ToxCast Revisited: A Comprehensive Statistical Analysis of the *In Vitro*-to-*In Vivo* Predictive Capacity of High-Throughput Toxicity Screens

Russ Wolfinger

Director of Scientific Discovery and Genomics
SAS Institute Inc., Cary, NC

SRP Meeting
October 24, 2012

TOXICOLOGICAL SCIENCES **128(2)**, 398–417 (2012)
doi:10.1093/toxsci/kfs159
Advance Access publication April 26, 2012

A Comprehensive Statistical Analysis of Predicting *In Vivo* Hazard Using High-Throughput *In Vitro* Screening

Russell S. Thomas,*[,1] Michael B. Black,* Lili Li,† Eric Healy,* Tzu-Ming Chu,† Wenjun Bao,† Melvin E. Andersen,* and Russell D. Wolfinger†

**The Hamner Institutes for Health Sciences, Research Triangle Park, North Carolina 27709; and †SAS Institute Inc., Cary, North Carolina 27513*

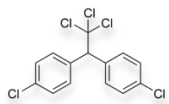
[1]To whom correspondence should be addressed at The Hamner Institutes for Health Sciences, 6 Davis Drive, Research Triangle Park, NC 27709. Fax: (919) 558-1300. E-mail: rthomas@thehamner.org.

Received February 24, 2012; accepted April 23, 2012

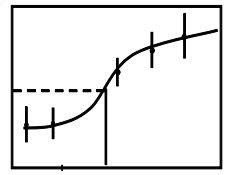
Over the past 5 years, increased attention has been focused on using high-throughput *in vitro* screening for identifying chemical hazards and prioritizing chemicals for additional *in vivo* testing. The U.S. Environmental Protection Agency's ToxCast program has generated a significant amount of high-throughput screening data allowing a broad-based assessment of the utility of these assays for predicting *in vivo* responses. In this study, a comprehensive cross-validation model comparison was performed to evaluate the predictive performance of the more than 600 *in vitro* assays from the ToxCast phase I screening effort across 60 *in vivo* endpoints using 84 different statistical classification methods. The predictive performance of the *in vitro* assays was compared and combined

hazards. In the United States, the Environmental Protection Agency's (EPA) ToxCast project and the Tox21 consortium have used high-throughput screening to characterize the *in vitro* biological activity of chemicals across multiple cellular pathways and biochemical targets with the intent to prioritize compounds for conventional toxicity testing (Dix *et al.*, 2007; Kavlock *et al.*, 2009). In Europe, multiple research consortia such as Sens-it-iv, ACuteTox, and carcinoGENOMICS have also examined the ability of *in vitro* systems for predicting *in vivo* responses (dos Santos *et al.*, 2009; Jennen *et al.*, 2010; Sjostrom *et al.*, 2008). To date, the research efforts have produced mixed results

ToxCast Data from the EPA



~200 Cellular Assays



***In Vivo* Endpoints**

~300 Biochemical Assays

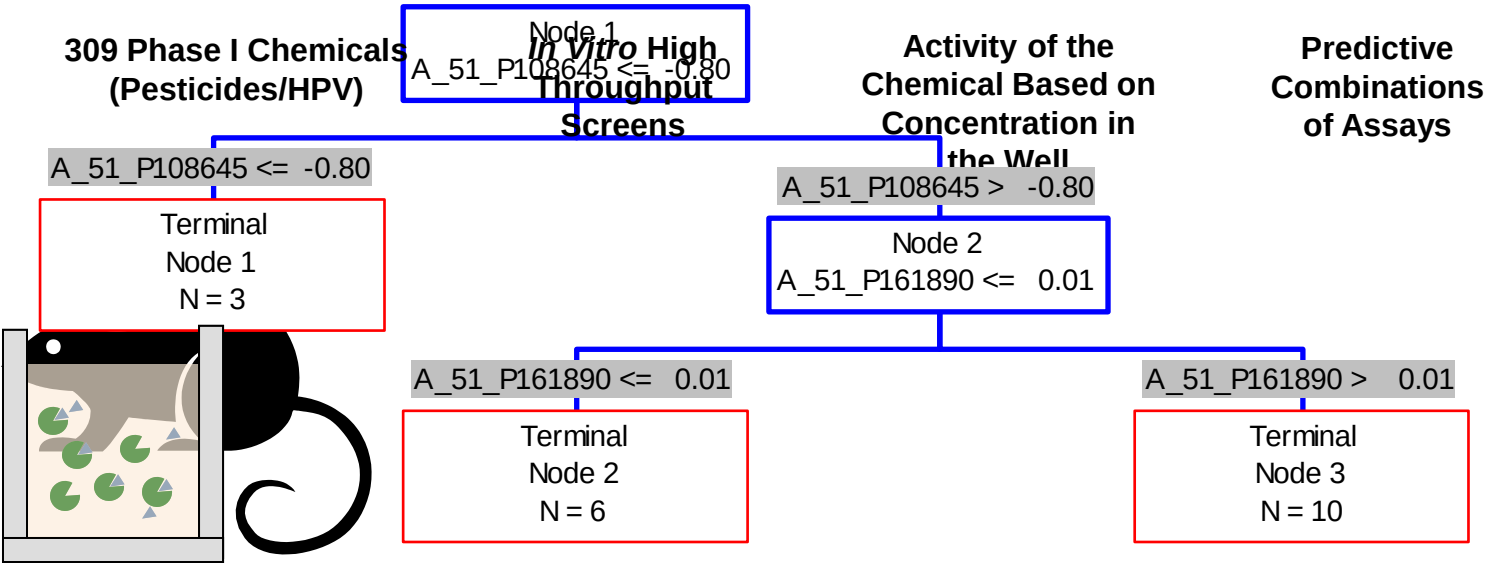
309 Phase I Chemicals (Pesticides/HPV)

***In Vitro* High Throughput Screens**

Activity of the Chemical Based on Concentration in the Well

Predictive Combinations of Assays

***In Vivo* Hazard Prediction and Prioritization**



Currently Published Work on Predictive Toxicity Signatures in ToxCast

Signature Development

BIOLOGY OF REPRODUCTION **85,** 327–339 (2011)
Published online before print 12 May 2011.
DOI 10.1095/biolreprod.111.090977

Predictive Model of Rat Reproductive Toxicity from ToxCast High Throughput Screening[1]

Matthew T. Martin,[2] Thomas B. Knudsen, David M. Reif, Keith A. Houck, Richard S. Judson, Robert J. Kavlock, and David J. Dix

National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina

- Reproductive toxicity signature
- 74% Balanced Accuracy
- **Pre-filtered assays and lumped subset into into 6 classes based on genes and functional grouping**
- **<u>Only study with external validation set</u>**

***In Vitro* Screening of Environmental Chemicals for Targeted Testing Prioritization: The ToxCast Project**

Richard S. Judson, Keith A. Houck, Robert J. Kavlock, Thomas B. Knudsen, Matthew T. Martin, Holly M. Mortensen, David M. Reif, Daniel M. Rotroff, Imran Shah, Ann M. Richard, and David J. Dix

National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA

- Rat liver tumor signature
- No formal classification statistical analysis (cross-validation)

TOXICOLOGICAL SCIENCES **124(1)**, 109–127 (2011)
doi:10.1093/toxsci/kfr220
Advance Access publication August 26, 2011

Predictive Models of Prenatal Developmental Toxicity from ToxCast High-Throughput Screening Data

Nisha S. Sipes,*,[1] Matthew T. Martin,* David M. Reif,* Nicole C. Kleinstreuer,* Richard S. Judson,* Amar V. Singh,† Kelly J. Chandler,‡ David J. Dix,* Robert J. Kavlock,* and Thomas B. Knudsen*

**National Center for Computational Toxicology, Office of Research & Development, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711; †Lockheed Martin, Research Triangle Park, North Carolina 27711; and ‡National Health and Environmental Effects Research Laboratory, Office of Research & Development, U. S. Environmental Protection Agency, Research Triangle Park, North Carolina 27711*

- Developmental toxicity signature
- 71% Balanced Accuracy
- **Pre-filtered assays and aggregated assays based on genes and GO categories**

Environmental Impact on Vascular Development Predicted by High Throughput Screening

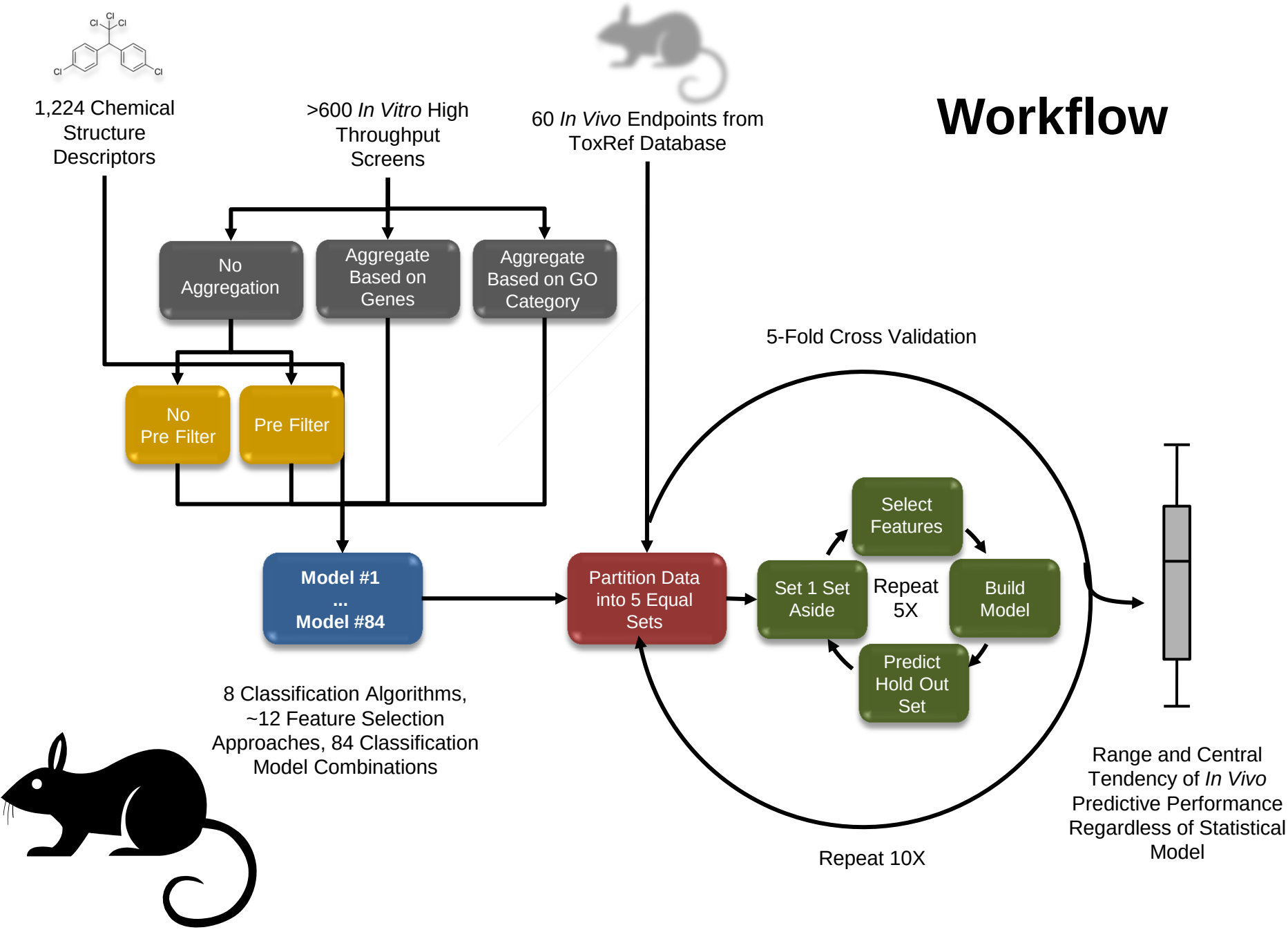
Nicole C. Kleinstreuer[1], Richard S. Judson[1], David M. Reif[4], Nisha S. Sipes[1], Amar V. Singh[2], Kelly J. Chandler[1,3], Rob DeWoskin[4], David J. Dix[1], Robert J. Kavlock[1] and Thomas B. Knudsen[1]

- Vascular development signature
- 80% Accuracy

Why is a Software Company Getting Involved with ToxCast?

