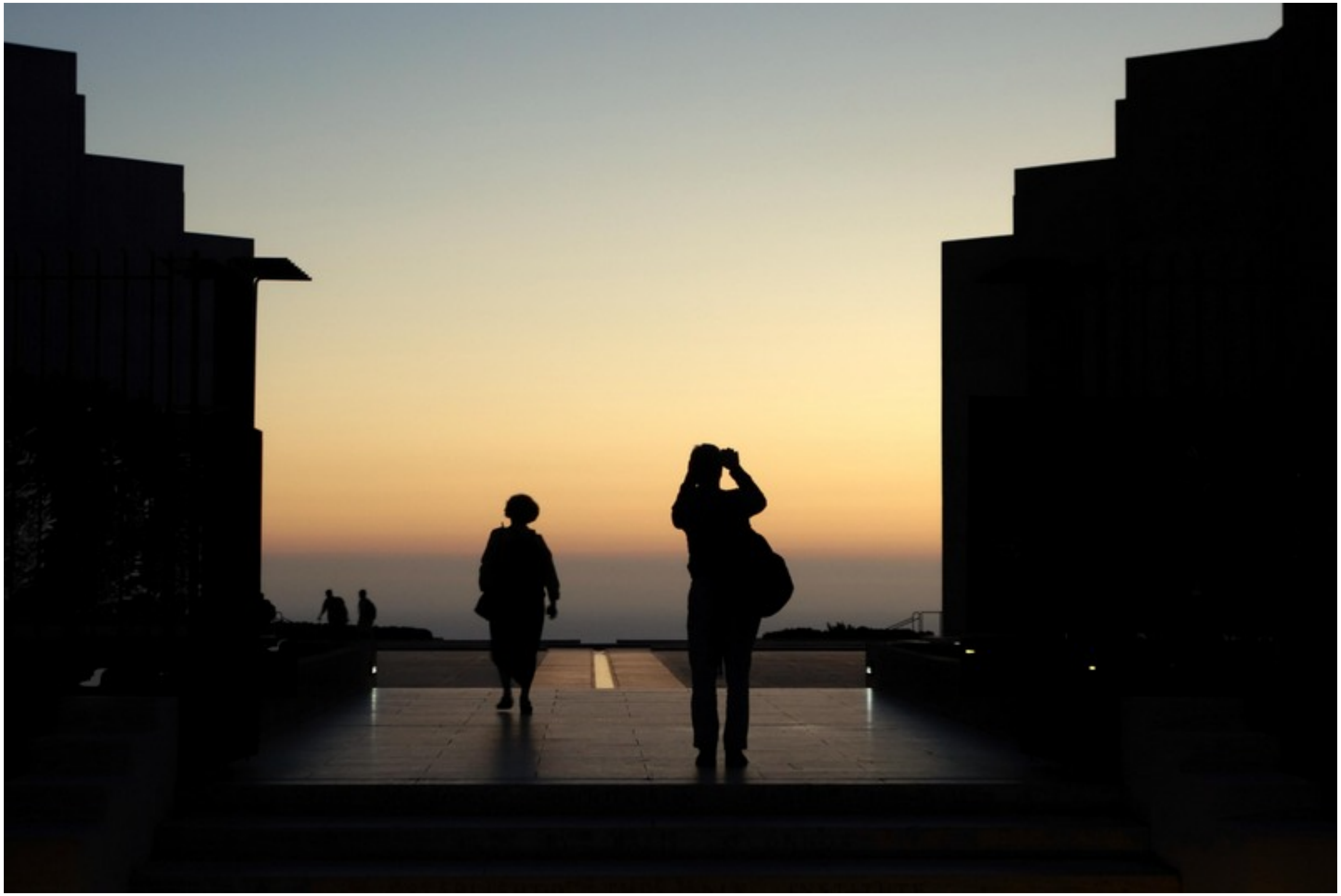
Brain Genomics:

Finding drug targets for neuropsychiatric disorders via deep-learning & Designing a predictor for the sensitivity of drugs to human population variation

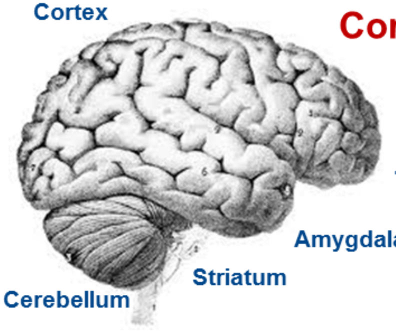


M Gerstein
Yale
(See last slide for more info.)

Slides freely downloadable
from **Lectures.GersteinLab.org**
& “tweetable” (via **@MarkGerstein**)

Sample Sources: >2,500 brains

Cross-disorder: ASD, SCZ, BP, Neurodevelopmental, Neurotypical



Anterior Cingulate Cortex

Dorsolateral Prefrontal Cortex

Orbitofrontal Cortex

Temporal Cortex

Subgenual Anterior Cingulate Cortex

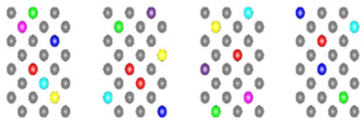
Amygdala

Striatum

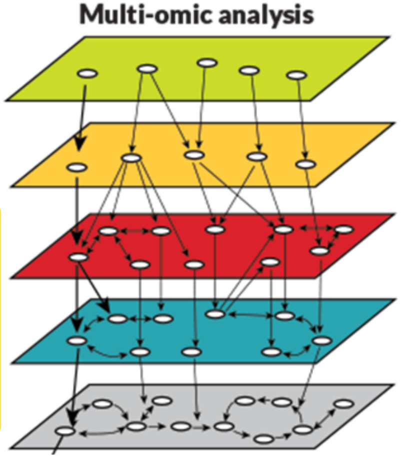
Cerebellum

Bulk Tissue

Neu –/Neu+ Sorted Cells



Single Cell

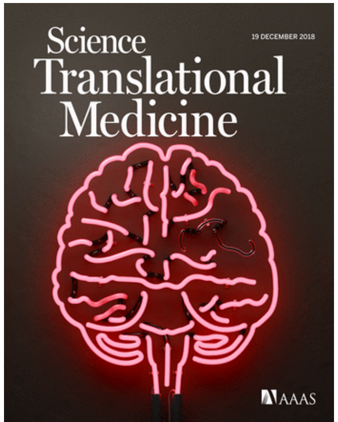


Genome:
WGS, genotype

Epigenome:
ChIP-seq, ATAC-seq, HiC, ERRBS, Array Methylation, NOMeSeq

Transcriptome:
RNA-seq, lncRNAseq,

Proteome:
MWP, LC-MS/MS



PsychENCODE

’18 rollout in Science

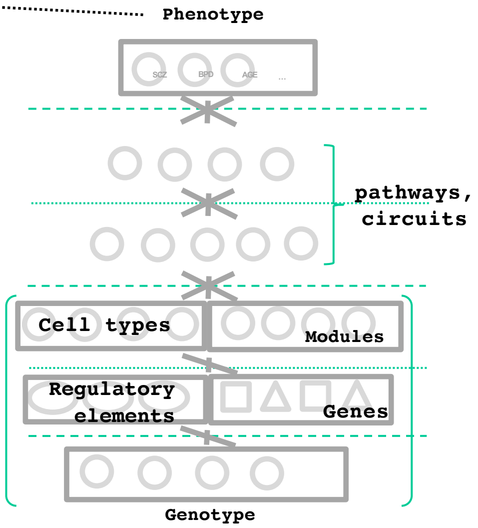
11 papers in total.
Major material in the 3 capstones:

Wang et al. (‘18), Li et al. (‘18), Gandal et al. ('18)

A core issue addressed by PsychENCODE: Using functional genomics to reveal molecular mechanisms between genotype and phenotype in brain disorders

Disease	Heritability*	Molecular **Mechanisms**
Schizophrenia	**81%**	**(C4A)**
Bipolar disorder	70%	-
Alzheimer's disease	58 - 79%	Apolipoprotein E (APOE), Tau
Hypertension	30%	Renin–angiotensin–aldosterone
Heart disease	34-53%	Atherosclerosis, VCAM-1
Stroke	32%	Reactive oxygen species (ROS), Ischemia
Type-2 diabetes	26%	Insulin resistance
Breast Cancer	25-56%	BRCA, PTEN

*https://www.snpedia.com/index.php/Heritability



Many psychiatric conditions are highly heritable

Schizophrenia: up to 80%

But we don’t understand basic molecular mechanisms underpinning this association

(in contrast to many other diseases such as cancer & heart disease)

Thus, interested in developing predictive models of psychiatric traits which:

Use observations at intermediate (molecular levels) levels to inform latent structure

Use the predictive features of these “molecular endo phenotypes” to begin to suggest actors involved in mechanism

Finding drug targets for neuropsychiatric disorders via deep-learning & Designing a predictor for the sensitivity of drugs to human population variation

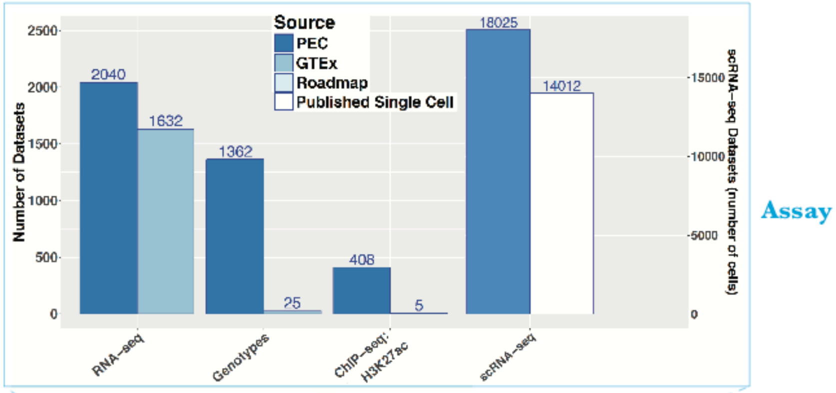
- **PsychENCODE:** Population-level analysis of functional genomics data related to neuropsychiatric disease
  - Construction of an adult brain resource with 1866 individuals
  - Large-scale processing **creates a comprehensive QTL** resource (~2.5M eQTLs).
  - Connecting QTLs, enhancer activity relationships & Hi-C contacts into a **brain regulatory network**
  - Embedding the reg. network in a **deep-learning model** to predict psychiatric disease from genotype & transcriptome. Using this to suggest specific pathways & genes, as potential drug targets.

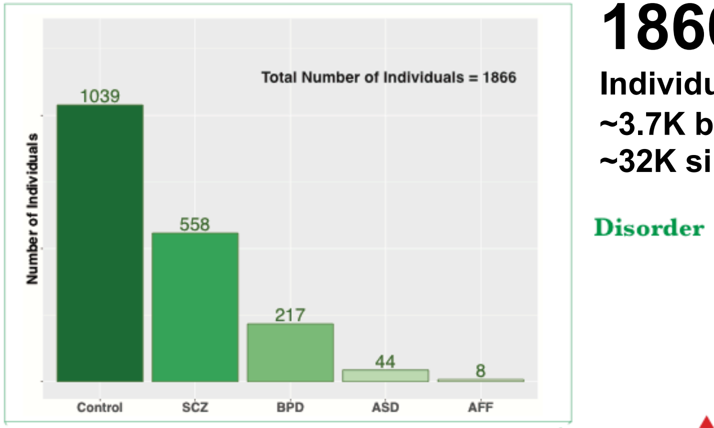
- **GenoDock:** Building a predictor for the sensitivity of drug binding to personal SNVs
  - Hybrid classifier connecting **physical modelling with statistical learning**
    - The modeling creates a pseudo gold-standard dataset, which is used to train the stat. classifier
  - **Classifier Results**
    - Independent validation on an expt. validation set
    - Gives higher disruption scores to cancer driver SNVs. Also, illustrates importance of different features (eg GERP).
    - Picks out certain drugs (eg imatinib) as being particularly sensitive to SNVs

Finding drug targets for neuropsychiatric disorders via deep-learning & Designing a predictor for the sensitivity of drugs to human population variation

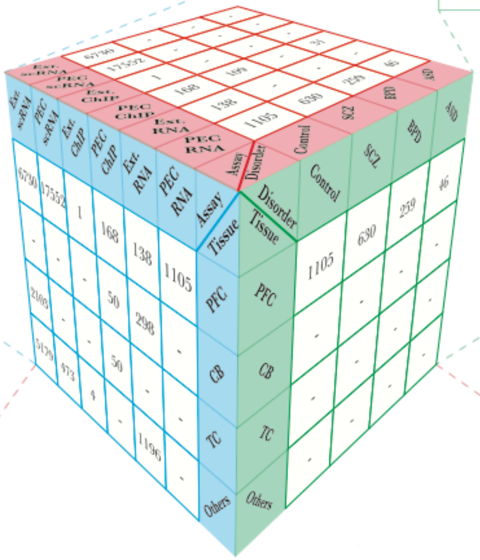
- **PsychENCODE**: Population-level analysis of functional genomics data related to neuropsychiatric disease
  - Construction of an adult brain resource with 1866 individuals
  - Large-scale processing **creates a comprehensive QTL** resource (~2.5M eQTLs).
  - Connecting QTLs, enhancer activity relationships & Hi-C contacts into a **brain regulatory network**
  - Embedding the reg. network in a **deep-learning model** to predict psychiatric disease from genotype & transcriptome. Using this to suggest specific pathways & genes, as potential drug targets.

