Cell Reports Medicine, Volume *3*

Supplemental information

A community challenge for a pancancer drug mechanism of action inference from perturbational profile data

Eugene F. Douglass Jr., Robert J. Allaway, Bence Szalai, Wenyu Wang, Tingzhong Tian, Adrià Fernández-Torras, Ron Realubit, Charles Karan, Shuyu Zheng, Alberto Pessia, Ziaurrehman Tanoli, Mohieddin Jafari, Fangping Wan, Shuya Li, Yuanpeng Xiong, Miquel Duran-Frigola, Martino Bertoni, Pau Badia-i-Mompel, Lídia Mateo, Oriol Guitart-Pla, Verena Chung, DREAM CTD-squared Pancancer Drug Activity Challenge Consortium, Jing Tang, Jianyang Zeng, Patrick Aloy, Julio Saez-Rodriguez, Justin Guinney, Daniela S. Gerhard, and Andrea Califano

Supporting Information:
A Community Challenge for Pancancer Drug Mechanism of Action Inference from Perturbational Profile Data

Figure S1. Potential Clinical Relevance of Problem Posed

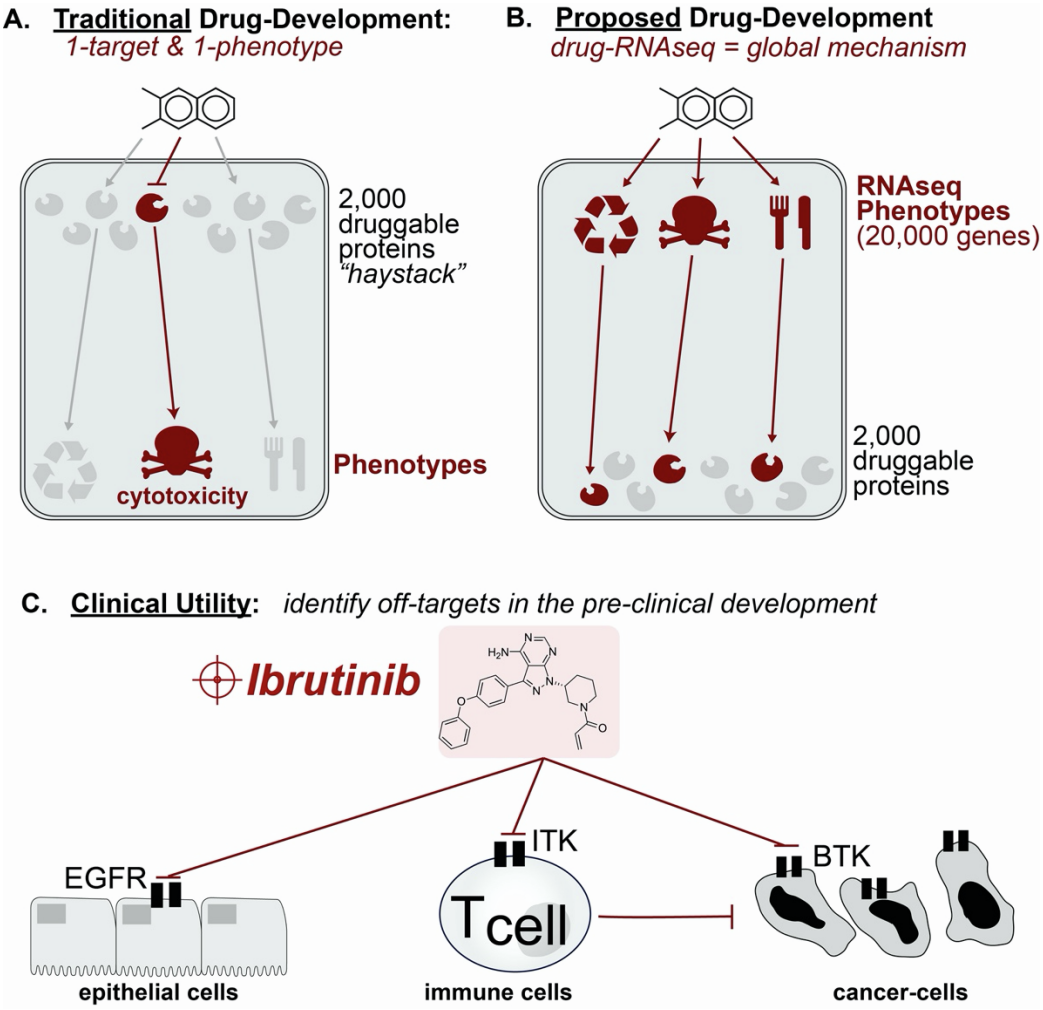

Figure S1. Proposed clinical utility of proposed work: global assessment of drug mechanism of action (A) Traditional drug development focuses on single-targets and phenotypes directly relevant to inhibiting cancer cell growth. This approach ignores other effects of the drug on cancer cells as metabolism and stress-response programs such as autophagy **(B)** drug-perturbed RNA sequencing of cancer cell-lines yields unbiased genome-scale information on a drugs mechanism of action. This information could be used to narrow the haystack of 2,000 druggable proteins that