- Our life sciences team has collaborated with the EPA, Hamner Institute, NIEHS, UNC, and NC State for many years, and has its roots in toxicology-based microarray data analysis.
- The ToxCast data is highly valuable and presents numerous analytical challenges, several of which JMP Genomics software can help address.
- We wanted to test and stretch the software in new directions and participate in the project by providing, as much as possible, a “neutral third party” assessment of the predictive performance of the assays.
- Previous work in collaboration with Fred Wright at UNC; current work in collaboration with Rusty Thomas at Hamner.

Workflow


1,224 Chemical Structure Descriptors
>600 *In Vitro* High Throughput Screens
60 *In Vivo* Endpoints from ToxRef Database
No Aggregation
Aggregate Based on Genes
Aggregate Based on GO Category
No Pre Filter
Pre Filter
Model #1
...
Model #84
8 Classification Algorithms, ~12 Feature Selection Approaches, 84 Classification Model Combinations
5-Fold Cross Validation
Partition Data into 5 Equal Sets
Set 1 Set Aside
Select Features
Build Model
Predict Hold Out Set
Repeat 5X
Repeat 10X
Range and Central Tendency of *In Vivo* Predictive Performance Regardless of Statistical Model

Workflow Details

- 60 Endpoints x 84 Models x 10 Iterations x 5 Folds = 252,000 separate model fits; computationally intensive.

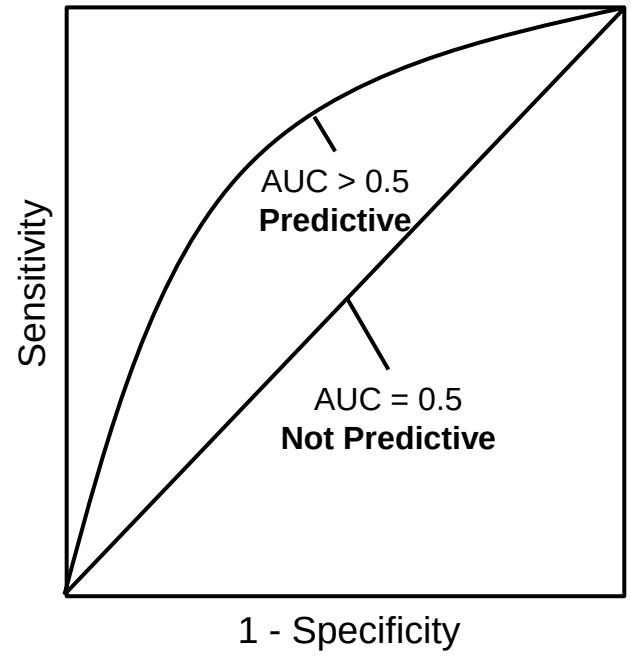
- Predictive models include discriminant, distance scoring, k-nearest neighbors, logistic regression, general linear model, partial least squares, partition trees, and radial basis machine.

- Various Sets of Predictor Variables:
  1. *In Vitro* Assays, optionally aggregated
  2. Chemical Structural Descriptors, computed by Dragon

- All endpoints are binary, so performance criteria include: Accuracy, Balanced Accuracy, Sensitivity, Specificity, Negative Predictive Value (NPV), Positive Predictive Value (PPV), and Area Under the Curve (AUC)

Predictive Performance Criteria for Binary Endpoints

		In Vivo Animal Response: Positive	*In Vivo* Animal Response: Negative	
Prediction Based on *In Vitro* Assays or Chemical Structure	Positive	TP	FP	PPV = TP / (TP + FP)
	Negative	FN	TN	NPV = TN / (FN + TN)
		Sensitivity = TP / (TP + FN)	Specificity = TN / (FP + TN)	



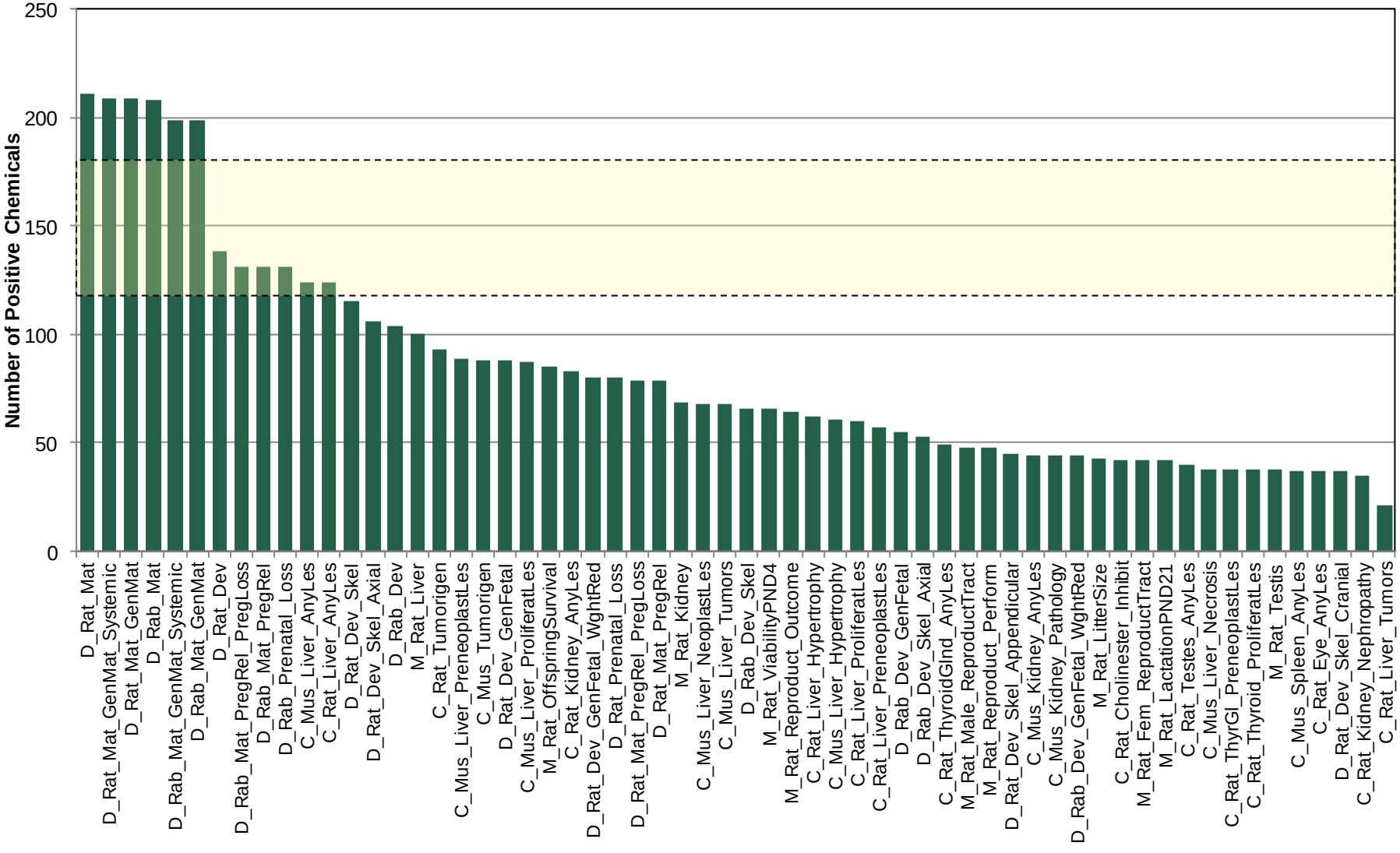
Accuracy = (TP + TN)/(TP + FP + FN + TN)

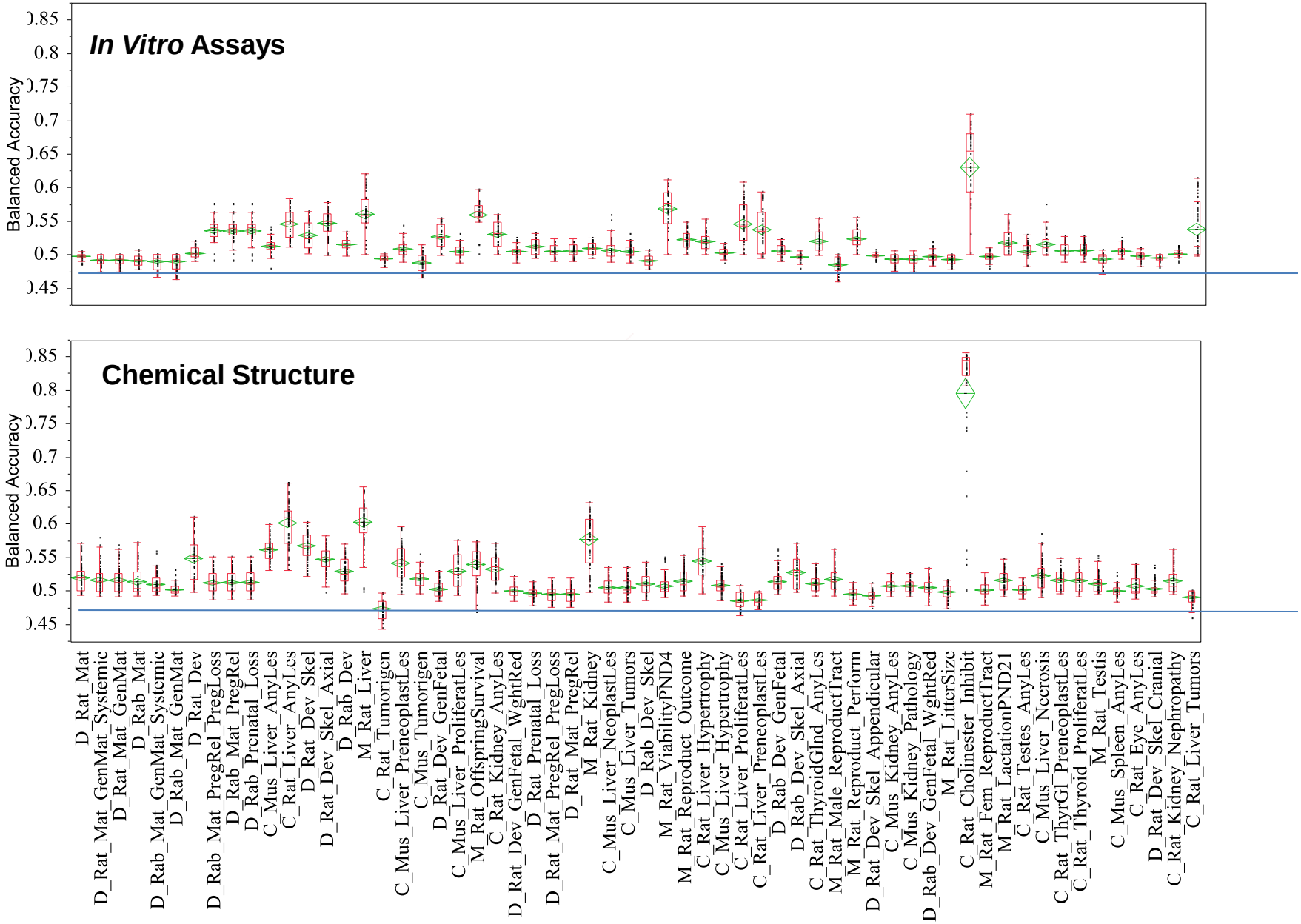
Balanced Accuracy = (Sensitivity + Specificity)/2, adjusts for prevalence