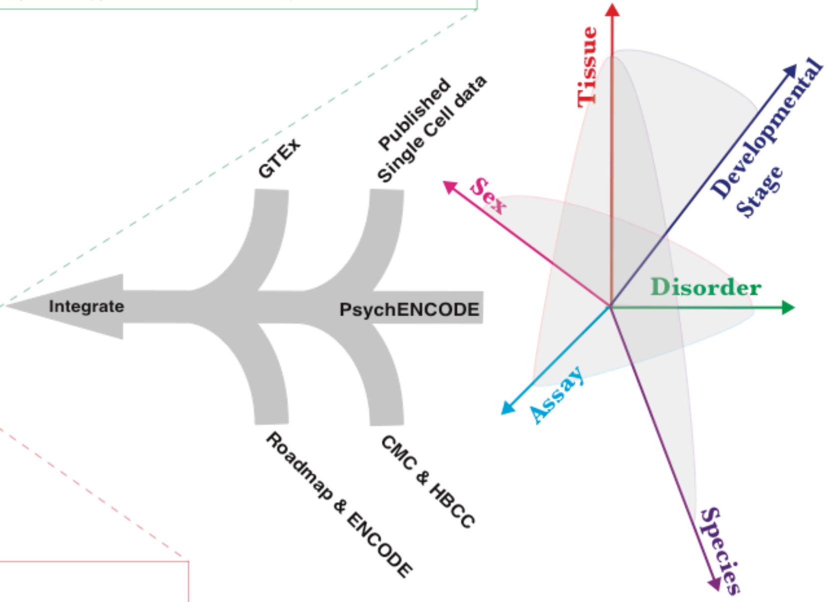
- **GenoDock**: Building a predictor for the sensitivity of drug binding to personal SNVs
  - Hybrid classifier connecting **physical modelling with statistical learning**
    - The modeling creates a pseudo gold-standard dataset, which is used to train the stat. classifier
  - **Classifier Results**
    - Independent validation on an expt. validation set
    - Gives higher disruption scores to cancer driver SNVs. Also, illustrates importance of different features (eg GERP).
    - Picks out certain drugs (eg imatinib) as being particularly sensitive to SNVs





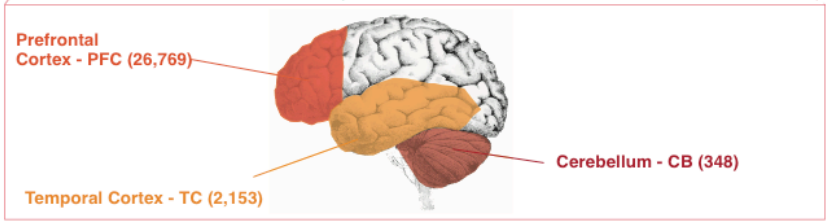
1866
Individuals
~3.7K bulk RNA-seq
~32K single-cells





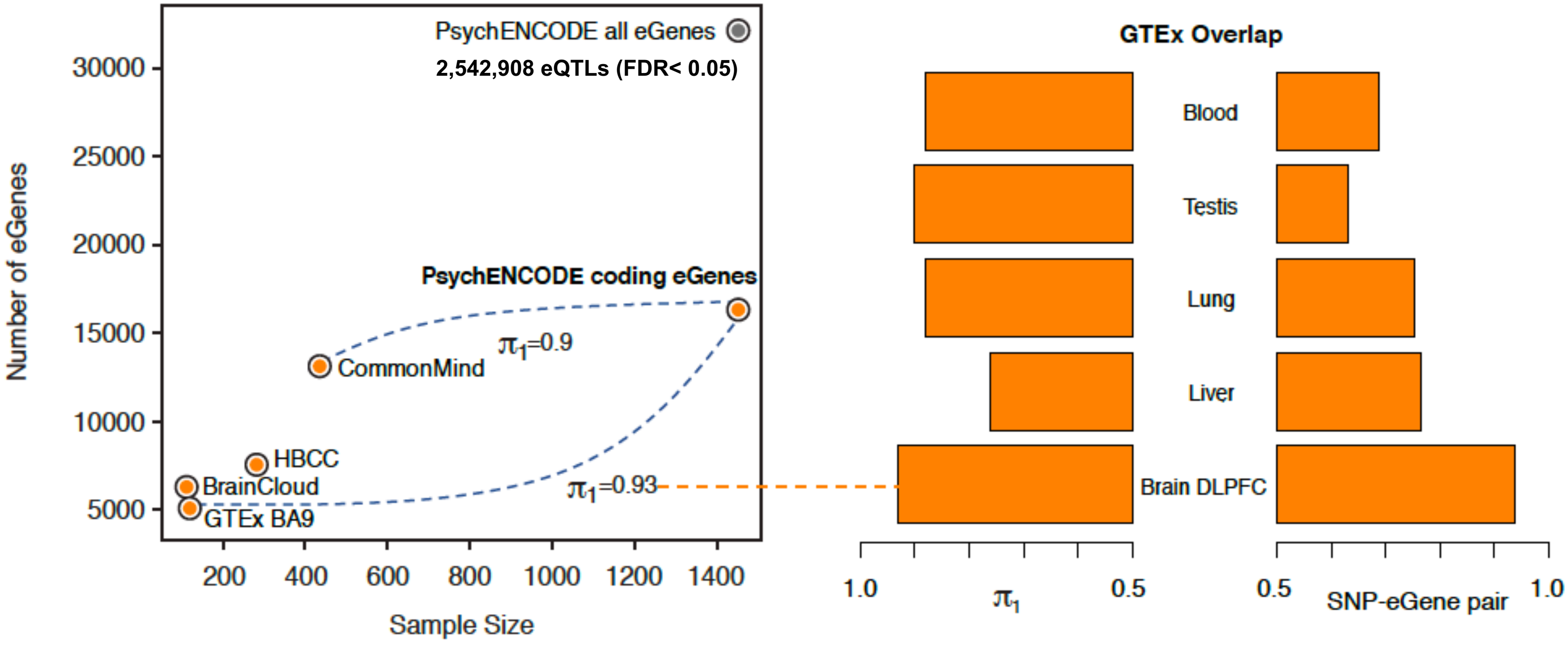
Collecting functional genomic datasets for the adult brain

from PsychENCODE, other large consortia & single cell studies



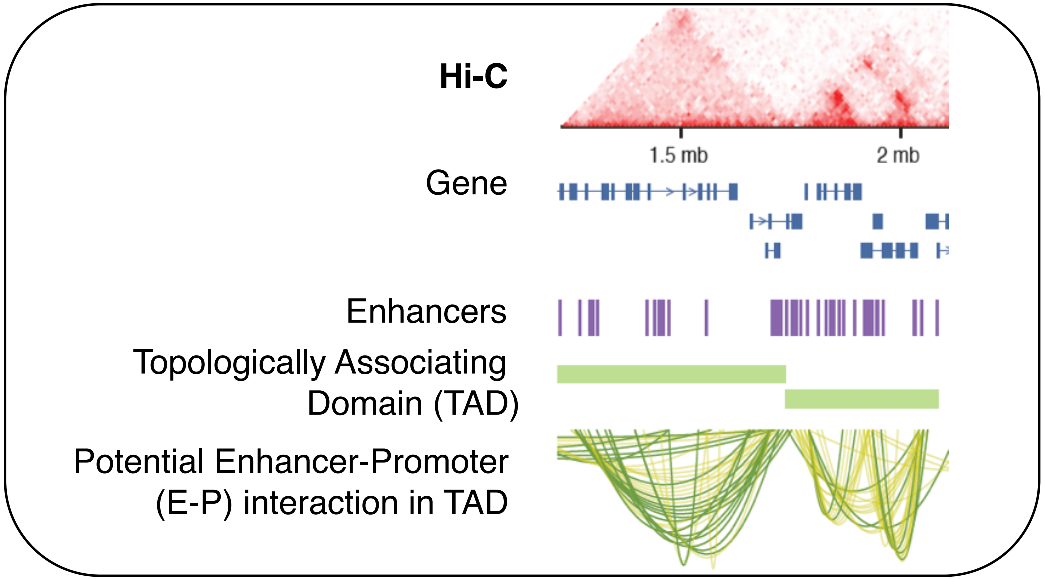
[Wang et al. ('18) Science]

Larger brain eQTL sets than previous studies, but strong overlap with them



[Wang et al. ('18) Science]

Gene regulatory network inference from Hi-C, QTLs & Activity Correlations



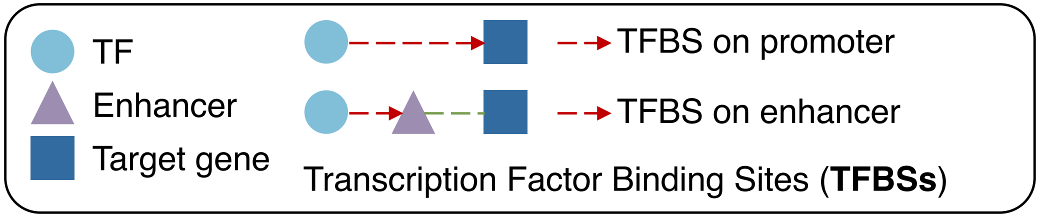
Hi-C

Gene

Enhancers

Topologically Associating Domain (TAD)

Potential Enhancer-Promoter (E-P) interaction in TAD



● TF

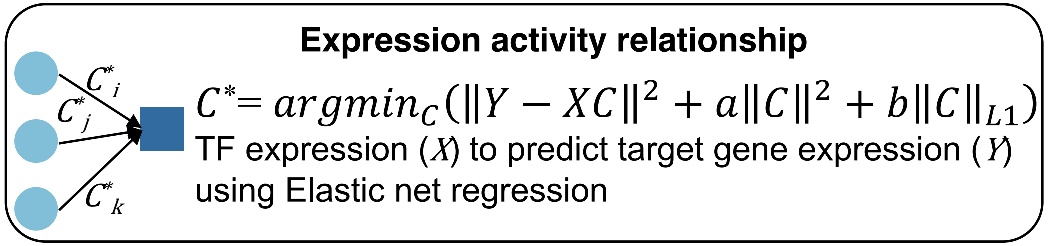
▲ Enhancer

■ Target gene

TFBS on promoter

TFBS on enhancer

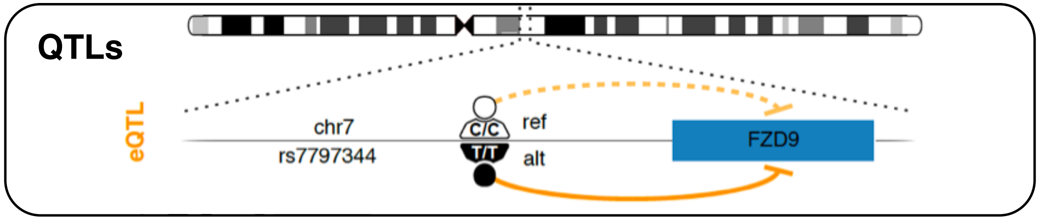
Transcription Factor Binding Sites (**TFBSs**)



Expression activity relationship

$$C^* = argmin_C(\|Y - XC\|^2 + a\|C\|^2 + b\|C\|_{L1})$$

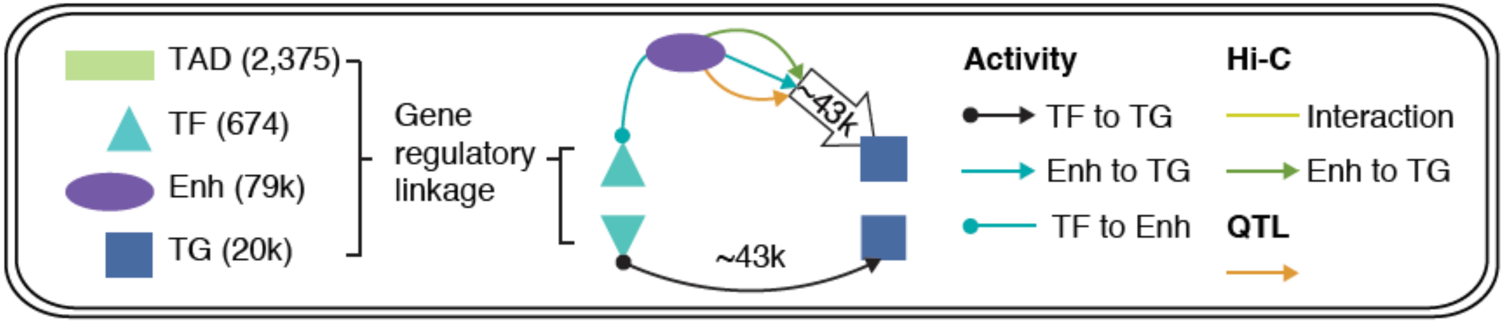
TF expression (X) to predict target gene expression (Y) using Elastic net regression

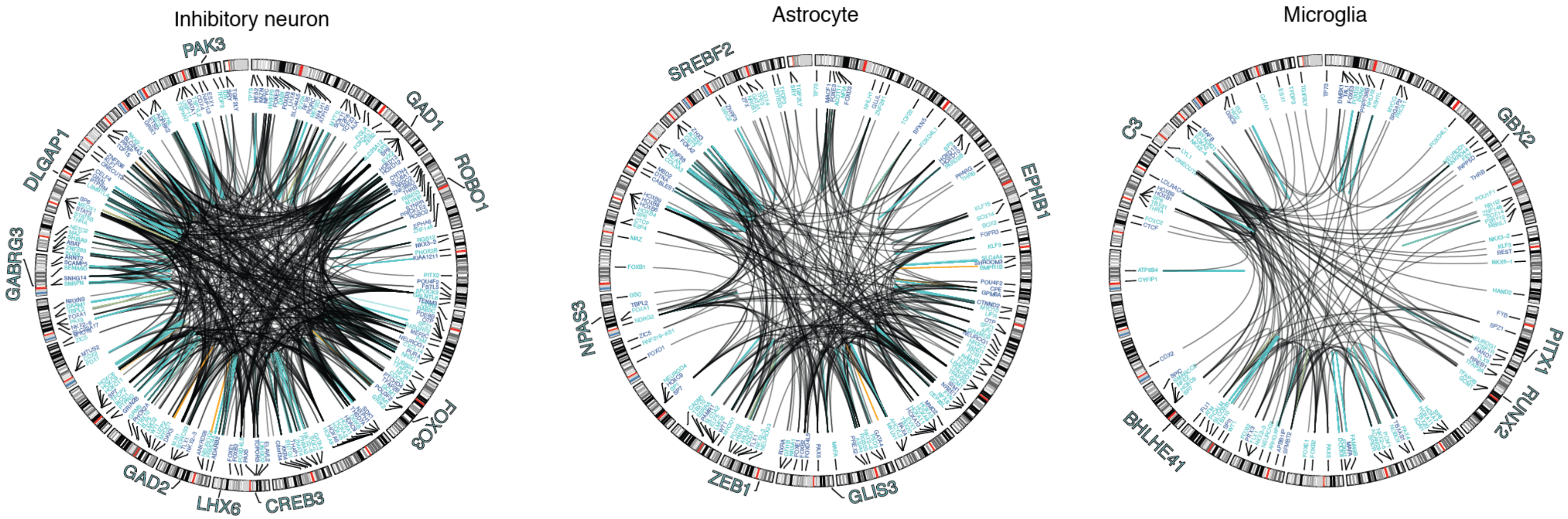


QTLs

[Wang et al. ('18) Science]

