D. C.; Deininger, M. W. N.; Druker, B. J.; Clackson, T. AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance. *Cancer Cell* **2009**, *16*, 401-412.

(64) Mingqiang, R.; Mei, H.; Gentao, L.; Hongjin, W.; Vijay, P.; Paul, B.; Jeane, S.; John, C.; Zhonglin, H. Novel FGFR inhibitor ponatinib suppresses the growth of non-small cell lung cancer cells overexpressing FGFR1. *Oncol. Rep.* **2013**, *29*, 2181-2190.

(65) Daud, A. I.; Krishnamurthi, S. S.; Saleh, M. N.; Gitlitz, B. J.; Borad, M. J.; Gold, P. J.; Chiorean, E. G.; Springett, G. M.; Abbas, R.; Agarwal, S. Phase I study of bosutinib, a src/abl tyrosine kinase inhibitor, administered to patients with advanced solid tumors. *Clin. Cancer Res.* **2012**, *18*, 1092-1100.

(66) Lee, S.-H.; Lee, J.-K.; Ahn, M.-J.; Kim, D.-W.; Sun, J.-M.; Keam, B.; Kim, T. M.; Heo, D. S.; Ahn, J. S.; Choi, Y.-L.; Min, H.-S.; Jeon, Y.-K.; Park, K. A phase II study of vandetanib in patients with non-small cell lung cancer harboring RET rearrangement. *J. Clin. Oncol.* **2016**, *34*, 9013-9013.

(67) Gao, W.; Wang, M.; Wang, L.; Lu, H.; Wu, S.; Dai, B.; Ou, Z.; Zhang, L.; Heymach, J. V.; Gold, K. A.; Minna, J.; Roth, J. A.; Hofstetter, W. L.; Swisher, S. G.; Fang, B. Selective antitumor activity of ibrutinib in EGFR-mutant non-small cell lung cancer cells. *J. Natl. Cancer Inst.* **2014**, *106*, dju204.

(68) Metz, J. T.; Johnson, E. F.; Soni, N. B.; Merta, P. J.; Lemma, K.; Hajduk, P. J. Navigating the kinome. *Nat. Chem. Biol.* **2011**, *7*, 200-202.

(69) He, H.; Garcia, E. A. Learning from imbalanced data. *IEEE Trans. Knowl. Data. Eng.* **2008**, 1263-1284.

(70) Junshui, M.; Sheridan, R. P.; Andy, L.; Dahl, G. E.; Vladimir, S. Deep neural nets as a method for quantitative structure-activity relationships. *J. Chem. Inf. Model.* **2015**, *55*, 263-274.

(71) Rogers, D.; Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742-754.

(72) Nair, V.; Hinton, G. E. In *Rectified linear units improve restricted boltzmann machines*, Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010; pp 807-814.

(73) Hinton, G. E., A Practical Guide to Training Restricted Boltzmann Machines. In *Neural Networks: Tricks of the Trade: Second Edition*, Montavon, G.; Orr, G. B.; Müller, K.-R., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; pp 599-619.

(74) Prechelt, L., Early Stopping — But When? In *Neural Networks: Tricks of the Trade: Second Edition*, Montavon, G.; Orr, G. B.; Müller, K.-R., Eds. Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; pp 53-67.

(75) Kingma, D. P.; Ba, J. Adam: A method for stochastic optimization. *Arxiv Preprint Arxiv:1412.6980* **2014**.

(76) Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929-1958.