traditional methods are agnostic too. **(C)** Clinically, off-targets are often identified post-hoc following the identification of clinical toxicity. **Related to Figure 1A which described the underlying assumptions of this DREAM competition.**

Figure S2. Rationale for Choice of Drug-Target Benchmarking Data

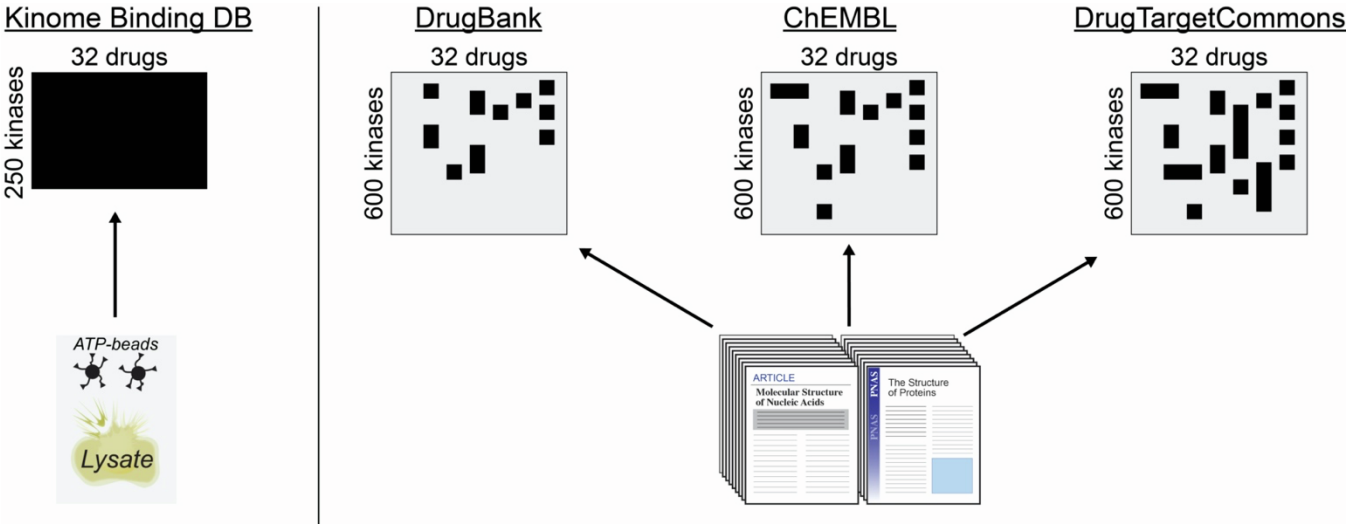

Figure S2. The Kinome Binding Resource (KBR) was chosen as a gold-standard due to concerns about literature-over fitting and overall sparsity. While the KBR only features 250 out of the total ~600 kinases in the Human genome, the fact that all drug-kinase pairs were annotated with a numeric value (K_d) that was obtained under the same conditions was very appealing as literature databases either have binary definitions of drug targets (DrugBank) or have inconsistent binding constant measurements. For example, within ChEMBL the Lapatinib-EGFR K_d is recorded as having a range of values from 0.92 to 3600 nM; these values were obtained using different experimental systems and conditions and cannot be easily reconciled. In addition, literature-curated databases are better known than the new KBR. As a result, the organizers were concerned that the use of one database in training would result in over-fitting as the ultimate source of data was the experimental literature. **Related to Figure 1C which illustrates the benchmarking data used and Figure 3 which compares different benchmarking datasets.**

Figure S3. Results of community challenge.