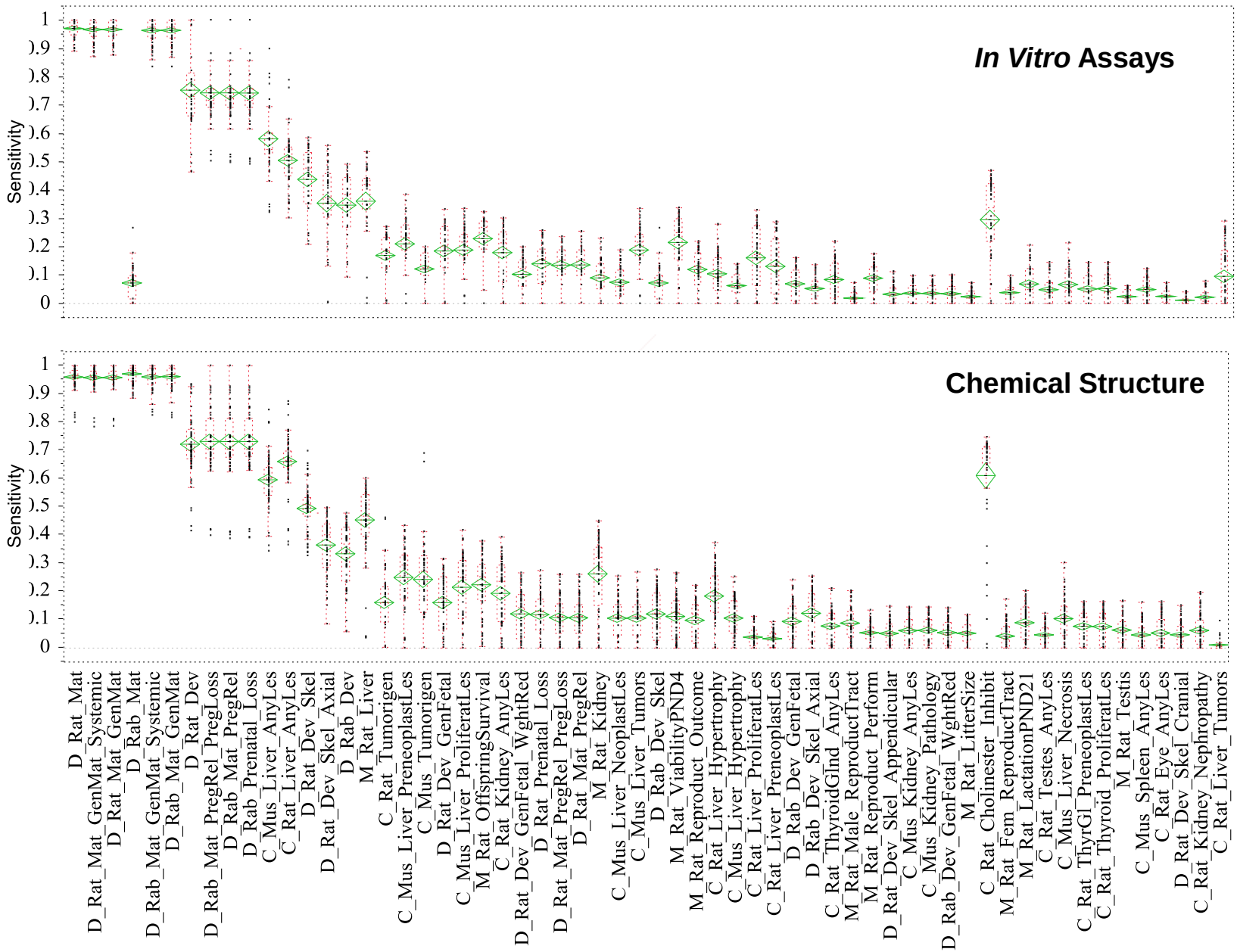
RMSE = square root (average (true value – predicted probability)^2)

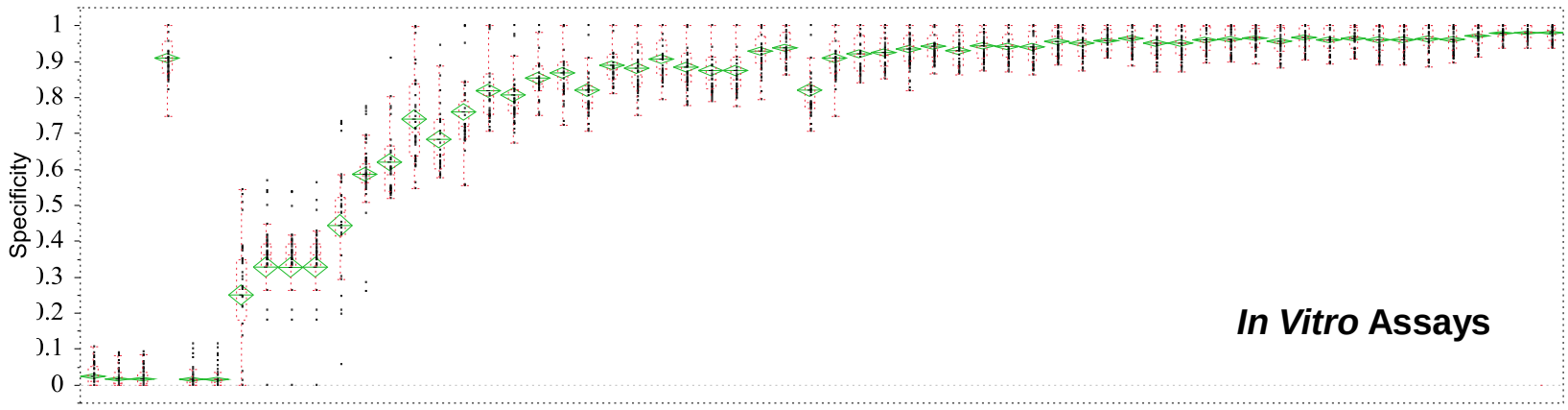
AUC is area under the Receiver Operating Characteristic (ROC) curve; it measures sorting efficiency and is directly related to the Mann-Whitney rank-sum statistic; a value of 1 indicates perfect sorting.

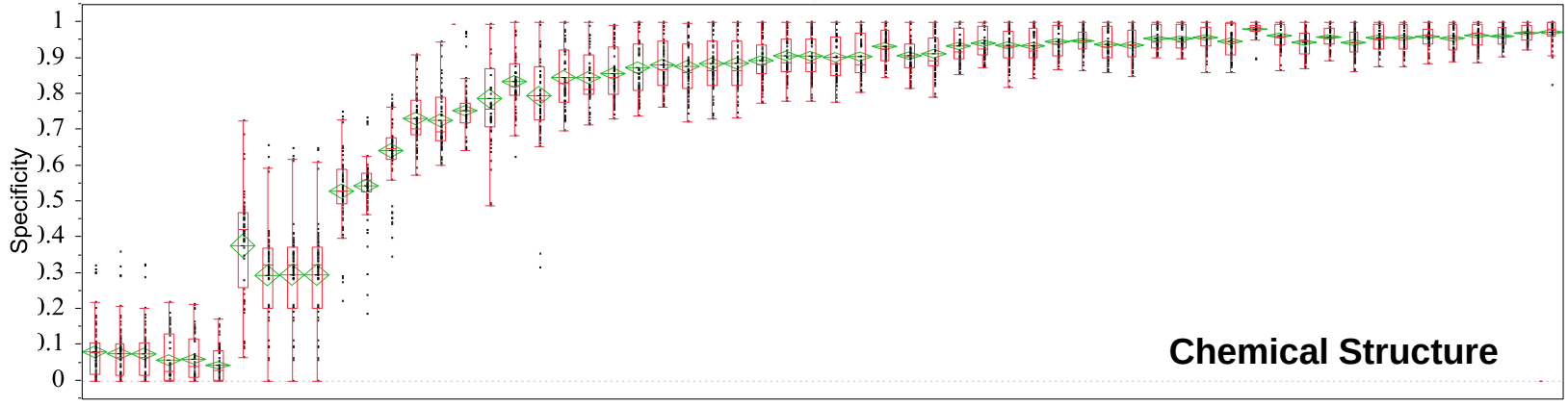
Prevalence of Positive Chemicals Among Endpoints Should be Kept in Mind



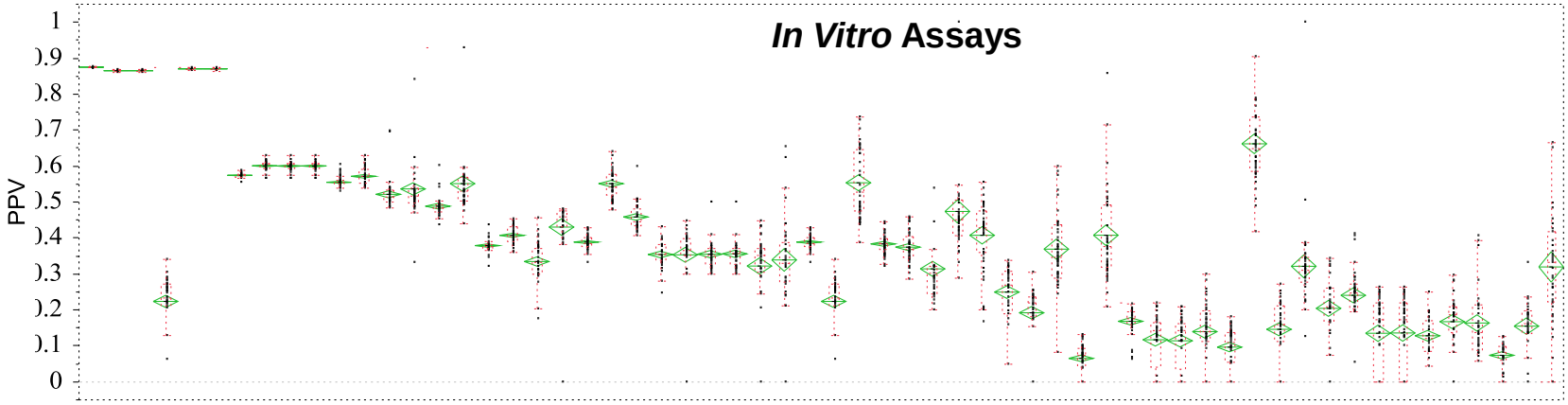

In Vitro Assays
Chemical Structure
Balanced Accuracy
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
D_Rat_Mat
D_Rat_Mat_GenMat_Systemic
D_Rat_Mat_GenMat
D_Rab_Mat
D_Rab_Mat_GenMat_Systemic
D_Rab_Mat_GenMat
D_Rat_Dev
D_Rab_Mat_PregRel_PregLoss
D_Rab_Mat_PregRel
D_Rab_Prenatal_Loss
C_Mus_Liver_AnyLes
C_Rat_Liver_AnyLes
D_Rat_Dev_Skel
D_Rat_Dev_Skel_Axial
D_Rab_Dev
M_Rat_Liver
C_Rat_Tumorigen
C_Mus_Liver_PreneoplastLes
C_Mus_Tumorigen
D_Rat_Dev_GenFetal
C_Mus_Liver_ProliferatLes
M_Rat_OffspringSurvival
C_Rat_Kidney_AnyLes
D_Rat_Dev_GenFetal_WghtRed
D_Rat_Prenatal_Loss
D_Rat_Mat_PregRel_PregLoss
D_Rat_Mat_PregRel
M_Rat_Kidney
C_Mus_Liver_NeoplastLes
C_Mus_Liver_Tumors
D_Rab_Dev_Skel
M_Rat_ViabilityPND4
M_Rat_Reproduct_Outcome
C_Rat_Liver_Hypertrophy
C_Mus_Liver_Hypertrophy
C_Rat_Liver_ProliferatLes
C_Rat_Liver_PreneoplastLes
D_Rab_Dev_GenFetal
D_Rab_Dev_Skel_Axial
C_Rat_ThyroidGlnd_AnyLes
M_Rat_Male_ReproductTract
M_Rat_Reproduct_Perform
D_Rat_Dev_Skel_Appendicular
C_Mus_Kidney_AnyLes
C_Mus_Kidney_Pathology
D_Rab_Dev_GenFetal_WghtRed
M_Rat_LitterSize
C_Rat_Cholinester_Inhibit
M_Rat_Fem_ReproductTract
M_Rat_LactationPND21
C_Rat_Testes_AnyLes
C_Mus_Liver_Necrosis
C_Rat_ThyrGl_PreneoplastLes
C_Rat_Thyroid_ProliferatLes
M_Rat_Testis
C_Mus_Spleen_AnyLes
C_Rat_Eye_AnyLes
D_Rat_Dev_Skel_Cranial
C_Rat_Kidney_Nephropathy
C_Rat_Liver_Tumors