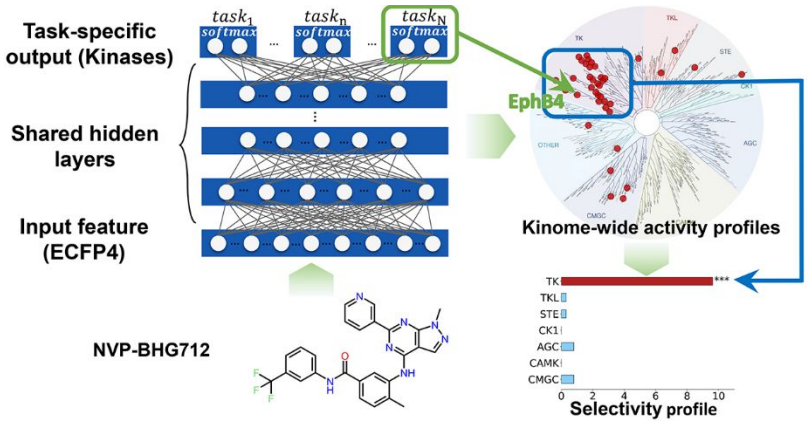
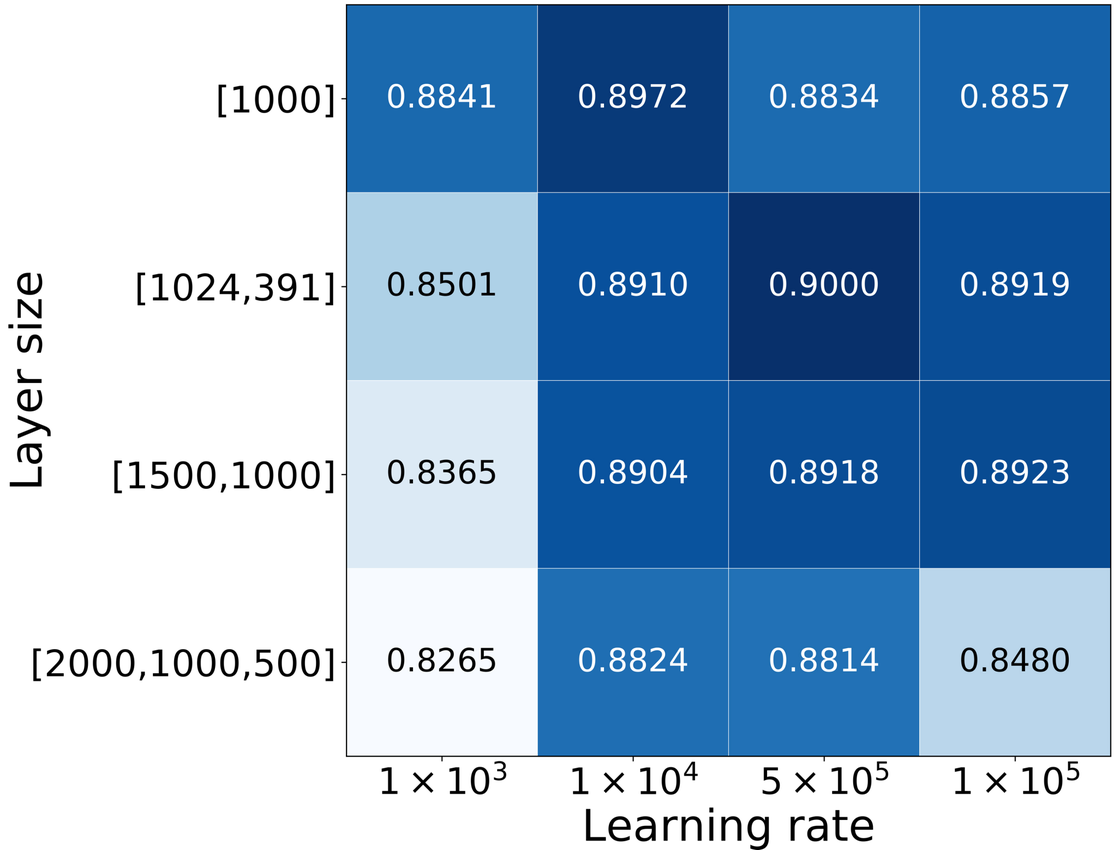
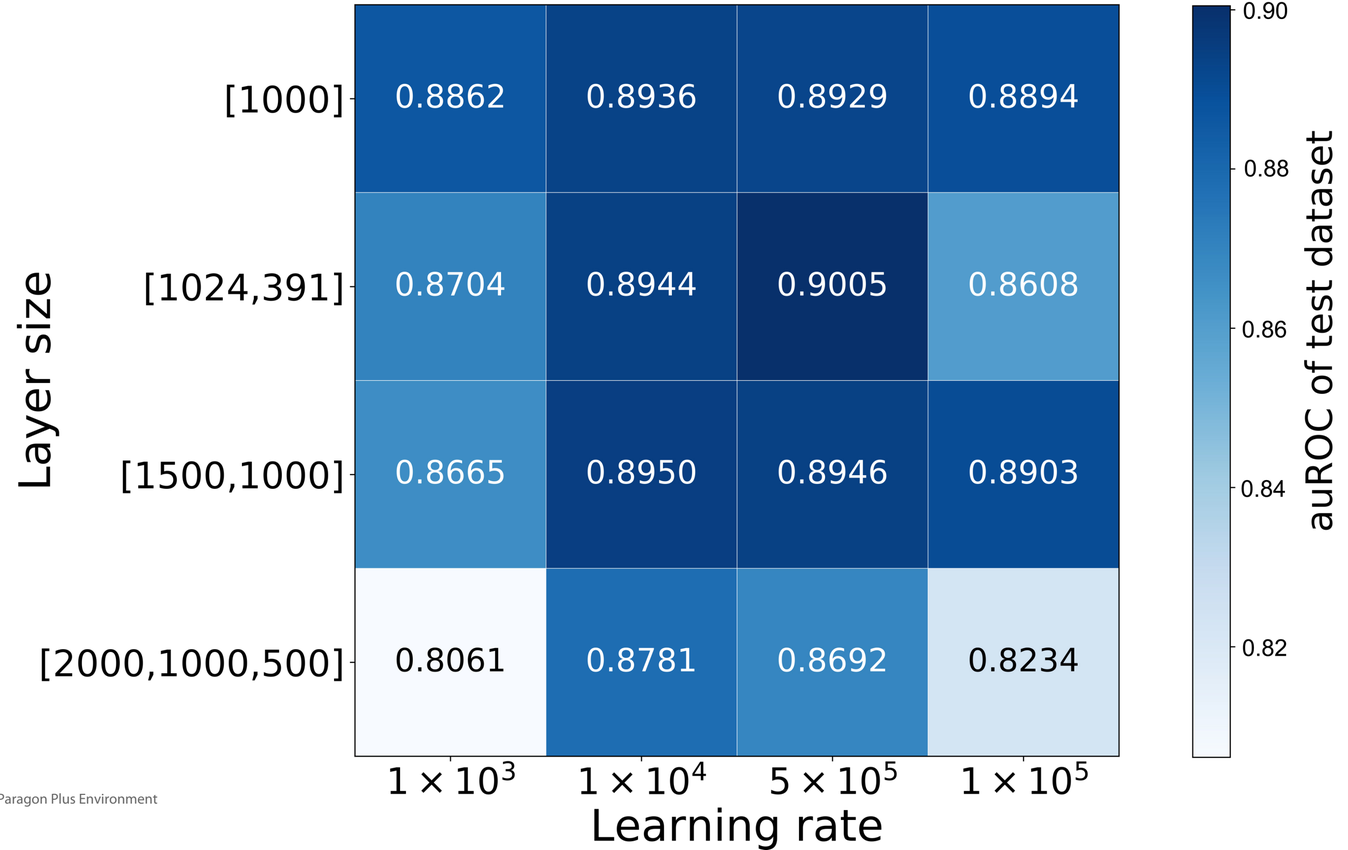