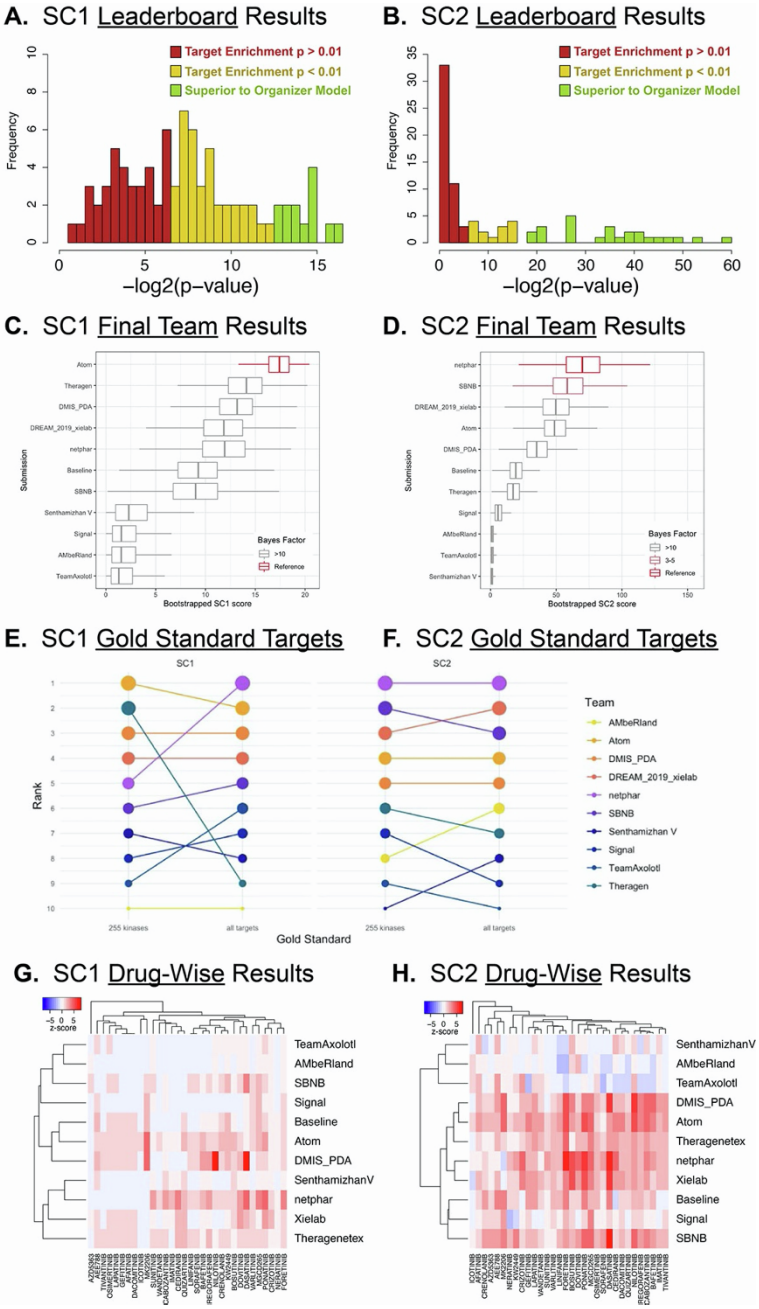

Figure S3. Results of community challenge. (A-B) In the Leaderboard phase, 39 out of 86 models showed statistically significant predictive power by both scoring criteria **(C-D)** In the final round, 6 models beat the base-line model of the organizers (E-F) We re-scored all teams removing the non-kinase filtering step to evaluate performance as a function of the rank of the gold-standard kinases within all 1259 "druggable" targets. **(G-H)** Unsupervised analysis of model performance across individual drugs**. Related to Figure 1F which described our competion scoring procedure (detailed here).**

Figure S4. Comparison of Model Performance on “canonical targets” vs “new targets”

