Elucidating synergistic dependencies in lung adenocarcinoma by proteome-wide signaling-network analysis

Mukesh Bansal[1,2,*#], Jing He[3,4,5,*], Michael Peyton[6*], Manjunath Kaustagi[3], Archana Iyer[3], Michael Comb[10], Michael White[6], John Minna[11], Andrea Califano[3,4,5,7,8,9,#]

[1] Psychogenics Inc. Tarrytown, New York, USA
[2] Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York
[3] Department of Systems Biology, Columbia University, New York, NY
[4] Center for Computational Biology and Bioinformatics (C2B2), Columbia University, New York, NY, USA
[5] Department of Biomedical Informatics (DBMI), Columbia University, New York, NY, USA
[6] Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA
[7] Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY
[8] Institute for Cancer Genetics, Columbia University, New York, NY
[9] Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY
[10] Cell Signaling Technology, 3 Trask Lane, Danvers, MA 01923, USA
[11] Hamon Center for Therapeutic Oncology Research, Simmons Comprehensive Cancer Center, Departments of Pharmacology, and Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas

* These authors contributed equally
[#] Correspondence should be address to M.B (mukesh.bansal@psychogenics.com); A.C. (ac2248@cumc.columbia.edu)

Summary

Signal transduction pathways are largely based on compilation of individual post-translational modification assays in heterogeneous cellular contexts. Indeed, *de novo* reconstruction of signaling interactions from large-scale molecular profiling is still lagging, compared to similar efforts in transcriptional and protein-protein interaction networks. To address this challenge, we present an novel algorithm - pARACNe - for systematic, computational reconstruction of tyrosine kinase (TK) signal transduction pathways, based on published mass spectrometry-based

proteomics profiling of phosphotyrosine-enriched peptides from 250 samples representative of lung adenocarcinoma. The network represents 43 TKs and their predicted substrates, which were validated at >60% accuracy by SILAC assays, including “novel’ substrates of the EGFR and c-MET TKs, which play a critical oncogenic role in lung cancer. Availability of the network allowed prediction of drug response in individual samples, including accurate prediction of synergistic EGFR and c-MET inhibitor activity in cells lacking mutations in either gene, which was experimentally validated, thus contributing to current precision oncology efforts.

Keywords: signaling network, lung cancer, tyrosine kinase, combination therapy, proteomics, EGFR, c-MET

41 # Introduction

42 Lung adenocarcinoma (LUAD) is a leading cause of cancer related deaths in United States,
43 representing 40% of the 225,500 new lung cancer cases every year and a 5-year survival rate of
44 only 16 % (1). Excluding immunotherapeutic agents, which have recently shown significant
45 success in a relatively small subset of patients (2), the most effective targeted therapies for this
46 diseases were designed to inhibit tyrosine kinases harboring genetic alterations inducing aberrant
47 activation of downstream signaling pathways (3-7). These events primarily include EGFR
48 mutations and ALK-EML4 fusion events, in ~15% and ~3-7% of LUAD patients, respectively (8,
49 9). Yet, while targeted therapy is initially effective in a significant fraction of patients harboring
50 these actionable alterations, the vast majority of them will either fail to respond or develop
51 resistance to mono-therapy (10, 11). In addition, most patient lack actionable alterations
52 altogether. This suggests that novel approaches are critically needed.

53 A possible alternative to minimize emergence of resistance is combination therapy, a strategy
54 that has been shown to be effective in many metastatic tumors, such as breast cancer and acute
55 myeloid leukemia (12-14). However, systematic identification of effective drug combinations on a
56 genetic alteration basis is difficult, because the number of patients presenting multiple actionable
57 events is extremely low. As a result, combination therapy is generally hypothesized and tested on
58 an empirical basis or based on elucidation of complex mechanisms of tumor cell adaptation. To
59 address these limitations, we and other have proposed that rational design of combination therapy
60 will require a more mechanistic and tumor-context-specific understanding of the molecular
61 interactions that underlie their potential synergistic activity, starting with tyrosine kinases, which
62 represent a critical class of pharmacological targets in cancer (15). Such an approach requires
63 methodologies for the accurate and systematic elucidation of tumor-specific signaling
64 transduction pathways.

65 Dissection of signal transduction networks represents a complex endeavor, requiring elucidation
66 of hundreds of thousands of tissue-specific molecular interactions that mediate the post-
67 translational modification of protein substrates. *In vitro* approaches generally fail to capture the
68 tissue-specific nature of these interactions, thus providing "average" signal transduction pathways
69 that are both incomplete and inaccurate. In addition, experimental approaches that have been
70 successful in accelerating the analysis of molecular interactions in transcriptional regulation and
71 protein-protein interaction in stable-complexes, such as those based on co-expression or yeast-
72 2-hybrid assays, do not easily translate to elucidating signaling interactions. Similarly, approaches
73 based on the use of phospho-specific antibodies, while elegant and effective, are limited to only
74 a handful of proteins. Computationally, compared to the many algorithms that have been
75 developed for the reverse engineering of transcriptional and protein-complex interactions (16, 17),
76 only a handful of experimentally validated algorithms are available for the dissection of signaling
77 networks, none of which works at the proteome-wide level or is tumor-context specific (16, 18,
78 19).

79 Recent availability of proteome-wide molecular profile data, characterizing the abundance of
80 phospho-tyrosine-enriched peptides by liquid chromatography coupled to tandem mass
81 spectrometry (LC-MS/MS), suggests that additional methodologies may be developed to extend
82 approaches that have been successfully applied to the dissection of transcriptional networks from
83 gene expression profiling. In this manuscript, we propose extending the Algorithm for the
84 Reconstruction of Accurate Cellular Networks (ARACNe) (20) for the reverse engineering of
85 signal transduction networks from large-scale phosphoproteomic profiles. The new method,
86 *pARACNe*, addresses critical issues that prevented the direct application of the original ARACNe
87 algorithm on phosphoproteomic datasets. Briefly, It overcomes the computation challenges
88 presented by LC-MS/MS and spectral counting, and incorporates enzymatic signaling
89 characteristics into the algorithm design. In particular, it has been modified to handle the highly

90 sparse nature of phosphopeptide abundance data, the large amount of noise and missing data,
91 and the degenerate peptides-to-protein mapping, as resulting from the use of LC-MS/MS-based
92 assays.

93 When used to analyze published genome-wide phosphoproteomic data from 245 lung cancer
94 samples, including fresh-frozen biopsies and cell lines for non-small-cell lung carcinoma, as well
95 as normal lung tissue, pARACNe produced a dense network connecting 46 tyrosine kinases (TK)
96 with 415 substrates (including 377 proteins without TK activity). This represents the first genome-
97 wide, tumor-context-specific model for a TK signal transduction network, capturing both protein-
98 specific and phospho-site specific events. We validated substrate predictions for two "hubs,"
99 whose activity may play a key role in determining sensitivity to Erlotinib and Crizotinib, two FDA-
100 approved drugs for LUAD, including the EGFR and c-MET tyrosine kinases by independent
101 SILAC assays and database analysis, with an accuracy of over 60%. Of particular note, the
102 pARACNe-predicted TK-substrate network provided unique information about tyrosine kinase
103 auto-phosphorylation, either direct (cis) or via a second kinase (trans). Analysis of the resulting
104 TK-network, via master regulator analysis (21), recapitulated established genetic determinants of
105 LUAD and identified predictive biomarkers of response to Erlotinib and Crizotinib combination
106 therapy, which were validated in an independent set of LUAD cell lines, the majority of which
107 harbored no genetic alterations in the corresponding genes. Furthermore, predictions based on
108 the analysis of the corresponding patient cohort were strongly supported by genomic information,
109 suggesting potential value in using these analyses for the identification of effective combination
110 therapies in precision oncology.

111

112 # Results

113 ***Overview of the pARACNe algorithm***

114 Enzymatic activity of tyrosine kinase (TK) proteins – as assessed by the ability to phosphorylate
115 their downstream substrates – is effectively determined by their phosphorylated isoform
116 abundance (**Fig. 1A, B**). Therefore, we reasoned that computational inference of TK-substrate
117 interactions (**TK→S**) could be effectively performed by measuring the dependency between their
118 respective phospho-states, as measured over a large compendium of samples **(Fig. 1C)**.
119 Unfortunately, due to signal transduction cascade complexity and pathway cross-talk, such
120 dependencies can manifest between protein pairs that are not involved in direct **TK→S**
121 interactions. The ARACNe algorithm, which we had previously designed for the reverse
122 engineering of transcriptional networks, effectively addresses this problem by leveraging the Data
123 Processing Inequality (22), a critical property of mutual information (23) , that effectively allows
124 disambiguating between direct and indirect interactions. This is accomplished by assessing
125 whether the information transfer on a candidate direct interaction (e.g., $\mathbf{TK_1 \rightarrow S}$) is greater than
126 on all other indirect paths (e.g., $\mathbf{TK_1 \rightarrow TK_2 \rightarrow S}$). ARACNe has been highly successful in the
127 experimentally validated dissection of transcriptional networks via analysis of large gene
128 expression profile compendia. ARACNe-inferred targets of transcription factors were validated in
129 multiple cellular contexts, with an accuracy of 70% to 80% (20, 24-27).

130 However, ARACNe relies on molecular profile data that is both continuous and non-sparse,
131 properties that are not always provided by quantitative proteomic data sets, which can be
132 generated by a variety of methods. Those based on LC-MS/MS represent the most popular
133 approaches (28), but different implementations have specific performance profiles in terms of
134 analyte throughput, consistency of measurement of peptides across samples and linear dynamic
135 range (29). Depending on the data acquisition method, one or both of these assumptions of

136 ARACNe are violated in proteome-wide datasets generated by the most popular methods based
137 on data-dependent acquisition. Particularly when employing quantification by spectral counting,
138 as is typically conducted for global protein-protein interaction studies (30, 31), phosphoproteomic
139 data is both discrete (i.e., generally represented by spectral counts), very sparse, with a majority
140 of peptides having zero spectral counts, and presenting a significantly skewed distribution for low-
141 abundance peptides.

142

143 To address these limitations, we developed an entirely novel algorithm, pARACNe (phospho-
144 ARACNe) (**Fig. 1C**), specifically designed to measure phospho-state dependencies between TKs
145 and their candidate substrates using published large-scale LC-MS/MS phosphoproteomic
146 profiles. pARACNe infers signaling cascades between each TK and all their cognate substrates
147 based on whether they are collectively phosphorylated under various cellular perturbations.
148 pARACNe thus extends the original ARACNe framework to allow systematic inference of **TK→S**
149 interactions. Specifically, to handle the highly discrete nature of the data, pARACNe replaces the
150 kernel-density and adaptive partitioning based mutual information estimators with the naive
151 method (**Fig 1C4**), using gold standard data to select the most effective number of bins [12] (see
152 Methods). Furthermore, to deal with the skewed spectral count distribution, we introduce an
153 iterative quantile discretization method, where samples are binned together, based on their
154 spectral counts, to produce a distribution as close to uniform as possible (**Fig 1C3,** Methods).