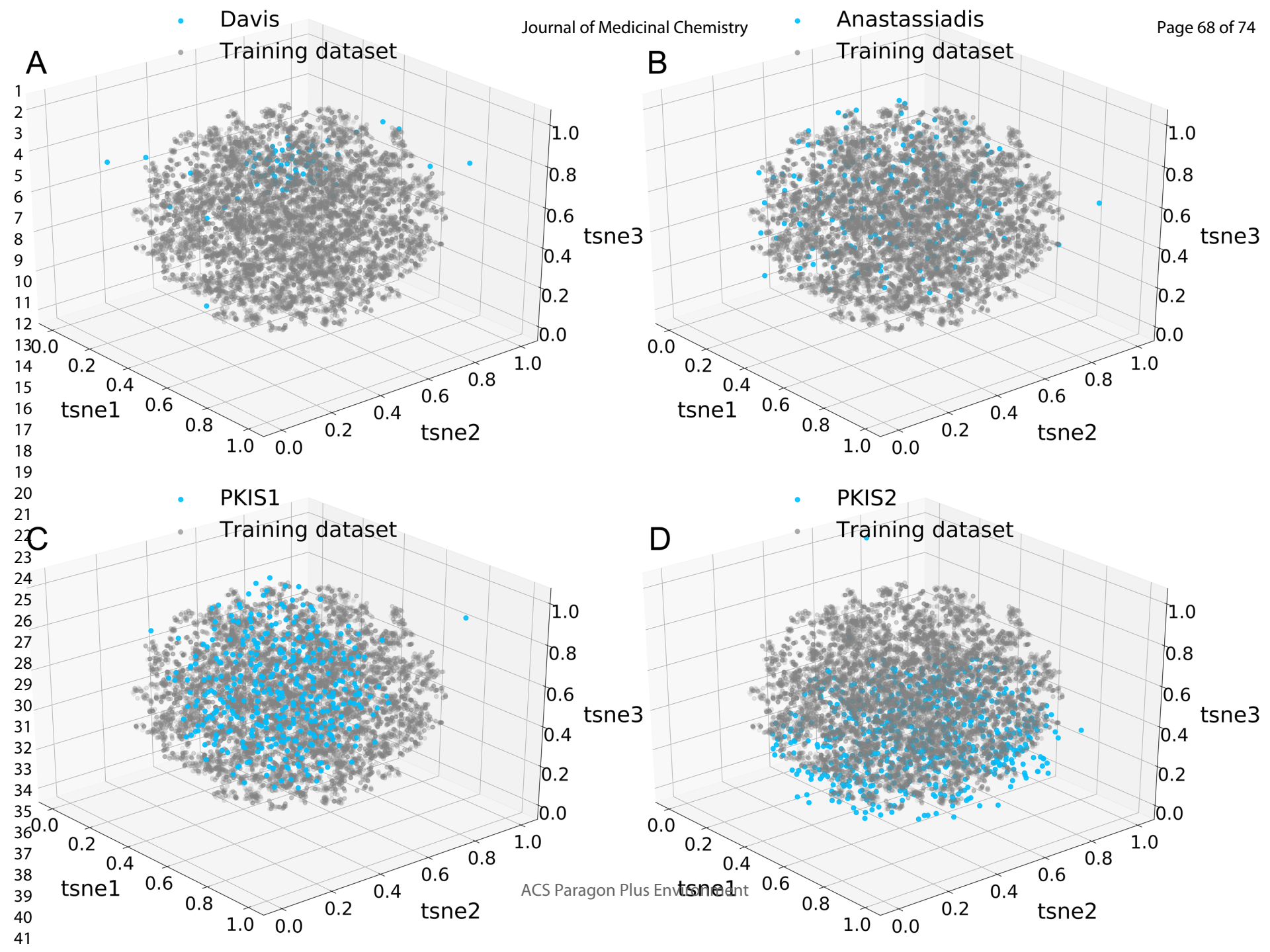
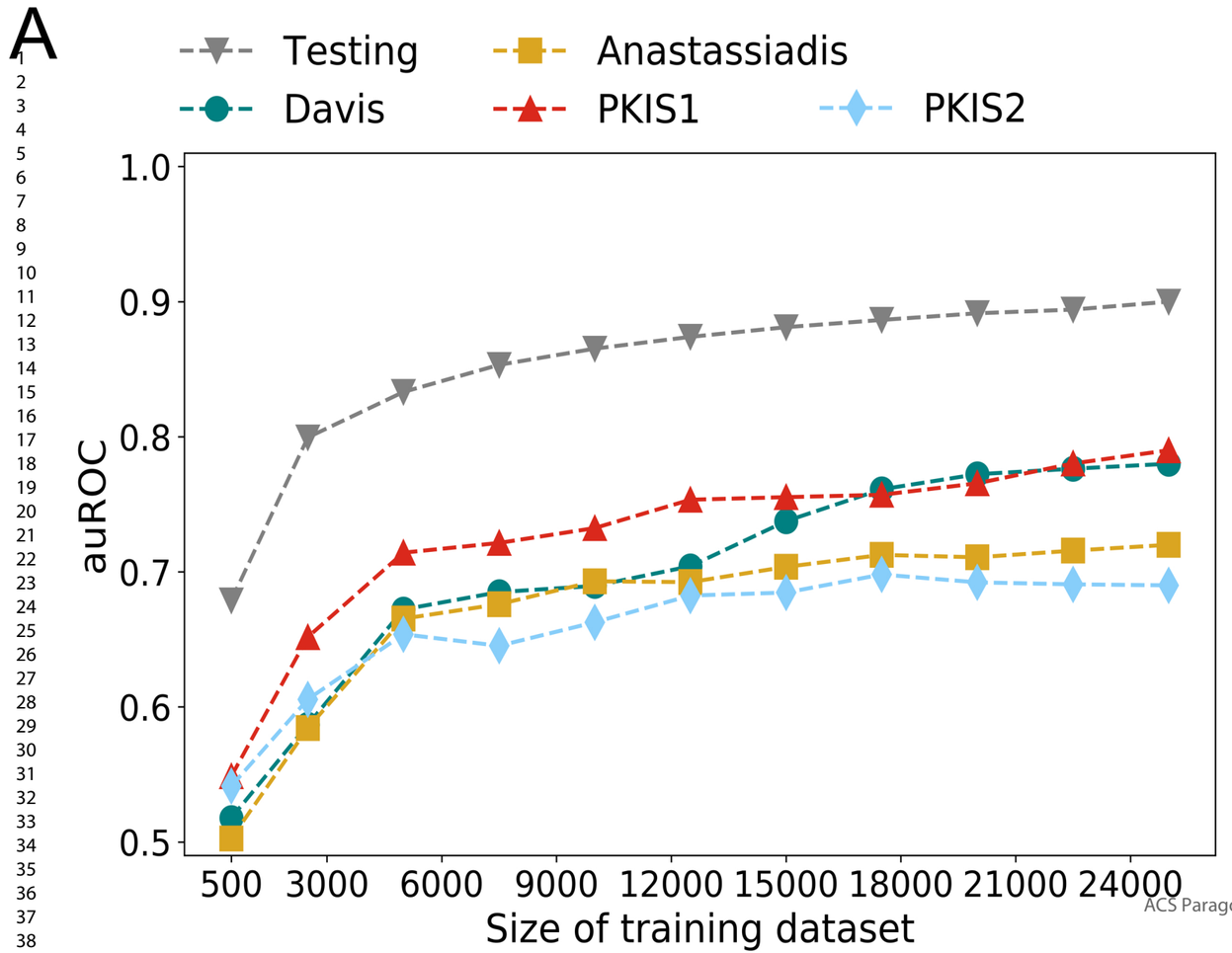