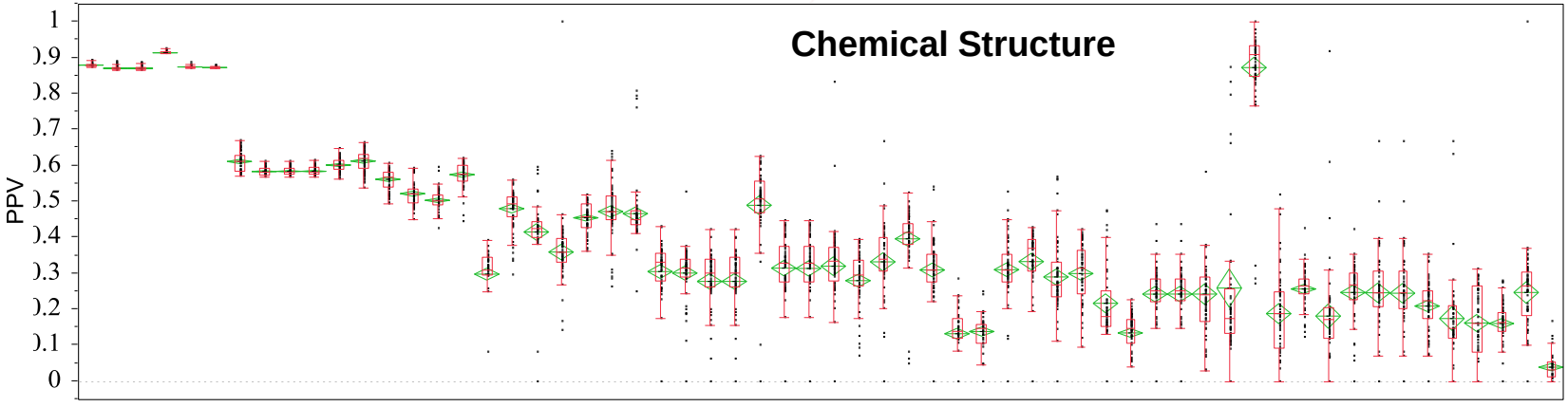

In Vitro Assays
Chemical Structure
Sensitivity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
D_Rat_Mat
D_Rat_Mat_GenMat_Systemic
D_Rat_Mat_GenMat
D_Rab_Mat
D_Rab_Mat_GenMat_Systemic
D_Rab_Mat_GenMat
D_Rat_Dev
D_Rab_Mat_PregRel_PregLoss
D_Rab_Mat_PregRel
D_Rab_Prenatal_Loss
C_Mus_Liver_AnyLes
C_Rat_Liver_AnyLes
D_Rat_Dev_Skel
D_Rat_Dev_Skel_Axial
D_Rab_Dev
M_Rat_Liver
C_Rat_Tumorigen
C_Mus_Liver_PreneoplastLes
C_Mus_Tumorigen
D_Rat_Dev_GenFetal
C_Mus_Liver_ProliferatLes
M_Rat_OffspringSurvival
C_Rat_Kidney_AnyLes
D_Rat_Dev_GenFetal_WghtRed
D_Rat_Prenatal_Loss
D_Rat_Mat_PregRel_PregLoss
D_Rat_Mat_PregRel
M_Rat_Kidney
C_Mus_Liver_NeoplastLes
C_Mus_Liver_Tumors
D_Rab_Dev_Skel
M_Rat_ViabilityPND4
M_Rat_Reproduct_Outcome
C_Rat_Liver_Hypertrophy
C_Mus_Liver_Hypertrophy
C_Rat_Liver_ProliferatLes
C_Rat_Liver_PreneoplastLes
D_Rab_Dev_GenFetal
D_Rab_Dev_Skel_Axial
C_Rat_ThyroidGlnd_AnyLes
M_Rat_Male_ReproductTract
M_Rat_Reproduct_Perform
D_Rat_Dev_Skel_Appendicular
C_Mus_Kidney_AnyLes
C_Mus_Kidney_Pathology
D_Rab_Dev_GenFetal_WghtRed
M_Rat_LitterSize
C_Rat_Cholinester_Inhibit
M_Rat_Fem_ReproductTract
M_Rat_LactationPND21
C_Rat_Testes_AnyLes
C_Mus_Liver_Necrosis
C_Rat_ThyrGl_PreneoplastLes
C_Rat_Thyroid_ProliferatLes
M_Rat_Testis
C_Mus_Spleen_AnyLes
C_Rat_Eye_AnyLes
D_Rat_Dev_Skel_Cranial
C_Rat_Kidney_Nephropathy
C_Rat_Liver_Tumors


Specificity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
In Vitro Assays

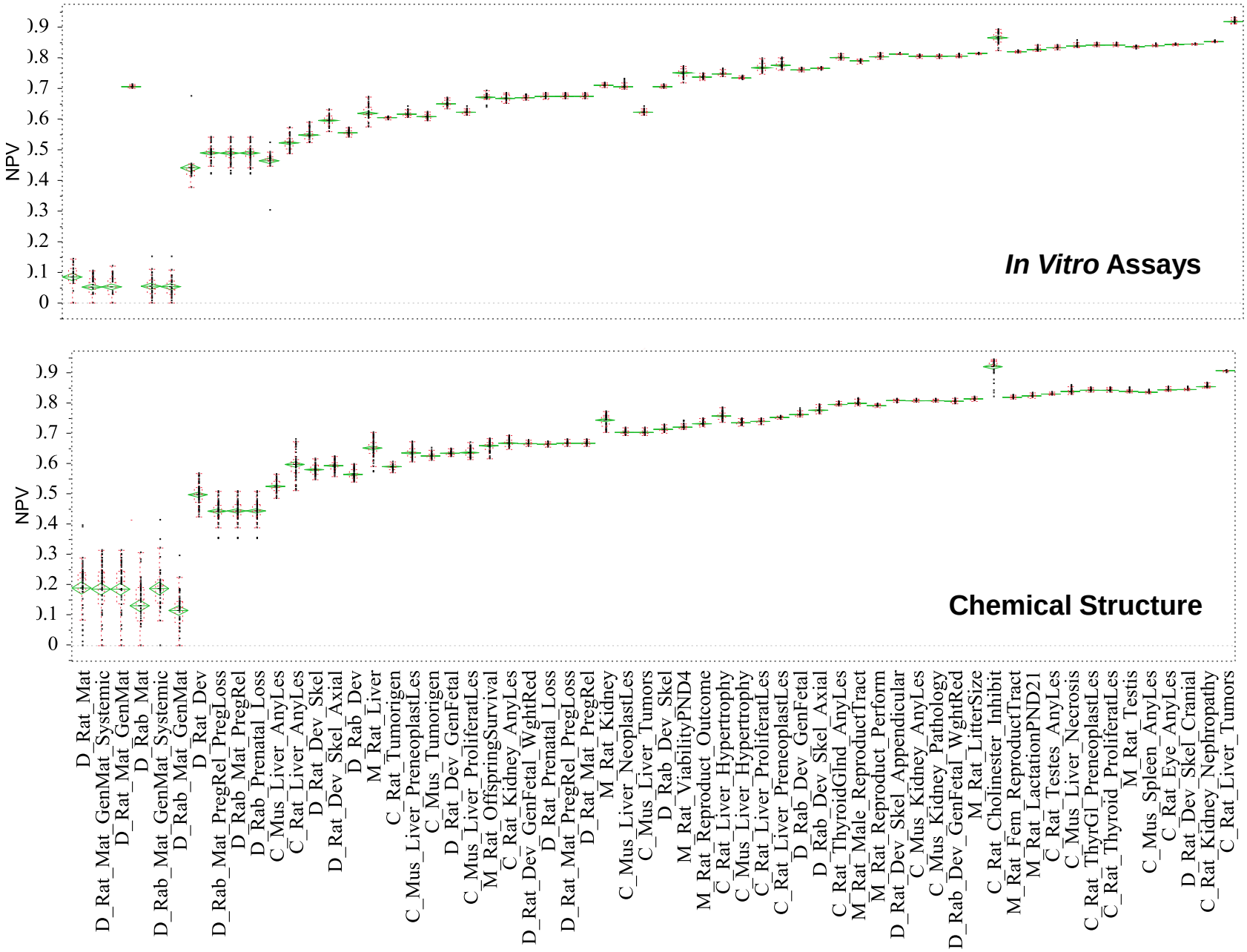

Specificity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Chemical Structure

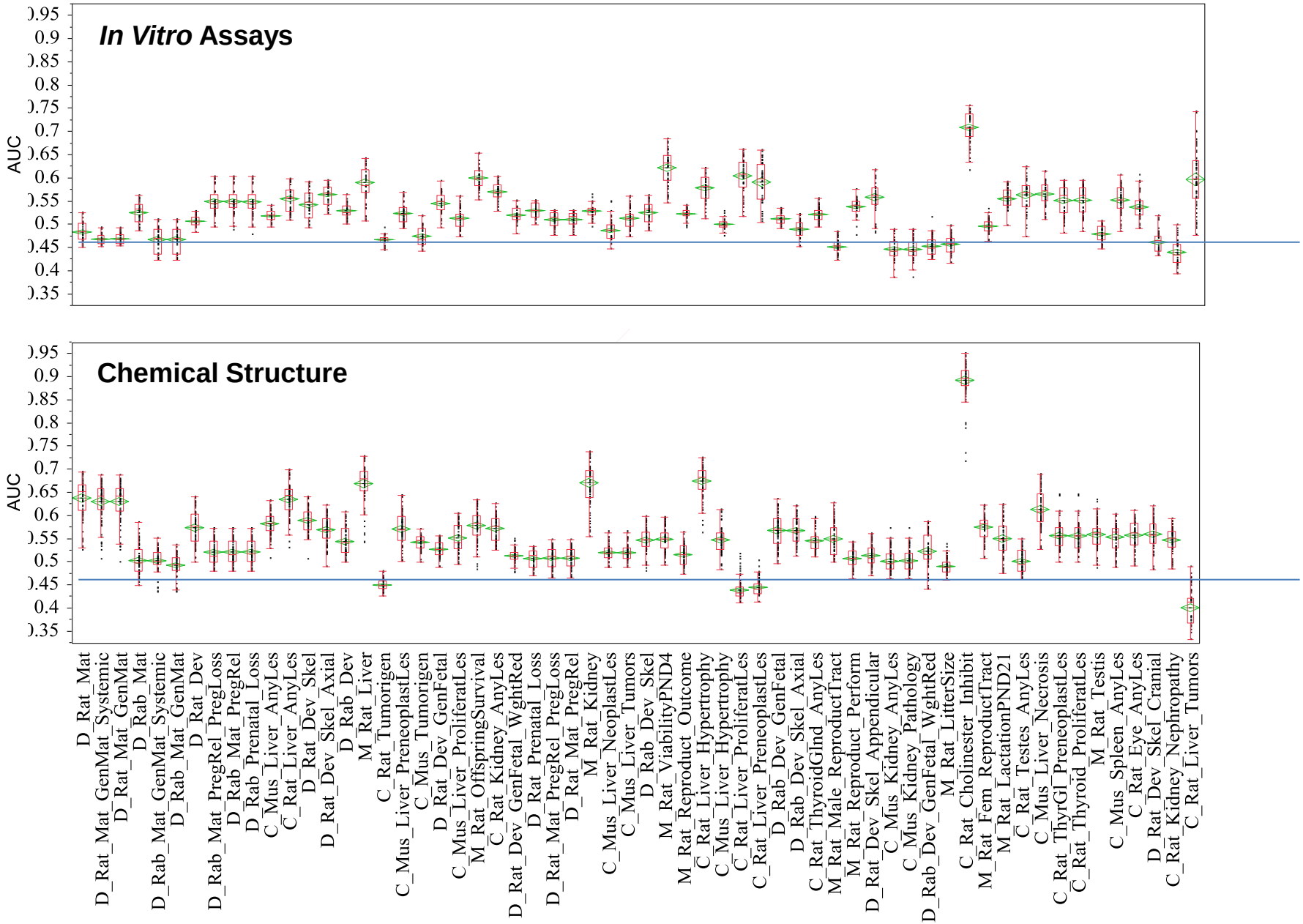
D_Rat_Mat
D_Rat_Mat_GenMat_Systemic
D_Rat_Mat_GenMat
D_Rab_Mat
D_Rab_Mat_GenMat_Systemic
D_Rab_Mat_GenMat
D_Rat_Dev
D_Rab_Mat_PregRel_PregLoss
D_Rab_Mat_PregRel
D_Rab_Prenatal_Loss
C_Mus_Liver_AnyLes
C_Rat_Liver_AnyLes
D_Rat_Dev_Skel
D_Rat_Dev_Skel_Axial
D_Rab_Dev
M_Rat_Liver
C_Rat_Tumorigen
C_Mus_Liver_PreneoplastLes
C_Mus_Tumorigen
D_Rat_Dev_GenFetal
C_Mus_Liver_ProliferatLes
M_Rat_OffspringSurvival
C_Rat_Kidney_AnyLes
D_Rat_Dev_GenFetal_WghtRed
D_Rat_Prenatal_Loss
D_Rat_Mat_PregRel_PregLoss
D_Rat_Mat_PregRel
M_Rat_Kidney
C_Mus_Liver_NeoplastLes
C_Mus_Liver_Tumors
D_Rab_Dev_Skel
M_Rat_ViabilityPND4
M_Rat_Reproduct_Outcome
C_Rat_Liver_Hypertrophy
C_Mus_Liver_Hypertrophy
C_Rat_Liver_ProliferatLes
C_Rat_Liver_PreneoplastLes
D_Rab_Dev_GenFetal
D_Rab_Dev_Skel_Axial
C_Rat_ThyroidGlnd_AnyLes
M_Rat_Male_ReproductTract
M_Rat_Reproduct_Perform
D_Rat_Dev_Skel_Appendicular
C_Mus_Kidney_AnyLes
C_Mus_Kidney_Pathology
D_Rab_Dev_GenFetal_WghtRed
M_Rat_LitterSize
C_Rat_Cholinester_Inhibit
M_Rat_Fem_ReproductTract
M_Rat_LactationPND21
C_Rat_Testes_AnyLes
C_Mus_Liver_Necrosis
C_Rat_ThyrGl_PreneoplastLes
C_Rat_Thyroid_ProliferatLes
M_Rat_Testis
C_Mus_Spleen_AnyLes
C_Rat_Eye_AnyLes
D_Rat_Dev_Skel_Cranial
C_Rat_Kidney_Nephropathy
C_Rat_Liver_Tumors

