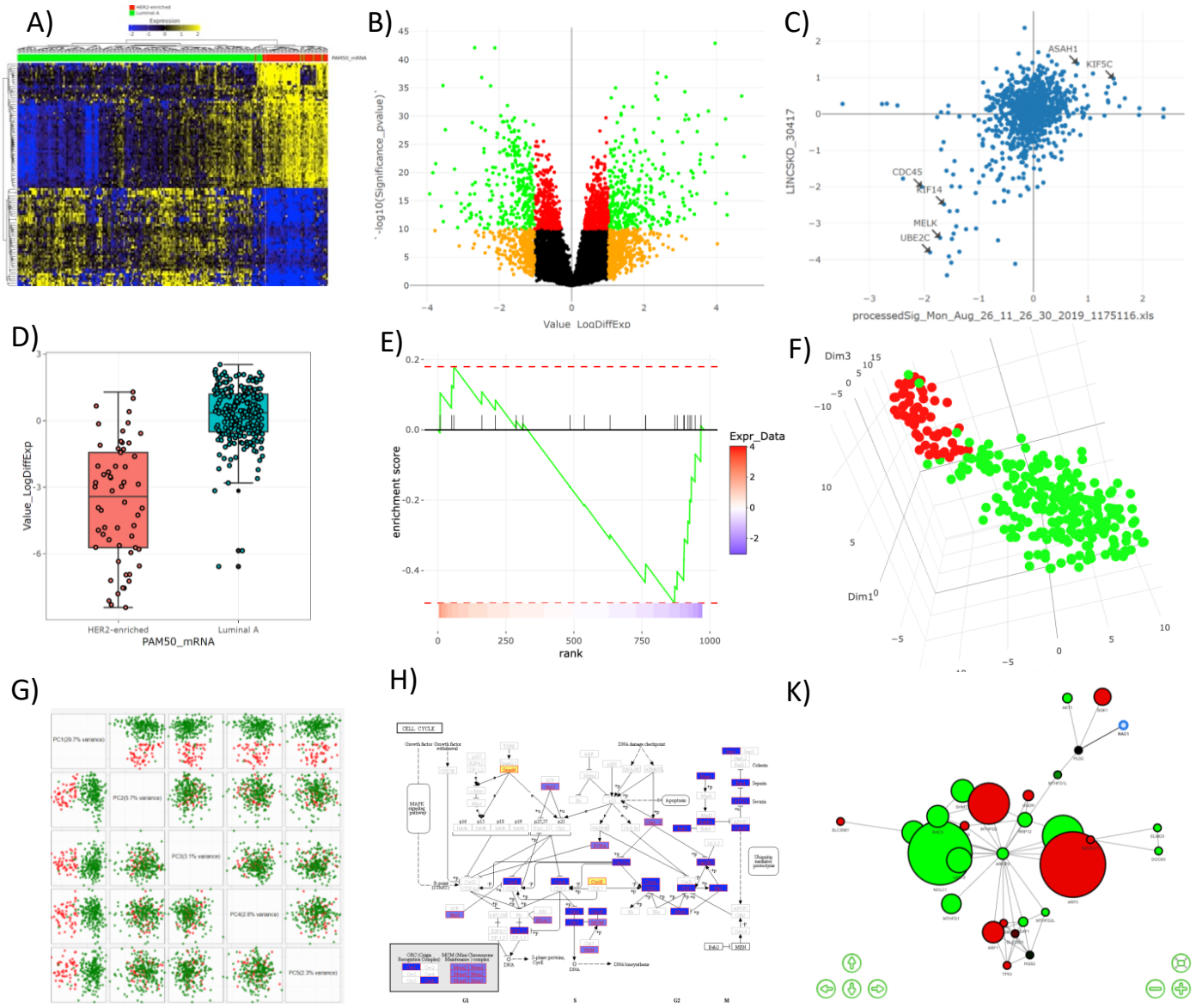
Contents

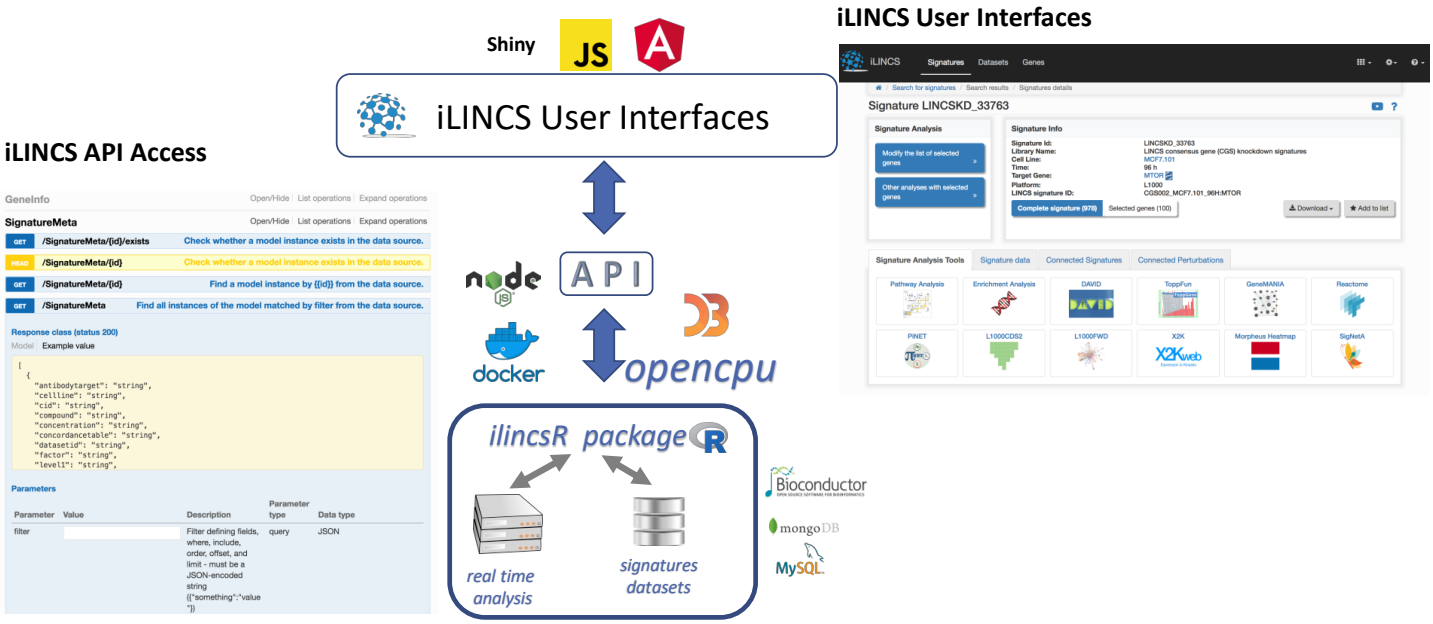
Supplemental Figure 1: Interactive visualization tools in iLINCS.

A) Interactive heatmaps: As a gold standard graphical display for visualizing high-dimensional data and relationships, interactive heatmaps are used throughout iLINCS via several different applications: Native Shiny heatmap, Java based FTreeView, Java script based Morpheus and Clustergrammer. Diversity of heatmap apps provide for diversity of functionalities and it facilitates use of iLINCS in different configurations of network speed vs the computer speed. B) Interactive volcano plots are used for visualizing and selecting informative (eg differentially expressed) genes/proteins in a signature; C) Interactive scatter plots are used to visualize relationships between two signatures and identifying genes/proteins driving the "connectivity"; D) Interactive box plots are used to visualize differential distribution of gene/proteins in different samples and up- and down-regulated genes/proteins in different signatures; E) Interactive GSEA plots serve to visualize strength of connectivity and identify genes/proteins driving the connectivity between a gene list and a signature; F) Interactive 3D scatter plots are used to visualize high-dimensional relationship in dimensionality reduction analysis (PCA and t-SNE); G) Connected 2D scatter plots for visualizing high-dimensional relationship in dimensionality reduction analysis (PCA and t-SNE); H) Interactive pathway visualizations allows for exploring the relationships between signatures and pathways; and K) Interactive network visualizations is used to integrate signatures with the global protein-protein interaction network.