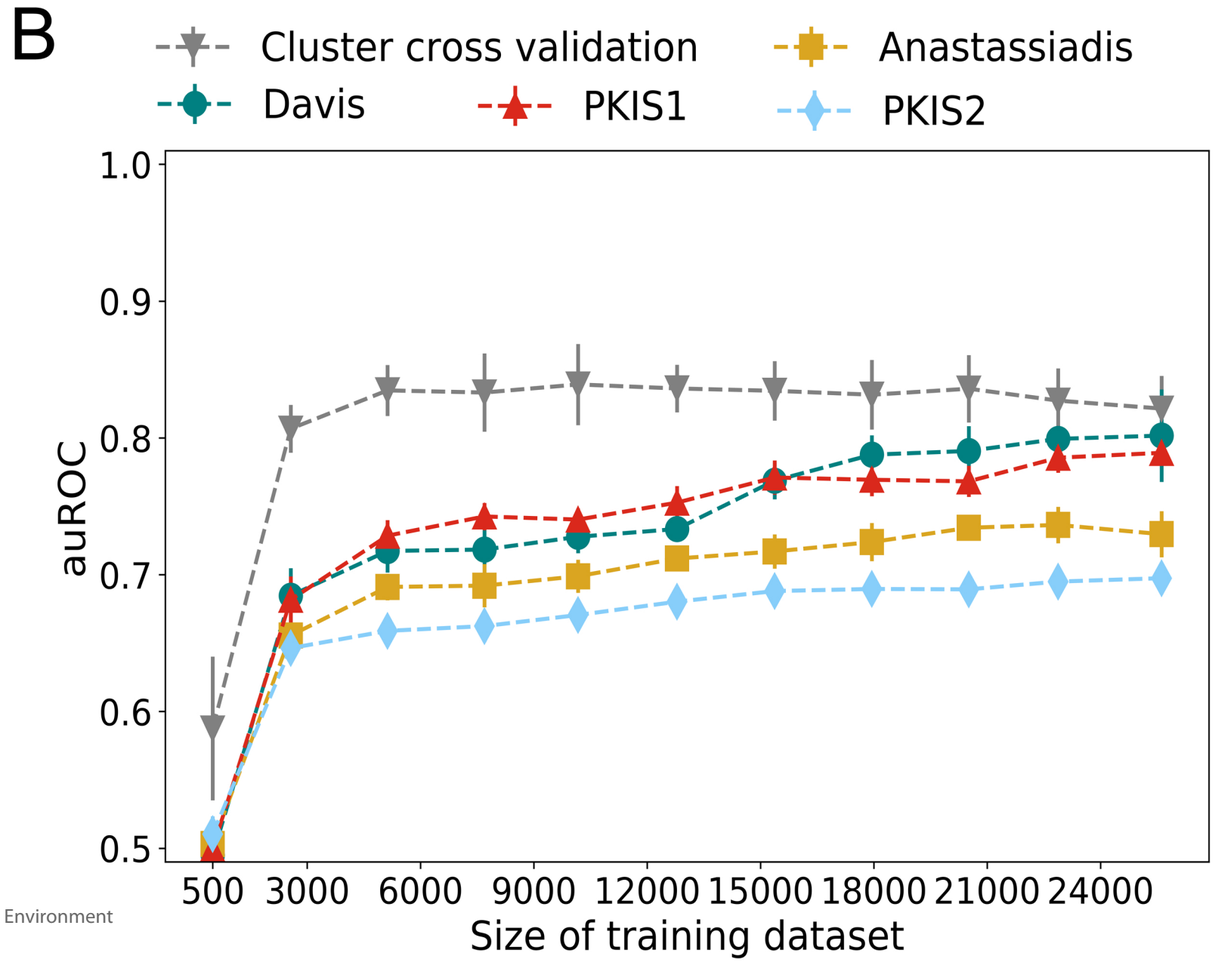
Table of Contents graphic

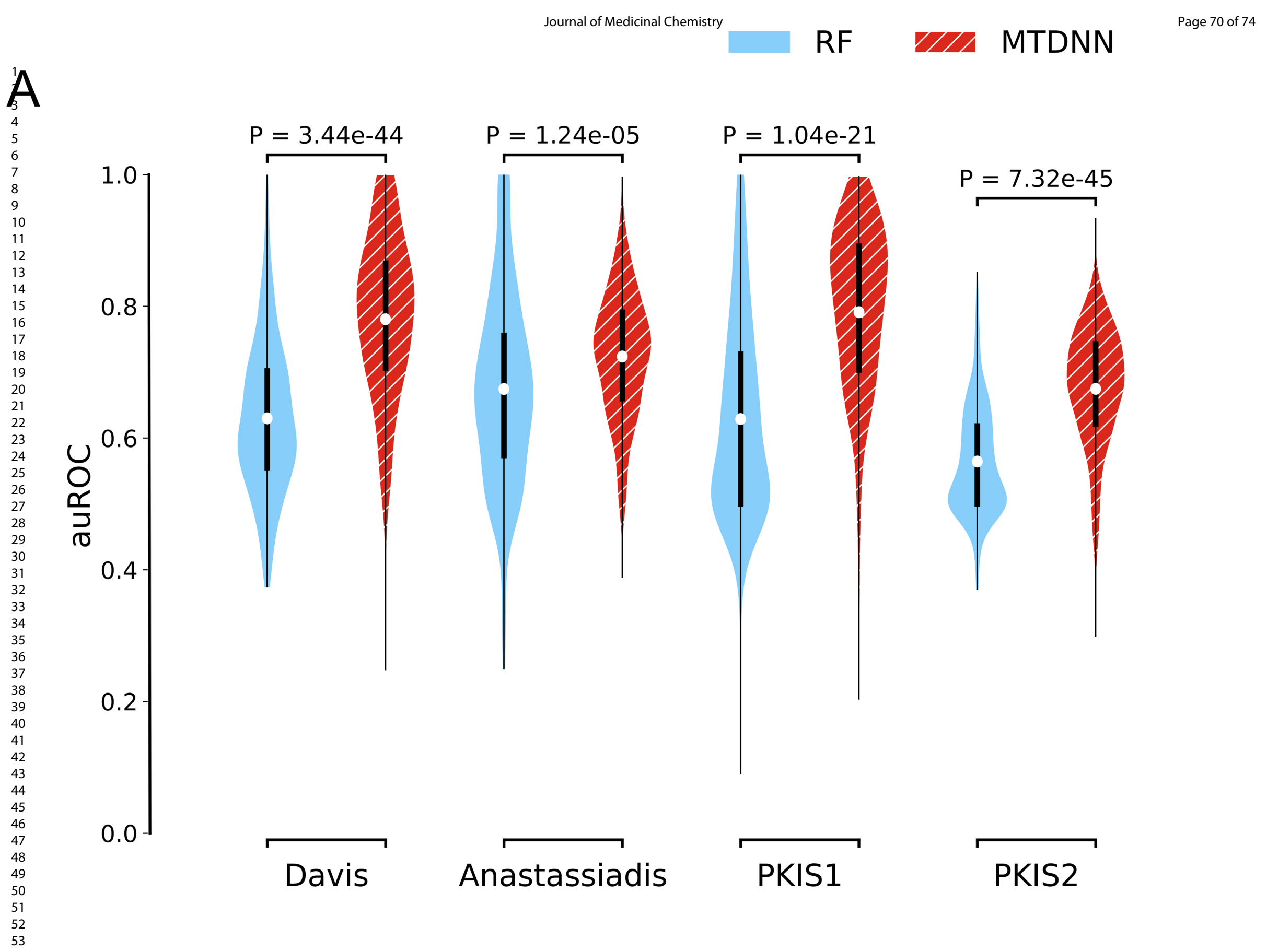

Batch size = 64
Layer size
[1000]
[1024,391]
[1500,1000]
[2000,1000,500]
0.8841
0.8972
0.8834
0.8857
0.8501
0.8910
0.9000
0.8919
0.8365
0.8904
0.8918
0.8923
0.8265
0.8824
0.8814
0.8480
1×10^3
1×10^4
5×10^5
1×10^5
Learning rate

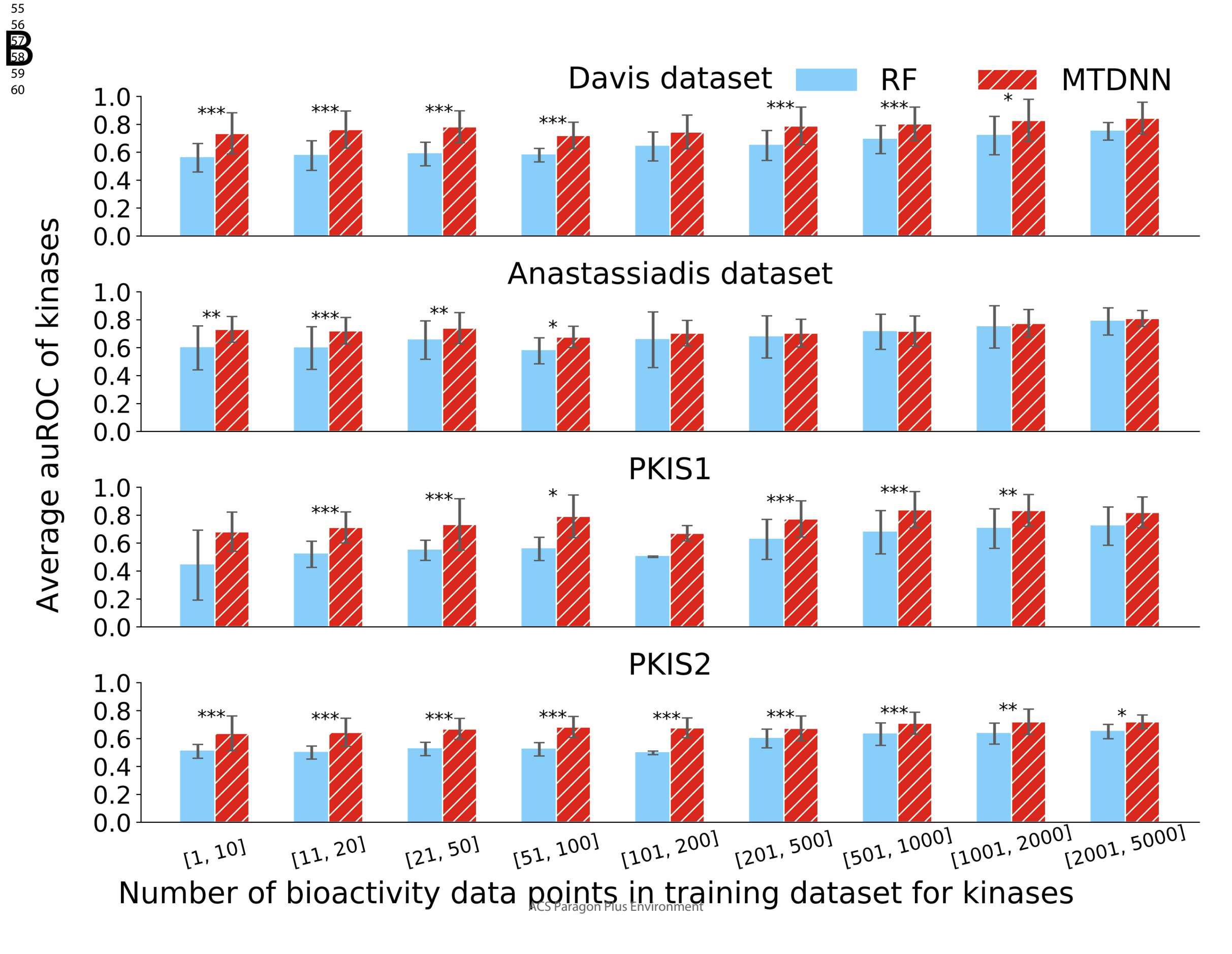

Batch size = 128
Layer size
[1000]
[1024,391]
[1500,1000]
[2000,1000,500]
0.8862
0.8936
0.8929
0.8894
0.8704
0.8944
0.9005
0.8608
0.8665
0.8950
0.8946
0.8903
0.8061
0.8781
0.8692
0.8234
1×10^3
1×10^4
5×10^5
1×10^5
Learning rate
0.90
0.88
0.86
0.84
0.82
auROC of test dataset