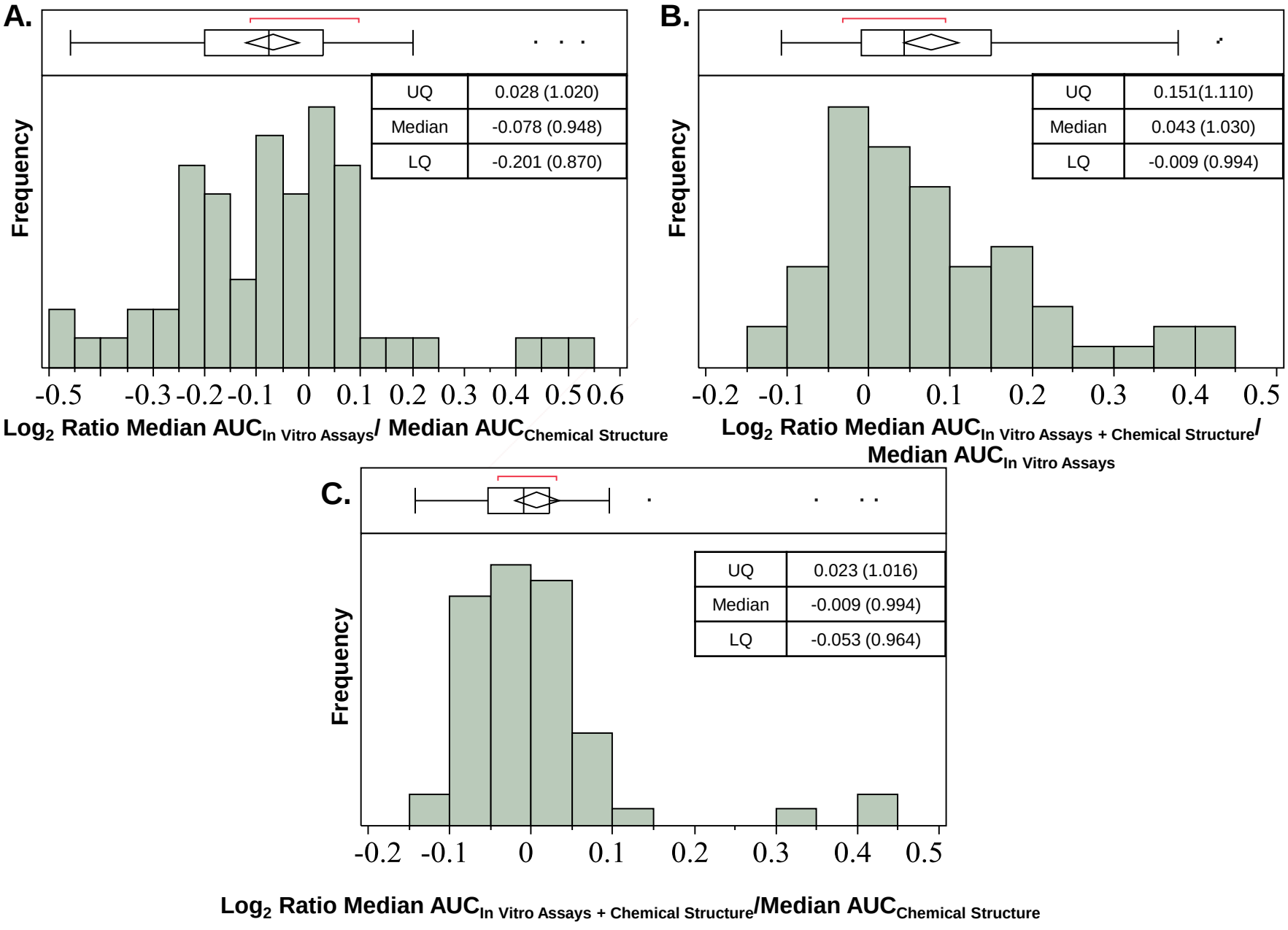


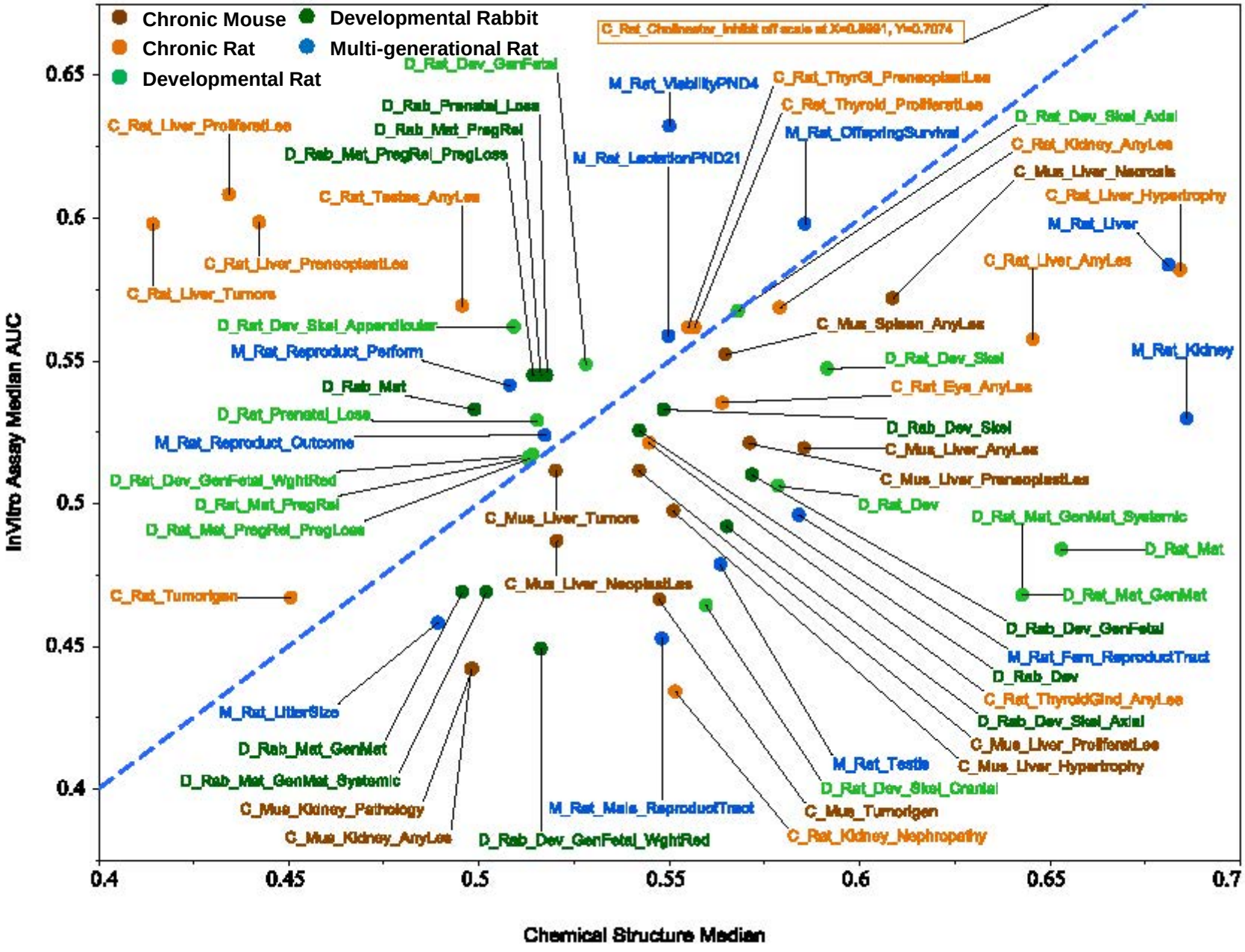


D_Rat_Mat
D_Rat_Mat_GenMat_Systemic
D_Rat_Mat_GenMat
D_Rab_Mat
D_Rab_Mat_GenMat_Systemic
D_Rab_Mat_GenMat
D_Rat_Dev
D_Rab_Mat_PregRel_PregLoss
D_Rab_Mat_PregRel
D_Rab_Prenatal_Loss
C_Mus_Liver_AnyLes
C_Rat_Liver_AnyLes
D_Rat_Dev_Skel
D_Rat_Dev_Skel_Axial
D_Rab_Dev
M_Rat_Liver
C_Rat_Tumorigen
C_Mus_Liver_PreneoplastLes
C_Mus_Tumorigen
D_Rat_Dev_GenFetal
C_Mus_Liver_ProliferatLes
M_Rat_OffspringSurvival
C_Rat_Kidney_AnyLes
D_Rat_Dev_GenFetal_WghtRed
D_Rat_Prenatal_Loss
D_Rat_Mat_PregRel_PregLoss
D_Rat_Mat_PregRel
M_Rat_Kidney
C_Mus_Liver_NeoplastLes
C_Mus_Liver_Tumors
D_Rab_Dev_Skel
M_Rat_ViabilityPND4
M_Rat_Reproduct_Outcome
C_Rat_Liver_Hypertrophy
C_Mus_Liver_Hypertrophy
C_Rat_Liver_ProliferatLes
C_Rat_Liver_PreneoplastLes
D_Rab_Dev_GenFetal
D_Rab_Dev_Skel_Axial
C_Rat_ThyroidGlnd_AnyLes
M_Rat_Male_ReproductTract
M_Rat_Reproduct_Perform
D_Rat_Dev_Skel_Appendicular
C_Mus_Kidney_AnyLes
C_Mus_Kidney_Pathology
D_Rab_Dev_GenFetal_WghtRed
M_Rat_LitterSize
C_Rat_Cholinester_Inhibit
M_Rat_Fem_ReproductTract
M_Rat_LactationPND21
C_Rat_Testes_AnyLes
C_Mus_Liver_Necrosis
C_Rat_ThyrGl_PreneoplastLes
C_Rat_Thyroid_ProliferatLes
M_Rat_Testis
C_Mus_Spleen_AnyLes
C_Rat_Eye_AnyLes
D_Rat_Dev_Skel_Cranial
C_Rat_Kidney_Nephropathy
C_Rat_Liver_Tumors