Imputed gene regulatory network for the human brain





subnetworks targeting single cell marker genes

[Wang et al. ('18) Science]

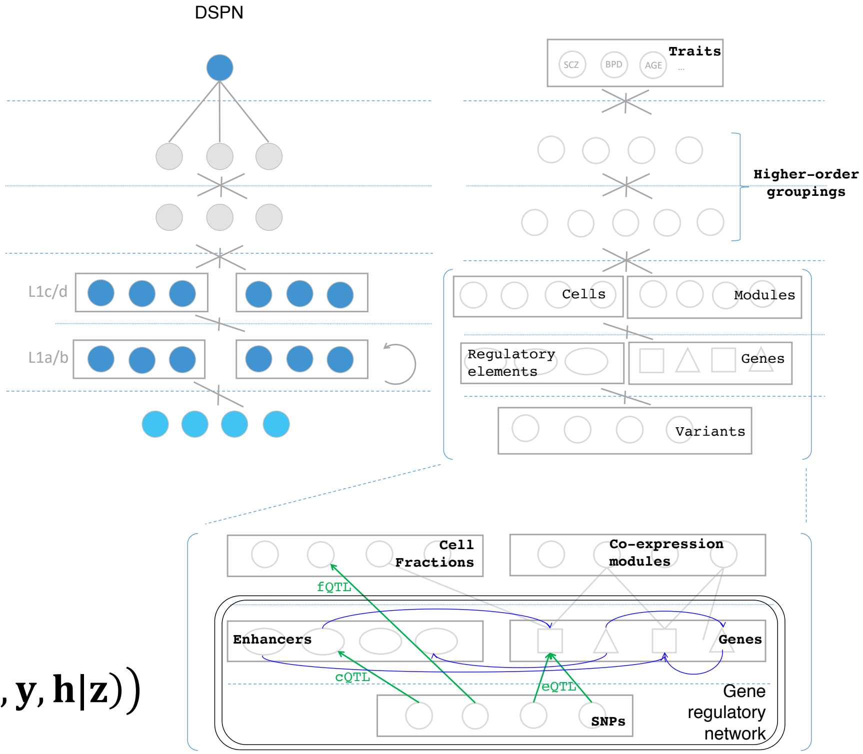
Finding drug targets for neuropsychiatric disorders via deep-learning & Designing a predictor for the sensitivity of drugs to human population variation

- **PsychENCODE**: Population-level analysis of functional genomics data related to neuropsychiatric disease
  - Construction of an adult brain resource with 1866 individuals
  - Large-scale processing **creates a comprehensive QTL** resource (~2.5M eQTLs).
  - Connecting QTLs, enhancer activity relationships & Hi-C contacts into a **brain regulatory network**
  - Embedding the reg. network in a **deep-learning model** to predict psychiatric disease from genotype & transcriptome. Using this to suggest specific pathways & genes, as potential drug targets.

- **GenoDock**: Building a predictor for the sensitivity of drug binding to personal SNVs
  - Hybrid classifier connecting **physical modelling with statistical learning**
    - The modeling creates a pseudo gold-standard dataset, which is used to train the stat. classifier
  - **Classifier Results**
    - Independent validation on an expt. validation set
    - Gives higher disruption scores to cancer driver SNVs. Also, illustrates importance of different features (eg GERP).
    - Picks out certain drugs (eg imatinib) as being particularly sensitive to SNVs

Deep Structured Phenotype Network (DSPN)

Gene regulatory network builds skeleton



Boltzmann machine

$\mathbf{y}$: phenotypes

$\mathbf{h}$: hidden units (e.g., circuits)

$\mathbf{x}$: intermediate phenotypes (e.g., genes, enhancers)

$\mathbf{z}$: genotypes (e.g., SNPs)

W: weights (e.g., regulatory network)

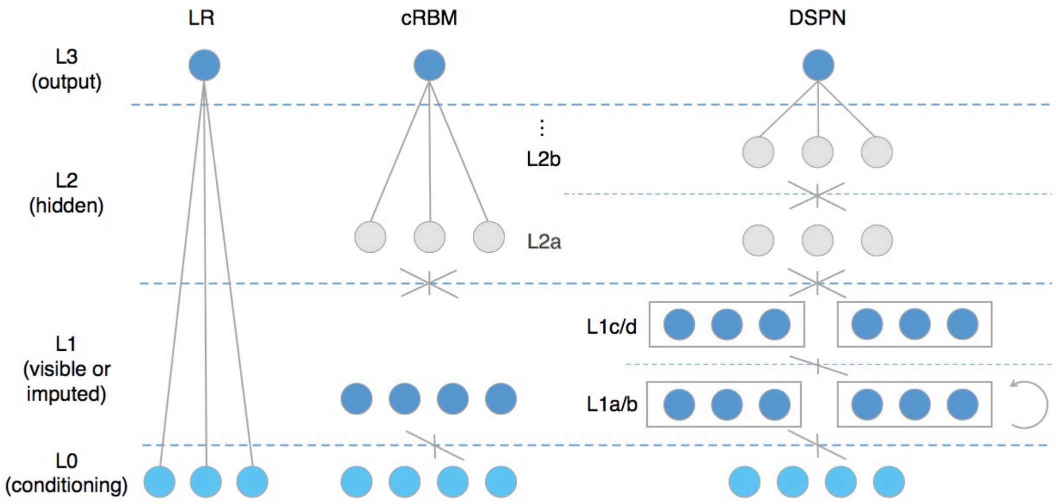
Energy model:

$$p(\mathbf{x},\mathbf{y},\mathbf{h}|\mathbf{z}) \propto \exp\big(-E(\mathbf{x},\mathbf{y},\mathbf{h}|\mathbf{z})\big)$$

$$E(\mathbf{x},\mathbf{y},\mathbf{h}|\mathbf{z}) = -\mathbf{z}^{\mathrm{T}}\mathbf{W}_1\mathbf{x} - \mathbf{x}^{\mathrm{T}}\,\mathbf{W}_2\mathbf{x} - \mathbf{x}^{\mathrm{T}}\mathbf{W}_3\mathbf{h} - \mathbf{h}^{\mathrm{T}}\mathbf{W}_4\mathbf{h} - \mathbf{h}^{\mathrm{T}}\mathbf{W}_5\mathbf{y} - \boldsymbol{Bias}$$

[Wang et al. ('18) Science]

DSPN improves brain disease prediction by adding deep layers



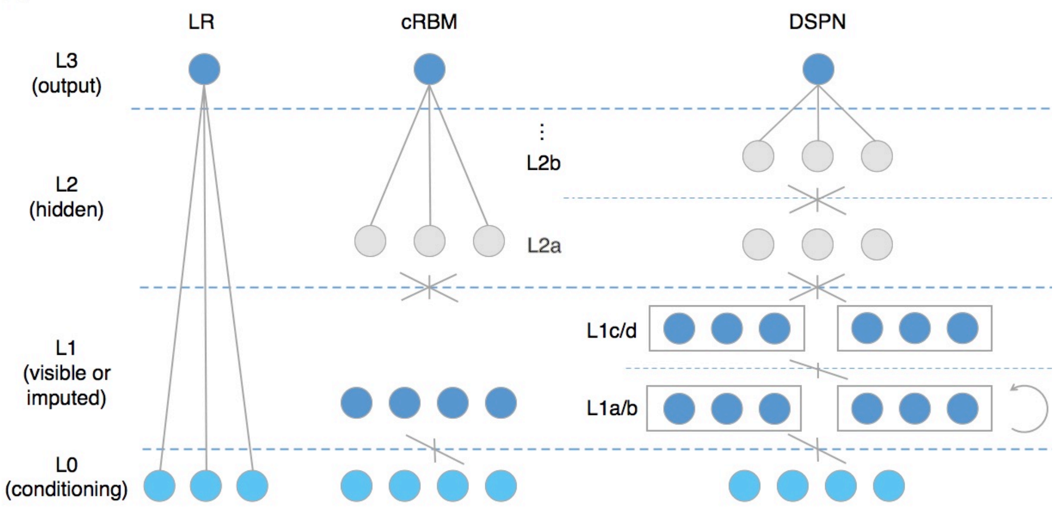
Method	LR-genotype	LR-transcriptome	cRBM	DSPN-imputation	DSPN-full
Schizophrenia	**54.6%**	63.0%	70.0%	59.0%	**73.6%**
Bipolar Disorder	**56.7%**	63.3%	71.1%	67.2%	**76.7%**
Autism Spectrum Disorder	**50.0%**	51.7%	67.2%	62.5%	**68.3%**

X 6.0

Accuracy = chance to correctly predict disease/health

[Wang et al. ('18) Science]

DSPN improves brain disease prediction by adding deep layers



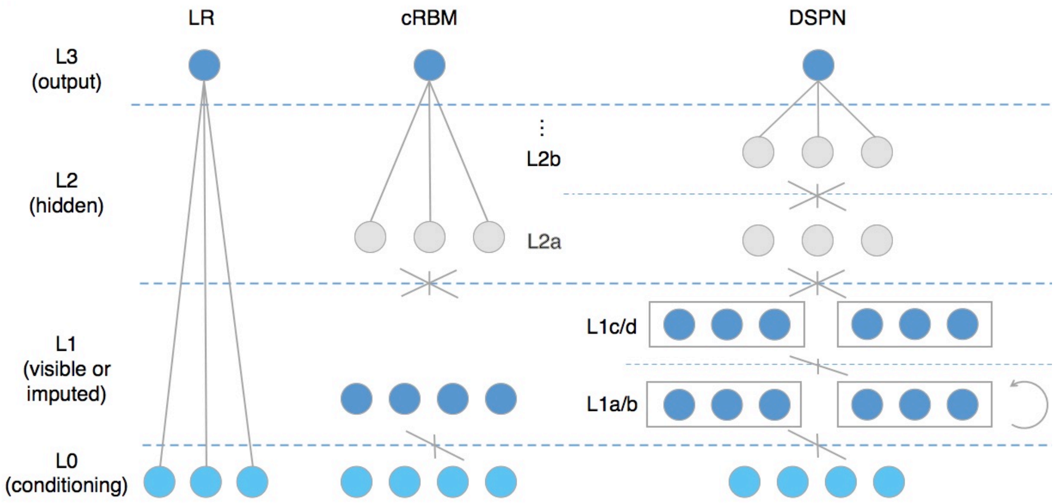
Method	LR-genotype	LR-transcriptome	cRBM	DSPN-imputation	DSPN-full
Schizophrenia	54.6%	63.0%	70.0%	59.0%	73.6%
Bipolar Disorder	56.7%	63.3%	71.1%	67.2%	76.7%
Autism Spectrum Disorder	50.0%	51.7%	67.2%	62.5%	68.3%

X 2.5

Accuracy = chance to correctly predict disease/health

[Wang et al. ('18) Science]

DSPN improves brain disease prediction by adding deep layers



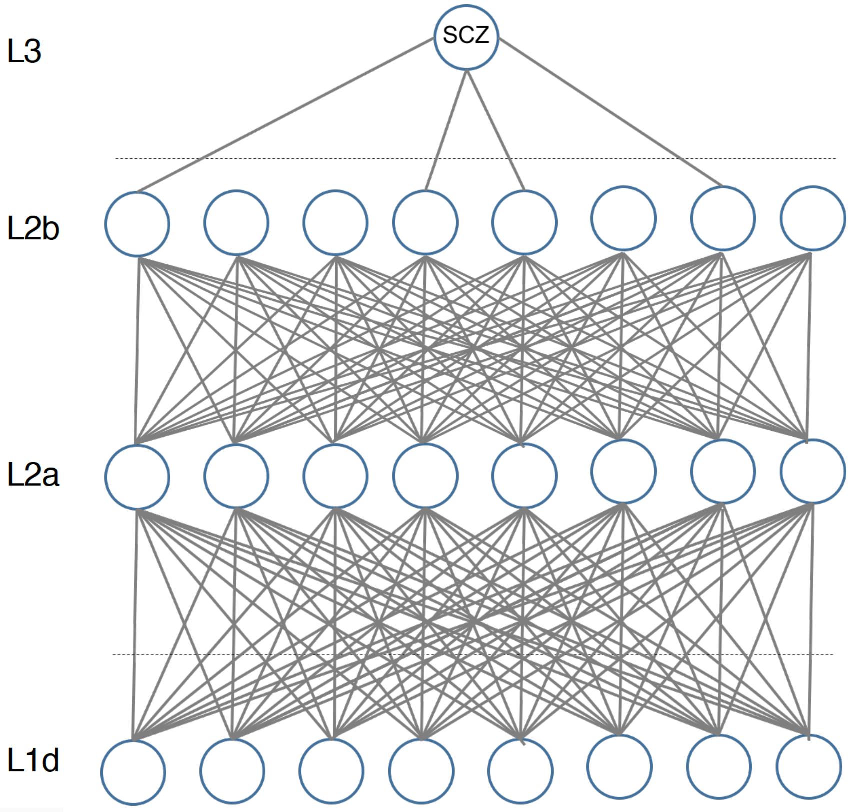
Method	LR-genotype	LR-transcriptome	cRBM	DSPN-imputation	DSPN-full
Schizophrenia	**54.6%**	63.0%	70.0%	**59.0%**	73.6%
Bipolar Disorder	**56.7%**	63.3%	71.1%	**67.2%**	76.7%
Autism Spectrum Disorder	**50.0%**	51.7%	67.2%	**62.5%**	68.3%