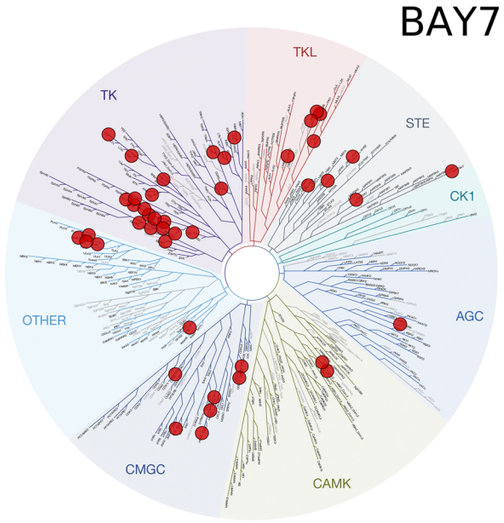
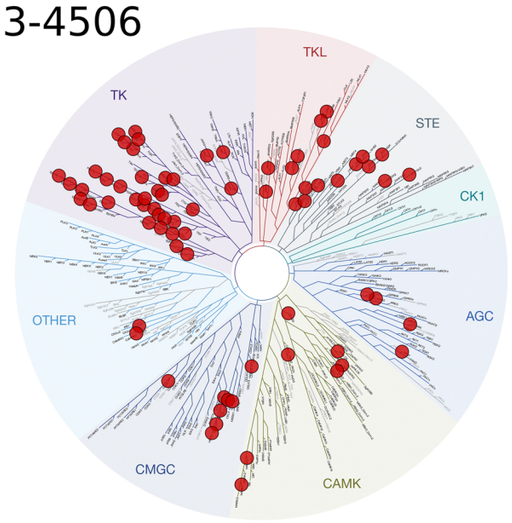
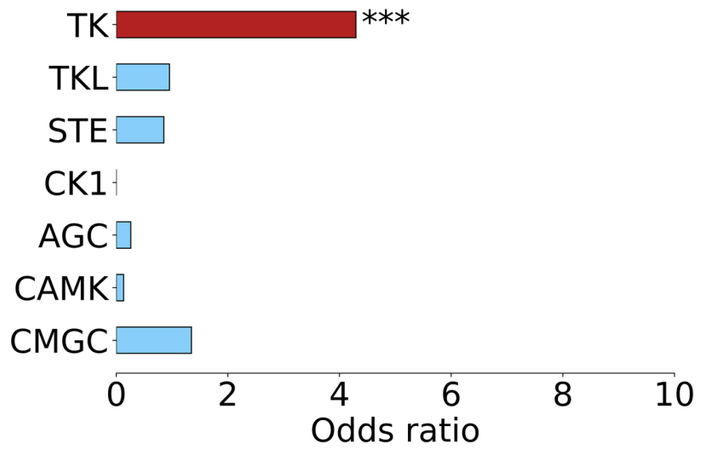
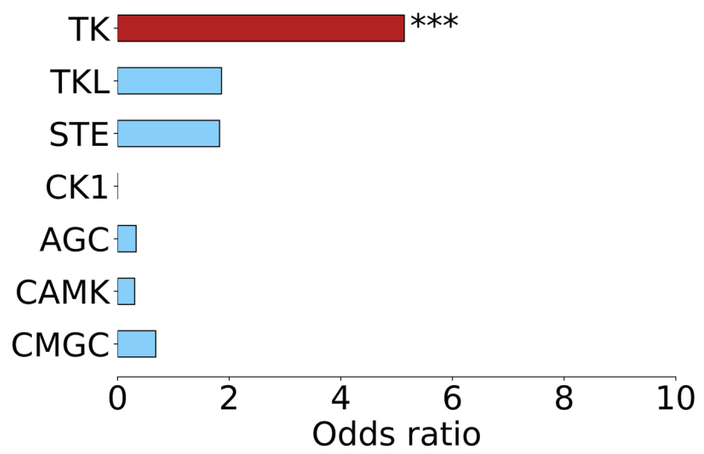
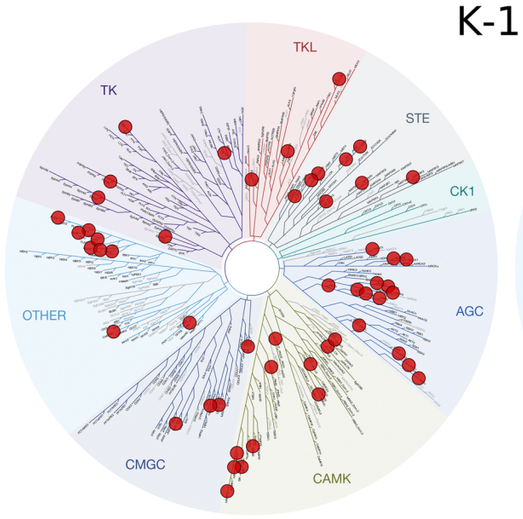
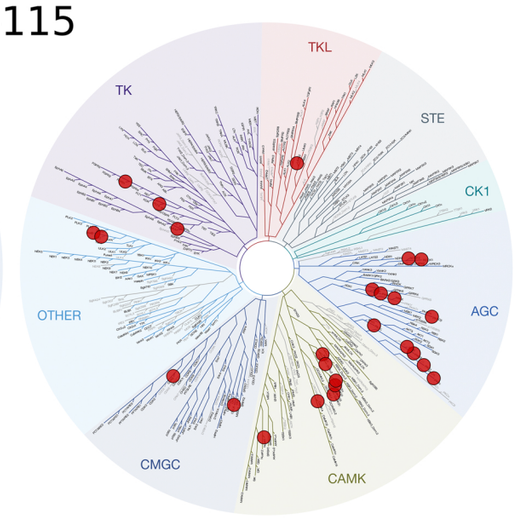
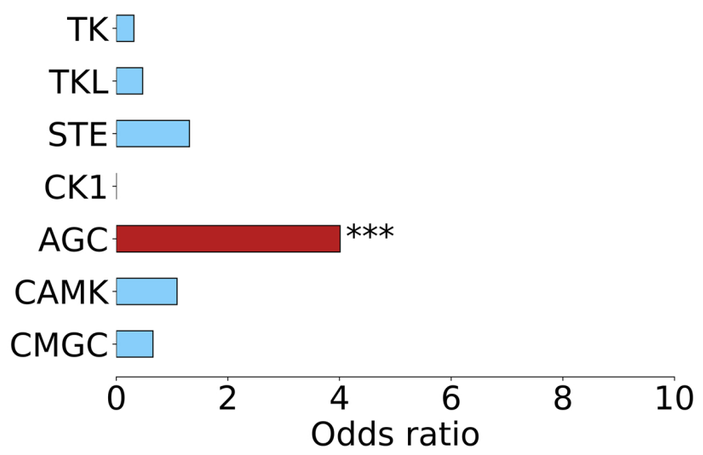
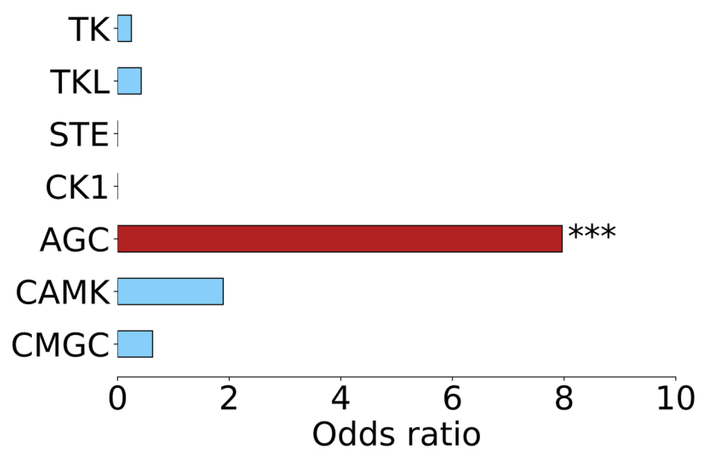
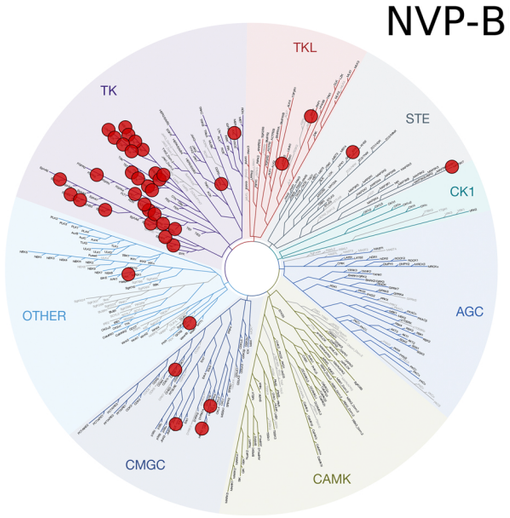
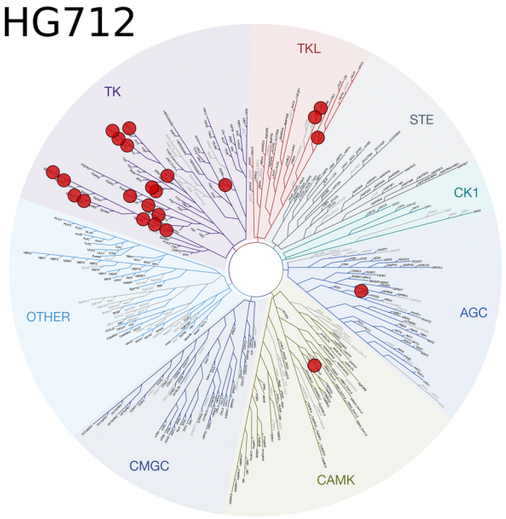