155

156 ***pARACNe-inferred LUAD-specific TK-phosphorylation network***

157 We used pARACNe to reconstruct a LUAD-specific TK-signaling network, by analyzing
158 phosphopeptide profiles obtained from 245 LUAD samples from Guo et al. (32). These data
159 represent the abundance of peptides containing at least one phospho-tyrosine, as obtained by

160 phosphoproteomic analysis of 46 LUAD cell lines, 151 LUAD tumors, and 48 adjacent normal
161 samples. LC-MS/MS profiling identified 3920 phospho-tyrosine residues on approximately 2600
162 different proteins. Based on these data, pARACNE identified a total of 2611 peptide-peptide
163 interactions, which could be further mapped to 2064 unique **TK→S** interactions (**Table S1, S2)**.
164 The latter represent interactions between 46 distinct TKs and their candidate substrates, including
165 174 $\mathbf{TK_1 \rightarrow TK_2}$ interactions between two TKs (**Fig 2A**), which represents a statistically significant
166 bias toward TK-TK interactions in the network ($p = 10^{-62}$). This suggests that TKs may form a more
167 densely inter-connected subnetwork within the complete TK signaling network than previously
168 assessed, thus potentially providing critical novel information about adaptive response, pathway
169 cross-talk, and auto-regulatory loops. Indeed, such highly interconnected structure provides
170 potential functional advantage compared to less connected or "flat" architectures, including the
171 ability to provide more fine-grain response to a highly heterogeneous variety of exogenous signals
172 and conditions, the ability to provide rapid adaptive response to changing stimuli, and the ability
173 to preserve cell state via autoregulatory feedback. Interestingly, in contrast to transcriptional
174 networks, the vast majority of pARACNe-inferred interactions have a positive Spearman
175 correlation, with higher counts of TK-mapped phosphopeptides corresponding to higher counts of
176 candidate substrate-mapped ones. This is consistent with the fact that, in contrast to
177 transcriptional networks where interactions may activate or suppress expression of a target gene,
178 TKs only phosphorylate their substrates, thus inducing positive phospho-state correlation. Only a
179 negligible number of interactions (0.5%) were associated with a negative Spearman correlation
180 ($N = 11$, $p \leqslant 0.05$). These may represent either indirect interactions where the TK activates a
181 substrate-specific phosphatase or direct interactions where phosphorylation of one phosphosite
182 may prevent phosphorylation of another site on the same protein.

183

184 ***LUAD Network Accuracy and Sensitivity Analysis***

185 To estimate the accuracy of the inferred TK-signaling network, we investigated the substrates of
186 two TK-proteins, EGFR and c-MET, representing high-affinity binding targets of existing FDA-
187 approved TK inhibitors for LUAD. Specifically, we compared their pARACNe-inferred substrates
188 to those reported in the phosphoDB database (33) and those supported by experimental
189 evidence, based on previously published SILAC assays, following cell line treatment with
190 associated, selective TK inhibitors. pARACNe inferred 123 EGFR substrates (**Fig 2B**). Of these,
191 5 (blue and cyan) were included as high-confidence EGFR substrates in phosphoDB, out of 13 in
192 total (38%), including the established EGFR auto-phosphorylation site. Moreover, 50 additional
193 proteins (45%, green) showed significant decrease (at least 2 fold) in the abundance of their
194 phosphorylated isoforms in SILAC assays (32), following treatment of H3255 cells with the EGFR
195 inhibitor Gefitinib. Similarly, pARACNe predicted 179 c-MET substrates (**Fig 2C)**. Notably, both
196 of the established substrates reported in PhosphoDB were identified by pARACNE (100%, blue).
197 Moreover, 126 additional proteins (71.5%, blue) showed significant decrease in the abundance
198 of their phosphorylated isoforms in SILAC assays(32), following treatment of MKN45 cells with
199 the first-generation c-MET-specific inhibitor Su11274.

200 We used MKN45, even though it represents a gastric cancer cell line, to assess the overall
201 conservation of predictions in different tissues. Indeed, while specific substrates are likely to be
202 highly tumor-context specific, signaling networks should be significantly more conserved than
203 transcriptional ones, since lineage-specific chromatin states represent a major determinant of
204 transcriptional regulation, while signal transduction is only affected indirectly by the availability of
205 specific partner proteins. As a result, it is reasonable to expect that an even greater overlap of
206 inferred vs. SILAC positive substrates may be achieved in native LUAD cells.

207

208 Taken together, these data suggest that pARACNe can identify a much larger subset of candidate
209 substrates, while both identifying a significant proportion of established substrates (46% on
210 average, based on phosphoDB) and maintaining high accuracy (~60% on average, by SILAC
211 assays). This also suggest that, similar to transcription factor targets reported in the literature, TK
212 substrates are poorly characterized, even for highly relevant and exceedingly well-studied kinases
213 such as EGFR and c-MET. As a result, pARACNe could provide significant novel hypotheses for
214 TK→S interactions that can be validated as required. We should also note that the reported
215 accuracy for pARACNe is estimated using SILAC data on a single cell line. SILAC assays have
216 significant false negatives and it would be reasonable to expect that, once tested in additional cell
217 lines, the accuracy of pARACNe could further increase. As a further performance benchmark, we
218 used the same SILAC benchmark to test predictions by NetworkIN, a reverse engineering method
219 based on protein sequence motif analysis and protein association networks (16). The analysis
220 found almost no consensus with SILAC assays, with only one out of 33 NetworkIN-predicted
221 EGFR substrate identified as significantly dephosphorylated following treatment with TK-specific
222 inhibitors.

223

224 ***Systematic, network-based inference of pharmacological dependencies***

225 Once an accurate model of signal transduction in LUAD cells was established by pARACNe
226 analysis, we interrogated the corresponding TK→S network using phosphoproteomic signatures
227 from 46 LUAD cell lines to identify key dependencies for experimental validation. For this purpose,
228 we extended the Virtual Proteomics by Enriched Regulon analysis (VIPER)(34), which had been
229 originally developed to identify key proteins controlling the transcriptional state of tumor-related
230 cells, on an individual sample basis. Specifically, VIPER was designed to identify these proteins
231 by assessing the enrichment of their transcriptional targets in differentially expressed genes, in a
232 tumor-specific gene expression signature. VIPER, which extends the Master Regulator Inference

233 algorithm (MARINa) to support analysis of single samples, has been instrumental in identifying
234 master regulator proteins representing key functional determinants of tumor-related phenotypes
235 in many cancer types, from glioblastoma (26, 27, 35), lymphoma (36, 37), and leukemia (38) to
236 prostate (39-41) and breast adenocarcinoma(42-44), among others. We thus reasoned that
237 VIPER could be modified to identify master regulator TKs, most likely to induce the differential
238 phosphorylation pattern observed in a specific tumor sample (**Fig. 3A**). A specific additional value
239 of the algorithm is that it could not only identify individual MR TK proteins, representing individual,
240 pharmacologically accessible dependencies of the tumor, but also TKs representing potential
241 synergistic MR-pair as candidate dependencies for combination therapy.

242 VIPER can be easily modified to analyze phosphoproteomic signatures (pVIPER). Specifically,
243 rather than assessing the enrichment of a protein's transcriptional targets (regulon) in differentially
244 expressed genes, pVIPER is designed to measure the enrichment of a TK's substrates (signalon)
245 in differentially phosphorylated proteins. We performed pVIPER analysis at phosphopeptide
246 specific level, where dependencies were assessed based on individual phosphopeptide ather
247 than whole protein state, and were finally combined the result of all phosphopeptide mapping to
248 the same protein. Consistent with VIPER's experimentally validated ability to identify synergistic
249 master regulators proteins by transcriptional network analysis (26, 37, 39), pVIPER can identify
250 candidate synergistic TKs by assessing the statistical significance of the enrichment of the
251 common substrates of both TKs over the enrichment of unique substrates of each individual TK
252 (see Method section). Systematic VIPER analysis of phosphoproteomic profiles from 46 LUAD
253 cell lines generated between 2 and 13 master regulator TKs or synergistic TK-pairs, as candidate

254 pharmacologically actionable dependencies, for each cell line, thus generating a plausible number
255 of hypothesis for each line (**Fig. 3B** and **Fig. 3C**).

256

257 **pVIPER identifies LUAD-specific dependencies**

258 pVIPER analysis inferred several TK proteins as highly conserved individual dependencies across
259 multiple cell lines, including the Ephrin type-A receptor 2 (*EPHA2*), epidermal growth factor
260 receptor (*EGFR*), c-Met proto-oncogene (*MET*), and HER2 receptor tyrosine kinase 2 (*ERBB2*),
261 suggesting a critical role of these proteins in the maintenance of LUAD cell line state. This is also
262 in agreement with the functional role of these genes and the use of inhibitors of these kinases
263 across a large panel of patients in multiple cancer types (45-49).

264 In contrast to these established LUAD cell line dependencies, we also identified several TKs as
265 dependencies of specific cell lines. This can either be the result of associated genetic or
266 epigenetic alterations in these cell lines or the result of field effects, where multiple genetic
267 alterations or alterations in upstream pathway contribute to the cell line dependency on a specific
268 TK activity. For instance, we identified ALK (Anaplastic Lymphoma Receptor Tyrosine Kinase) to
269 be addiction point only in H2228 cell line. ALK is a conserved trans-membrane receptor tyrosine
270 kinase (RTK) protein in the insulin-receptor super family. Chromosomal alterations involving *ALK*
271 translocations and fusion events have been identified in several cancer types including LUAD (50,
272 51), diffuse large B-cell lymphomas (52), neuroblastoma (53), and inflammatory myofibroblastic
273 tumors (54), among others. Additionally, ALK fusion events with other genes, including EML4
274 (Echinoderm Microtubule-associated protein Like 4) in LUAD lead to aberrant protein activity
275 eliciting “oncogene addiction” (51). Presence of ALK-EML4 fusion transcripts, in ~3–7% of LUAD
276 patients (55-57), is a strong predictor of response to ALK inhibitors, such as Crizotinib, among
277 others (58, 59). Interestingly, among all available LUAD cell lines for which a phosphoproteomic

278 profile was available, H2228 was the only one with an established ALK-EML4 fusion event and
279 with established sensitivity to ALK inhibitor (60, 61). This further reflects the specificity of our
280 analysis as this was the only cell line predicted to depend on ALK activity. Interestingly, we
281 identified 4 additional H2228 dependencies, namely EGFR, Epha2, c-MET, and PTK2. H2228
282 sensitivity to EGFR inhibitors, in combination with ALK inhibitors, was already established (61).
283 In addition, we confirmed dependency of H2228 cells on the other 3 TKs (Epha2, c-MET and
284 PTK2) by cell viability assay following siRNA-mediated silencing of the corresponding gene (**Fig**
285 **S2A**). BrdU assays confirmed H2228 sensitivity to c-MET and PTK2 silencing (**Fig S2B).**
286 Surprisingly, in contrast to inhibition by small molecule inhibitor, siRNA-mediated EGFR silencing
287 did not show decrease in cell viability, while ALK silencing showed little effect. This may be due
288 to the lower degree of siRNA-mediated inhibition, compared to specific inhibitors, or to the multi-
289 kinase effect of the latter. For instance, consistent with their identification as cell line specific
290 dependencies, Crizotinib is both an ALK and a c-MET inhibitor. Indeed, although siRNA-mediated
291 c-MET inhibition was modest, BrdU incorporation assays showed significant effect on cell viability.
292

293 ***EGFR and c-MET are predicted as dependencies in multiple LUAD cell lines***

294 As discussed, pVIPER analysis revealed several TK-pairs as candidate synergistic dependencies
295 across several cell lines, such as Epha2/c-MET, EGFR/PTK2, EGFR/Epha2, Epha2/c-MET, and
296 EGFR/c-MET. Among these the EGFR/c-MET pair emerged as the most conserved synergistic
297 TK-pair across the available cell lines. In addition, EGFR and c-MET were also identified as
298 candidate TK MRs in several of these cell lines, suggesting either a complementary or synergistic
299 role for these proteins and a potential therapeutic opportunity for combination therapy in LUAD
300 (62, 63).
301

302 To validate pVIPER-predicted, cell line specific EGFR/c-MET synthetic lethality, we selected a
303 panel of 14 cell lines, 11 of which were predicted to be synergistically dependent on EGFR/c-MET