NPV
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
In Vitro Assays
Chemical Structure
D_Rat_Mat
D_Rat_Mat_GenMat_Systemic
D_Rat_Mat_GenMat
D_Rab_Mat
D_Rab_Mat_GenMat_Systemic
D_Rab_Mat_GenMat
D_Rat_Dev
D_Rab_Mat_PregRel_PregLoss
D_Rab_Mat_PregRel
D_Rab_Prenatal_Loss
C_Mus_Liver_AnyLes
C_Rat_Liver_AnyLes
D_Rat_Dev_Skel
D_Rat_Dev_Skel_Axial
D_Rab_Dev
M_Rat_Liver
C_Rat_Tumorigen
C_Mus_Liver_PreneoplastLes
C_Mus_Tumorigen
D_Rat_Dev_GenFetal
C_Mus_Liver_ProliferatLes
M_Rat_OffspringSurvival
C_Rat_Kidney_AnyLes
D_Rat_Dev_GenFetal_WghtRed
D_Rat_Prenatal_Loss
D_Rat_Mat_PregRel_PregLoss
D_Rat_Mat_PregRel
M_Rat_Kidney
C_Mus_Liver_NeoplastLes
C_Mus_Liver_Tumors
D_Rab_Dev_Skel
M_Rat_ViabilityPND4
M_Rat_Reproduct_Outcome
C_Rat_Liver_Hypertrophy
C_Mus_Liver_Hypertrophy
C_Rat_Liver_ProliferatLes
C_Rat_Liver_PreneoplastLes
D_Rab_Dev_GenFetal
D_Rab_Dev_Skel_Axial
C_Rat_ThyroidGlnd_AnyLes
M_Rat_Male_ReproductTract
M_Rat_Reproduct_Perform
D_Rat_Dev_Skel_Appendicular
C_Mus_Kidney_AnyLes
C_Mus_Kidney_Pathology
D_Rab_Dev_GenFetal_WghtRed
M_Rat_LitterSize
C_Rat_Cholinester_Inhibit
M_Rat_Fem_ReproductTract
M_Rat_LactationPND21
C_Rat_Testes_AnyLes
C_Mus_Liver_Necrosis
C_Rat_ThyrGl_PreneoplastLes
C_Rat_Thyroid_ProliferatLes
M_Rat_Testis
C_Mus_Spleen_AnyLes
C_Rat_Eye_AnyLes
D_Rat_Dev_Skel_Cranial
C_Rat_Kidney_Nephropathy
C_Rat_Liver_Tumors


In Vitro Assays
Chemical Structure
AUC
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
D_Rat_Mat
D_Rat_Mat_GenMat_Systemic
D_Rat_Mat_GenMat
D_Rab_Mat
D_Rab_Mat_GenMat_Systemic
D_Rab_Mat_GenMat
D_Rat_Dev
D_Rab_Mat_PregRel_PregLoss
D_Rab_Mat_PregRel
D_Rab_Prenatal_Loss
C_Mus_Liver_AnyLes
C_Rat_Liver_AnyLes
D_Rat_Dev_Skel
D_Rat_Dev_Skel_Axial
D_Rab_Dev
M_Rat_Liver
C_Rat_Tumorigen
C_Mus_Liver_PreneoplastLes
C_Mus_Tumorigen
D_Rat_Dev_GenFetal
C_Mus_Liver_ProliferatLes
M_Rat_OffspringSurvival
C_Rat_Kidney_AnyLes
D_Rat_Dev_GenFetal_WghtRed
D_Rat_Prenatal_Loss
D_Rat_Mat_PregRel_PregLoss
D_Rat_Mat_PregRel
M_Rat_Kidney
C_Mus_Liver_NeoplastLes
C_Mus_Liver_Tumors
D_Rab_Dev_Skel
M_Rat_ViabilityPND4
M_Rat_Reproduct_Outcome
C_Rat_Liver_Hypertrophy
C_Mus_Liver_Hypertrophy
C_Rat_Liver_ProliferatLes
C_Rat_Liver_PreneoplastLes
D_Rab_Dev_GenFetal
D_Rab_Dev_Skel_Axial
C_Rat_ThyroidGlnd_AnyLes
M_Rat_Male_ReproductTract
M_Rat_Reproduct_Perform
D_Rat_Dev_Skel_Appendicular
C_Mus_Kidney_AnyLes
C_Mus_Kidney_Pathology
D_Rab_Dev_GenFetal_WghtRed
M_Rat_LitterSize
C_Rat_Cholinester_Inhibit
M_Rat_Fem_ReproductTract
M_Rat_LactationPND21
C_Rat_Testes_AnyLes
C_Mus_Liver_Necrosis
C_Rat_ThyrGl_PreneoplastLes
C_Rat_Thyroid_ProliferatLes
M_Rat_Testis
C_Mus_Spleen_AnyLes
C_Rat_Eye_AnyLes
D_Rat_Dev_Skel_Cranial
C_Rat_Kidney_Nephropathy
C_Rat_Liver_Tumors


A.
Frequency
UQ 0.028 (1.020)
Median -0.078 (0.948)
LQ -0.201 (0.870)
-0.5 -0.3 -0.2 -0.1 0 0.1 0.2 0.3 0.4 0.5 0.6
Log2 Ratio Median AUCIn Vitro Assays/ Median AUCChemical Structure
B.
Frequency
UQ 0.151(1.110)
Median 0.043 (1.030)
LQ -0.009 (0.994)
-0.2 -0.1 0 0.1 0.2 0.3 0.4 0.5
Log2 Ratio Median AUCIn Vitro Assays + Chemical Structure/ Median AUCIn Vitro Assays
C.
Frequency
UQ 0.023 (1.016)
Median -0.009 (0.994)
LQ -0.053 (0.964)
-0.2 -0.1 0 0.1 0.2 0.3 0.4 0.5
Log2 Ratio Median AUCIn Vitro Assays + Chemical Structure/Median AUCChemical Structure


Chronic Mouse
Chronic Rat
Developmental Rat
Developmental Rabbit
Multi-generational Rat
C_Rat_Cholinester_Inhibit cf/ acute at X=0.8881, Y=0.7074
D_Rat_Dev_GenFetal
M_Rat_ViabilityPND4
C_Rat_ThyrGl_PreneoplastLes
C_Rat_Thyroid_ProliferatLes
D_Rat_Dev_Skel_Axial
C_Rat_Liver_ProliferatLes
D_Rab_Prenatal_Loss
D_Rab_Mat_PregRel
M_Rat_OffspringSurvival
C_Rat_Kidney_AnyLes
D_Rab_Mat_PregRel_PregLoss
M_Rat_LactationPND21
C_Mus_Liver_Necrosis
C_Rat_Testes_AnyLes
C_Rat_Liver_Hypertrophy
M_Rat_Liver
C_Rat_Liver_PreneoplastLes
C_Rat_Liver_AnyLes
C_Rat_Liver_Tumors
D_Rat_Dev_Skel_Appendicular
C_Mus_Spleen_AnyLes
M_Rat_Reproduct_Perform
D_Rat_Dev_Skel
M_Rat_Kidney
D_Rab_Mat
C_Rat_Eye_AnyLes
D_Rat_Prenatal_Loss
D_Rab_Dev_Skel
M_Rat_Reproduct_Outcome
C_Mus_Liver_AnyLes
D_Rat_Dev_GenFetal_WghtRed
C_Mus_Liver_PreneoplastLes
D_Rat_Mat_PregRel
D_Rat_Dev
D_Rat_Mat_PregRel_PregLoss
C_Mus_Liver_Tumors
D_Rat_Mat_GenMat_Systemic
D_Rat_Mat
C_Rat_Tumorigen
C_Mus_Liver_NeoplastLes
D_Rat_Mat_GenMat
D_Rab_Dev_GenFetal
M_Rat_Fem_ReproductTract
D_Rab_Dev
C_Rat_ThyroidGlnd_AnyLes
M_Rat_LitterSize
D_Rab_Dev_Skel_Axial
C_Mus_Liver_ProliferatLes
D_Rab_Mat_GenMat
M_Rat_Testis
C_Mus_Liver_Hypertrophy
D_Rab_Mat_GenMat_Systemic
D_Rat_Dev_Skel_Cranial
C_Mus_Kidney_Pathology
M_Rat_Male_ReproductTract
C_Mus_Tumorigen
C_Mus_Kidney_AnyLes
D_Rab_Dev_GenFetal_WghtRed
C_Rat_Kidney_Nephropathy
0.65
0.6
0.55
0.5
0.45
0.4
0.4
0.45
0.5
0.55
0.6
0.65
0.7
InVitro Assay Median AUC
Chemical Structure Median