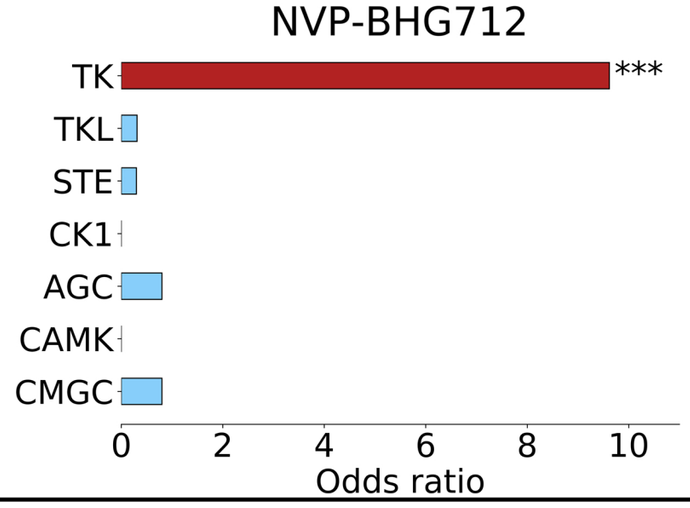
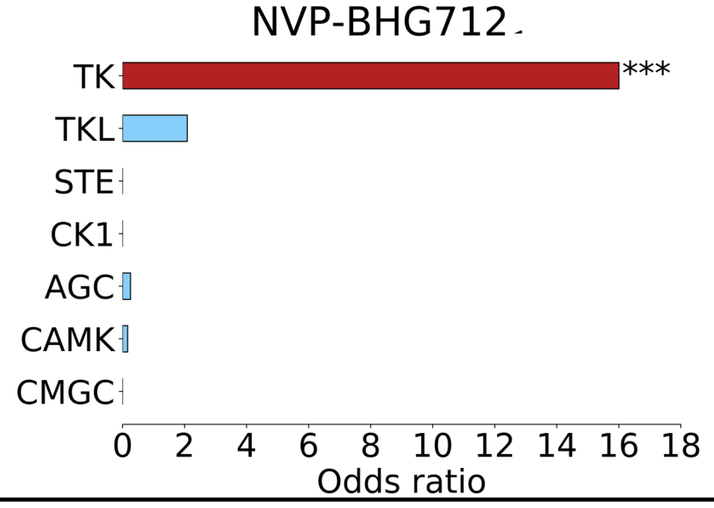
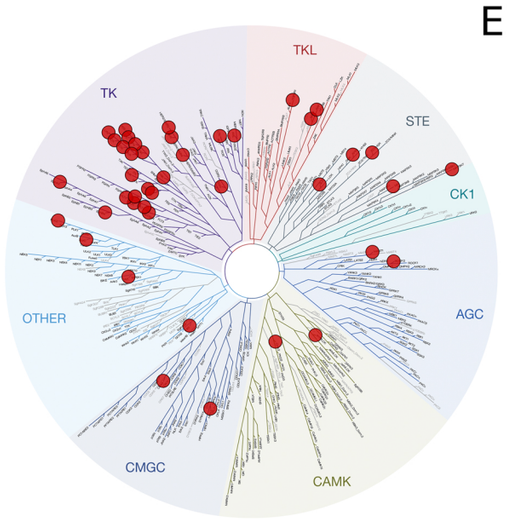
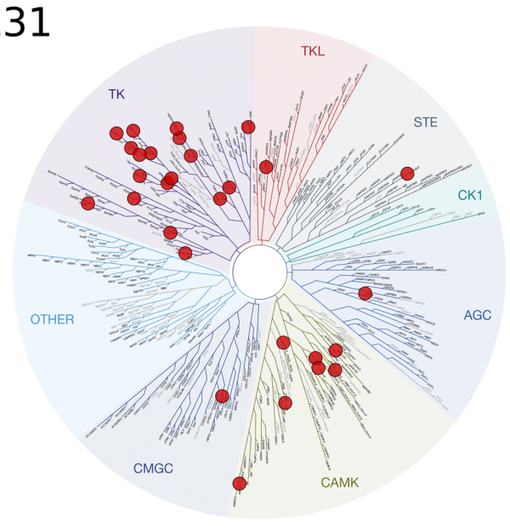
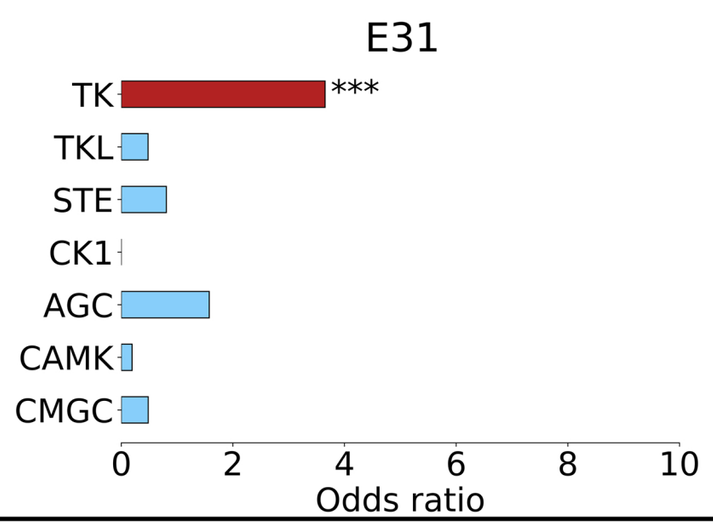
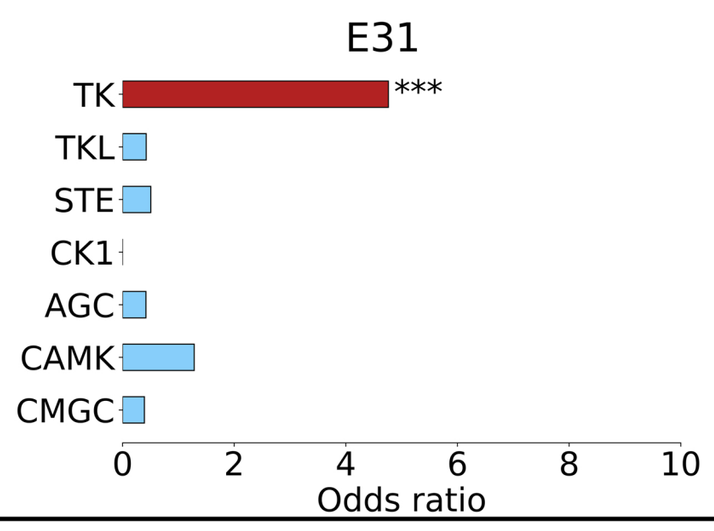
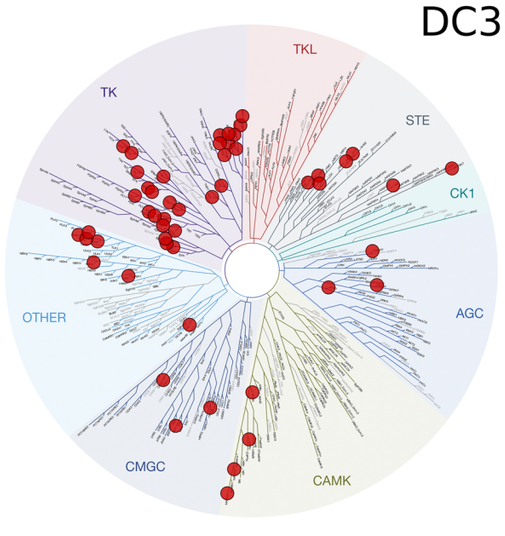
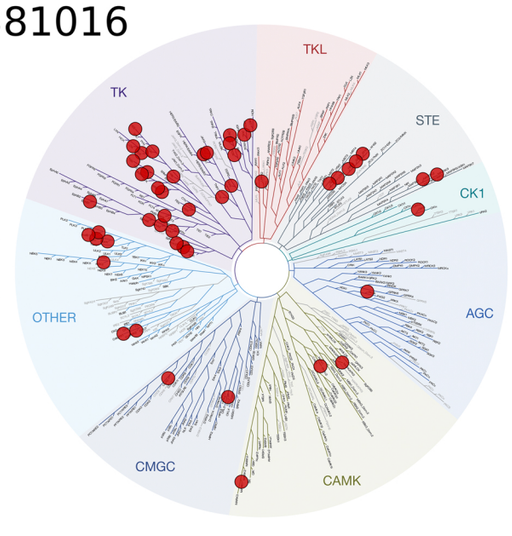
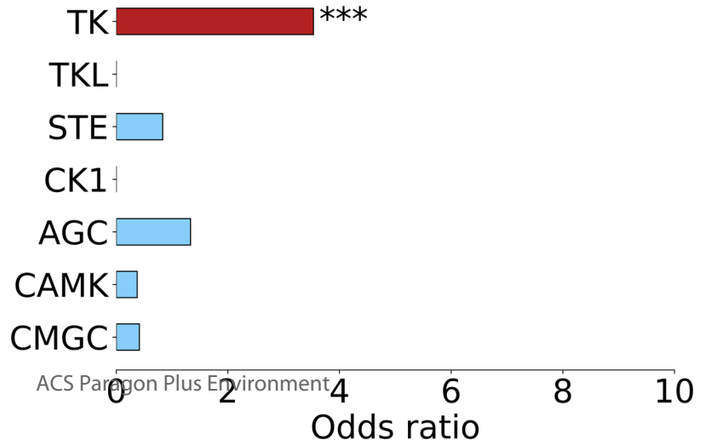
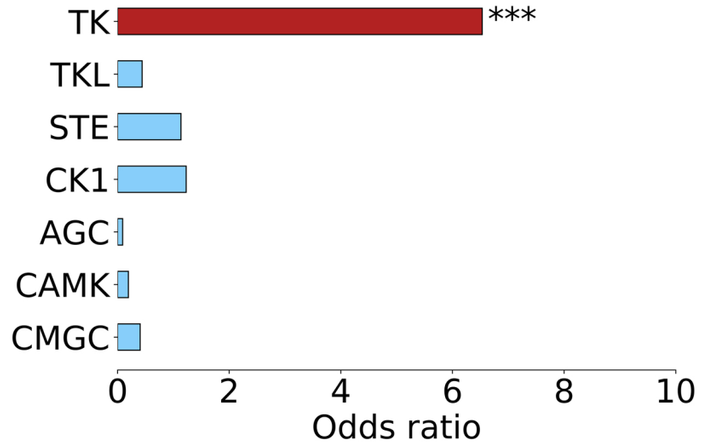