304 (H226, H2122, H1666, H2172, Cal-12T, H2023, H1568, Calu-3, H1650, HCC78, and A549), as
305 well as 3 negative controls with no predicted synergistic or individual dependencies on the two
306 TKs (H2170, H460, and H520). To measure sensitivity to these agents, we used two different and
307 complementary assays, including: (a) colony formation assay to assess long term sensitivity (**Fig.**
308 **4A and Methods**) and (b) 3-[4,5-dimethylthiazol-2-yl]-2,5 diphenyl tetrazolium bromide (MTS)
309 assay for short term sensitivity analysis (**Fig. 5A, Table S3 and Methods**). For colony formation
310 assays, cells were treated with either an EGFR inhibitor (Erlotinib, 1uM) or a c-MET/ALK inhibitor
311 (Crizotinib, 0.1uM), either individually or in combination (see Methods). To evaluate synergistic
312 dependency on EGFR/c-MET we used Excess Over Bliss (64), which measures the difference
313 between the observed effect on colony formation and the effect expected from a purely additive
314 model. For MTT assay, first, cells were treated with EGFR inhibitor (Erlotinib) or MET inhibitor
315 (Crizotinib) individually at various concentrations to identify IC50 (concentration resulting in 50%
316 cell death). Next, cells were treated with 1 uM of Erlotinib and varying concentrations of Crizotinib
317 to identify combinations resulting in IC50 and used the combination index (CI) statistic (65) to
318 measure interaction between the two drugs.
319
320 Across all 11 cell lines tested in colony formation assay, 8 showed significant sensitivity to either
321 individual inhibitors ($H226_{E,C}$, $H2122_{E}$, $H1666_{E,C}$, $Cal\text{-}12T_{E}$, $Calu\text{-}3_{E,C}$, $H1650_{E}$) or synergistic
322 sensitivity to the combination ($HCC78_{E+C}$ and $H2023_{E+C}$) (**Fig 4B-C)**. Surprisingly, all of these cell
323 lines were $EGFR^{WT}$, ALK^{WT}, and $c\text{-}MET^{WT}$, except for H1650, which was $EGFR^{Mut}$. Thus, based
324 on standard of care criteria, 7 out of 8 cell lines would not have been considered as sensitive to
325 either EGFR or ALK/c-MET inhibitors. Several cell lines presented striking sensitivity to either one
326 ($H2122_{E}$, $Cal\text{-}12T_{E}$, $H1650_{E}$) or both inhibitors ($H226_{E,C}$, $H1666_{E,C}$, $Calu\text{-}3_{E,C}$) in isolation, thus
327 making the assessment of synergistic drug sensitivity difficult. In addition, three $EGFR^{WT}$ cell lines
328 harboring BRAF (Cal-12T and H1666) or KRAS (H2122) mutations were also highly sensitive to
329 Erlotinib as a single agent, as predicted by pVIPER, despite the fact that KRAS pathway mutations

330 are mutually exclusive with EGFR mutations and predictive of Erlotinib resistance (**Fig. 4C**).
331 Finally, none of these cell lines was predicted to be sensitive to ALK inhibitors, suggesting that
332 Crizotinib sensitivity is mediated by c-MET specific dependencies. Of the negative controls, only
333 one (H2170) showed high sensitivity to Erlotinib. Taken together, 8/11 cell lines (73%) predicted
334 as sensitive to the inhibitors were validated long term colony formation assays, while only 1/3
335 negative controls showed sensitivity to these agents (33%).

336

337 To evaluate the short-term interaction between EGFR and c-MET, we performed MTT assay
338 across 11 cell lines (HCC78, H2023, H1650, Calu-3, H2172, H2122, H1568, A549, H1666, H520
339 and H2170) including 2 negative control cell lines (H520 and H2170). Similar to colony formation
340 assay, we found synergistic sensitivity to EGFR and c-MET inhibitors in 6/9 cell lines (67%), with
341 5 cell lines showing strong synergy (CI ≤ 0.8) (**Fig. 5B-C**) and 1 borderline synergy (CI = 0.82),
342 showing the consistency between two assays. However, for two cell lines, H1666 and H2170 (a
343 negative control), results were inconsistent between long term colony formation and MTT assays.
344 For both H1666 and H2170 cell lines, colony formation and MTT assay to showed sensitivity to
345 Erlotinib alone, where colony formation assay has complete abrogation of colonies at 1 μM of
346 Erlotinib, and later had IC_{50} =1.25 μM and 3.7 μM for H1666 and H2170 respectively. However,
347 in combination therapy, MTT assay showed antagonism (CI >1), despite the fact that colony
348 formation assay still showed complete abrogation which could be associated to the accumulation
349 of new mutations in these cell lines. However, this is just hypotheses and needs to be verified by
350 further experiments such as sequencing of these cell lines pre-and post-treatment.

351

352 ***Phosphosite specific phosphorylation predicts EGFR/c-MET inhibitor synergy.***

353 In previous section, we assessed the pVIPER predictions after consolidating the result at protein
354 level. Following the results from MTT and colony formation assays, we reanalyzed the pVIPER

355 predictions at the phosphopeptide level. Interestingly, this revealed that whenever synergistic
356 EGFR/c-MET dependencies were predicted from phosphosite $EGFR_{1197}$ and phosphosites other
357 than $c\text{-}MET_{1003}$ (H1666, Cal-12T, H1650), cell lines responded to Erlotinib in isolation, while when
358 predictions were based on phosphosites $EGFR_{1197}$ and $c\text{-}MET_{1003}$ (HCC78, H2023, and Calu-3),
359 cells exhibited bona fide synergistic sensitivity to the two inhibitors, with the only possible
360 exception of Calu-3, which showed synergistic sensitivity in MTT assays and additive sensitivity
361 to both inhibitors in colony formation assays. Conversely, when predictions were not based on
362 either phosphotyrosine, cells exhibited no sensitivity to the individual inhibitors or the combination
363 (H2172, H226, A549, H460, H520, H1568). Thus, predictions based on these two phosphosites
364 produced no false positives (6 out of 6 predicted and validated as non-sensitive) and only 2 false
365 negatives (H2170 and H2122), resulting in an error rate of only 2 out of 14 cell lines (14%, $p =$
366 0.0093 using fisher exact test).
367
368 This finding is in agreement with the established role of $EGFR_{1197}$ as a predictor of EGFR inhibitor
369 sensitivity (66). Intriguingly, when sensitivity was predicted using phosphosites other than
370 $EGFR_{1197}$ and $c\text{-}MET_{1003}$, cell lines did not respond to the inhibitors, either individually or in
371 combination. For these two peptides, we found their common substrates to be hyper
372 phosphorylated in the sensitive cell lines **(Fig. 6A**) compared to the specific substrates of each of
373 them, whereas cell line responding only to EGFR inhibitors showed more hyper phosphorylation
374 of EGFR only substrates (**Fig. 6B**). Cell lines resistant to both EGFR and c-MET inhibitors either
375 showed no change in the phosphorylation status or hypo-phosphorylation compared to the normal
376 samples (**Fig. 6C**). Hence, either the common substrates of EGFR and c-MET, or the
377 phosphorylation status of $EGFR_{1197}$ and $c\text{-}MET_{1003}$ could potentially be used as biomarkers for
378 predicting therapy with the dual inhibitors. However, this conclusion is based on a very limited
379 number of observations and lacks the statistical power. This finding needs a re-

380 evaluation/validation using larger cohort of samples to establish an appropriate biomarker for
381 combination therapy.
382
383 ***Addictions points in patient samples***
384 Similar to cell lines, when applied on patient data (32), pVIPER identified EGFR as one of the
385 most common addiction points (**Fig. 6**). We inferred EGFR dependency in 12 patients. Of these,
386 5 harbored EGFR mutation, while the remaining 7 patients had not been tested for this mutation,
387 showing a high consistency between our predictions and the genetic predisposition for sensitivity
388 to EGFR inhibitors. In the entire cohort, there were only 3 patients with EGFR mutations that were
389 not identified as EGFR dependent by pVIPER, resulting in an overall sensitivity of 62.5% (5/8).
390 However, it is well known that >50% of patients harboring EGFR mutations do not respond to
391 Erlotinib, suggesting that these may not be false negatives but rather patients with low activation
392 of downstream EGFR pathways, despite their $EGFR^{Mut}$ state. Similarly, our analysis identified
393 candidate ALK dependencies in 4 patients. Of these one had an established TFG-ALK fusion,
394 whereas the others had not been tested for ALK fusion events.

395 Across all patient samples, we observed Discoidin Domain Receptor-1 (DDR1) to be the most
396 frequent addiction point, which was not predicted for any of the 46 cell lines. One reason for the
397 difference is that DDR1 is collagen dependent and there may be differences in the 3D structure
398 of the tumor and the cell lines growing on the plate. An independent study (67) in a cohort of 83
399 lung cancer specimens found that silencing of DDR1 in these samples leads to the hampering of
400 cell survival, reduced invasiveness in collagen matrices, increased apoptosis in basal condition
401 and decreased metastatic activity in model of tumor metastasis to bone, signifying it as a potential
402 novel therapeutic target.

403 ## Discussion

404 In this paper, we developed pARACNe to infer Tyrosine Kinase (TK) signaling network using
405 published genome-wide phosphoproteomic data from lung cancer. The network prediction was
406 validated using SILAC experiments, with high accuracy. Interrogation of the predicted TK-
407 substrate network generated biologically meaningful hypotheses, followed by experimental
408 validations illustrating the effectiveness of predicted kinase inhibitor combination, EGFR and c-
409 MET combination inhibitors, in treating lung cancer cell lines. Furthermore, Master Regulator
410 Analysis using patient proteomics data provides implications for using targeted agent
411 combinations to treat patients based on patient proteomic data.

412

413 Notably, pARACNe is significant and powerful as of its genome-wide scale and context-specificity
414 in discovering global signaling cascading relationships, which were missing by previous methods.
415 For example, methods proposed by Linding et al. (16) combine motif-based phospho-site
416 predictions with information of physical association, co-occurrence, and co-expression to identify
417 substrates with high specificity and accuracy, but with low coverage and lack of contextual
418 specificity. Bender et al (17) used reverse phase protein assay data after various stimulations to
419 cells and inferred signaling network using hidden Markov models and genetic algorithms. Even
420 though the resulting networks are context specific, they lack genomic-scale coverage. There have
421 been methods which used existing large-scale protein networks and prune them using
422 transcriptomic information to identify signaling pathways (68-70). In addition, attempts have been
423 made to reconstruct signaling network using gene expression data (71, 72). However, as signaling
424 complexity lies mostly in upper level of cellular processes, inferring the cascades from
425 downstream gene expression fails to capture all the dynamics. Also, PrePPI proposed by Zhang
426 et al. (73) used protein structure-based methods to infer global protein-protein interaction, but this
427 approach fails to address phosphorylation context specificity. Innovative uses of multiplex and

428 microarray-based approaches, where multiple antibodies can be used to probe an ensemble of
429 phosphoproteins, are finally becoming sufficiently mature to allow characterization of small
430 pathways. Yet, these methods are still far from providing an unbiased, genome-wide view of
431 signal-transduction processes and continue to be completely dependent on antibody specificity
432 and availability. Similarly, assays developed specifically to monitor phosphorylation pathways,
433 such as Stable Isotope Labeling with Amino acids in Cell culture (SILAC), provides a simple and
434 straightforward approach to detect differential protein abundance. Coupled with phosphorylation
435 enriched assays, it can provide high quality quantification for post-translation phosphorylation
436 changes in cell lines. However, these methods are 1) laborious and costly; 2) can only be
437 performed to dissect the substrates of a single enzyme at a time and 3) do not differentiate
438 between direct and indirect targets.
439
440 To be noted, as the LC-MS/MS experiments used here was generated based on Tyrosine-kinase
441 enrichment, which is only about ~2% of whole phosphoproteome. pARACNe is shown only on
442 TK-substrates network. The current methodology could be extended to whole phosphoproteomic
443 data based signaling network reconstruction where the data is available. In addition to label free
444 based LC-MS/MS proteomics data used in this work, label based approaches, such as ITRAQ or
445 TMT, could generate higher throughput whole proteomic profiles which might require future
446 redesign of ARACNe to incorporate both kinases and phosphatases in regulating their
447 downstream substrates. It is reasonable to expect that a version of ARACNe developed
448 specifically to dissect signaling networks should work at least as well as its transcriptional
449 counterpart. Since the relationship between the mRNA abundance of a gene encoding a
450 transcription factor (TF) and the activity of the corresponding protein is much looser than that
451 between the abundance of a phospho-isoform of a kinase and its enzymatic activity.
452
453

Even though research has attempted to identify addiction points based on gene expression data (74), predictions based on phosphoproteomic data appear superior in a way that they can directly reflect contextual specific signaling activity and are able to be directly targeted by kinase inhibitors. It is important to note that clinically, only patients with base-pair deletion at exon 19 (del746_A750) or a point mutation at exon 21 mutation (L858R) in EGFR shows sensitivity to EGFR inhibitor such as Cedirinib or Erlotinib (75).