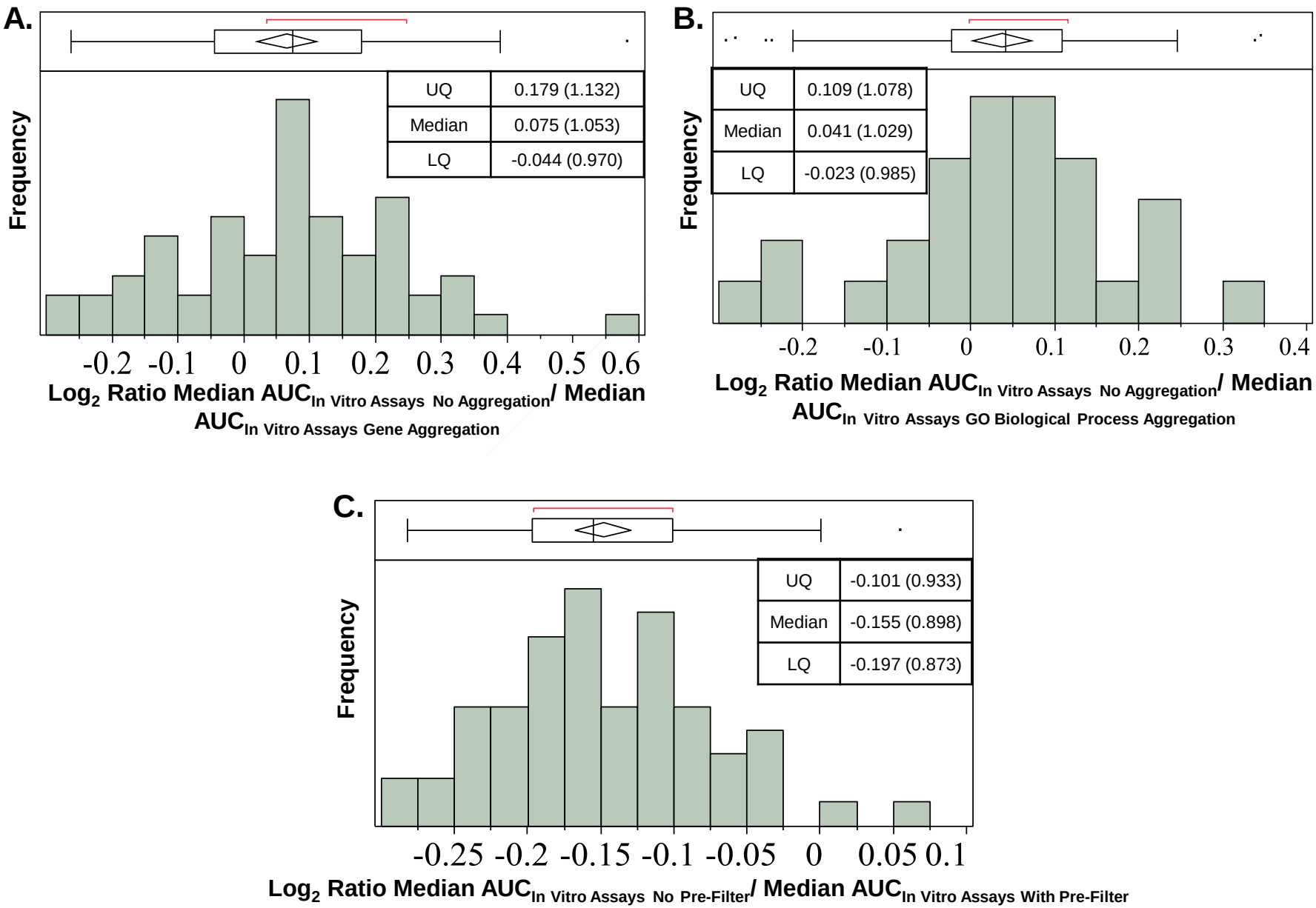
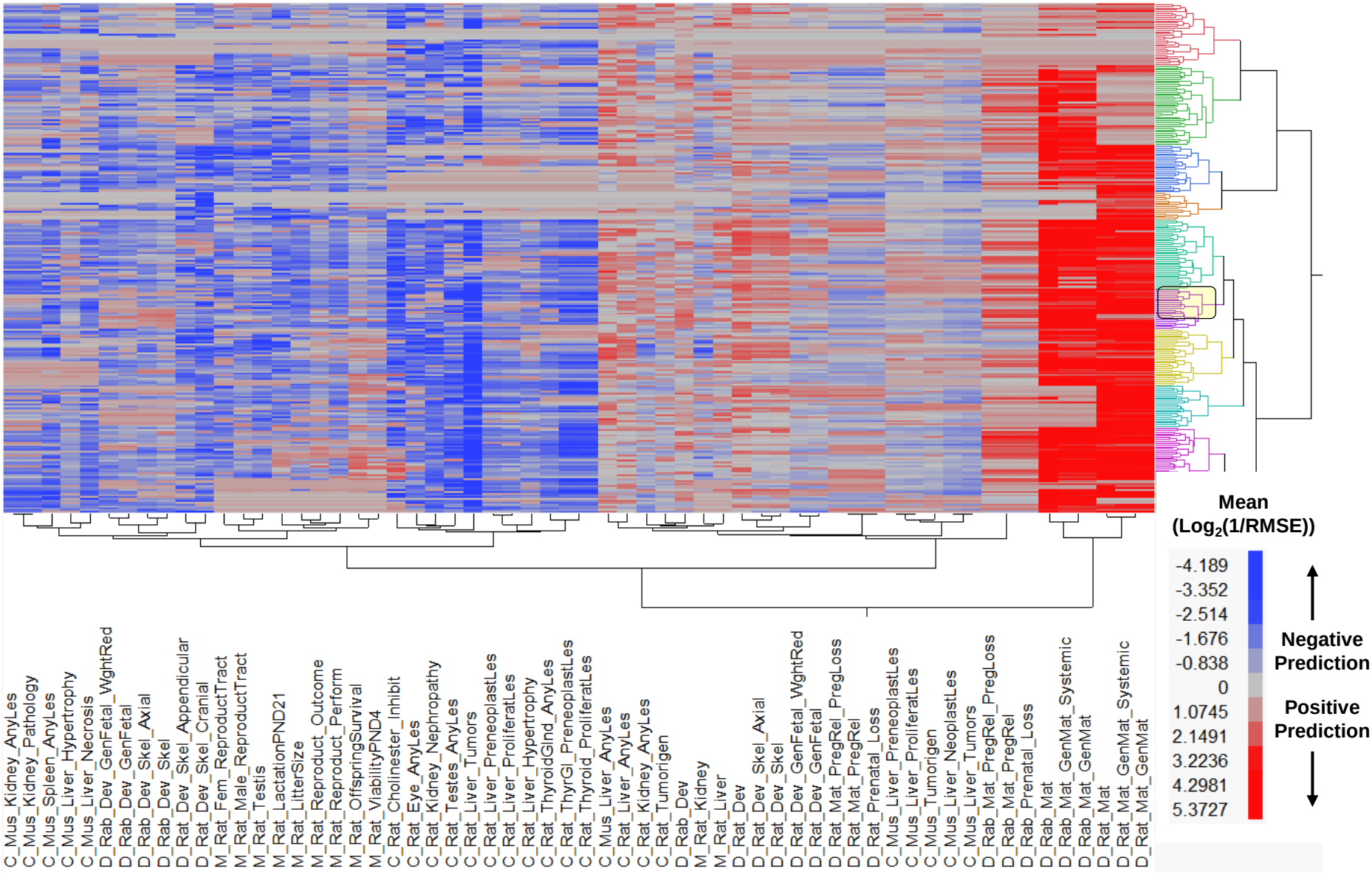
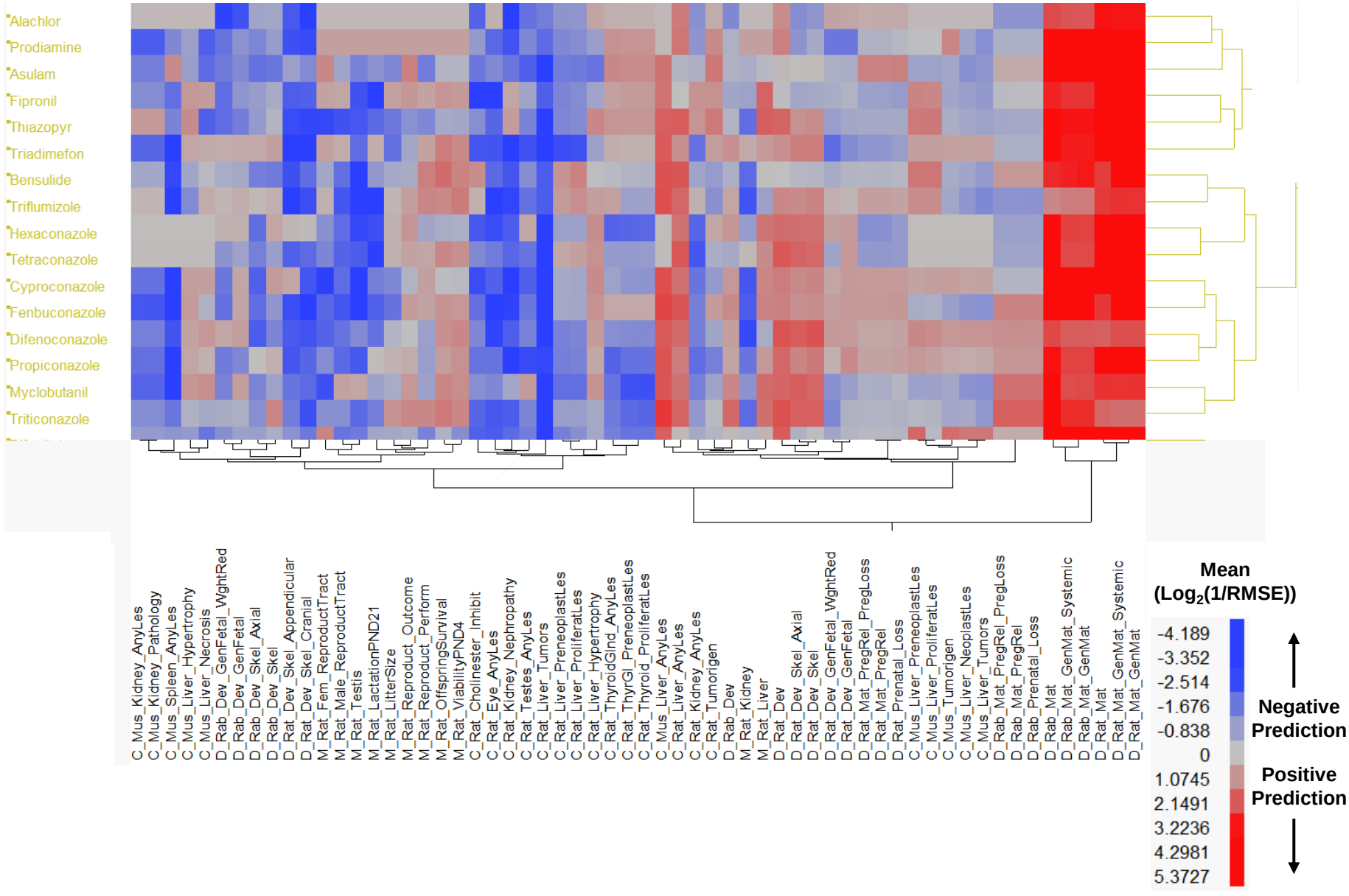

A.
Frequency
UQ 0.179 (1.132)
Median 0.075 (1.053)
LQ -0.044 (0.970)
-0.2 -0.1 0 0.1 0.2 0.3 0.4 0.5 0.6
Log2 Ratio Median AUCIn Vitro Assays No Aggregation/ Median AUCIn Vitro Assays Gene Aggregation
B.
Frequency
UQ 0.109 (1.078)
Median 0.041 (1.029)
LQ -0.023 (0.985)
-0.2 -0.1 0 0.1 0.2 0.3 0.4
Log2 Ratio Median AUCIn Vitro Assays No Aggregation/ Median AUCIn Vitro Assays GO Biological Process Aggregation
C.
Frequency
UQ -0.101 (0.933)
Median -0.155 (0.898)
LQ -0.197 (0.873)
-0.25 -0.2 -0.15 -0.1 -0.05 0 0.05 0.1
Log2 Ratio Median AUCIn Vitro Assays No Pre-Filter/ Median AUCIn Vitro Assays With Pre-Filter