X 3.1

Accuracy = chance to correctly predict disease/health

[Wang et al. ('18) Science]

Multilevel Network Interpretation

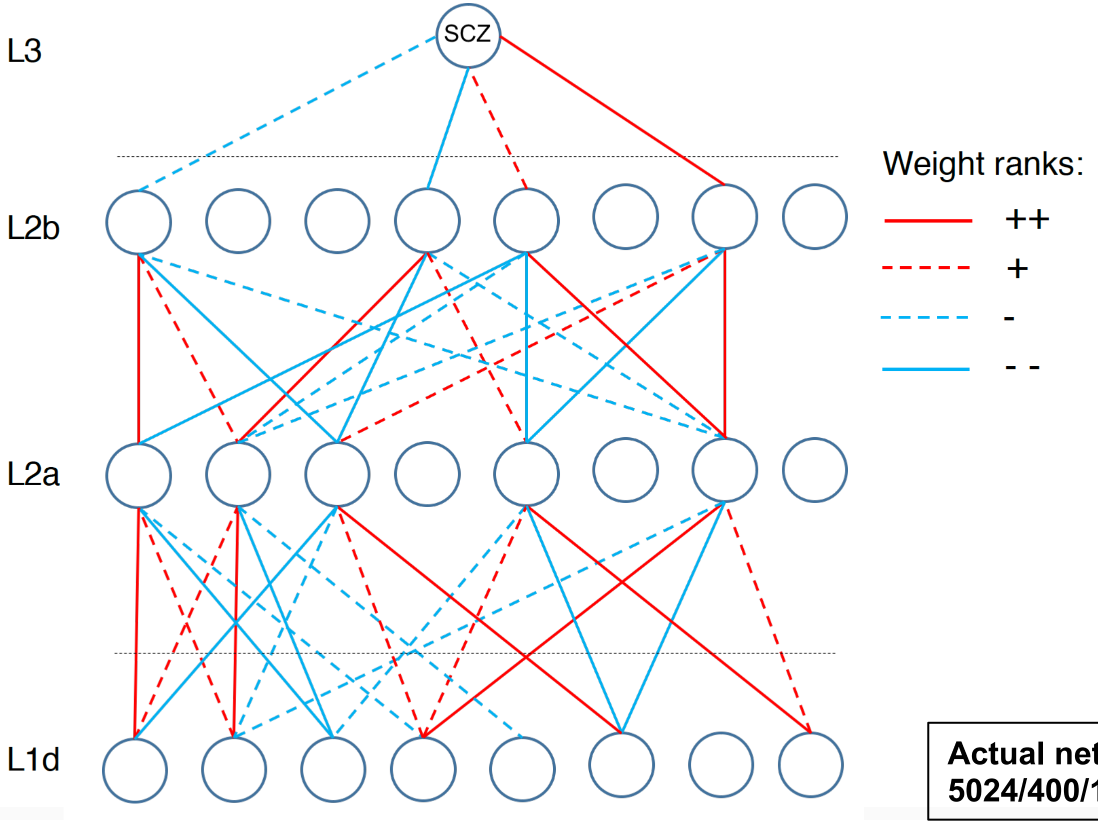


Actual network size:
5024/400/100/1 nodes

- Start with a fully connected trained network

[Wang et al. ('18) Science]

Multilevel Network Interpretation

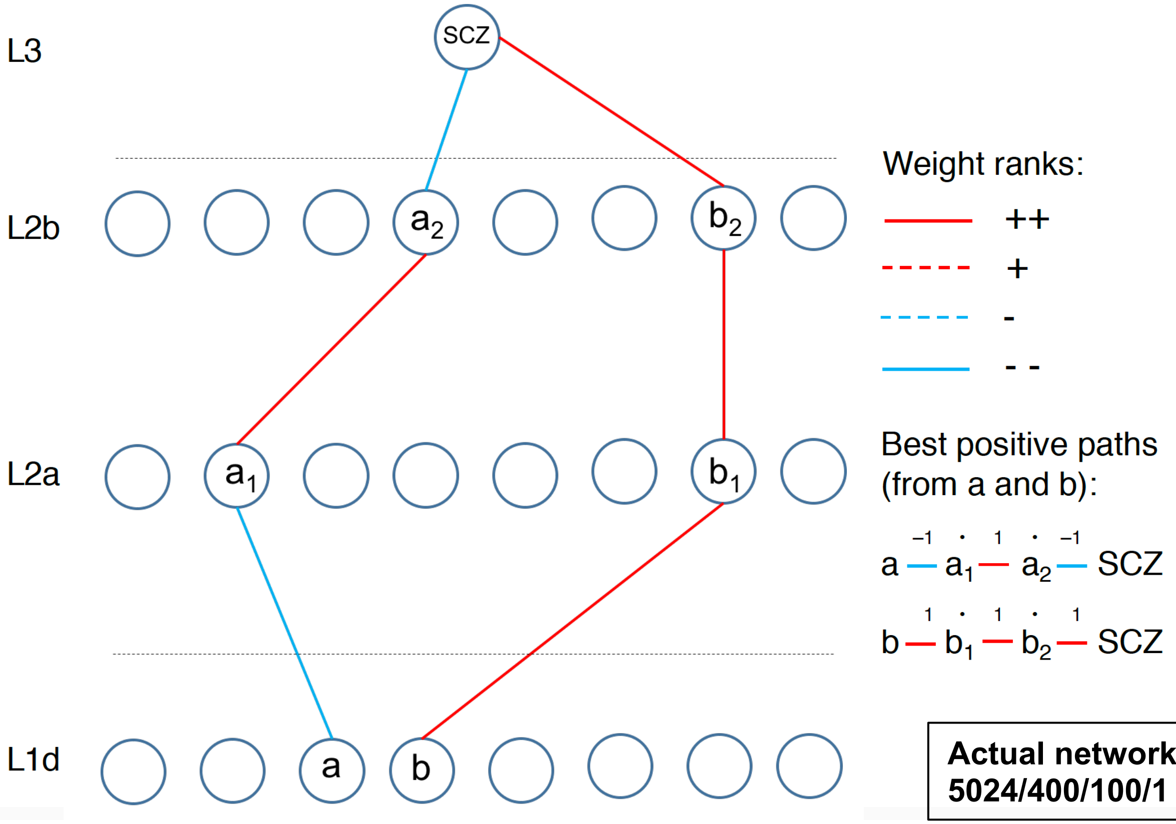


Actual network size: 5024/400/100/1 nodes

- Start with a fully connected trained network
- Sparsify network using edges with largest absolute weights (+/-)

[Wang et al. ('18) Science]

Multilevel Network Interpretation



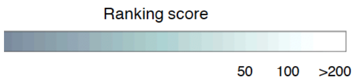
Actual network size:
5024/400/100/1 nodes

- Start with a fully connected trained network
- Sparsify network using edges with largest absolute weights (+/-)
- Extract 'best positive paths' to each prioritized module
  (e.g. a-a_1-a_2-SCZ) by summing weights and multiplying signs

[Wang et al. ('18) Science]

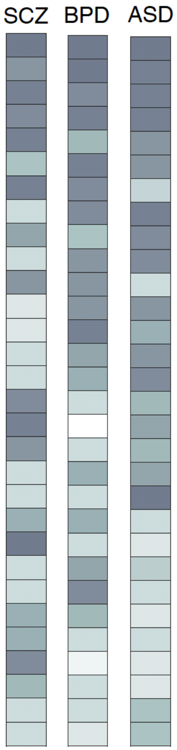
DSPN discovers enriched pathways and linkages to genetic variation

Cross-disorder MOD/HOG enrichment ranking

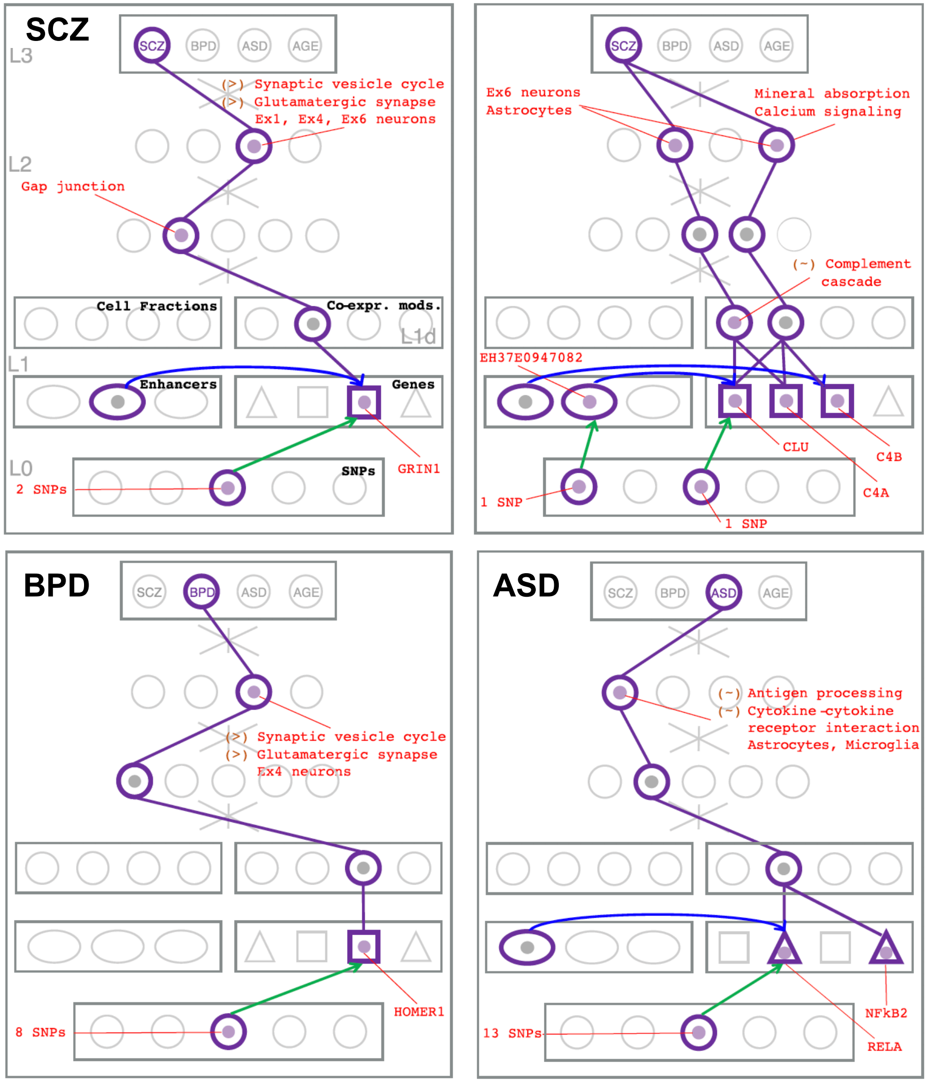


Functional categories

(*)	RNA proc.	(~)	Immune
(>)	Synaptic	(#)	Metabolic



- (*) Spliceosome / RNA splicing
- (>) Synaptic vesicle cycle
- (~) Antigen proc. and presentation
- Vesicle localization
- Proteasome
- (*) mRNA processing
- Chromatin modification
- (#) Oxidative phosphorylation
- Retrograde endocannabinoid sig.
- (>) Chemical synaptic transmission
- Peptidyl-lysine modification
- Endocytosis
- Ubiquitin mediated proteolysis
- (>) Anterograde trans-synaptic sig.
- (*) mRNA transport
- Phosphatidylinositol signaling
- Hippo signaling pathway
- (~) Staph./ Epstein-Barr virus inf.
- (>) Synaptic signaling
- Autophagy
- (>) Dop./GABA/Glutamatergic synapse
- (>) Calcium signaling
- (>) Endocrine calcium reabsorption
- (*) RNA degradation / transport
- (#) Ribosome
- Neuron projection morphogenesis
- (~) Fc receptor signaling pathway
- cGMP-PKG signaling pathway
- (~) mTOR signaling pathway
- (~) Cytokine-cytokine receptor int.



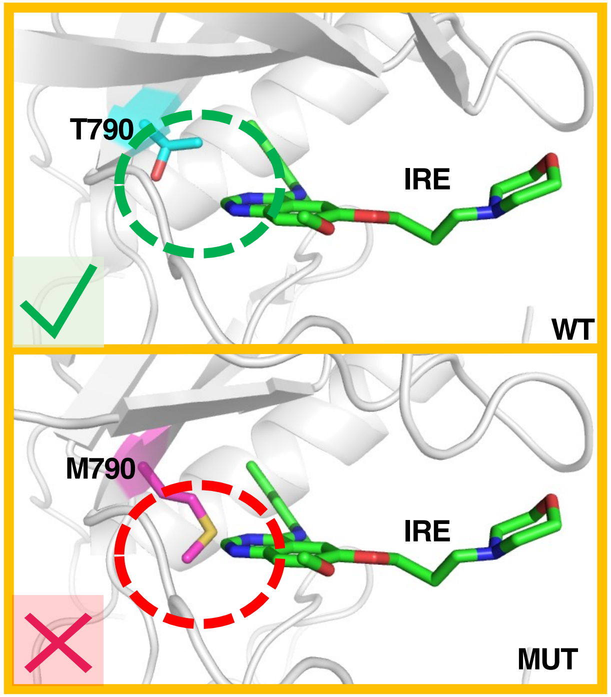
[Wang et al. ('18) Science]

Finding drug targets for neuropsychiatric disorders via deep-learning & Designing a predictor for the sensitivity of drugs to human population variation

- **PsychENCODE**: Population-level analysis of functional genomics data related to neuropsychiatric disease
  - Construction of an adult brain resource with 1866 individuals
  - Large-scale processing **creates a comprehensive QTL** resource (~2.5M eQTLs).
  - Connecting QTLs, enhancer activity relationships & Hi-C contacts into a **brain regulatory network**
  - Embedding the reg. network in a **deep-learning model** to predict psychiatric disease from genotype & transcriptome. Using this to suggest specific pathways & genes, as potential drug targets.