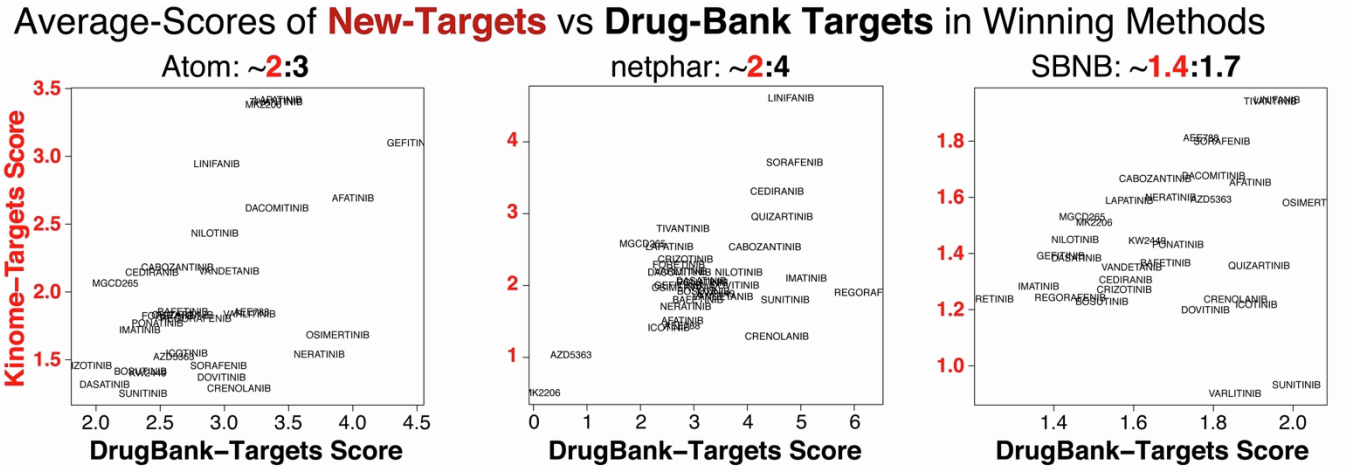

Figure S4. Comparison method scores for DrugBank-targets (x-axis) versus Unique-Kinome-targets (y-axis) for each drug within the predictions of the winning teams. **Related to Figure 3 which described the difference between benchmarking datasets.**