Supplemental Figure 2: iLINCS architecture.

iLINCS is based on software stack: MySQL, MongoDB, R, NodeJS, and AngularJS. MySQL and MongoDB backend databases contain pre-processed genomics datasets, signatures and their connections, and all associated metadata. For external users, the database is accessed through a powerful API, which can be explored and tested with the Swagger UI interface. The iLINCS API was created in nodeJS using ExpressJS and Loopback frameworks. The API is documented and can be explored through swagger UI interface, which allows users to explore available API models and methods. The iLINCS internal analytical engine is written in R utilizing a range of specialized R packages. The iLINCS API is based on nodeJS using ExpressJS and Loopback frameworks. The nodeJS API connects with the R analytical engine using the openCPU framework which provides HTTP API for executing R functions. The user interfaces are written in Java Script using AngularJS framework. iLINCS facilitates submission of signatures and intermediate analysis results via redirection APIs to a range of third-party task specific bioinformatics web tools and services.



Supplemental Results 1: Comparison of iLINCS and clue.io query results

iLINCS and clue.io query tools were compared based on the "connected" chemical perturbagens (CPs) returned by the query with the MTOR CRISPR consensus gene signature (CGS) used in the Use Case 1 of the main text (Fig 2A). For submission to the clue.io, the signature is summarized as the list of 50 most up-regulated and the 50 most down-regulated genes. The default clue.io connectivity metric, median τ across all cell-lines, was used to identify 20 most connected CPs and results, along with results obtained from iLINCS for the same CPs are shown in Table S1. Most of the connected CPs are also deemed significant by iLINCS analysis. Interestingly, none of the CPs implicated by clue.io, but not picked up as connected by iLINCS target the core elements of the mTOR signaling. On the other hand, all top 20 signatures (lowest pValue) implicated by iLINCS (Table S2) are known to target core elements of mTOR signaling (mTOR, PI3K and AKT proteins). This includes four perturbagens not deemed connected by clue.io (highlighted). Two CPs that show in the iLINCS results, but were missed by clue.io are not in the Touchstone set used by clue.io [1] (is_touchstone=0). Finally, 6 CPs are missing connectivity information in clue.io because they were part of the Phase II LINCS L1000 dataset (GSE70138), which are included in iLINCS, but not used by clue.io. The CPs in red were missing MOA information in the released data, and we established their MOA by manual literature and online databases searches.

In summary, clue.io and iLINCS provide qualitatively similar results with connected CPs targeting mTOR pathway. Some differences were introduced by the different statistical measure of similarity used by the two system, but the biggest differences most likely came from the scope of the query, where clue.io searches over a space of 2,837 Touchstone CPs and iLINCS searches over a space of 15,349 CPs with at least one high quality signature.