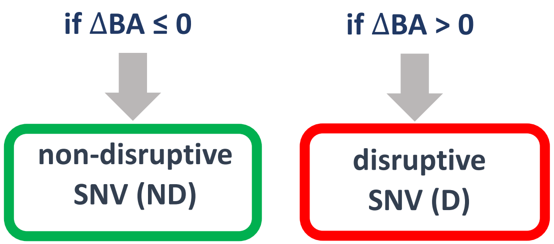
- **GenoDock**: Building a predictor for the sensitivity of drug binding to personal SNVs
  - Hybrid classifier connecting **physical modelling with statistical learning**
    - The modeling creates a pseudo gold-standard dataset, which is used to train the stat. classifier
  - **Classifier Results**
    - Independent validation on an expt. validation set
    - Gives higher disruption scores to cancer driver SNVs. Also, illustrates importance of different features (eg GERP).
    - Picks out certain drugs (eg imatinib) as being particularly sensitive to SNVs

An Example of Binding Affinity Change between Protein & Drug Ligand under the Impact of Single Nucleotide Variants (SNV)



human EGFR & gefitinib (IRE)
PDB: 2ity, Chain A, amino acid 790
Modeling and Visualization: Modeller & PyMol

Epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (EGFR-TKIs) are used in the treatments of non-small cell lung cancer (NSCLC)



For protein-drug binding upon point mutation,

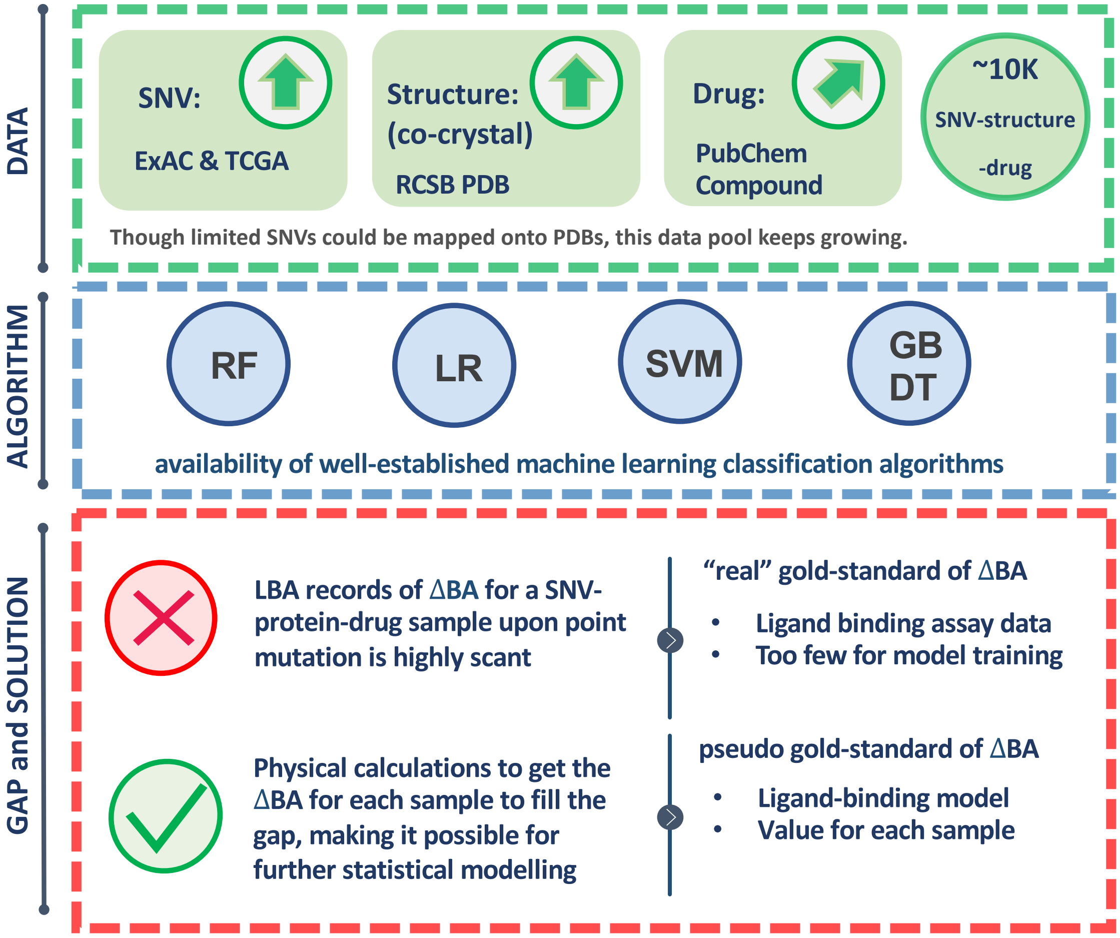
if $\Delta BA \leq 0$ → **non-disruptive SNV (ND)**

if $\Delta BA > 0$ → **disruptive SNV (D)**

Is there any method that could predict the effects of SNVs to drug binding (D or ND)?

Wang et al. Structure, 2019

Assessment of feasibility to build a supervised-learning classifier for binding-disruptive SNVs



SNV: ⬆

ExAC & TCGA

Structure: (co-crystal) ⬆

RCSB PDB

Drug: ↗

PubChem Compound

~10K SNV-structure -drug

Though limited SNVs could be mapped onto PDBs, this data pool keeps growing.

availability of well-established machine learning classification algorithms

✗ LBA records of ΔBA for a SNV-protein-drug sample upon point mutation is highly scant

“real” gold-standard of ΔBA

- Ligand binding assay data
- Too few for model training

✓ Physical calculations to get the ΔBA for each sample to fill the gap, making it possible for further statistical modelling

pseudo gold-standard of ΔBA

- Ligand-binding model
- Value for each sample

Wang et al. Structure, 2019

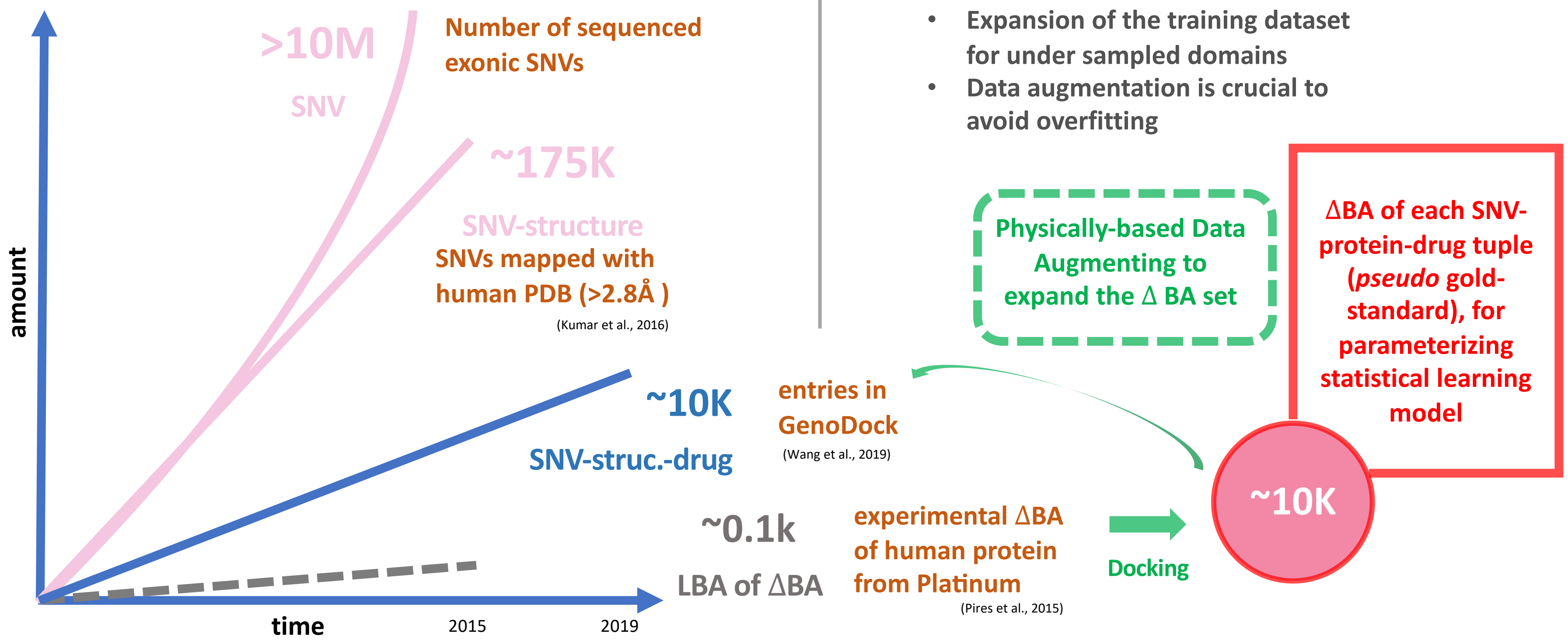
A Hot Topic in Machine Learning is "Hybrid" Model Integrating Physical & Statistical Calculations

The Major Hurdle:
Highly Scant Ligand Binding Assay Data for ΔBA

The Physically-based Data Augmentation Approach:
Leveraging Physical Calculations of ΔBA to Fill the Gap

(Reichstein *et al.*, **Nature**, 2019 & Xie et al., preprint, 2018)

- Expansion of the training dataset for under sampled domains
- Data augmentation is crucial to avoid overfitting



Number of sequenced exonic SNVs

SNVs mapped with human PDB (>2.8Å)

(Kumar et al., 2016)

entries in GenoDock

(Wang et al., 2019)

experimental ΔBA of human protein from Platinum

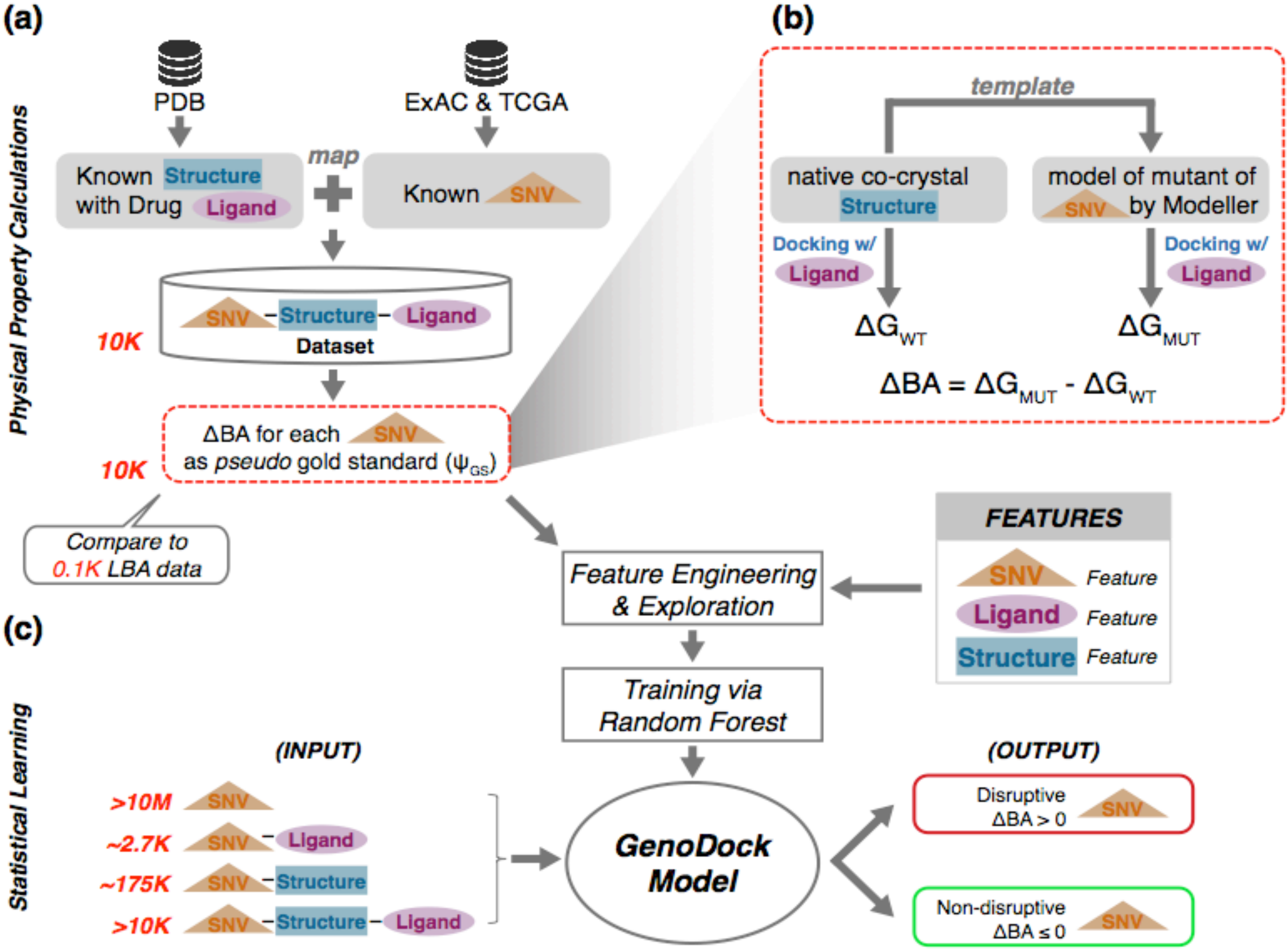
(Pires et al., 2015)

Physically-based Data Augmenting to expand the Δ BA set

ΔBA of each SNV-protein-drug tuple (*pseudo* gold-standard), for parameterizing statistical learning model

Wang et al. Structure, 2019

Framework for GenoDock: from Dataset Preparation to Model Construction



Wang et al. Structure, 2019

3 Feature Groups as Predictor, with 4 Application Cases Based on Info Availability

What are features are effective for prioritization of disruptive SNVs?