Figure S5. Team ranking as a function of gold standard dataset

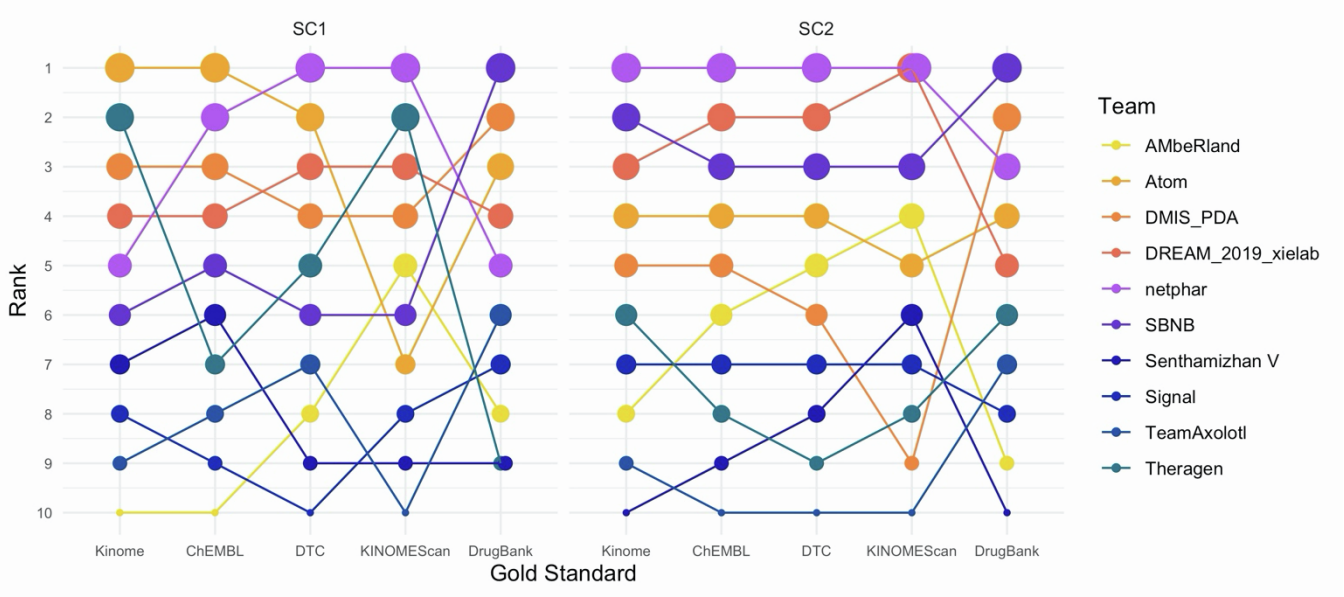

Figure S5. Team ranking as a function of gold standard dataset. "Kinome" ranks were determined using the Challenge scoring metrics, gold standard, and null modes generated from the possible kinome targets. Kinome (KBR, the dataset used in the challenge), ChEMBL, DTC (DrugTargetCommons), KinomeScan, and DrugBank ranks were calculated using the DREAM challenge scoring metrics, compound target profiles from each of the respective databases, and using a null model generated from 1259 possible "druggable" targets, to consider targets beyond the kinome. Team ranks are relatively consistent across different gold standards, but in some cases affects the relative ranking of different teams. **Related to Figure 3 which described the difference between benchmarking datasets.**