A
Davis
Training dataset
B
Anastassiadis
Training dataset
C
PKIS1
Training dataset
D
PKIS2
Training dataset
tsne1
tsne2
tsne3
0.0
0.2
0.4
0.6
0.8
1.0

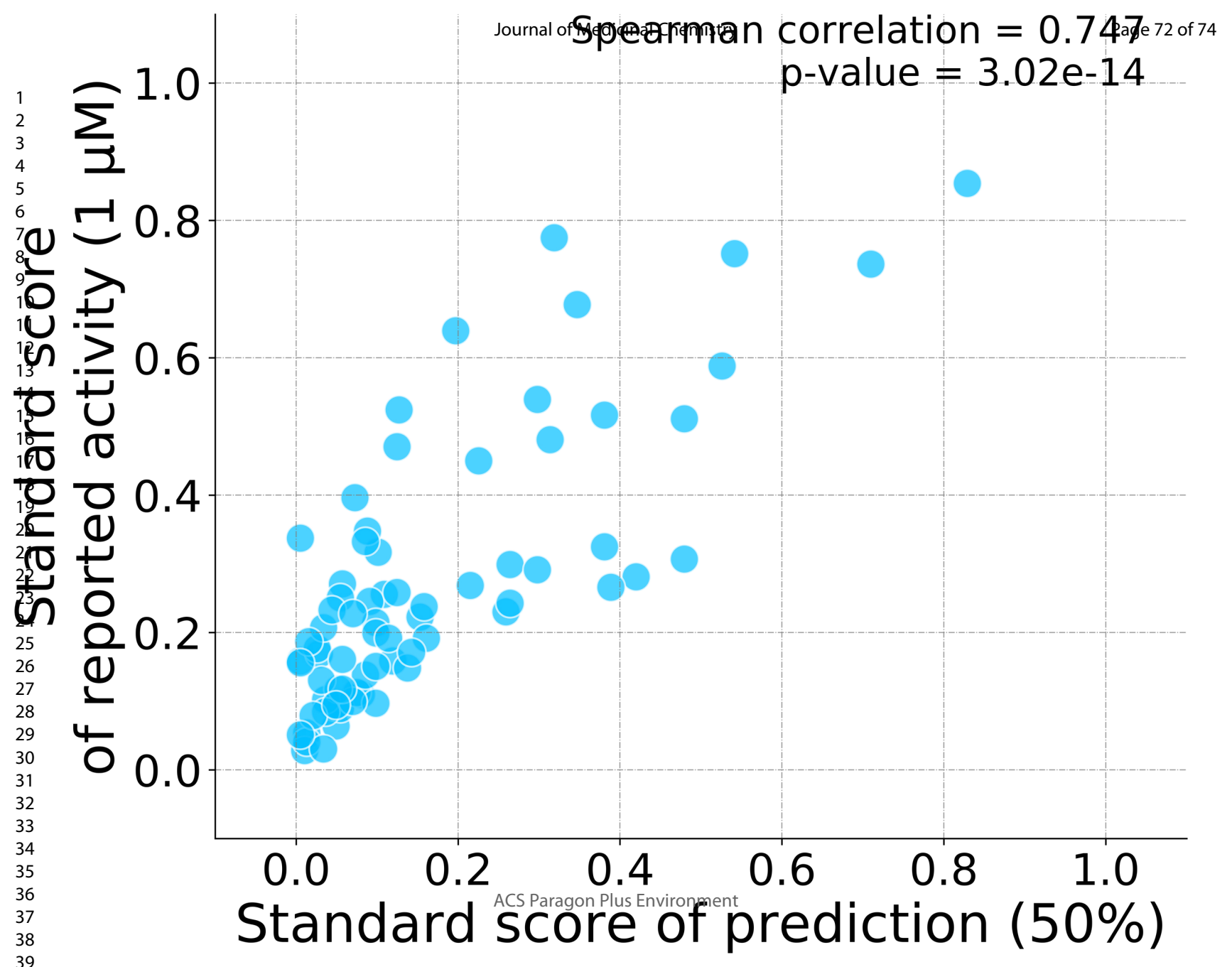

A
Testing
Anastassiadis
Davis
PKIS1
PKIS2
auROC
1.0
0.9
0.8
0.7
0.6
0.5
500 3000 6000 9000 12000 15000 18000 21000 24000
Size of training dataset