Table S1: Top 20 CPs returned by the clue.io

Perturbagen ID	Perturbagen Name	ilincsMoa	is_touchstone	median_tau	iLINCS	pValue
BRD-K67566344	KU-0063794	MTOR inhibitor	1	97.35	Connected	3.83E-89
BRD-K84937637	sirolimus	MTOR inhibitor	1	97.31	Connected	7.21E-55
BRD-A84045418	calpeptin	Calpain inhibitor	1	97.23	Connected	2.08E-33
BRD-K12184916	NVP-BEZ235	MTOR inhibitor	1	96.02	Connected	6.94E-161
BRD-K71726959	BRD-K71726959	CDK inhibitor	1	95.33	Not Connected	
BRD-K27305650	LY-294002	MTOR inhibitor	1	95.32	Connected	3.35E-45
BRD-K92577649	GBR-13069	Dopamine uptake inhibitor	1	95.14	Not Connected	
BRD-K30677119	PP-30	RAF inhibitor	1	94.71	Connected	3.48E-18
BRD-K99818283	PIK-90	PI3K inhibitor	1	94.42	Connected	1.27E-54
BRD-K69932463	AZD-8055	MTOR inhibitor	1	93.54	Connected	1.17E-218
BRD-A62025033	temsirolimus	MTOR inhibitor	1	93.49	Connected	3.46E-125
BRD-K77008974	WYE-354	MTOR inhibitor	1	92.05	Connected	4.03E-19
BRD-K06593056	LE-135	Retinoid receptor agonist	1	91.85	Not Connected	
BRD-K21350491	phenamil	TRPV antagonist	1	91.54	Not Connected	
BRD-K64835161	BRD-K64835161	CLK inhibitor,DYRK inhibitor	1	91.53	Connected	1.18E-04
BRD-A77299732	salubrinal	Eukaryotic translation initiation factor inhi	1	91.47	Not Connected	
BRD-K13800121	parecoxib	Cyclooxygenase inhibitor	1	90.63	Connected	5.08E-06
BRD-K67868012	PI-103	PI3K inhibitor,MTOR inhibitor	1	90.59	Connected	7.37E-144
BRD-K71879491	tretinoin	Retinoid receptor agonist	1	90.33	Not Connected	
BRD-A75409952	wortmannin	PI3K inhibitor	1	90.30	Connected	2.14E-110

Table S2: Top 20 CPs returned by the iLINCS

Perturbagen ID	Perturbagen Name	ilincsMoa	is_touchstone	median_tau	iLINCS	pValue
BRD-K02708799	GSK 1059615	PI3K inhibitor	0		Connected	0
BRD-A79768653	sirolimus	MTOR inhibitor	1	83.07	Connected	1.6E-293
BRD-K69932463	AZD-8055	MTOR inhibitor	1	93.54	Connected	1.2E-218
BRD-K12184916	NVP-BEZ235	MTOR inhibitor	1	96.02	Connected	6.9E-161
BRD-K59317601	MLN-0128	MTOR inhibitor			Connected	4.6E-144
BRD-K67868012	PI-103	PI3K inhibitor,MTOR inhibitor	1	90.59	Connected	7.4E-144
BRD-K94012289	936890-98-1	MTOR inhibitor	1		Connected	9.4E-137
BRD-K40175214	torin-1	MTOR inhibitor,PI3K inhibitor	1	71.90	Connected	1.9E-135
BRD-A45498368	WYE-125132	MTOR inhibitor	1	79.43	Connected	4.3E-127
BRD-A62025033	temsirolimus	MTOR inhibitor	1	93.49	Connected	3.5E-125
BRD-K52911425	GDC-0941	PI3K inhibitor	1	82.75	Connected	5.2E-125
BRD-A75409952	wortmannin	PI3K inhibitor	1	90.30	Connected	2.1E-110
BRD-K13154216	Everolimus	MTOR inhibitor			Connected	4.4E-109
BRD-K42898655	Temsirolimus	MTOR inhibitor			Connected	8E-104
BRD-K99023089	AZD5363	AKT inhibitor			Connected	3.2E-100
BRD-K09078998	Ridaforolimus	MTOR inhibitor			Connected	8.83E-97
BRD-K63068307	ZSTK-474	PI3K inhibitor	1	78.95	Connected	7.61E-94
BRD-K67566344	KU-0063794	MTOR inhibitor	1	97.35	Connected	3.83E-89
BRD-K72636697	SCHEMBL17052537	MTOR inhibitor	0		Connected	2.91E-86
BRD-A25736793	Everolimus	MTOR inhibitor			Connected	4.02E-79

Reference List

1 Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**, 1437-1452.e1417, doi:10.1016/j.cell.2017.10.049 (2017).