Acknowledgments

We thank George Rosenberger for manuscript advice. This work was supported by the National Cancer Institute (NCI) Cancer Target Discovery and Development program (1U01CA168426 to A.C., and 1U01CA176284 to JDM), NCI Research Centers for Cancer Systems Biology Consortium (1U54CA209997 to A.C.), NCI Outstanding Investigator Award (R35CA197745-02 to A.C.), NCI SPORE in Lung Cancer (P50CA70907 to JDM), and CPRIT (Award RP110708 to JDM).

Author Contributions

M.B. and M.K. developed the algorithm. M.B. and J.H. performed the analysis. M.P. performed the experimental validation. A.I., M.B., M. W., J.M and A.C. guided the experiments. M.B. J.H., A.I. and A.C wrote the manuscript. M.B. and A.C. provided supervision and guidance. A.C. conceived, funded and administrated the project.

Declaration of Interests

A.C.is a founder and shareholder of DarwinHealth Inc. and a member of the Tempus Inc. SAB and shareholder. Columbia University is a shareholder of DarwinHealth Inc.

476 # References

477 1. Ettinger DS, Akerley W, Borghaei H, Chang AC, Cheney RT, Chirieac LR, et al. Non-small cell lung
478 cancer. Journal of the National Comprehensive Cancer Network : JNCCN. 2012;10(10):1236-71.
479 2. Karasaki T, Nagayama K, Kawashima M, Hiyama N, Murayama T, Kuwano H, et al. Identification of
480 Individual Cancer-Specific Somatic Mutations for Neoantigen-Based Immunotherapy of Lung Cancer. J
481 Thorac Oncol. 2016;11(3):324-33.
482 3. Pasqualucci L, Kitaura Y, Gu H, Dalla-Favera R. PKA-mediated phosphorylation regulates the
483 function of activation-induced deaminase (AID) in B cells. Proc Natl Acad Sci U S A. 2006;103(2):395-400.
484 4. Fiedler D, Braberg H, Mehta M, Chechik G, Cagney G, Mukherjee P, et al. Functional organization
485 of the S. cerevisiae phosphorylation network. Cell. 2009;136(5):952-63.
486 5. Gainor JF, Varghese AM, Ou SH, Kabraji S, Awad MM, Katayama R, et al. ALK rearrangements are
487 mutually exclusive with mutations in EGFR or KRAS: an analysis of 1,683 patients with non-small cell lung
488 cancer. Clin Cancer Res. 2013;19(15):4273-81.
489 6. Janne PA, Shaw AT, Pereira JR, Jeannin G, Vansteenkiste J, Barrios C, et al. Selumetinib plus
490 docetaxel for KRAS-mutant advanced non-small-cell lung cancer: a randomised, multicentre, placebo-
491 controlled, phase 2 study. Lancet Oncol. 2013;14(1):38-47.
492 7. Mak KS, Gainor JF, Niemierko A, Oh KS, Willers H, Choi NC, et al. Significance of targeted therapy
493 and genetic alterations in EGFR, ALK, or KRAS on survival in patients with non-small cell lung cancer treated
494 with radiotherapy for brain metastases. Neuro Oncol. 2014.
495 8. Vendrell JA, Taviaux S, Beganton B, Godreuil S, Audran P, Grand D, et al. Detection of known and
496 novel ALK fusion transcripts in lung cancer patients using next-generation sequencing approaches. Sci Rep.
497 2017;7(1):12510.
498 9. Peters S, Camidge DR, Shaw AT, Gadgeel S, Ahn JS, Kim DW, et al. Alectinib versus Crizotinib in
499 Untreated ALK-Positive Non-Small-Cell Lung Cancer. N Engl J Med. 2017;377(9):829-38.
500 10. Tsvetkova E, Goss GD. Drug resistance and its significance for treatment decisions in non-small-
501 cell lung cancer. Curr Oncol. 2012;19(Suppl 1):S45-51.
502 11. Wilson FH, Johannessen CM, Piccioni F, Tamayo P, Kim JW, Van Allen EM, et al. A functional
503 landscape of resistance to ALK inhibition in lung cancer. Cancer Cell. 2015;27(3):397-408.
504 12. O'Shaughnessy J, Miles D, Vukelja S, Moiseyenko V, Ayoub JP, Cervantes G, et al. Superior survival
505 with capecitabine plus docetaxel combination therapy in anthracycline-pretreated patients with advanced
506 breast cancer: phase III trial results. J Clin Oncol. 2002;20(12):2812-23.
507 13. Pegram MD, Slamon DJ. Combination therapy with trastuzumab (Herceptin) and cisplatin for
508 chemoresistant metastatic breast cancer: evidence for receptor-enhanced chemosensitivity. Semin Oncol.
509 1999;26(4 Suppl 12):89-95.
510 14. Ravandi F, Cortes JE, Jones D, Faderl S, Garcia-Manero G, Konopleva MY, et al. Phase I/II study of
511 combination therapy with sorafenib, idarubicin, and cytarabine in younger patients with acute myeloid
512 leukemia. J Clin Oncol. 2010;28(11):1856-62.
513 15. Wu P, Nielsen TE, Clausen MH. Small-molecule kinase inhibitors: an analysis of FDA-approved
514 drugs. Drug Discov Today. 2016;21(1):5-10.
515 16. Linding R, Jensen LJ, Ostheimer GJ, van Vugt MA, Jorgensen C, Miron IM, et al. Systematic
516 discovery of in vivo phosphorylation networks. Cell. 2007;129(7):1415-26.
517 17. Bender C, Henjes F, Frohlich H, Wiemann S, Korf U, Beissbarth T. Dynamic deterministic effects
518 propagation networks: learning signalling pathways from longitudinal protein array data. Bioinformatics.
519 2010;26(18):i596-602.

520 18. Papin JA, Hunter T, Palsson BO, Subramaniam S. Reconstruction of cellular signalling networks and
521 analysis of their properties. Nature reviews. 2005;6(2):99-111.
522 19. Hashemikhabir S, Ayaz ES, Kavurucu Y, Can T, Kahveci T. Large-scale signaling network
523 reconstruction. IEEE/ACM Trans Comput Biol Bioinform. 2012;9(6):1696-708.
524 20. Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, Califano A. Reverse engineering of
525 regulatory networks in human B cells. Nat Genet. 2005;37(4):382-90.
526 21. Lefebvre C, Rajbhandari P, Alvarez MJ, Bandaru P, Lim WK, Sato M, et al. A human B-cell
527 interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. Mol Syst
528 Biol. 2010;6:377.
529 22. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, et al. ARACNE: an
530 algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC
531 Bioinformatics. 2006;7 Suppl 1:S7.
532 23. MacKay DJC. Information theory, inference, and learning algorithms. Cambridge, UK ; New York:
533 Cambridge University Press; 2003. xii, 628 p. p.
534 24. Basso K, Saito M, Sumazin P, Margolin AA, Wang K, Lim WK, et al. Integrated biochemical and
535 computational approach identifies BCL6 direct target genes controlling multiple pathways in normal
536 germinal center B cells. Blood. 2010;115(5):975-84.
537 25. Della Gatta G, Palomero T, Perez-Garcia A, Ambesi-Impiombato A, Bansal M, Carpenter ZW, et al.
538 Reverse engineering of TLX oncogenic transcriptional networks identifies RUNX1 as tumor suppressor in
539 T-ALL. Nat Med. 2012;18(3):436-40.
540 26. Carro MS, Lim WK, Alvarez MJ, Bollo RJ, Zhao X, Snyder EY, et al. The transcriptional network for
541 mesenchymal transformation of brain tumours. Nature. 2010;463(7279):318-25.
542 27. Sonabend AM, Bansal M, Guarnieri P, Lei L, Amendolara B, Soderquist C, et al. The transcriptional
543 regulatory network of proneural glioma determines the genetic alterations selected during tumor
544 progression. Cancer Res. 2014;74(5):1440-51.
545 28. Aebersold R, Mann M. Mass-spectrometric exploration of proteome structure and function.
546 Nature. 2016;537(7620):347-55.
547 29. Gillet LC, Leitner A, Aebersold R. Mass Spectrometry Applied to Bottom-Up Proteomics: Entering
548 the High-Throughput Era for Hypothesis Testing. Annu Rev Anal Chem (Palo Alto Calif). 2016;9(1):449-72.
549 30. Huttlin EL, Ting L, Bruckner RJ, Gebreab F, Gygi MP, Szpyt J, et al. The BioPlex Network: A
550 Systematic Exploration of the Human Interactome. Cell. 2015;162(2):425-40.
551 31. Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, et al. Architecture of the human
552 interactome defines protein communities and disease networks. Nature. 2017;545(7655):505-9.
553 32. Guo A, Villen J, Kornhauser J, Lee KA, Stokes MP, Rikova K, et al. Signaling networks assembled by
554 oncogenic EGFR and c-Met. Proc Natl Acad Sci U S A. 2008;105(2):692-7.
555 33. Giansanti P, Aye TT, van den Toorn H, Peng M, van Breukelen B, Heck AJ. An Augmented Multiple-
556 Protease-Based Human Phosphopeptide Atlas. Cell Rep. 2015;11(11):1834-43.
557 34. Alvarez MJ, Shen Y, Giorgi FM, Lachmann A, Ding BB, Ye BH, et al. Functional characterization of
558 somatic mutations in cancer using network-based inference of protein activity. Nat Genet.
559 2016;48(8):838-47.
560 35. Chudnovsky Y, Kim D, Zheng SY, Whyte WA, Bansal M, Bray MA, et al. ZFHX4 Interacts with the
561 NuRD Core Member CHD4 and Regulates the Glioblastoma Tumor-Initiating Cell State. Cell Rep.
562 2014;6(2):313-24.
563 36. Ying CY, Dominguez-Sola D, Fabi M, Lorenz IC, Hussein S, Bansal M, et al. MEF2B mutations lead
564 to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. Nat Immunol.
565 2013;14(10):1084-+.