Figure S6. Contribution of Drug-Sensitivity to Model Performance

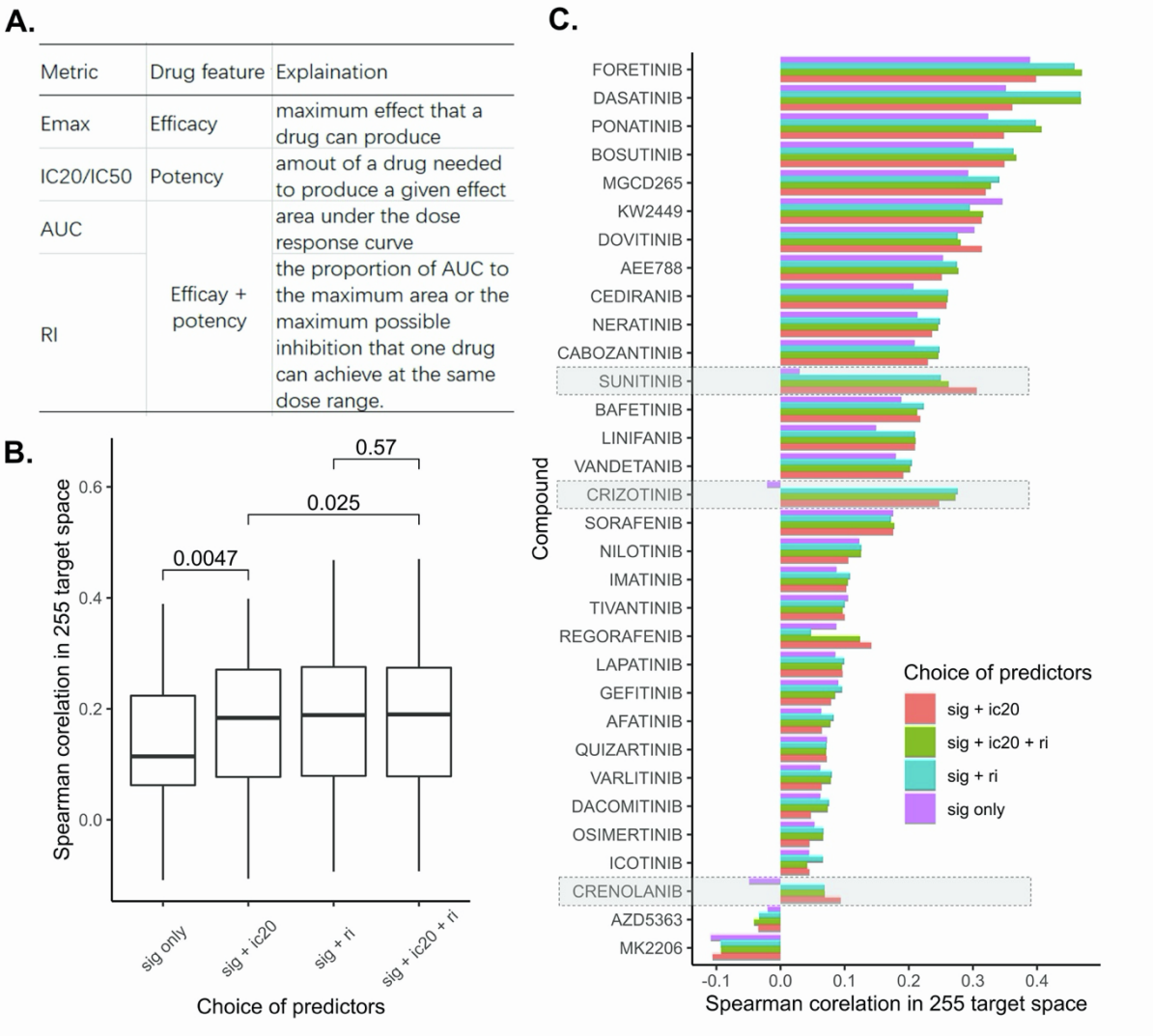

Figure S6 Assessment of the contributions of drug-sensitivity and drug-perturbed RNAseq to model performance a. Comparison of different drug sensitivity metrics. **b.** Overall prediction accuracy for the weighted averaging method using different types of predictors. P-values are determined using paired Wilcoxon tests. **c.** Drug-wise prediction accuracy for the weighted averaging method using different predictors. Drugs highlighted in grey are those predicted poorly using signature data alone. **Related to Figure 5 which describes winning methodologies.**