Fig. 12

Predictability

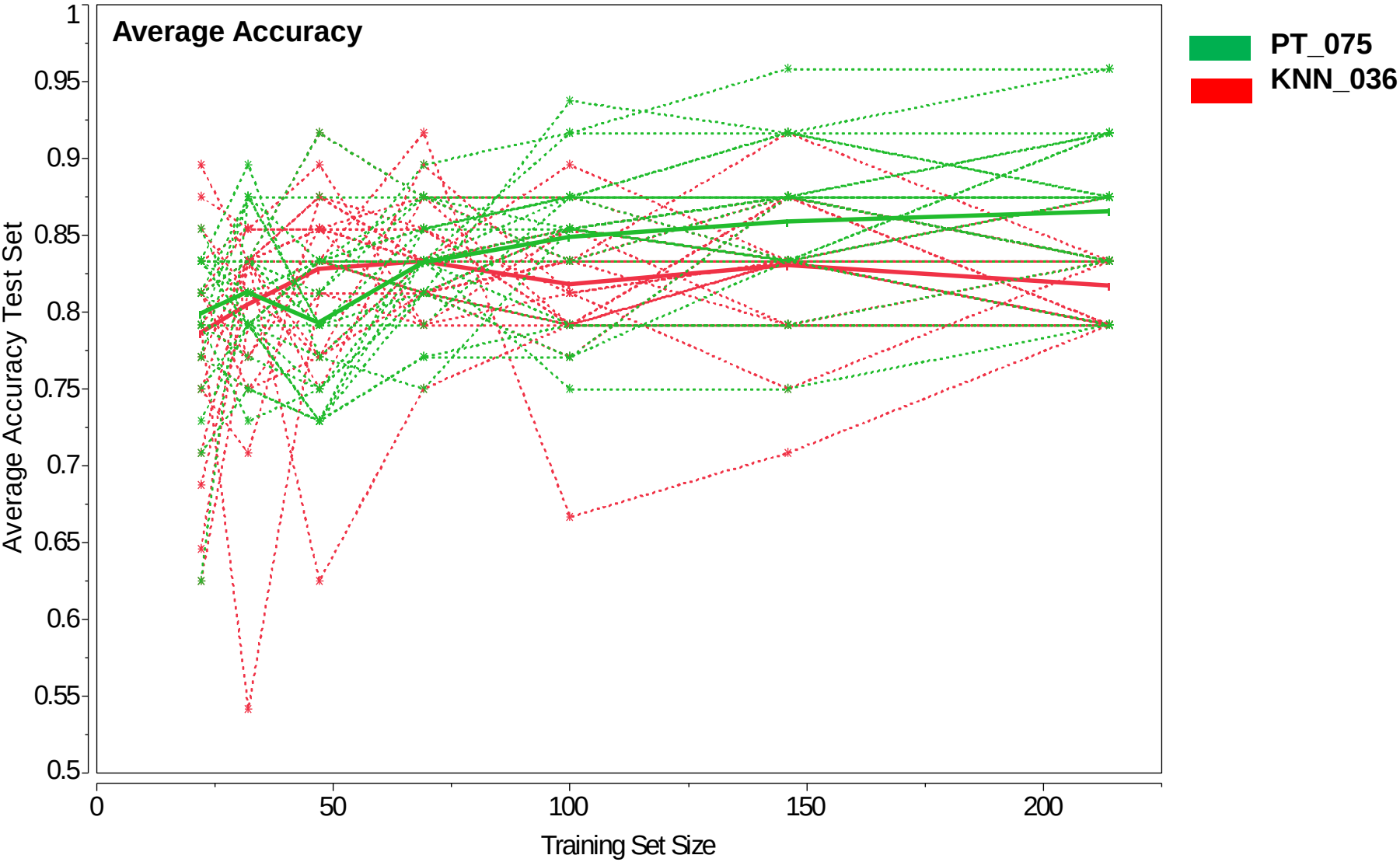



Alachlor
Prodiamine
Asulam
Fipronil
Thiazopyr
Triadimefon
Bensulide
Triflumizole
Hexaconazole
Tetraconazole
Cyproconazole
Fenbuconazole
Difenoconazole
Propiconazole
Myclobutanil
Triticonazole
C_Mus_Kidney_AnyLes
C_Mus_Kidney_Pathology
C_Mus_Spleen_AnyLes
C_Mus_Liver_Hypertrophy
C_Mus_Liver_Necrosis
D_Rab_Dev_GenFetal_WghtRed
D_Rab_Dev_GenFetal
D_Rab_Dev_Skel_Axial
D_Rab_Dev_Skel
D_Rat_Dev_Skel_Appendicular
D_Rat_Dev_Skel_Cranial
M_Rat_Fem_ReproductTract
M_Rat_Male_ReproductTract
M_Rat_Testis
M_Rat_LactationPND21
M_Rat_LitterSize
M_Rat_Reproduct_Outcome
M_Rat_Reproduct_Perform
M_Rat_OffspringSurvival
M_Rat_ViabilityPND4
C_Rat_Cholinester_Inhibit
C_Rat_Eye_AnyLes
C_Rat_Kidney_Nephropathy
C_Rat_Testes_AnyLes
C_Rat_Liver_Tumors
C_Rat_Liver_PreneoplastLes
C_Rat_Liver_ProliferatLes
C_Rat_Liver_Hypertrophy
C_Rat_ThyroidGlnd_AnyLes
C_Rat_ThyrGl_PreneoplastLes
C_Rat_Thyroid_ProliferatLes
C_Mus_Liver_AnyLes
C_Rat_Liver_AnyLes
C_Rat_Kidney_AnyLes
C_Rat_Tumorigen
D_Rab_Dev
M_Rat_Kidney
M_Rat_Liver
D_Rat_Dev
D_Rat_Dev_Skel_Axial
D_Rat_Dev_Skel
D_Rat_Dev_GenFetal_WghtRed
D_Rat_Dev_GenFetal
D_Rat_Mat_PregRel_PregLoss
D_Rat_Mat_PregRel
D_Rat_Prenatal_Loss
C_Mus_Liver_PreneoplastLes
C_Mus_Liver_ProliferatLes
C_Mus_Tumorigen
C_Mus_Liver_NeoplastLes
C_Mus_Liver_Tumors
D_Rab_Mat_PregRel_PregLoss
D_Rab_Mat_PregRel
D_Rab_Prenatal_Loss
D_Rab_Mat
D_Rab_Mat_GenMat_Systemic
D_Rab_Mat_GenMat
D_Rat_Mat
D_Rat_Mat_GenMat_Systemic
D_Rat_Mat_GenMat
Mean
(Log2(1/RMSE))
-4.189
-3.352
-2.514
-1.676
-0.838
0
1.0745
2.1491
3.2236
4.2981
5.3727
Negative
Prediction
Positive
Prediction

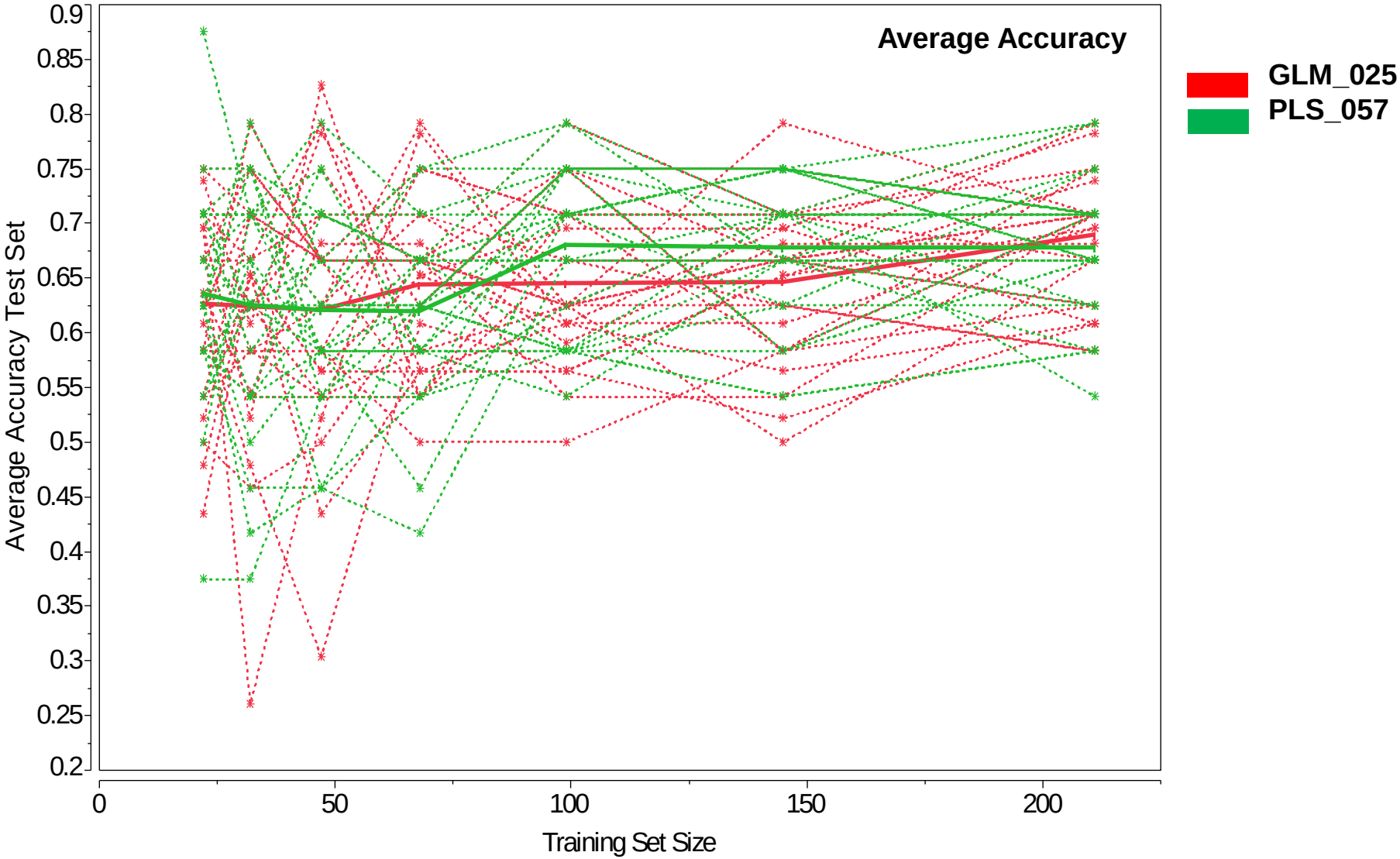
Learning Curves

- An appropriate way to assess adequacy of sample size in a predictive modeling context
- Completely different from classic power and sample size calculations, which are designed for hypothesis testing, not predictive modeling
- Constructed by taking subsets of different sizes and performing cross-validation model comparison on each, then plot results with sample size as the x-axis
- An upward trend indicates that results would likely improve with more samples; a flat curve indicates otherwise.

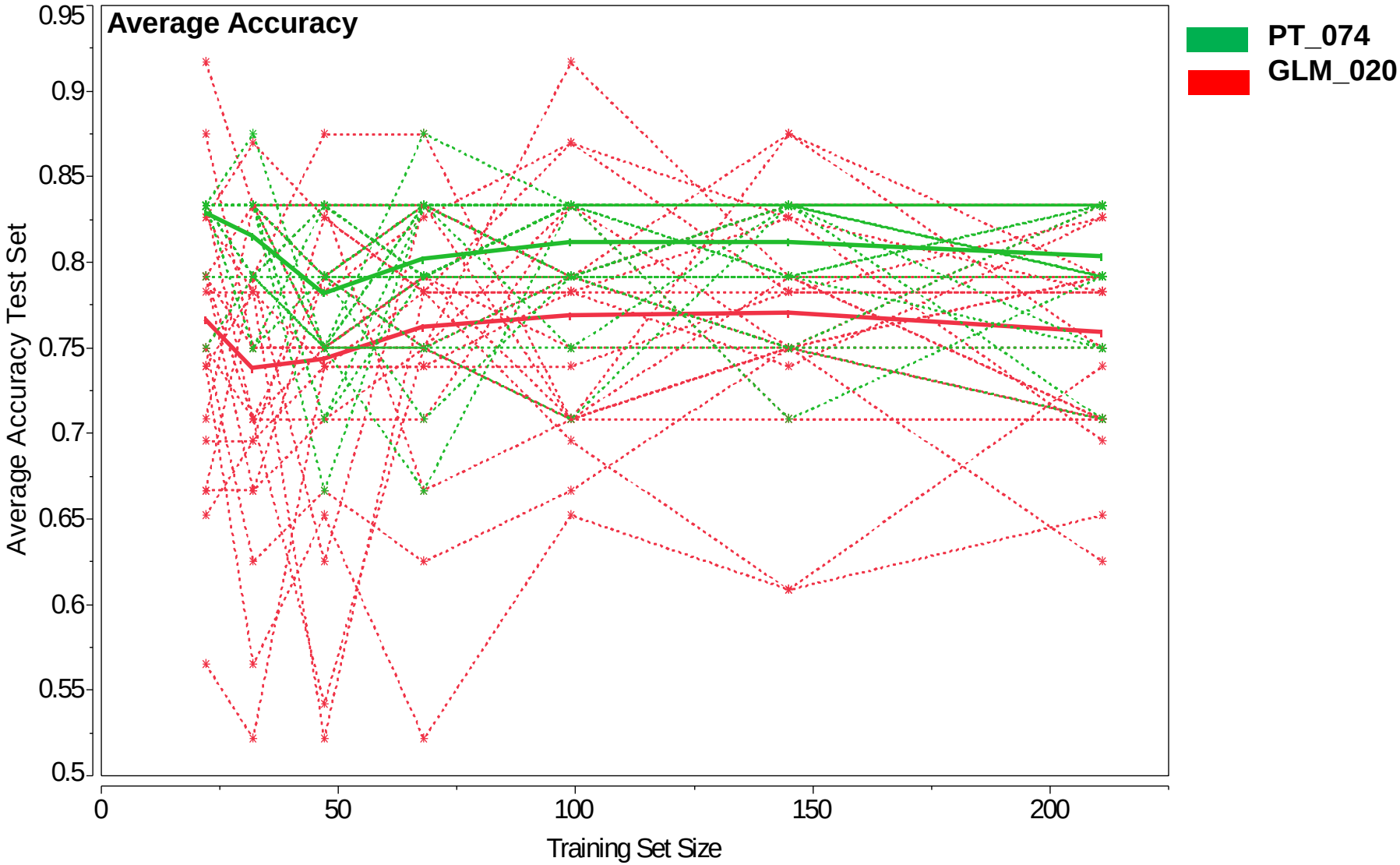
CHR_Rat_CholinesteraseInhibition
Best model from CVMC = PT_075 (AUC = 0.75583)
Worst model from CVMC = KNN_036 (AUC = 0.61631)



MGR_Rat_ReproductiveOutcome
Best model from CVMC = GLM_025 (AUC = 0.54170)
Worst model from CVMC = PLS_057 (AUC = 0.49296)



MGR_Rat_Testis
Best model from CVMC = PT_074_025 (AUC = 0.50620)
Worst model from CVMC = GLM_020 (AUC = 0.44754)



Summary

- The current ToxCast *in vitro* high-throughput screening assays provide somewhat limited ability to predict *in vivo* toxic responses.

- Sensitivity and specificity of the *in vitro* assays is related to the balance of positive and negative chemicals, but even for balanced endpoints, the overall predictive performance is relatively low.

- The *in vitro* assays provide somewhat lower predictive performance than chemical structure.

- Aggregating the assays based on gene and biological processes did not appear to improve predictive performance.

- Pre-filtering the in vitro assay data, as has been done in previous studies, can significantly bias estimates of cross-validation performance in an optimistic direction.

- Analysis of predictability and learning curves can help in prioritizing chemicals for further study and in assessing adequacy of sample sizes.