566 37. Bisikirska B, Bansal M, Shen Y, Teruya-Feldstein J, Chaganti R, Califano A. Elucidation and
567 Pharmacological Targeting of Novel Molecular Drivers of Follicular Lymphoma Progression. Cancer Res.
568 2016;76(3):664-74.
569 38. Piovan E, Yu J, Tosello V, Herranz D, Ambesi-Impiombato A, Da Silva AC, et al. Direct reversal of
570 glucocorticoid resistance by AKT inhibition in acute lymphoblastic leukemia. Cancer Cell. 2013;24(6):766-
571 76.
572 39. Aytes A, Mitrofanova A, Lefebvre C, Alvarez MJ, Castillo-Martin M, Zheng T, et al. Cross-species
573 regulatory network analysis identifies a synergistic interaction between FOXM1 and CENPF that drives
574 prostate cancer malignancy. Cancer Cell. 2014;25(5):638-51.
575 40. Mitrofanova A, Aytes A, Shen C, Abate-Shen C, Califano A. A systems biology approach to predict
576 drug response for human prostate cancer based on in vivo preclinical analyses of mouse models. Cell
577 Reports. 2015;12:1–12.
578 41. Talos F, Mitrofanova A, Bergren SK, Califano A, Shen MM. A computational systems approach
579 identifies synergistic specification genes that facilitate lineage conversion to prostate tissue. Nature
580 communications. 2017;8:14662.
581 42. Walsh LA, Alvarez MJ, Sabio EY, Reyngold M, Makarov V, Mukherjee S, et al. An Integrated Systems
582 Biology Approach Identifies TRIM25 as a Key Determinant of Breast Cancer Metastasis. Cell Rep.
583 2017;20(7):1623-40.
584 43. Rodriguez-Barrueco R, Yu J, Saucedo-Cuevas LP, Olivan M, Llobet-Navas D, Putcha P, et al.
585 Inhibition of the autocrine IL-6-JAK2-STAT3-calprotectin axis as targeted therapy for HR-/HER2+ breast
586 cancers. Genes Dev. 2015;29(15):1631-48.
587 44. Putcha P, Yu J, Rodriguez-Barrueco R, Saucedo-Cuevas L, Villagrasa P, Murga-Penas E, et al. HDAC6
588 activity is a non-oncogene addiction hub for inflammatory breast cancers. Breast Cancer Res.
589 2015;17(1):149.
590 45. Davoli A, Hocevar BA, Brown TL. Progression and treatment of HER2-positive breast cancer.
591 Cancer Chemother Pharmacol.65(4):611-23.
592 46. Guo L, Kozlosky CJ, Ericsson LH, Daniel TO, Cerretti DP, Johnson RS. Studies of ligand-induced site-
593 specific phosphorylation of epidermal growth factor receptor. Journal of the American Society for Mass
594 Spectrometry. 2003;14(9):1022-31.
595 47. Okamoto I. Epidermal growth factor receptor in relation to tumor development: EGFR-targeted
596 anticancer therapy. FEBS J.277(2):309-15.
597 48. Wang S, Placzek WJ, Stebbins JL, Mitra S, Noberini R, Koolpe M, et al. Novel targeted system to
598 deliver chemotherapeutic drugs to EphA2-expressing cancer cells. J Med Chem. 2012;55(5):2427-36.
599 49. Birtwistle MR, Hatakeyama M, Yumoto N, Ogunnaike BA, Hoek JB, Kholodenko BN. Ligand-
600 dependent responses of the ErbB signaling network: experimental and modeling analyses. Mol Syst Biol.
601 2007;3:144.
602 50. Rikova K, Guo A, Zeng Q, Possemato A, Yu J, Haack H, et al. Global survey of phosphotyrosine
603 signaling identifies oncogenic kinases in lung cancer. Cell. 2007;131(6):1190-203.
604 51. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the
605 transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature. 2007;448(7153):561-6.
606 52. Laurent C, Do C, Gascoyne RD, Lamant L, Ysebaert L, Laurent G, et al. Anaplastic lymphoma kinase-
607 positive diffuse large B-cell lymphoma: a rare clinicopathologic entity with poor prognosis. J Clin Oncol.
608 2009;27(25):4211-6.
609 53. Janoueix-Lerosey I, Lequin D, Brugieres L, Ribeiro A, de Pontual L, Combaret V, et al. Somatic and
610 germline activating mutations of the ALK kinase receptor in neuroblastoma. Nature. 2008;455(7215):967-
611 70.
612 54. Sokai A, Enaka M, Sokai R, Mori S, Mori S, Gunji M, et al. Pulmonary Inflammatory Myofibroblastic
613 Tumor Harboring EML4-ALK Fusion Gene. Jpn J Clin Oncol. 2014;44(1):93-6.

614 55. Bender C, Henjes F, Frohlich H, Wiemann S, Korf U, Beissbarth T. Dynamic deterministic effects
615 propagation networks: learning signalling pathways from longitudinal protein array data. Bioinformatics.
616 2010;26(18):i596-i602.
617 56. Koivunen JP, Mermel C, Zejnullahu K, Murphy C, Lifshits E, Holmes AJ, et al. EML4-ALK fusion gene
618 and efficacy of an ALK kinase inhibitor in lung cancer. Clin Cancer Res. 2008;14(13):4275-83.
619 57. TCGA-Consortium. Comprehensive genomic characterization defines human glioblastoma genes
620 and core pathways. Nature. 2008;455(7216):1061-8.
621 58. Crystal AS, Shaw AT. Variants on a theme: a biomarker of crizotinib response in ALK-positive non-
622 small cell lung cancer? Clin Cancer Res. 2012;18(17):4479-81.
623 59. Shaw AT, Solomon B. Targeting anaplastic lymphoma kinase in lung cancer. Clin Cancer Res.
624 2011;17(8):2081-6.
625 60. . !!! INVALID CITATION !!! [24].
626 61. Voena C, Di Giacomo F, Panizza E, D'Amico L, Boccalatte FE, Pellegrino E, et al. The EGFR family
627 members sustain the neoplastic phenotype of ALK+ lung adenocarcinoma via EGR1. Oncogenesis.
628 2013;2:e43.
629 62. Chan BA, Hughes BG. Targeted therapy for non-small cell lung cancer: current standards and the
630 promise of the future. Transl Lung Cancer Res. 2015;4(1):36-54.
631 63. Landi L, Minuti G, D'Incecco A, Cappuzzo F. Targeting c-MET in the battle against advanced
632 nonsmall-cell lung cancer. Curr Opin Oncol. 2013;25(2):130-6.
633 64. Borisy AA, Elliott PJ, Hurst NW, Lee MS, Lehar J, Price ER, et al. Systematic discovery of
634 multicomponent therapeutics. Proc Natl Acad Sci U S A. 2003;100(13):7977-82.
635 65. Chou TC. Drug Combination Studies and Their Synergy Quantification Using the Chou-Talalay
636 Method. Cancer research. 2010;70(2):440-6.
637 66. Zhang G, Fang B, Liu RZ, Lin H, Kinose F, Bai Y, et al. Mass spectrometry mapping of epidermal
638 growth factor receptor phosphorylation related to oncogenic mutations and tyrosine kinase inhibitor
639 sensitivity. J Proteome Res. 2011;10(1):305-19.
640 67. Valencia K, Ormazabal C, Zandueta C, Luis-Ravelo D, Anton I, Pajares MJ, et al. Inhibition of
641 collagen receptor discoidin domain receptor-1 (DDR1) reduces cell survival, homing, and colonization in
642 lung cancer bone metastasis. Clin Cancer Res. 2012;18(4):969-80.
643 68. Scott J, Ideker T, Karp RM, Sharan R. Efficient algorithms for detecting signaling pathways in
644 protein interaction networks. Journal of computational biology : a journal of computational molecular cell
645 biology. 2006;13(2):133-44.
646 69. Steffen M, Petti A, Aach J, D'Haeseleer P, Church G. Automated modelling of signal transduction
647 networks. BMC bioinformatics. 2002;3:34.
648 70. Kelley BP, Sharan R, Karp RM, Sittler T, Root DE, Stockwell BR, et al. Conserved pathways within
649 bacteria and yeast as revealed by global protein network alignment. Proceedings of the National Academy
650 of Sciences of the United States of America. 2003;100(20):11394-9.
651 71. Shimoni Y, Fink MY, Choi SG, Sealfon SC. Plato's cave algorithm: inferring functional signaling
652 networks from early gene expression shadows. PLoS Comput Biol. 2010;6(6):e1000828.
653 72. Budak G, Eren Ozsoy O, Aydin Son Y, Can T, Tuncbag N. Reconstruction of the temporal signaling
654 network in Salmonella-infected human cells. Front Microbiol. 2015;6:730.
655 73. Zhang QC, Petrey D, Deng L, Qiang L, Shi Y, Thu CA, et al. Structure-based prediction of protein-
656 protein interactions on a genome-wide scale. Nature. 2012;490(7421):556-60.
657 74. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al. The Connectivity Map: using
658 gene-expression signatures to connect small molecules, genes, and disease. Science.
659 2006;313(5795):1929-35.
660 75. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, et al. Gefitinib or carboplatin-paclitaxel
661 in pulmonary adenocarcinoma. The New England journal of medicine. 2009;361(10):947-57.

76. Steuer R, Kurths J, Daub CO, Weise J, Selbig J. The mutual information: detecting and evaluating dependencies between variables. Bioinformatics. 2002;18 Suppl 2:S231-40.
77. Gandhi J, Zhang JL, Xie Y, Soh J, Shigematsu H, Zhang W, et al. Alterations in Genes of the EGFR Signaling Pathway and Their Relationship to EGFR Tyrosine Kinase Inhibitor Sensitivity in Lung Cancer Cell Lines. PloS one. 2009;4(2).

Figure Legends

Figure 1. Framework for the reverse engineering of TK signaling networks from phosphoproteomic profiles.

(A) Schematic diagram of a **TK→S** interaction. The non-phosphorylated kinase is inactive in terms of phosphorylating a substrate, while the active isoform successfully phosphorylates the substrate. **(B)** Schematic diagram showing the correlation between TK phosphorylation and that of its potential substrates. The first two rows in the heatmap show proteins representing candidate TK substrates **(C)** Illustration of the pARACNe framework including 6 steps. Step-1 depicts peptides collection from primary lung cancer tissue and cell lines for whole phosphrtyrosine proteomics quantification. Step-2 depicts inferences of TK→S interactions using Mutual Information by Step-3 Naïve-Bayes estimator and Step-4 of the iterative quantile discretization methods. Step-5 and 6 depict network pruning and bootstrapping to construct final network. **(C)**. Workflow of pARACNe from LC-MS/MS data normalization, IQD process, MI calculation, DPI process, bootstrapping to network consolidation.

Figure 2. Predicted TK-TK network and validation of EGFR and c-MET prediction

(A) pARACNe-inferred densely inter-connected TK-TK network, with red nodes representing candidate TKs involved in auto-phosphorylation, where the phospho-state of a tyrosine is correlated with the phospho-state of a different tyrosine on the same TK protein. **(B)** pARACNe-inferred EGFR and **(C)** c-MET substrate overlap with SILAC-based and Database reported substrates, respectively.

688 **Figure 3. Inference of Master Regulator and combination**

689 **(A)**. Schema of Master Regulator analysis in lung cancer using pVIPER. Prioritized Master
690 Regulators **(B)** and Prioritized Master Regulator Pairs **(C)** as significantly activated (red circle) or
691 de-activated (blue) in different lung cancer cell lines (column). Red color represents an enrichment
692 of substrates hyper-phosphorylation by a Master Regulator or Master Regulator Pairs. Blue color
693 represents that of hypo-phosphorylation.

694 **Figure 4. Experimental validation of EGFR and c-MET combination by colony**
695 **formation assay**

696 **(A)** Colony formation assay schema. **(B)** shows the image of long-term EGFR and c-MET double
697 inhibition effects in HCC78 cell line with different treatments. **(C)** shows long-term clony formation
698 data for 14 cell lines with different EGFR, BRAF and KRAS genomic mutation status.

699 **Figure 5. MTT Assay validation of EGFR and c-MET combination**

700 **(A)**. MTT assay experimental schema. **(B)** MTT assay of HCC78 cell line shows synergistic effects
701 of Crizotinib and Erlotinib treatment. **(C)** shows short-term effects of EGFR and c-MET inhibitors'
702 combination index in 11 cell lines include 2 control cell lines (red).

703 **Figure 6. Master Regulating peptides in primary lung cancer samples**

704 EGFR and c-MET co-regulate in three scenarios **(A)** when their common substrates are
705 hyperphosphrylated, the patient responds to combination treatment well; **(B)** when most EGFR
706 substrates are hyper-phosphorylated, the patient responds to EGFR inhibitor; **(C)** when
707 substrates of both EGFR and c-MET are mostly hypophosphorylated, the patient does not
708 respond. **(D)** show the Master Regulator and Master Regulator Pairs regulating hyper/hypo-
709 phosphorylation of their network substrates in each primary samples.

710

711 # Methods and Data

712 ***Phosphoproteomic data***

713 The phosphoproteomic data used to reconstruct signaling network was download from (50),
714 representing the abundance of phospho-tyrosine containing peptides, was obtained by tandem
715 mass spectrometry analysis of 46 non-small cell lung cancer (NSCLC) cell lines, 151 NSCLC
716 tumors, and 48 normal lung tissue samples. Immunohistochemistry and a phospho-tyrosine
717 specific antibody were used to screen 96 paraffin-embedded, formalin fixed tissue samples from
718 NSCLC patients as described by Rikova et al.(50). About 30% of tumors showed high-levels of
719 phospho-tyrosine expression. Immunoblotting of 46 NSCLC cell lines with a phospho-tyrosine
720 specific antibody also showed heterogeneous reactivity especially in the molecular weight range
721 characteristic of receptor tyrosine kinases.

722 Since phospho-tyrosine represents less than 1% of the cellular phosphoproteome, as determined
723 by tandem mass spectrometry (MS/MS), and is difficult to analyze by conventional methods,
724 immuno-affinity purification was performed with a phospho-tyrosine antibody to enrich for
725 phospho-tyrosine containing peptides prior to tandem mass spectrometry. All tumors were
726 identified as NSCLC based on standard pathology. Only those tumors with greater than 50% of
727 cancer cells were considered for further analysis. NSCLC cell lines were grown overnight in low
728 serum to reduce background phosphorylation from culture conditions.