③ groups of features as predictors

SNV Feature

- Allele Frequency
- SIFT
- PolyPhen-2
- GERP
- Germline/Somatic

Ligand Feature

- Molecular Weight
- H-bond donor
- H-bond acceptor
- Rotatable Bond #
- Polar Surface Area

Structure Feature

- Distance
- Binding Site
- Polarity Change
- Volume Change

Will SNV of interest disrupt protein-ligand binding

④ random forest model trained based on information available

- SNV + Structure + Ligand
- SNV + Structure
- SNV + Ligand
- SNV only

validate the "full feature" case

then, expand the model to 3 more "feature poor" cases

Wang et al. Structure, 2019

List of Models & Datasets in the Study

Model 1: statistical model (GenoDock)
Model 2: ligand binding model (to calculate ΔBA)

Model	Role	Parameterization	Validation	Description
1	Core Model	Statistical model from Ψ_{GS}	Platinum	Supervised learning model using the pseudo gold-standard set as target feature. The direct validation of this model is to apply the model to an independent, experiment-based validation dataset.
2	Auxillery Model	Physically based	-	A physical-based, previously published computational ligand-docking model to calculate binding affinity change for the pseudo gold standard set.

Dataset	Role	Size	Source	Description
Ψ_{GS}	Trains 1	~10k	Built from 2	Core dataset constructed for training the statistical model. Contains pseudo gold standard set as the target feature.
Platinum	Validates 1	86	Experiment	The human protein subset from Platinum. used as direct validation dataset of our statistical method.

KEY TAKE-AWAY

- **The statistical model and ligand binding model are the two models for this study;**
- **The validation of the statistical model and the assessment of rigor of the ligand binding model are two independent process.**

Wang et al. Structure, 2019

Finding drug targets for neuropsychiatric disorders via deep-learning & Designing a predictor for the sensitivity of drugs to human population variation

- **PsychENCODE**: Population-level analysis of functional genomics data related to neuropsychiatric disease
  - Construction of an adult brain resource with 1866 individuals
  - Large-scale processing **creates a comprehensive QTL** resource (~2.5M eQTLs).
  - Connecting QTLs, enhancer activity relationships & Hi-C contacts into a **brain regulatory network**
  - Embedding the reg. network in a **deep-learning model** to predict psychiatric disease from genotype & transcriptome. Using this to suggest specific pathways & genes, as potential drug targets.

- **GenoDock**: Building a predictor for the sensitivity of drug binding to personal SNVs
  - Hybrid classifier connecting **physical modelling with statistical learning**
    - The modeling creates a pseudo gold-standard dataset, which is used to train the stat. classifier
  - **Classifier Results**
    - Independent validation on an expt. validation set
    - Gives higher disruption scores to cancer driver SNVs. Also, illustrates importance of different features (eg GERP).
    - Picks out certain drugs (eg imatinib) as being particularly sensitive to SNVs

The *pseudo* Gold-Standard as Self-Constructed Prediction Target: Physical Calculations for Binding Affinity Score Change (ΔBA)

Dataset

ExAC: 8565 SNV-PDB native-mutant pairs
TCGA: 1718 SNV-PDB native-mutant pairs

Equation

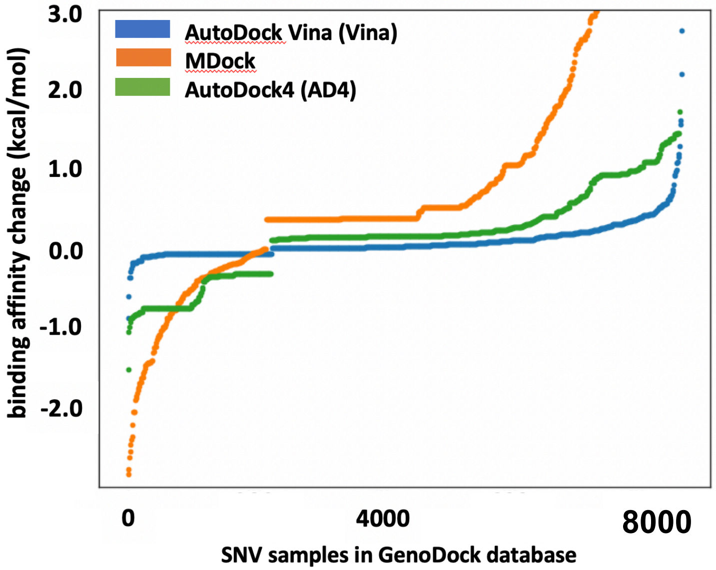
$$\Delta\Delta G(SNV) = \Delta G(SNV)\text{-}\Delta G(WT)$$

$\Delta G(WT)$: BA of WT protein-drug complex
$\Delta G(SNV)$: BA of point mutated protein-drug complex
$\Delta\Delta G(SNP)$: BA change

Approach

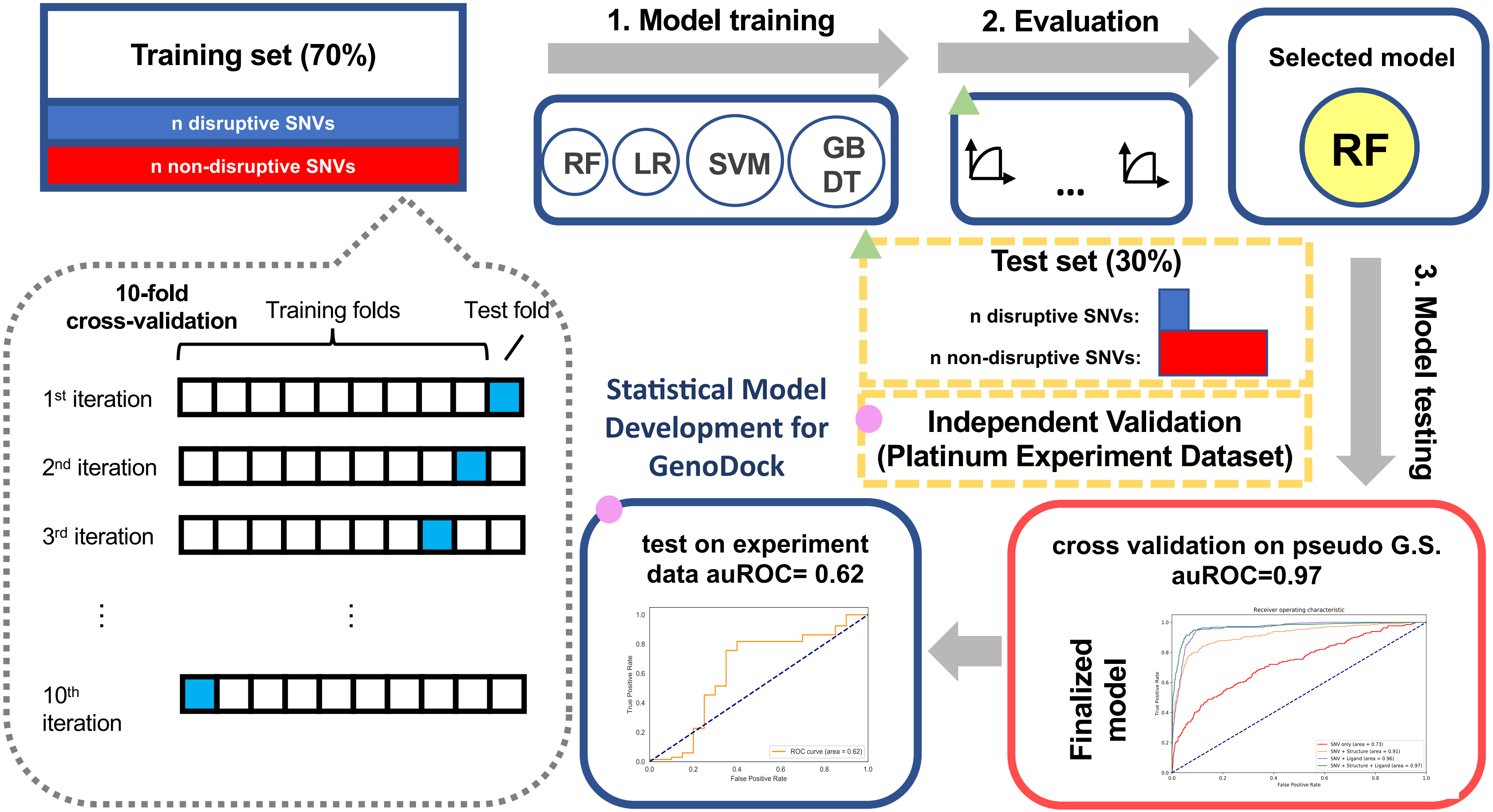
- MDock
- AutoDock Vina
- AutoDock 4

- Pearson Product-Moment Correlation (PMCC) reveals good consistency of different docking calculations
- PMCC (Vina & AD4) = 0.89
- PMCC (Vina & MDock) = 0.94



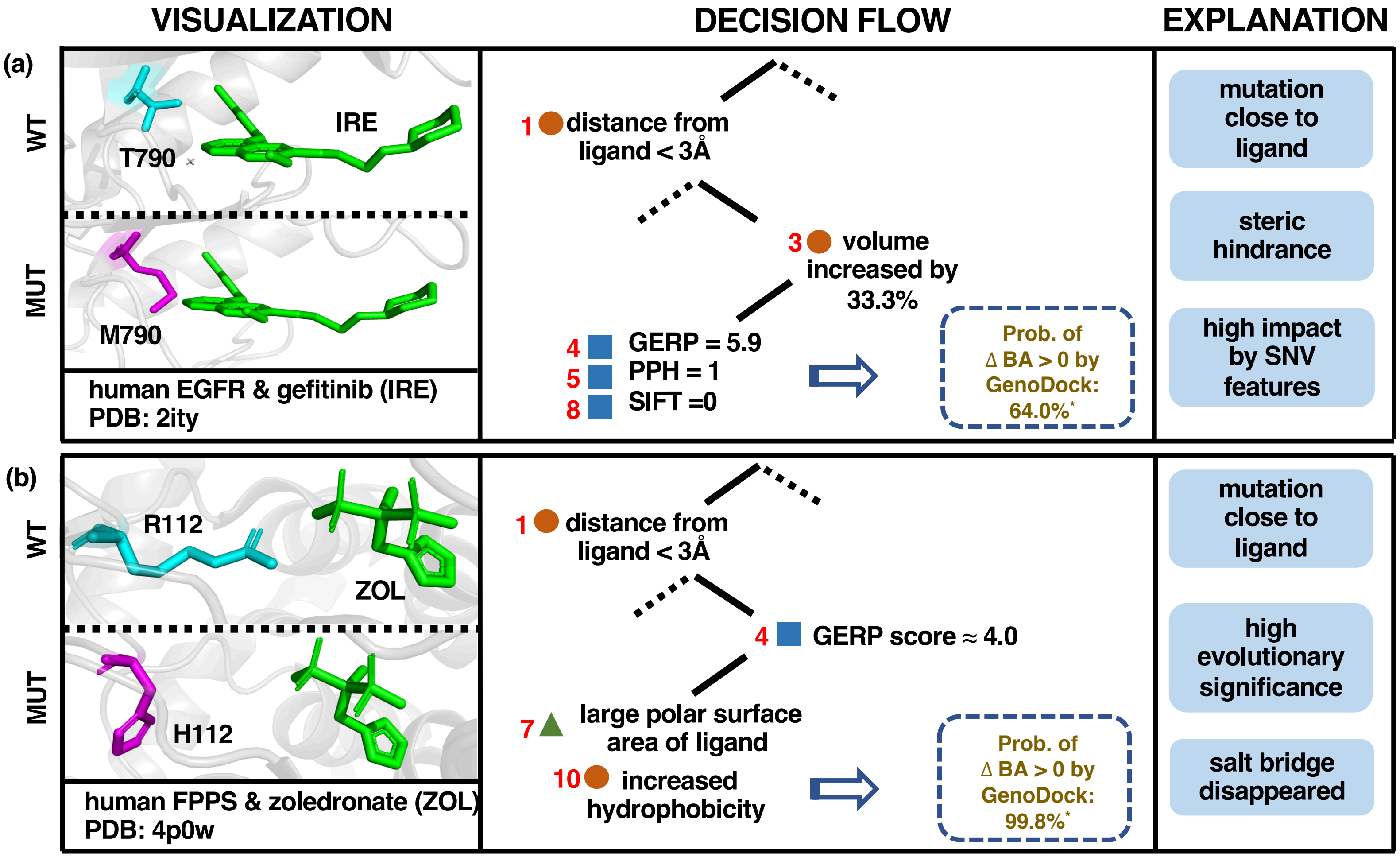
Wang et al. Structure, 2019

Given the pseudo Gold-Standard, the Workflow for Building the Statistical Model & its Performance in Cross-validation & Independent Testing



Wang et al. Structure, 2019

Example of the Output of the Classifier: GenoDock Helps Characterize Known & Unknown SNVs that Disrupt Protein-Ligand Binding