Figure S7. Overview of Team Netphars modeling Strategy

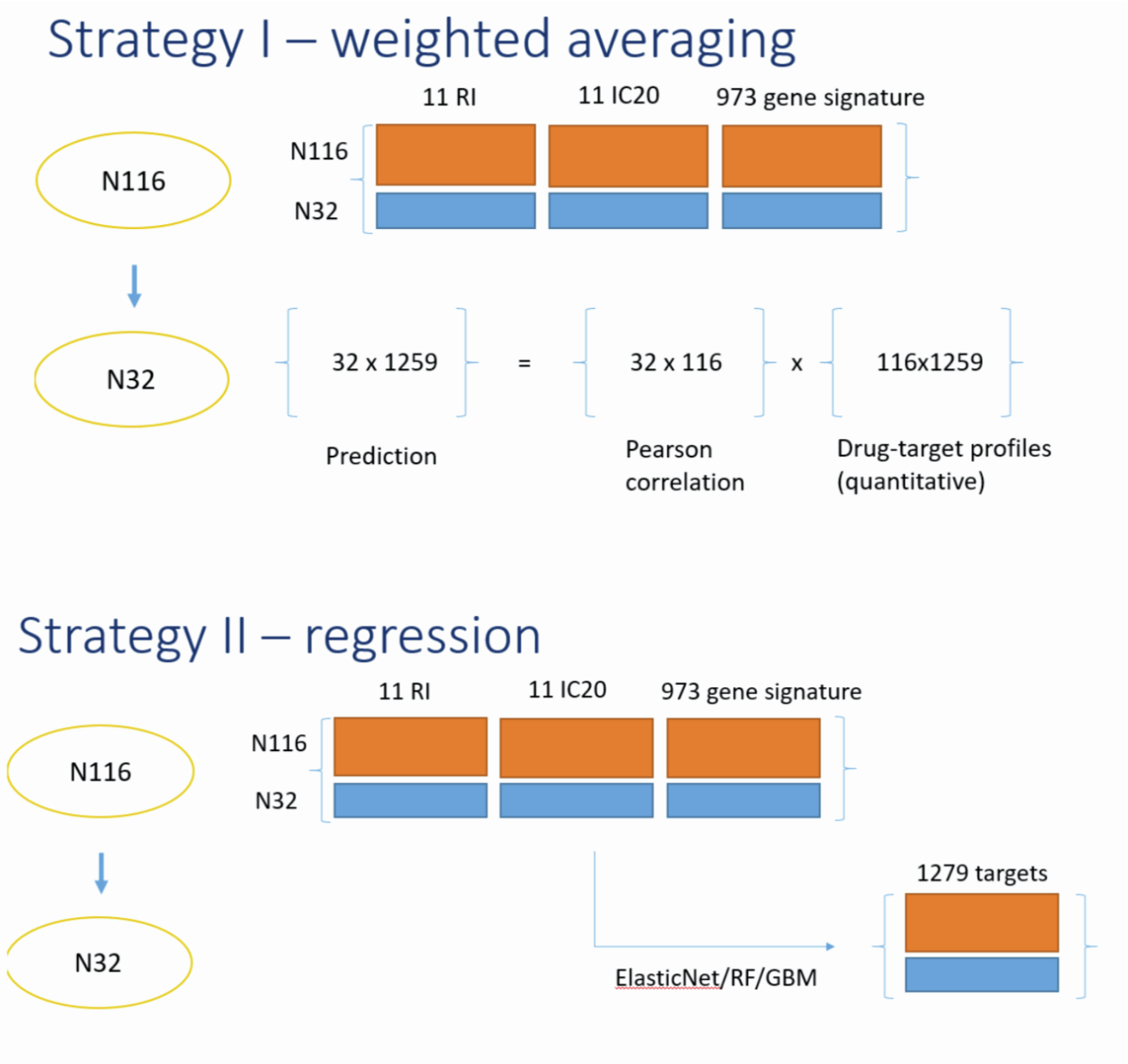

Figure S7. Overview of Team Netphar's modeling strategy. **Related to Figure 5 which describes winning methodologies.**