729 Tandem MS profiling identified 3920 tyrosine phosphorylation sites on approximately 2600
730 different proteins. 85% of these sites appeared to be novel when compared against PhosphoSite
731 (http://www.phosphosite.org), a comprehensive resource of known phosphorylation sites.

732 ***pARACNe algorithm***

733 ARACNe is originally designed for gene expression data, where expression of genes is usually
734 continuous and non-sparse. Quantitative data obtained from label-free LC-MS/MS by data-
735 dependent acquisition via spectral counting is discrete and very sparse, with many
736 phosphopeptides counts not observed for multiple peptides in each sample causing the current
737 version of ARACNe to be not suitable for this data, which thus required major modifications to
738 handle discrete data. To handle discrete abundances, we modified the mutual information
739 computation approach from a kernel density estimation based method to a Naïve based
740 estimation of mutual information, which is a histogram based technique(76). Briefly, consider a
741 collection of *N* simultaneous measurements of two genes X and Y. Data is partitioned into M
742 discrete bins a_i, and k_i denotes the number of measurements that lie within the bin a_i. The
743 probabilities $p(a_i)$ are then approximated by the corresponding relative frequencies of
744 occurrence $p(a_i) \rightarrow \frac{k_i}{N}$ and the mutual information $I(X,Y)$ between datasets X and Y is
745 expressed as

746 $$I(X,Y) = \log N + \frac{1}{N}\sum_{ij} k_{ij} \log \frac{k_{ij}}{k_i k_j}$$

747 Here k_{ij} denotes the number of measurements where X lies in a_i and Y in a_j and *N* total number
748 of samples.

749 Accuracy of mutual information is dependent on correct numbers of bins, M. To find the optimal
750 number of bins we applied ARACNe on the whole dataset by varying M from 1 to 20 and testing
751 the connections in predicted sub-network against the set of known connections (gold standard)
752 from databases (phosphoDB) (33).

753 In case of continuous data, partitioning can be achieved by dividing the range of data into M
754 equally spaced distance bins. Our data being discrete, equally spaced partitioning was not
755 possible. So, to overcome this problem, we used an iterative approach of partitioning **(Fig. 1C,**
756 **Fig. S1A)**. The basic idea is to divide the number of N data points into M with each bin containing
757 equal number of data point. If the data point(s) with the same value falls into consecutive bin(s),
758 we put those data point(s) into current bin and repartition the remaining points into remaining
759 number of bins. We keep iterating this till we finish either the bins or there are no more data points
760 to bin. For example, in **Fig. S1A**, we initially partition N points into 4 bins. The data points with 0
761 value does not fit into first bin and falls into subsequent bins, so we assign all data points with 0
762 value into first bin and repartition the remaining points into 3 bins. We keep on doing this process
763 till we achieve 4 bins.

764 To evaluate initial performance and decide number of bins, we computed the network among all
765 tyrosine kinases and substrates, parsed the sub-network between 49 tyrosine kinases and 114
766 substrates which were present in PhosphoSite database and compared the results with the
767 connections present in database. From our analysis, we found that M=10 to be an optimal number
768 **(Fig. S1B)** which gave us precision of 14% and sensitivity of 24%. This precision is an
769 underestimate of real precision as in the gold standard many interactions are not present.

770 ***Master Regulator analysis***

771 To discover the master regulator in various cell lines, we interrogated the network obtained from
772 pARACNe using a novel algorithm, MARINa (21) (Master Regulator Inference Analysis), designed
773 to infer kinases that are key players in a particular cell line. Protein activity is a good indicator of
774 key kinases in a particular phenotype but often phosphorylated state of a protein is not sufficient
775 to determine its activity both due to measurement noise in phosphorylated state as well as LC-
776 MS/MS technique noise. To overcome this problem, MARINa infers kinase activity from the global

777 kinase substrate relationship and its biological relevance by overlapping this information in a
778 particular phenotype-specific program.

779 MARINa requires a network model and signature of the phenotype transition (i.e., all genes
780 ranked by their differential phosphorylation in two phenotypes). Here, the signature, S_{kin}, was
781 obtained by t-test analysis by comparing each cell line against all normal samples. First, we
782 associate each kinase with positive and negative activity targets, by computing the correlation
783 between each kinase and its predicted substrates and selecting only those substrates which had
784 a significant correlation (p-value ≤ 0.05, Bonferroni corrected). Second, for each kinase we
785 computed an *activity* by measuring the enrichment of the S_{kin} signature in predicted substrates
786 list, separately for both positive and negative correlated, (S_{kin}-*enrichment*). Enrichment was
787 computed by Gene Set Enrichment Analysis (GSEA). Since very small percentage of kinases are
788 found to have negative correlation, we did not use those interactions to evaluate enrichment.

789 ***Cell Culture***

790 All cell lines were grown in RPMI-1640 with 5% fetal bovine serum and incubated at 37°C in a
791 humidified atmosphere containing 5% CO_2. Cell lines were fingerprinted using the Perplex 1.2
792 system (Promega, Madison, WI). Fingerprints were compared to those generated at ATCC and/or
793 our internal database.

794 ***MTS Assay***

795 Short term MTS assays (CellTiter 96® AQ_{ueous} One Solution Cell Proliferation Assay, Promega,
796 Madison WI) were performed as previously described in (77). Specifically, each drug
797 concentration is octuplicated and the mean with standard deviation of all replicates were used to
798 generate a curve to allow calculation of the drug IC50 (Inhibitory Concentration of 50%) value.
799 The assays were repeated at least 3 times and the IC50s are the average of all replicates.

800

801 ***Colony formation assay***

802 Long term colony formation assays were performed in triplicate in 6-well plates. Cells were added
803 to media containing drug and incubated for 1-2 weeks such that control wells (no drug) contained
804 colonies of 50-70 cells each. At such time media was removed and all wells stained with a solution
805 containing 0.5% crystal violet and 6% glutaraldehyde for 1 hour. The plates were then rinsed,
806 dried, and colonies were manually counted.

807 ***BrdU Cell Proliferation Assay and Cell viability Assay for H2228***

808 The cell proliferation assay was performed following previously described methods (50). H2228
809 cells are treated with BrdU, a thymidine analog that is incorporated into the DNA during cell
810 proliferation. The BrdU was detected with a special antibody, and high rates of BrdU incorporation
811 correlate with high cell proliferation rates.

812

813 ***SILAC Experiment***

814 EGFR SILAC experiment was performed in H3255 cell line by treating samples with Gefitinib. c-
815 MET SILAC experiment was performed in c-MET-driven gastric cancer cell line, MKN45, by using
816 c-MET inhibitor Su11274. For both genes, cells were treated with inhibitors for 3 and 24hr. For
817 control, cells were grown in same conditions but were not treated with the drug. For our
818 comparison we combined the peptides, which were differentially obtained between treated and
819 untreated samples, for 3 and 24 hr. More details about the experiment can be obtained from Guo
820 et al (32).

821 **Gold standard**: In PhosphoSite database, there were 282 connections between 49 tyrosine
822 kinases and 114 substrates.

823

824 ## Supplemental Information

825 ***Fig. S1. Performance of pARACNe algorithm.***

826 **(A).** To select optimal bin number in pARACNe algorithm, precision and recall curves for various
827 number of bins were computed. Black curve is when no binning of data is done. When using 10
828 bins, the algorithm achieved the best performance.

829 ***Fig. S2. Validation of Master Regulators and Master Regulator combination.***

830 **(A)** Cell viability assay for top Master Regulators including ALK, EGFR, EPHA2, MET, PTK2, and
831 UBB in H2228 cell line. **(B)** BrdU assay for MET and PTK2 and its combination.

832 ***Table S1. pARACNe inferred TK peptides – substrate peptides represented signaling***
833 ***network.***

834 ***Table S2. pARACNe inferred TK protein – substrate protein represented signaling***
835 ***network.***

836 ***Table S3. Clony formation assay and MTS assay results.***

837

Figure 1

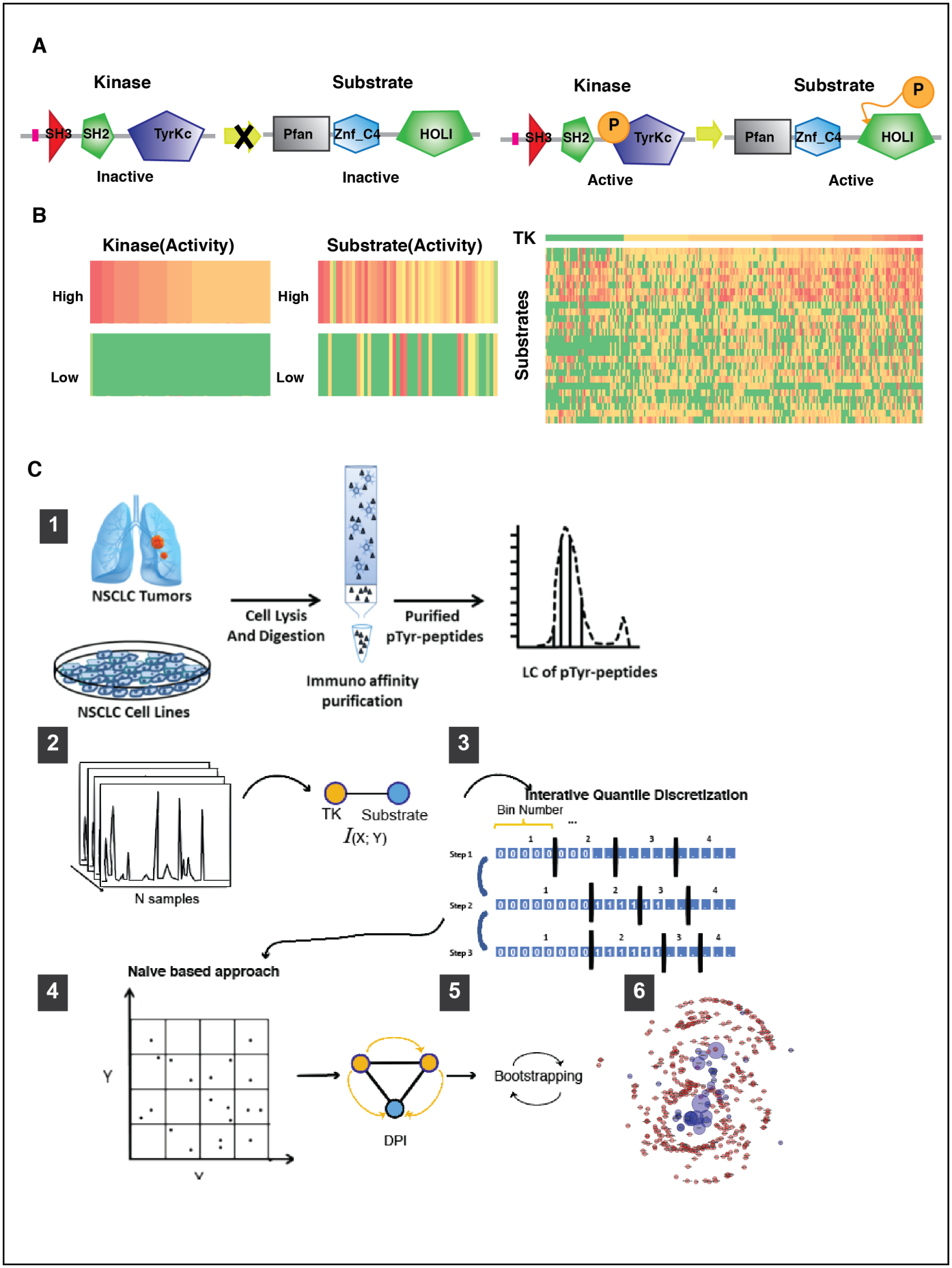

A
Kinase
Substrate
SH3
SH2
TyrKc
Pfan
Znf_C4
HOLI
Inactive
Inactive
Kinase
Substrate
P
Active
Active
B
Kinase(Activity)
Substrate(Activity)
High
Low
High
Low
TK
Substrates
C
1
NSCLC Tumors
NSCLC Cell Lines
Cell Lysis
And Digestion
Immuno affinity
purification
Purified
pTyr-peptides
LC of pTyr-peptides
2
N samples
TK
Substrate
I(X; Y)
3
Interative Quantile Discretization
Bin Number
Step 1
Step 2
Step 3
4
Naive based approach
Y
X
DPI
5
Bootstrapping
6

