Supporting information

Kinase inhibitor data modeling and de novo inhibitor design with fragment approaches.

Michal Vieth, Jon Erickson, Jibo Wang, Yue Webster, Mary Mader, Richard Higgs, Ian Watson

Content:

Biological assay description

Biological profiling data and compound data in fingerprint format are available from authors upon request.

Table 1SP.

General statistics for models, historical hit rates and prospective validation results

Target name	historical #ACTIVES >70 %	Historical #TESTED	Historical hit rate	Prospective # tested	# actives	Prospective model validation hit rate %	Enrichment
ABL1	2044	18351	11.1	1895	597	31.5	2.8
CHK2	1778	16983	10.5	1895	597	31.5	3.0
FLT3	4108	16915	24.3	1895	1022	53.9	2.2
MET	1490	16291	9.1	1392	509	36.6	4.0
P70S6K	2229	16983	13.1	1895	693	36.6	2.8
ROCK2	2004	16732	12.0	1895	762	40.2	3.4
Any/best target	6422	19313	33.3	1895	1600	84.4	2.5

Enrichment is computed as a ratio of prospective validation hit rate and historical hit rate.

Historical numbers refer to the data available before the prospective validation.

Table 2SP

Model Performance metrics for prospective validation (1895 compounds) and libraries (179 compounds).

a) Prospective and library mean prediction values, standard deviations of predictions, Positive predicted values and sensitivities for various targets

Model	Prospective mean	Library mean	Prospective STD	Library STD	Prospective PPV	Library PPV	Prospective Sens	Library SENS
ABL	49	58	26	10	0.66	0.94	0.55	0.11
CHK2	50	58	19	10	0.76	0.58	0.38	0.15
FLT3	69	67	25	12	0.73	0.96	0.90	0.43
MET	38	58	22	14	0.70	0.61	0.30	0.21
P70S6K	53	67	21	11	0.74	0.67	0.48	0.46
ROCK2	53	45	25	12	0.77	1.00	0.63	0.01

b) Model prediction Root Mean Square errors, R2 values for prospective and library compounds. Hit rates (percentage of actives >70%) for historical, prospective validation and library compounds.

Model	Prospective RMS	Library RMS	Prospective R2	Library R2	Historical hit rate	Prospective model validation hit rate %	Library hit rate %
ABL	28	23	0.35	0.01	11	32	84
CHK2	22	28	0.40	0.05	10	32	54
FLT3	29	20	0.40	0.02	24	54	91
MET	27	36	0.30	0.00	9	37	61
P70S6K	24	25	0.42	0.02	13	37	60
ROCK2	23	24	0.54	0.39	12	40	56

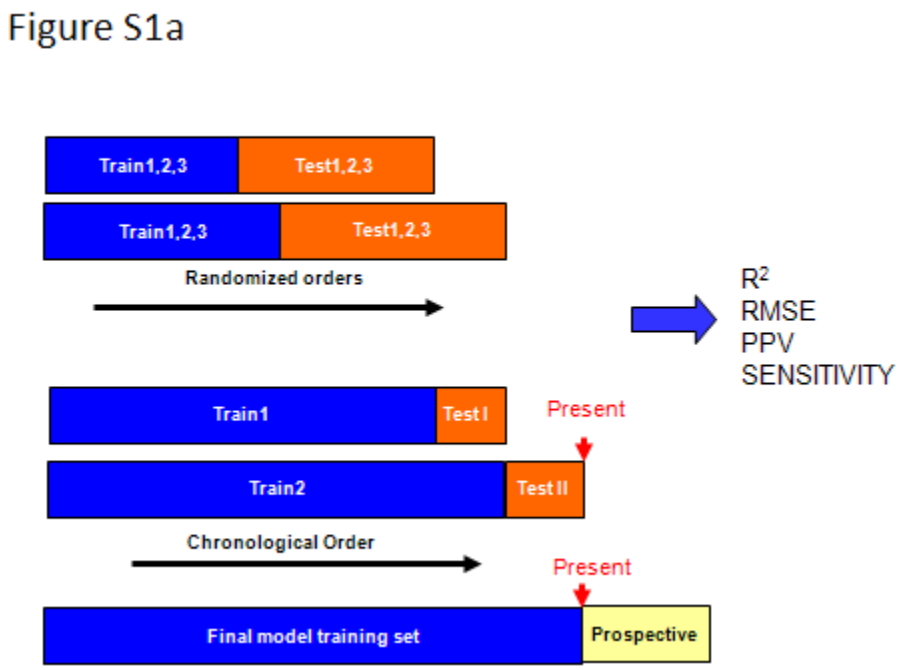
Table 3SP.

Target pair 20uM SP data correlations for compounds used in prospective studies with sequence identity between target pairs.

Target1	Target2	Experimental Prospective SP R2	Experimental Library SP R2	Predicted prospective Model SP R2	Predicted library cpds model correlation R2	Catalytic domain Seuqence ID	Alignment length	26 ATP site sequence ID
CHK2	ABL	0.15	0.05	0.19	0.40	0.183	251	0.423
FLT3	ABL	0.23	0.50	0.48	0.43	0.305	320	0.577
FLT3	CHK2	0.17	0.07	0.27	0.72	0.141	320	0.462
MET	ABL	0.16	0.10	0.42	0.21	0.377	246	0.462
MET	CHK2	0.19	0.17	0.17	0.02	0.160	246	0.346
MET	FLT3	0.11	0.24	0.34	0.06	0.287	320	0.423
P70S6K	ABL	0.04	0.16	0.01	0.41	0.198	241	0.423
P70S6K	CHK2	0.35	0.36	0.20	0.50	0.299	241	0.462
P70S6K	FLT3	0.03	0.30	0.01	0.55	0.151	320	0.500
P70S6K	MET	0.12	0.62	0.05	0.12	0.165	241	0.385
ROCK2	ABL	0.00	0.16	0.03	0.24	0.178	244	0.462
ROCK2	CHK2	0.05	0.00	0.00	0.25	0.288	244	0.423
ROCK2	FLT3	0.00	0.30	0.03	0.22	0.144	320	0.500
ROCK2	MET	0.06	0.44	0.01	0.00	0.196	244	0.538
ROCK2	P70S6K	0.24	0.37	0.26	0.43	0.328	241	0.440

Figure 1SP

a) Graphical demonstration of random split, chronological and prospective validation



b) 2x2 table demonstrating active/inactive for predicted and true/observed class

Model Hit Rate or PPV= P(truly active | predicted active)=TP/(TP+FP), where TP are true positives, FP are false positives

Sensitivity = P(predict active | truly active)=TP/(TP+FN), where FN are false negatives

Predicted Class \ True/Experimental Class	Active	Inactive
Active	TP (True Positives)	FP (False Positives)
Inactive	FN (False Negatives)	TN (True Negatives)

Biochemical assays

Recombinant protein (5-10 mU) is incubated with 0.2 mM EDTA, 8 mM MOPS having pH 7.0, 10 mM MgAcetate, 50 µM substrate, 20 uM concentration of test substance and γ-^{33}P-ATP, having concentration equal to K_m for the enzyme. The final reaction volume is 25uL. Addition of MgATP initiates the reaction, which is followed by a 40 minute incubation at room temperature. The reaction is stopped by the addition of 5 µl of a 3% phosphoric acid solution. A 10 µl sample is plotted onto a P30 filtermat and washed 3 times for 5 minutes in 75 mM phosphoric acid and once in methanol before drying and scintillation counting.