(a)

human EGFR & gefitinib (IRE)
PDB: 2ity

1 ● distance from ligand < 3Å

3 ● volume increased by 33.3%

4 🟦 GERP = 5.9
5 🟦 PPH = 1
8 🟦 SIFT =0

⇨ Prob. of Δ BA > 0 by GenoDock: 64.0%*

- mutation close to ligand
- steric hindrance
- high impact by SNV features

(b)

human FPPS & zoledronate (ZOL)
PDB: 4p0w

1 ● distance from ligand < 3Å

4 🟦 GERP score ≈ 4.0

7 ▲ large polar surface area of ligand
10 ● increased hydrophobicity

⇨ Prob. of Δ BA > 0 by GenoDock: 99.8%*

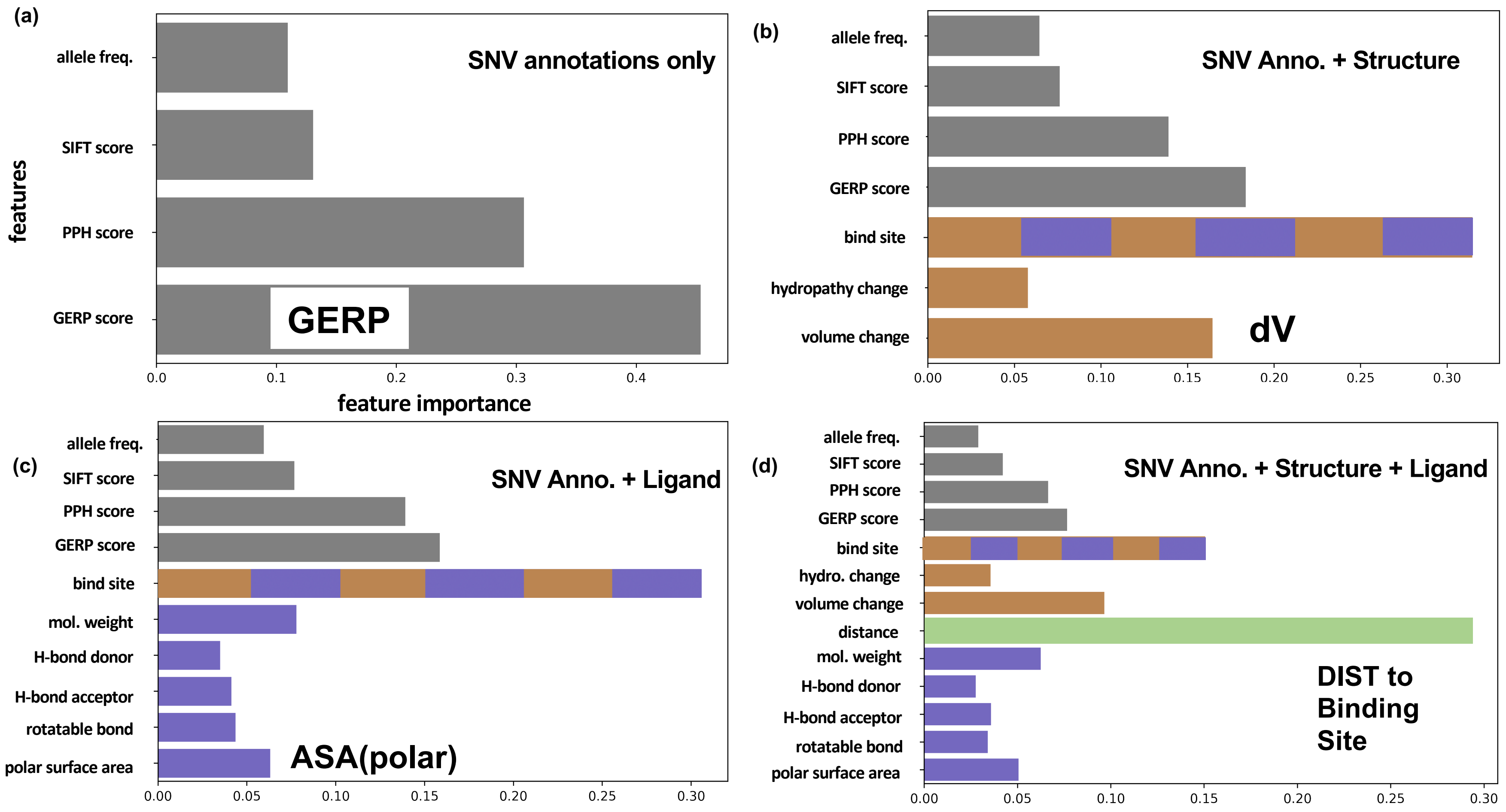
- mutation close to ligand
- high evolutionary significance
- salt bridge disappeared

● Structure Feature 🟦 SNV Feature ▲ Ligand Feature 1-10: Feature significance rank by Gini Distance for selected features

* Δ BA > 0 validated by docking calculations

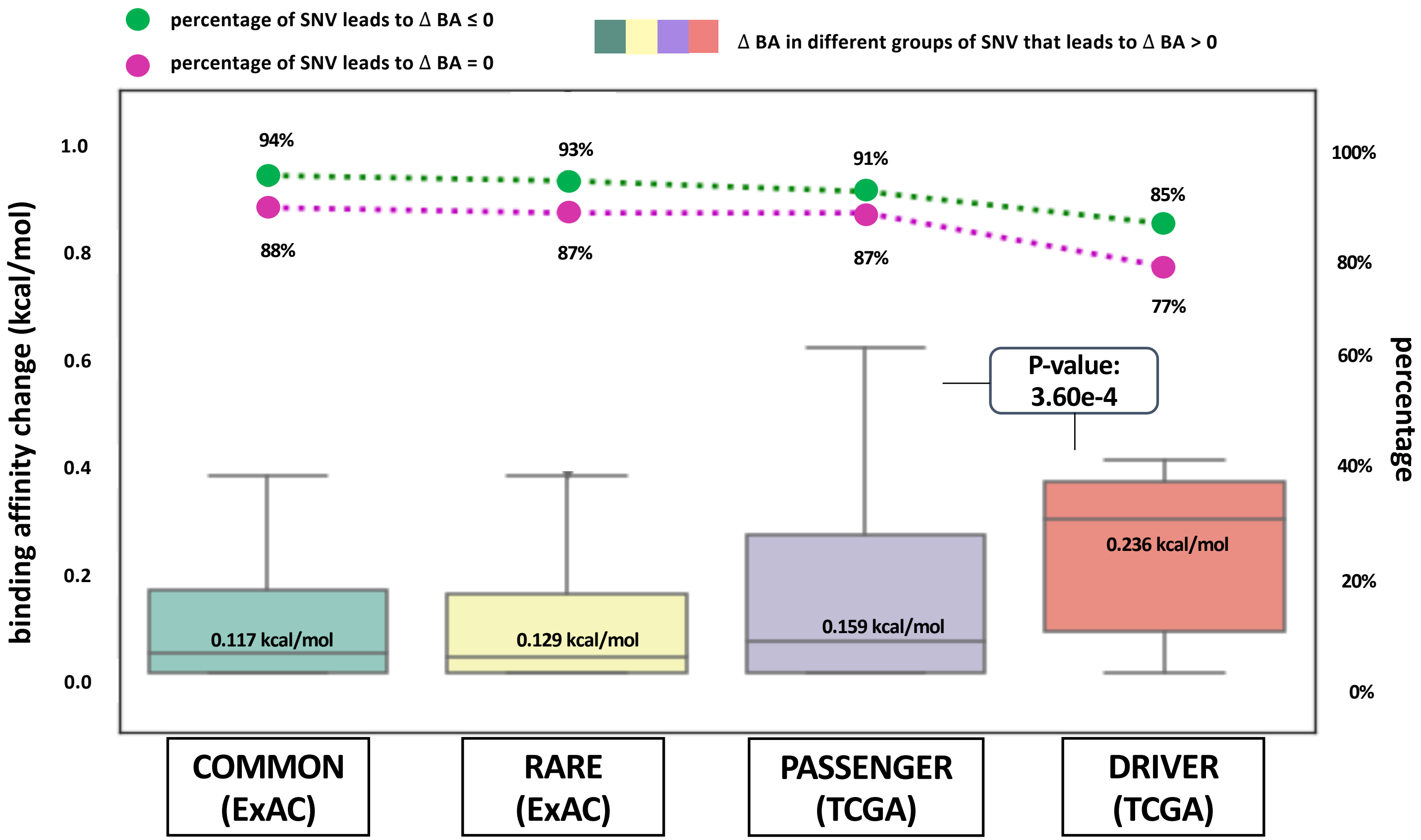
Wang et al. Structure, 2019

Gini Distance for Relative Feature Importance in 4 Models



Wang et al. Structure, 2019

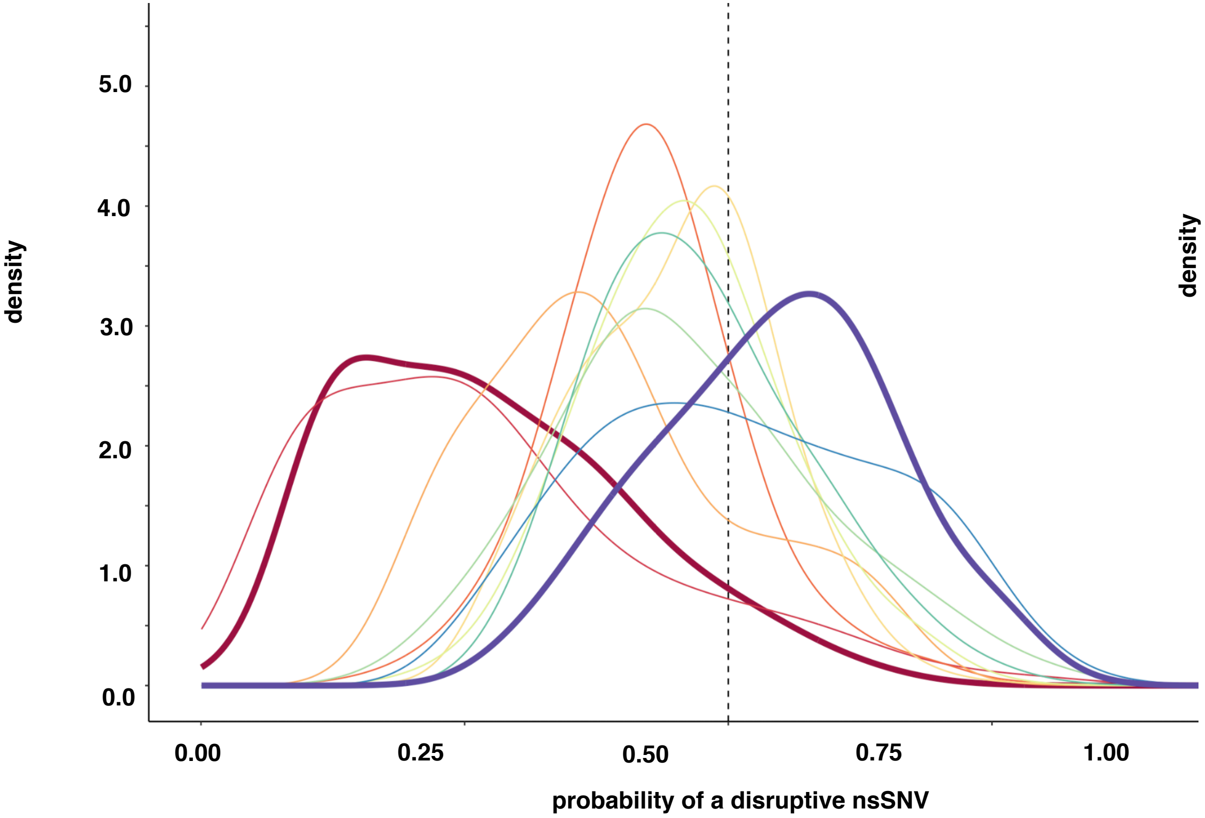
Boxplot of Overall Ligand Binding Affinity Changes for Different Types of SNVs in GenoDock

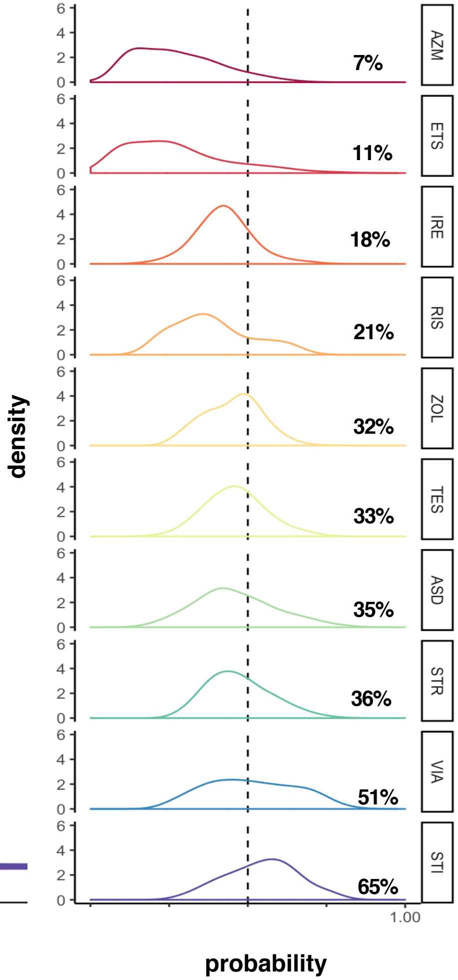


The more an SNV is considered disease-associated, the greater chance that this SNV would destabilize binding affinity of the protein and drug ligand.

Wang et al. Structure, 2019

Application of GenoDock to large-scale screening of disruptive SNVs for Drug Ligand interactions





Aceta-zolamide (glaucoma)

Imatinib (cancer)

Wang et al. Structure, 2019

Finding drug targets for neuropsychiatric disorders via deep-learning & Designing a predictor for the sensitivity of drugs to human population variation

- **PsychENCODE**: Population-level analysis of functional genomics data related to neuropsychiatric disease
  - Construction of an adult brain resource with 1866 individuals
  - Large-scale processing **creates a comprehensive QTL** resource (~2.5M eQTLs).
  - Connecting QTLs, enhancer activity relationships & Hi-C contacts into a **brain regulatory network**
  - Embedding the reg. network in a **deep-learning model** to predict psychiatric disease from genotype & transcriptome. Using this to suggest specific pathways & genes, as potential drug targets.

- **GenoDock**: Building a predictor for the sensitivity of drug binding to personal SNVs
  - Hybrid classifier connecting **physical modelling with statistical learning**
    - The modeling creates a pseudo gold-standard dataset, which is used to train the stat. classifier
  - **Classifier Results**
    - Independent validation on an expt. validation set
    - Gives higher disruption scores to cancer driver SNVs. Also, illustrates importance of different features (eg GERP).
    - Picks out certain drugs (eg imatinib) as being particularly sensitive to SNVs

Finding drug targets for neuropsychiatric disorders via deep-learning & Designing a predictor for the sensitivity of drugs to human population variation

- **PsychENCODE:** Population-level analysis of functional genomics data related to neuropsychiatric disease
  - Construction of an adult brain resource with 1866 individuals
  - Large-scale processing **creates a comprehensive QTL** resource (~2.5M eQTLs).
  - Connecting QTLs, enhancer activity relationships & Hi-C contacts into a **brain regulatory network**
  - Embedding the reg. network in a **deep-learning model** to predict psychiatric disease from genotype & transcriptome. Using this to suggest specific pathways & genes, as potential drug targets.