Figure 2

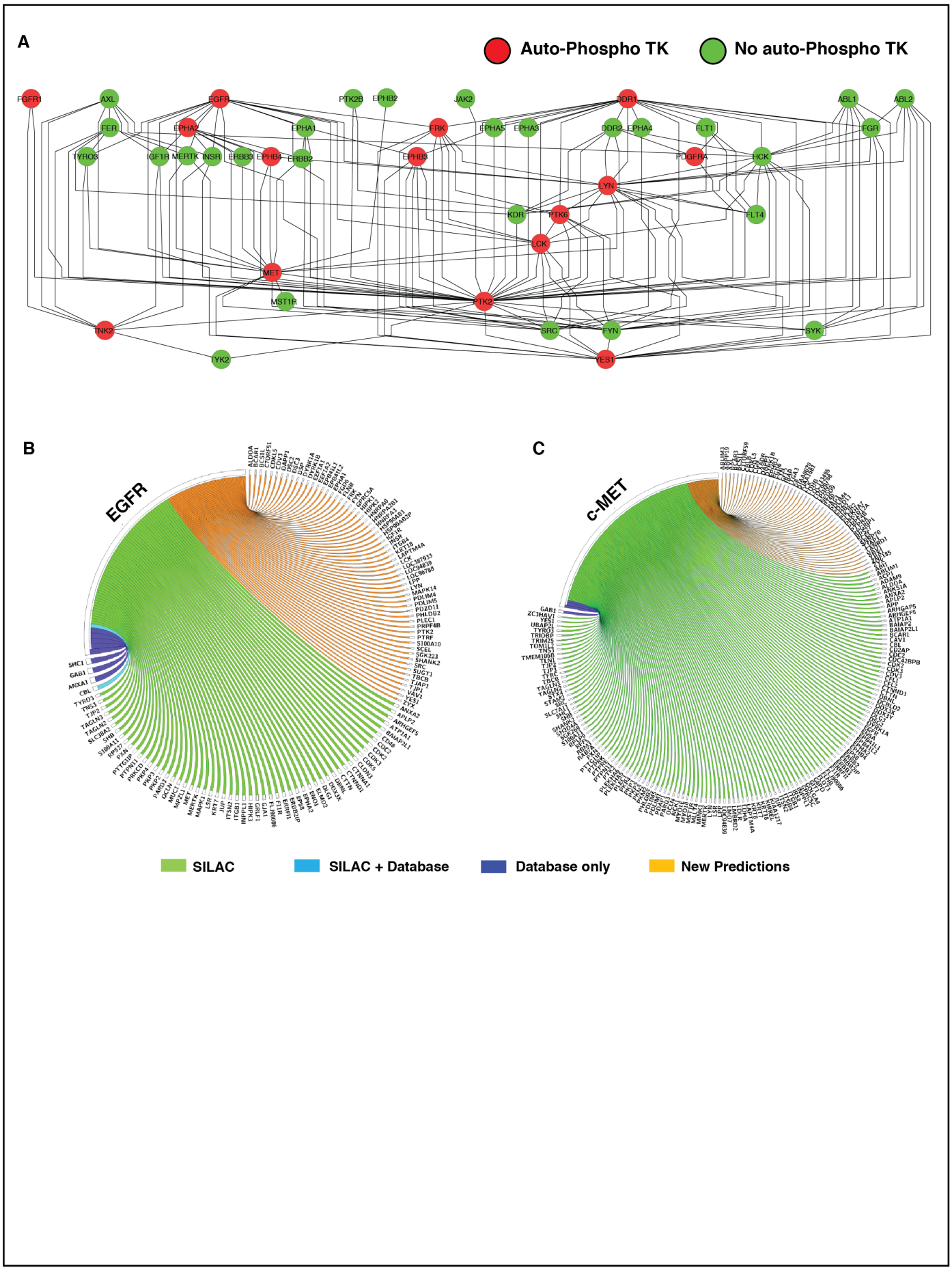

A
Auto-Phospho TK
No auto-Phospho TK
FGFR1
AXL
EGFR
PTK2B
EPHB2
JAK2
DDR1
ABL1
ABL2
FER
EPHA2
EPHA1
FRK
EPHA5
EPHA3
DDR2
EPHA4
FLT1
FGR
TYRO3
IGF1R
MERTK
INSR
ERBB3
EPHB4
ERBB2
EPHB3
PDGFRA
HCK
LYN
KDR
PTK6
FLT4
LCK
MET
MST1R
PTK2
TNK2
SRC
FYN
SYK
TYK2
YES1
B
EGFR
C
c-MET
SILAC
SILAC + Database
Database only
New Predictions

Figure 3

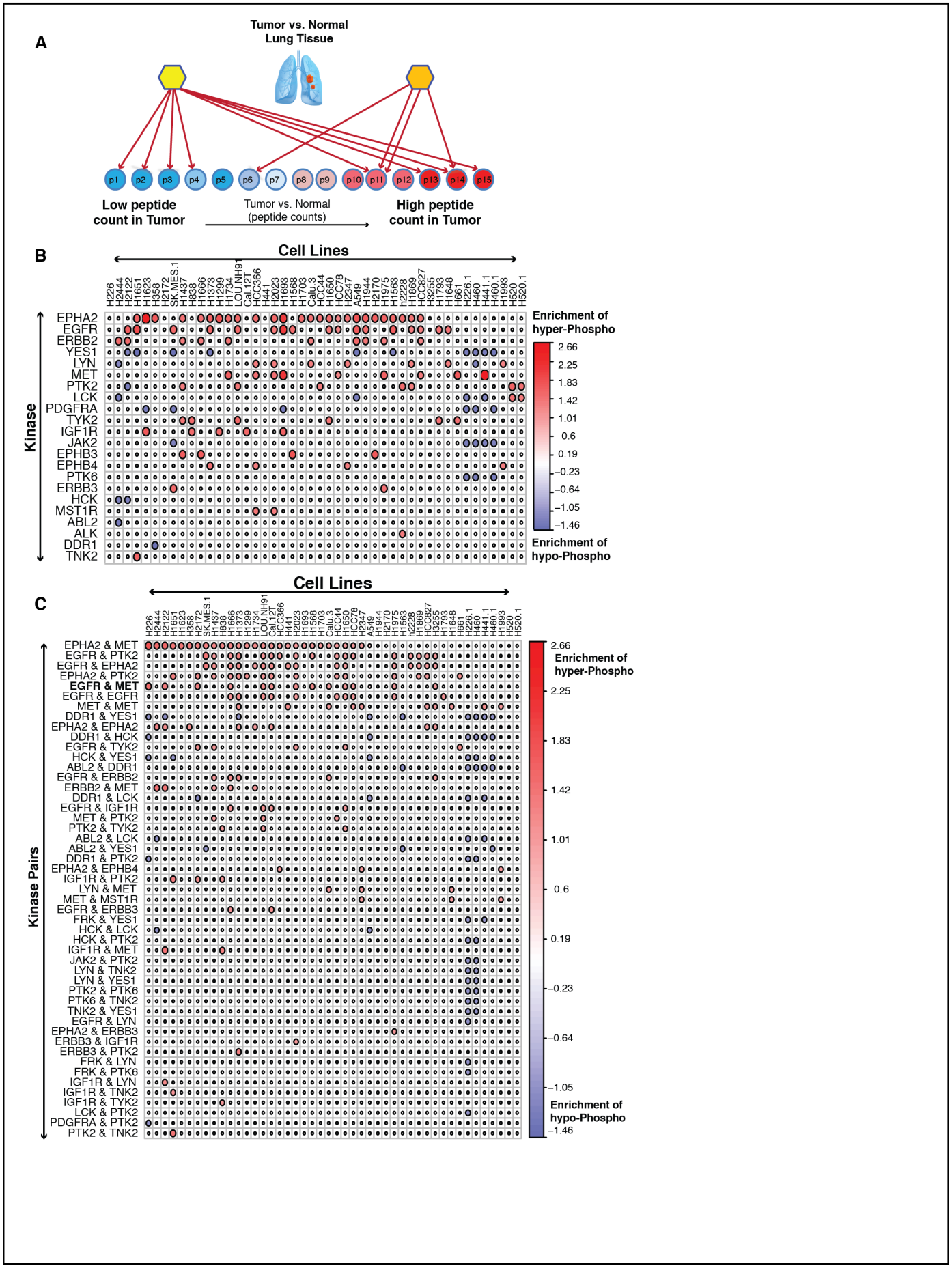

A
Tumor vs. Normal
Lung Tissue
p1 p2 p3 p4 p5 p6 p7 p8 p9 p10 p11 p12 p13 p14 p15
Low peptide count in Tumor
Tumor vs. Normal (peptide counts)
High peptide count in Tumor
B
Cell Lines
Kinase
EPHA2
EGFR
ERBB2
YES1
LYN
MET
PTK2
LCK
PDGFRA
TYK2
IGF1R
JAK2
EPHB3
EPHB4
PTK6
ERBB3
HCK
MST1R
ABL2
ALK
DDR1
TNK2
Enrichment of hyper-Phospho
2.66
2.25
1.83
1.42
1.01
0.6
0.19
−0.23
−0.64
−1.05
−1.46
Enrichment of hypo-Phospho
C
Cell Lines
Kinase Pairs
EPHA2 & MET
EGFR & PTK2
EGFR & EPHA2
EPHA2 & PTK2
EGFR & MET
EGFR & EGFR
MET & MET
DDR1 & YES1
EPHA2 & EPHA2
DDR1 & HCK
EGFR & TYK2
HCK & YES1
ABL2 & DDR1
EGFR & ERBB2
ERBB2 & MET
DDR1 & LCK
EGFR & IGF1R
MET & PTK2
PTK2 & TYK2
ABL2 & LCK
ABL2 & YES1
DDR1 & PTK2
EPHA2 & EPHB4
IGF1R & PTK2
LYN & MET
MET & MST1R
EGFR & ERBB3
FRK & YES1
HCK & LCK
HCK & PTK2
IGF1R & MET
JAK2 & PTK2
LYN & TNK2
LYN & YES1
PTK2 & PTK6
PTK6 & TNK2
TNK2 & YES1
EGFR & LYN
EPHA2 & ERBB3
ERBB3 & IGF1R
ERBB3 & PTK2
FRK & LYN
FRK & PTK6
IGF1R & LYN
IGF1R & TNK2
IGF1R & TYK2
LCK & PTK2
PDGFRA & PTK2
PTK2 & TNK2
Enrichment of hyper-Phospho
2.66
2.25
1.83
1.42
1.01
0.6
0.19
−0.23
−0.64
−1.05
−1.46
Enrichment of hypo-Phospho