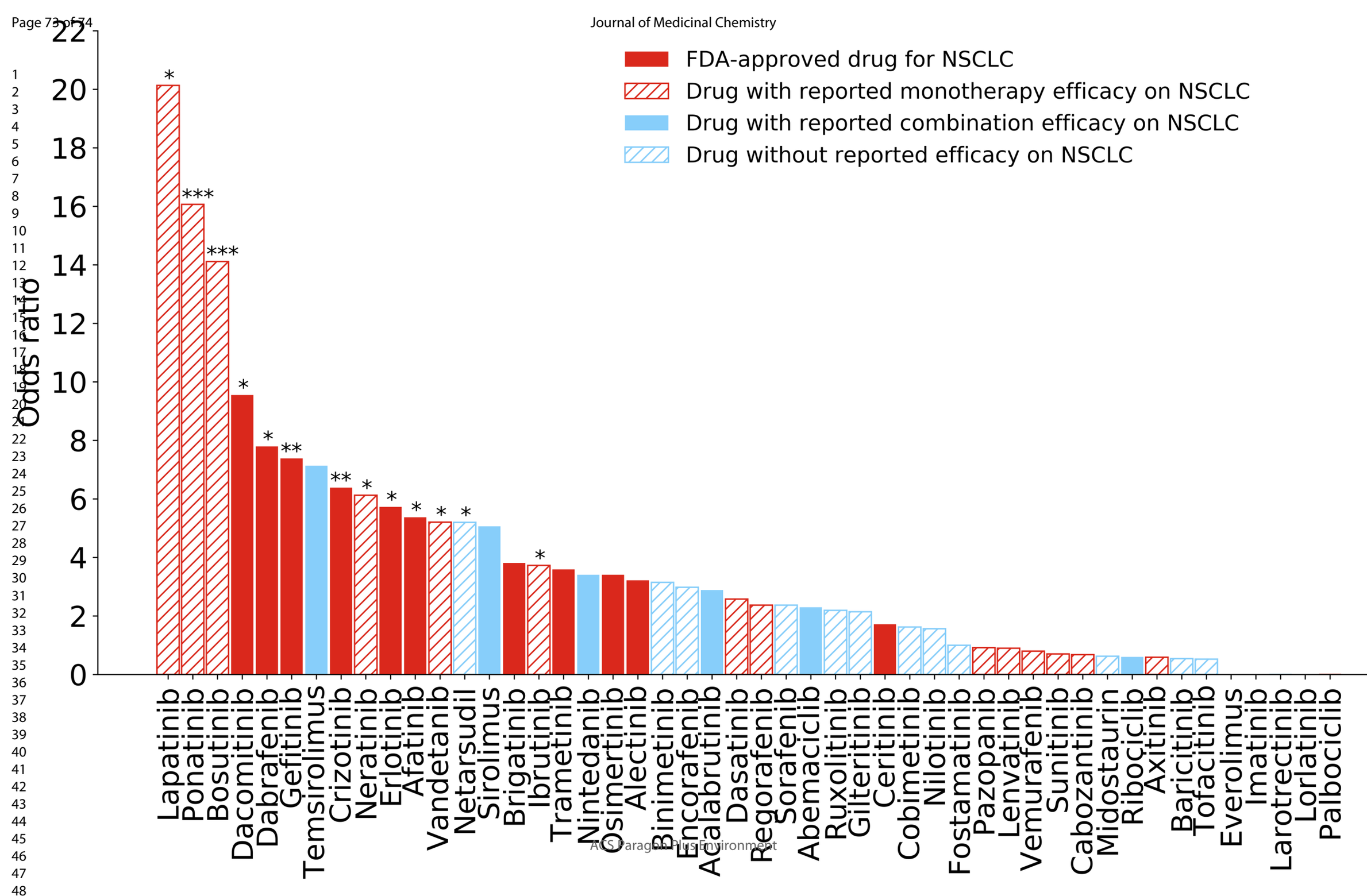

B
Cluster cross validation
Anastassiadis
Davis
PKIS1
PKIS2
auROC
1.0
0.9
0.8
0.7
0.6
0.5
500 3000 6000 9000 12000 15000 18000 21000 24000
Size of training dataset

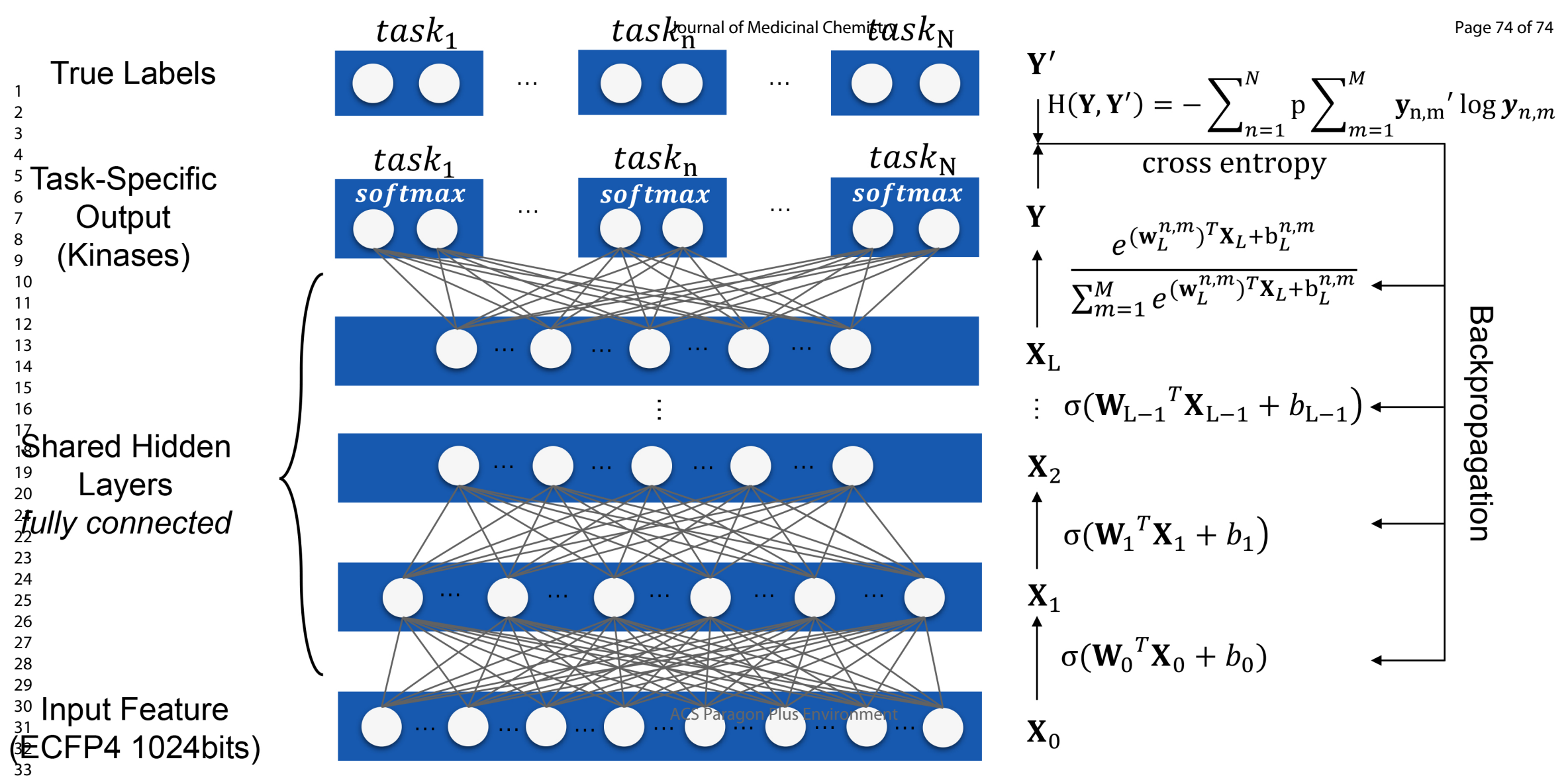

A
RF
MTDNN
P = 3.44e-44
P = 1.24e-05
P = 1.04e-21
P = 7.32e-45
auROC
1.0
0.8
0.6
0.4
0.2
0.0
Davis
Anastassiadis
PKIS1
PKIS2


B
Davis dataset
RF
MTDNN
Anastassiadis dataset
PKIS1
PKIS2
Average auROC of kinases
1.0
0.8
0.6
0.4
0.2
0.0
[1, 10]
[11, 20]
[21, 50]
[51, 100]
[101, 200]
[201, 500]
[501, 1000]
[1001, 2000]
[2001, 5000]
Number of bioactivity data points in training dataset for kinases

Predicted probability > 0.5	Experimental activity < 50%	Predicted selectivity	Experimental selectivity
BAY73-4506			
			
K-115			
			
NVP-BHG712			
			
E31			
			
DC381016			
			


Spearman correlation = 0.747
p-value = 3.02e-14
Standard score of reported activity (1 µM)
Standard score of prediction (50%)
0.0
0.2
0.4
0.6
0.8
1.0





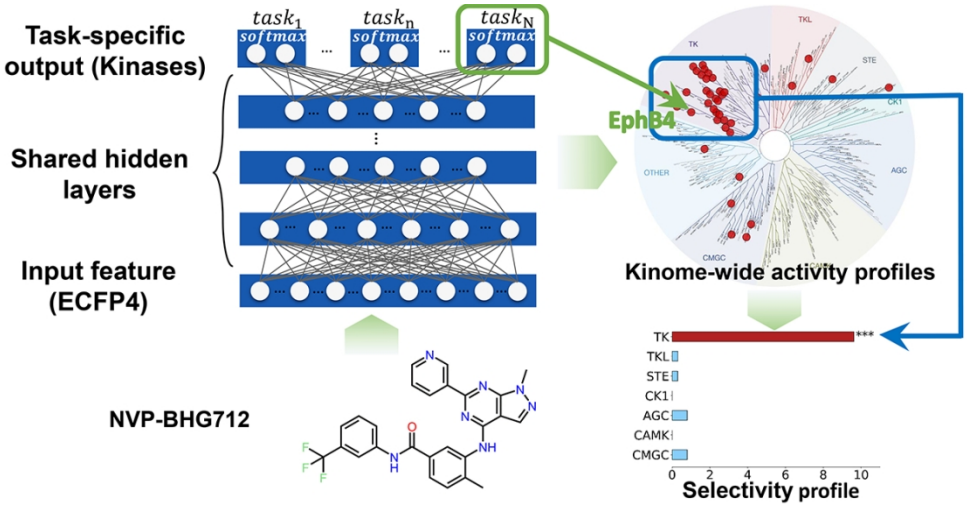


Table of Contents graphic