- **GenoDock:** Building a predictor for the sensitivity of drug binding to personal SNVs
  - Hybrid classifier connecting **physical modelling with statistical learning**
    - The modeling creates a pseudo gold-standard dataset, which is used to train the stat. classifier
  - **Classifier Results**
    - Independent validation on an expt. validation set
    - Gives higher disruption scores to cancer driver SNVs. Also, illustrates importance of different features (eg GERP).
    - Picks out certain drugs (eg imatinib) as being particularly sensitive to SNVs

PsychENCODE Acknowledgment

NIH National Institute of Mental Health

- Geetha Senthil
- Lora Bingaman
- David Panchision
- Alexander Arguello
- Thomas Lehner

“Adult Capstone” Team – 1 of 3 capstones

Daifeng Wang, Shuang Liu, Jonathan Warrell, Hyejung Won, Xu Shi, Fabio Navarro, Declan Clarke, Mengting Gu, Prashant Emani, Yucheng T. Yang, Min Xu, Michael Gandal, Shaoke Lou, Jing Zhang, Jonathan J. Park, Chengfei Yan, Suhn Kyong Rhie, Kasidet Manakongtreecheep, Holly Zhou, Aparna Nathan, Mette Peters, Eugenio Mattei, Dominic Fitzgerald, Tonya Brunetti, Jill Moore, Yan Jiang, Kiran Girdhar, Gabriel Hoffman, Selim Kalayci, Zeynep Hulya Gumus, Greg Crawford,

PsychENCODE Consortium,

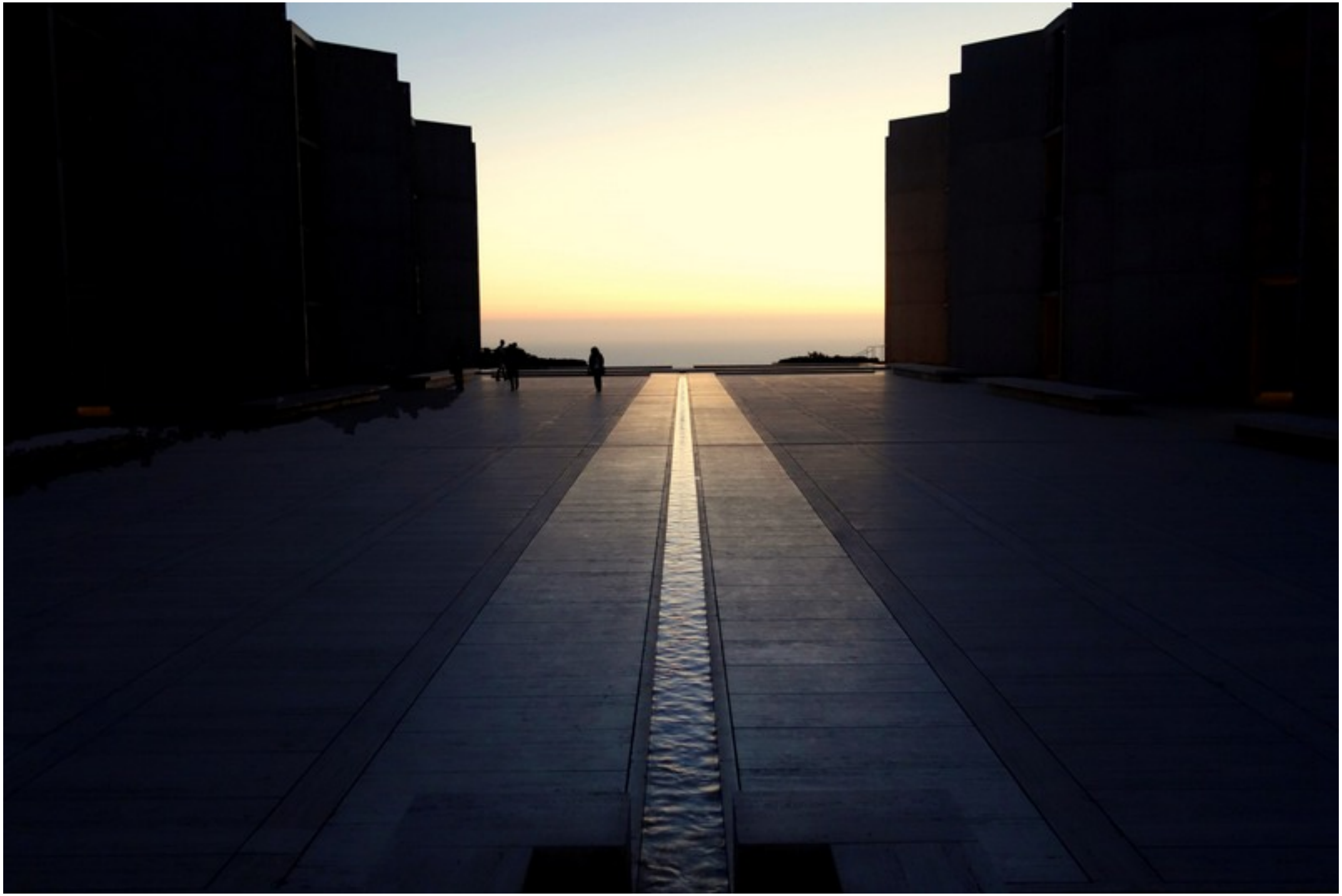
Panos Roussos, Schahram Akbarian, Andrew E. Jaffe, Kevin White, Zhiping Weng, Nenad Sestan,

Daniel H. Geschwind, James A. Knowles

Dedicated to **Pamela Sklar**

Resource.psychencode.org

The PsychENCODE Consortium: Allison E Ashley-Koch, Duke University; Gregory E Crawford, Duke University; Melanie E Garrett, Duke University; Lingyun Song, Duke University; Alexias Safi, Duke University; Graham D Johnson, Duke University; Gregory A Wray, Duke University; Timothy E Reddy, Duke University; Fernando S Goes, Johns Hopkins University; Peter Zandi, Johns Hopkins University; Julien Bryois, Karolinska Institutet; Andrew E Jaffe, Lieber Institute for Brain Development; Amanda J Price, Lieber Institute for Brain Development; Nikolay A Ivanov, Lieber Institute for Brain Development; Leonardo Collado-Torres, Lieber Institute for Brain Development; Thomas M Hyde, Lieber Institute for Brain Development; Emily E Burke, Lieber Institute for Brain Development; Joel E Kleiman, Lieber Institute for Brain Development; Ran Tao, Lieber Institute for Brain Development; Joo Heon Shin, Lieber Institute for Brain Development; Schahram Akbarian, Icahn School of Medicine at Mount Sinai; Kiran Girdhar, Icahn School of Medicine at Mount Sinai; Yan Jiang, Icahn School of Medicine at Mount Sinai; Marija Kundakovic, Icahn School of Medicine at Mount Sinai; Leanne Brown, Icahn School of Medicine at Mount Sinai; Bibi S Kassim, Icahn School of Medicine at Mount Sinai; Royce B Park, Icahn School of Medicine at Mount Sinai; Jennifer R Wiseman, Icahn School of Medicine at Mount Sinai; Elizabeth Zharovsky, Icahn School of Medicine at Mount Sinai; Rivka Jacobov, Icahn School of Medicine at Mount Sinai; Olivia Devillers, Icahn School of Medicine at Mount Sinai; Elie Flatow, Icahn School of Medicine at Mount Sinai; Gabriel E Hoffman, Icahn School of Medicine at Mount Sinai; Barbara K Lipska, Human Brain Collection Core, National Institutes of Health, Bethesda, MD; David A Lewis, University of Pittsburgh; Vahram Haroutunian, Icahn School of Medicine at Mount Sinai and James J Peters VA Medical Center; Chang-Gyu Hahn, University of Pennsylvania; Alexander W Charney, Mount Sinai; Stella Dracheva, Mount Sinai; Alexey Kozlenkov, Mount Sinai; Judson Belmont, Icahn School of Medicine at Mount Sinai; Diane DelValle, Icahn School of Medicine at Mount Sinai; Nancy Francoeur, Icahn School of Medicine at Mount Sinai; Evi Hadjimichael, Icahn School of Medicine at Mount Sinai; Dalila Pinto, Icahn School of Medicine at Mount Sinai; Harm van Bakel, Icahn School of Medicine at Mount Sinai; Panos Roussos, Mount Sinai; John F Fullard, Mount Sinai; Jaroslav Bendl, Mount Sinai; Mads E Hauberg, Mount Sinai; Lara M Mangravite, Sage Bionetworks; Mette A Peters, Sage Bionetworks; Yooree Chae, Sage Bionetworks; Junmin Peng, St. Jude Children's Hospital; Mingming Niu, St. Jude Children's Hospital; Xusheng Wang, St. Jude Children's Hospital; Maree J Webster, Stanley Medical Research Institute; Thomas G Beach, Banner Sun Health Research Institute; Chao Chen, Central South University; Yi Jiang, Central South University; Rujia Dai, Central South University; Annie W Shieh, SUNY Upstate Medical University; Chunyu Liu, SUNY Upstate Medical University; Kay S. Grennan, SUNY Upstate Medical University; Yan Xia, SUNY Upstate Medical University/Central South University; Ramu Vadukapuram, SUNY Upstate Medical University; Yongjun Wang, Central South University; Dominic Fitzgerald, The University of Chicago; Lijun Cheng, The University of Chicago; Miguel Brown, The University of Chicago; Mimi Brown, The University of Chicago; Tonya Brunetti, The University of Chicago; Thomas Goodman, The University of Chicago; Majd Alsayed, The University of Chicago; Michael J Gandal, University of California, Los Angeles; Daniel H Geschwind, University of California, Los Angeles; Hyejung Won, University of California, Los Angeles; Damon Polioudakis, University of California, Los Angeles; Brie Wamsley, University of California, Los Angeles; Jiani Yin, University of California, Los Angeles; Tarik Hadzic, University of California, Los Angeles; Luis De La Torre Ubieta, UCLA; Vivek Swarup, University of California, Los Angeles; Stephan J Sanders, University of California, San Francisco; Matthew W State, University of California, San Francisco; Donna M Werling, University of California, San Francisco; Joon-Yong An, University of California, San Francisco; Brooke Sheppard, University of California, San Francisco; A Jeremy Willsey, University of California, San Francisco; Kevin P White, The University of Chicago; Mohana Ray, The University of Chicago; Gina Giase, SUNY Upstate Medical University; Amira Kefi, University of Illinois at Chicago; Eugenio Mattei, University of Massachusetts Medical School; Michael Purcaro, University of Massachusetts Medical School; Zhiping Weng, University of Massachusetts Medical School; Jill Moore, University of Massachusetts Medical School; Henry Pratt, University of Massachusetts Medical School; Jack Huey, University of Massachusetts Medical School; Tyler Borrman, University of Massachusetts Medical School; Patrick F Sullivan, University of North Carolina - Chapel Hill; Paola Giusti-Rodriguez, University of North Carolina - Chapel Hill; Yunjung Kim, University of North Carolina - Chapel Hill; Patrick Sullivan, University of North Carolina - Chapel Hill; Jin Szatkiewicz, University of North Carolina - Chapel Hill; Suhn Kyong Rhie, University of Southern California; Christoper Armoskus, University of Southern California; Adrian Camarena, University of Southern California; Peggy J Farnham, University of Southern California; Valeria N Spitsyna, University of Southern California; Heather Witt, University of Southern California; Shannon Schreiner, University of Southern California; Oleg V Evgrafov, SUNY Downstate Medical Center; James A Knowles, SUNY Downstate Medical Center; Mark Gerstein, Yale University; Shuang Liu, Yale University; Daifeng Wang, Stony Brook University; Fabio C. P. Navarro, Yale University; Jonathan Warrell, Yale University; Declan Clarke, Yale University; Prashant S. Emani, Yale University; Mengting Gu, Yale University; Xu Shi, Yale University; Min Xu, Yale University; Yucheng T. Yang, Yale University; Robert R. Kitchen, Yale University; Gamze Gürsoy, Yale University; Jing Zhang, Yale University; Becky C Carlyle, Yale University; Angus C Nairn, Yale University; Mingfeng Li, Yale University; Sirisha Pochareddy, Yale University; Nenad Sestan, Yale University; Mario Skarica, Yale University; Zhen Li, Yale University; Andre M.M. Sousa, Yale University; Gabriel Santpere, Yale University; Jinmyung Choi, Yale University; Ying Zhu, Yale University; Tianliuyun Gao, Yale University; Daniel J Miller, Yale University; Adriana Cherskov, Yale University; Mo Yang, Yale University; Anahita Amiri, Yale University; Gianfilippo Coppola, Yale University; Jessica Mariani, Yale University; Soraya Scuderi, Yale University; Anna Szekely, Yale University; Flora M Vaccarino, Yale University; Feinan Wu, Yale University; Sherman Weissman, Yale University; Tanmoy Roychowdhury, Mayo Clinic Rochester; Alexej Abyzov, Mayo Clinic Rochester;.



See

JOBS.gersteinlab.org

Hiring Postdocs

GenoDock.molmovdb.org

B Wang, C Yan,

S Lou, P Emani, B Li, M Xu, X Kong, W Meyerson, Y Yang, D Lee

Extra

Info about content in this slide pack

- General PERMISSIONS
  - This Presentation is copyright Mark Gerstein, Yale University, 2019.
  - Please read permissions statement at www.**gersteinlab.org/misc/permissions.html** .
  - Feel free to use slides & images in the talk with PROPER acknowledgement (via citation to relevant papers or link to gersteinlab.org).
  - Paper references in the talk were mostly from Papers.GersteinLab.org.

- PHOTOS & IMAGES. For thoughts on the source and permissions of many of the photos and clipped images in this presentation see http://streams.gerstein.info .
  - In particular, many of the images have particular EXIF tags, such as kwpotppt , that can be easily queried from flickr, viz: http://www.flickr.com/photos/mbgmbg/tags/kwpotppt