Figure S8. Overview of Team SBNB's modeling Strategy

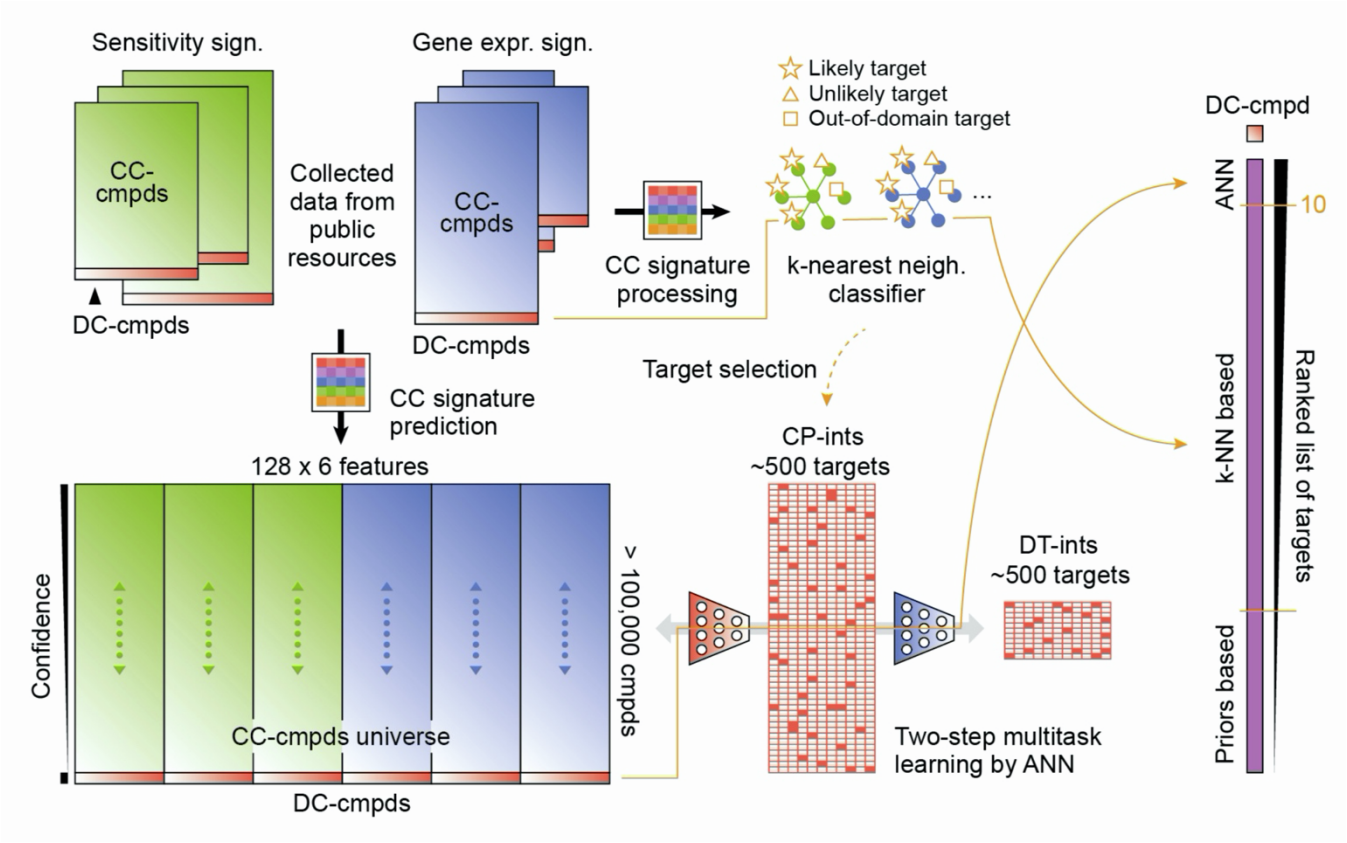

Figure S8. Overview of Team SBNB's modeling strategy. **Related to Figure 5 which describes winning methodologies.**

Figure S9. Overview of Team Atoms modeling Strategy

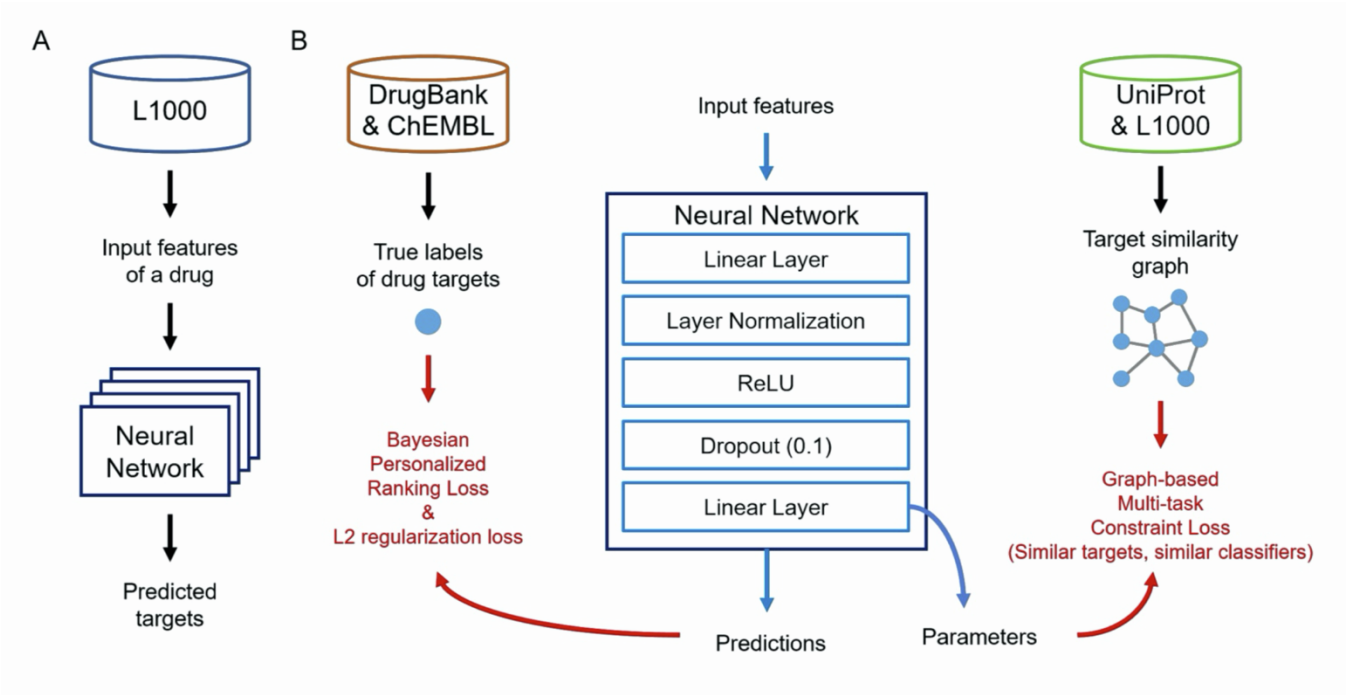

Figure S9. Model Architecture of Team Atom. **Related to Figure 5 which describes winning methodologies.**