Figure 4

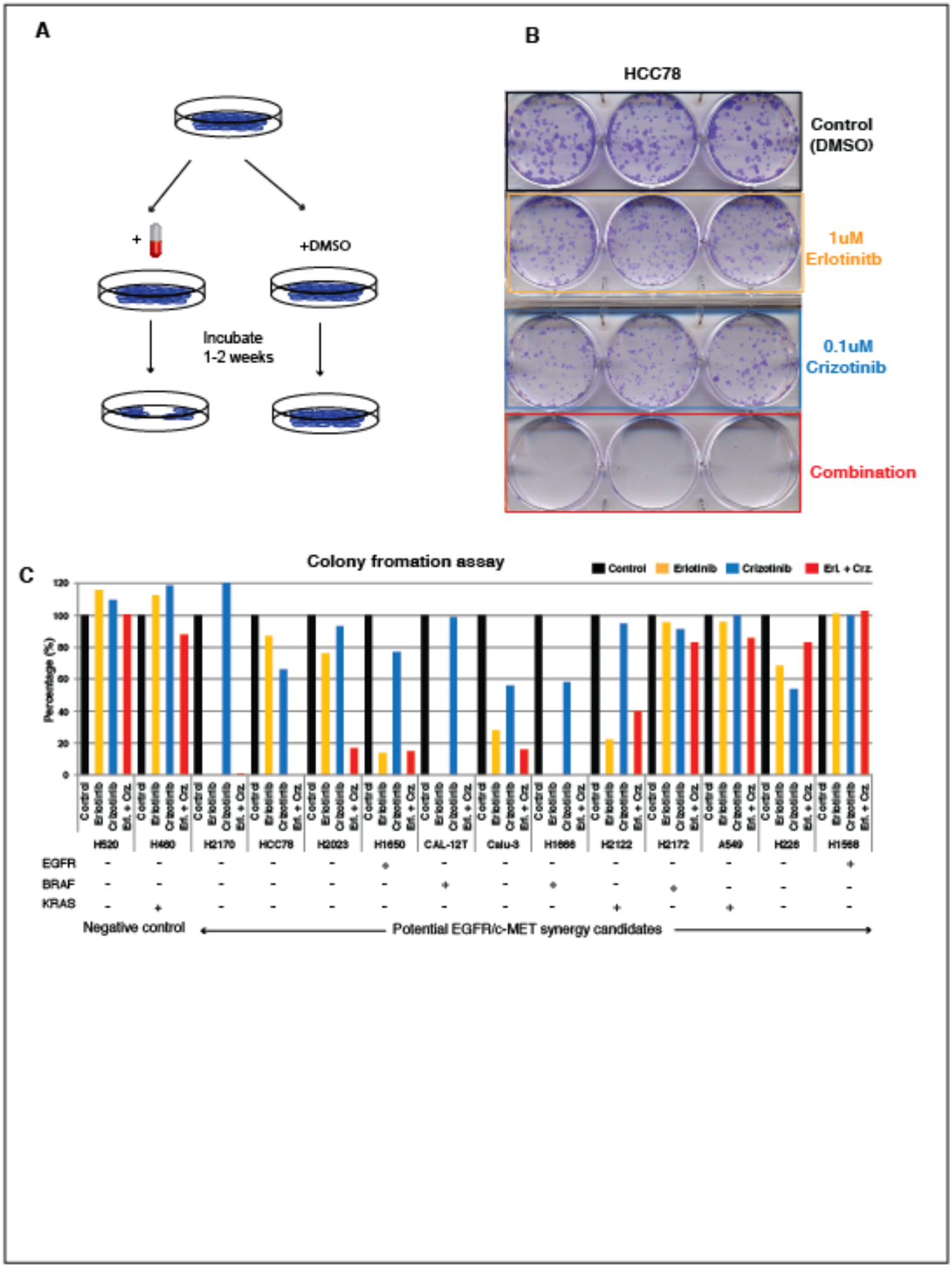

A
+
+DMSO
Incubate
1-2 weeks
B
HCC78
Control
(DMSO)
1uM
Erlotinitb
0.1uM
Crizotinib
Combination
C
Colony fromation assay
Control
Erlotinib
Crizotinib
Erl. + Crz.
Percentage (%)
120
100
80
60
40
20
0
H520
H460
H2170
HCC78
H2023
H1650
CAL-12T
Calu-3
H1666
H2122
H2172
A549
H226
H1568
EGFR
BRAF
KRAS
Negative control
Potential EGFR/c-MET synergy candidates

Figure 5

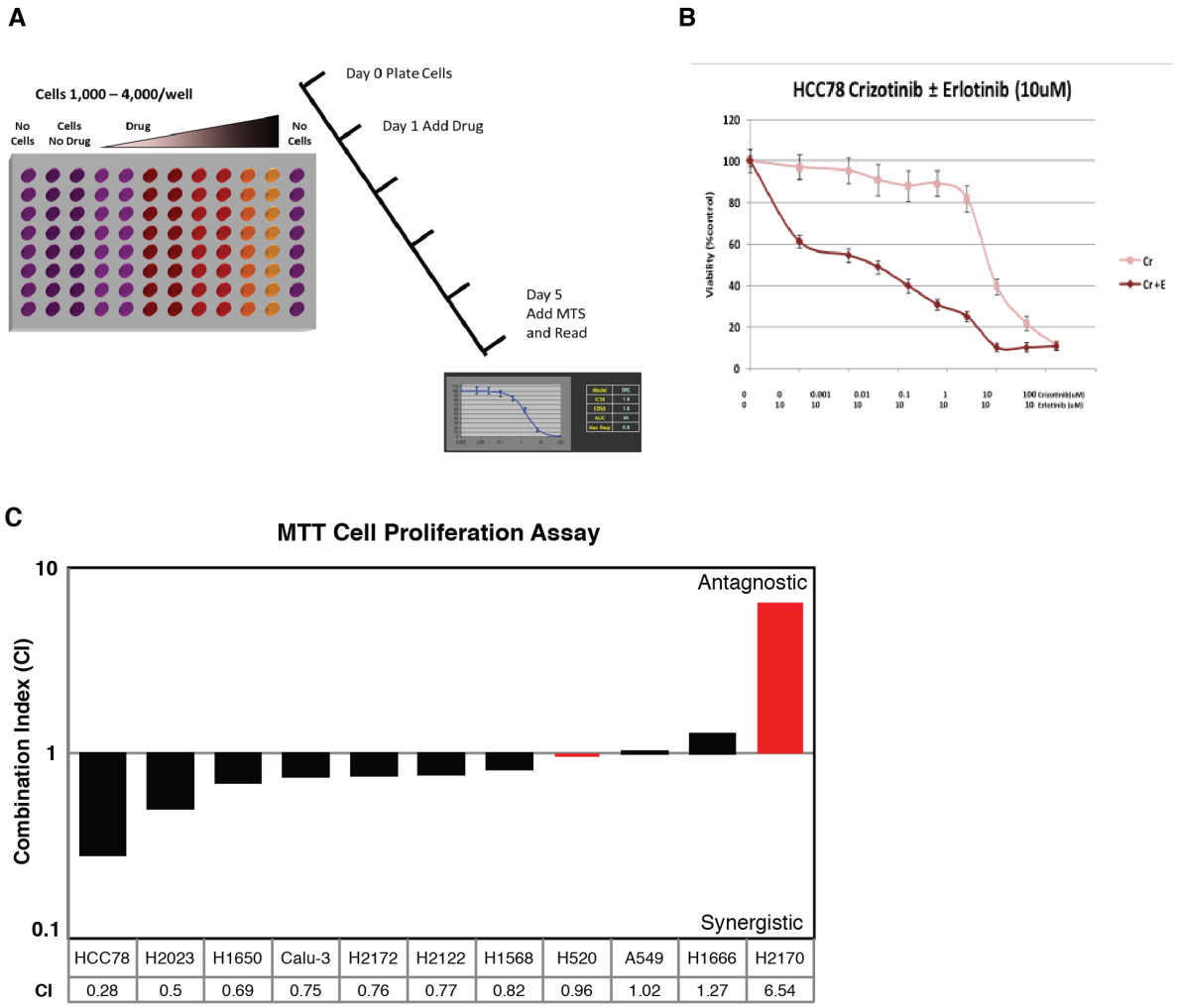
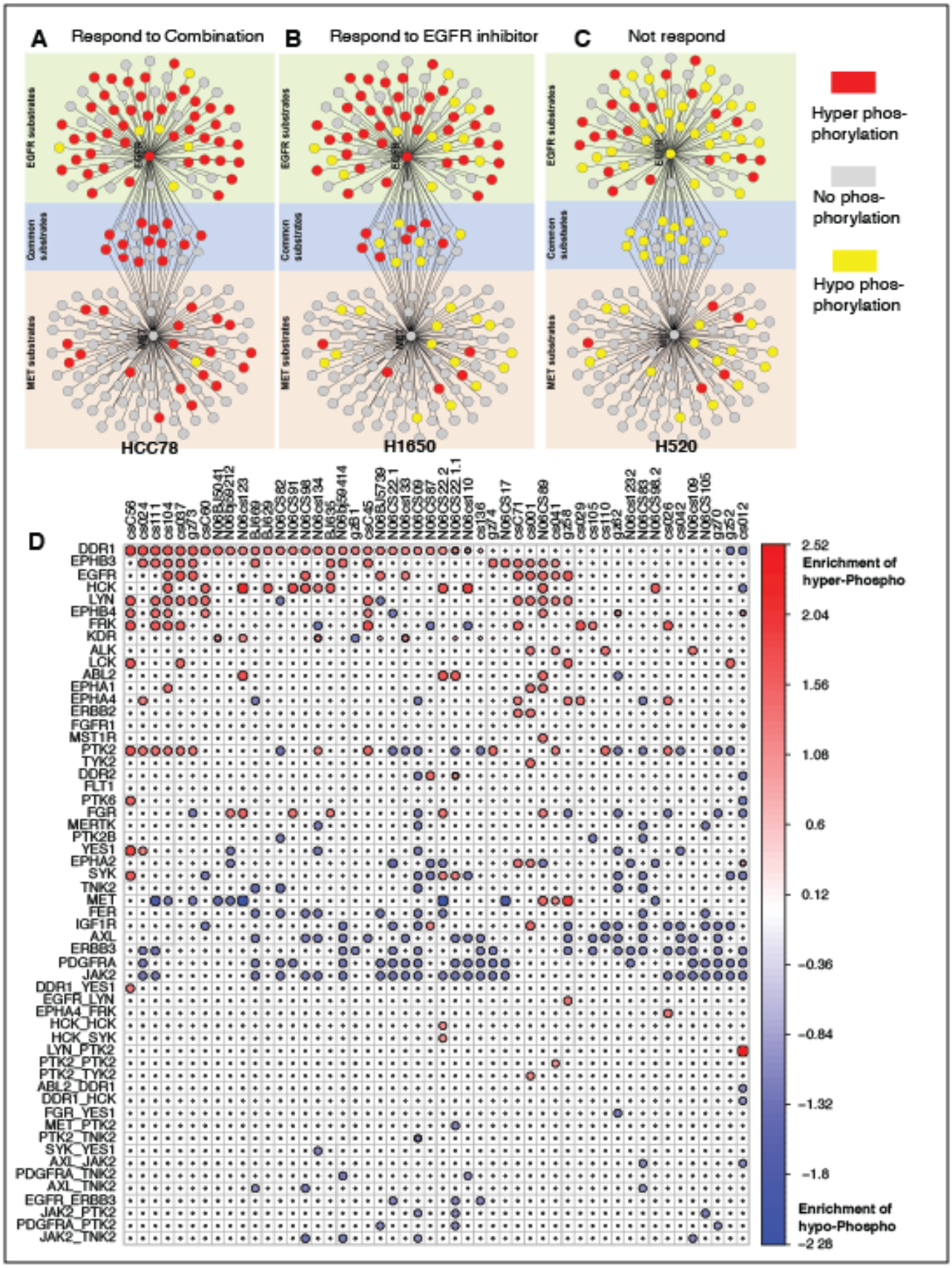

A
Cells 1,000 – 4,000/well
No Cells
Cells No Drug
Drug
No Cells
Day 0 Plate Cells
Day 1 Add Drug
Day 5 Add MTS and Read
B
HCC78 Crizotinib ± Erlotinib (10uM)
Viability (%control)
120
100
80
60
40
20
0
Cr
Cr+E
C
MTT Cell Proliferation Assay
Combination Index (CI)
10
1
0.1
Antagnostic
Synergistic
HCC78 H2023 H1650 Calu-3 H2172 H2122 H1568 H520 A549 H1666 H2170
CI 0.28 0.5 0.69 0.75 0.76 0.77 0.82 0.96 1.02 1.27 6.54

Figure 6


A
Respond to Combination
B
Respond to EGFR inhibitor
C
Not respond
EGFR substrates
Common substrates
MET substrates
HCC78
H1650
H520
Hyper phos-phorylation
No phos-phorylation
Hypo phos-phorylation
D
DDR1
EPHB3
EGFR
HCK
LYN
EPHB4
FRK
KDR
ALK
LCK
ABL2
EPHA1
EPHA4
ERBB2
FGFR1
MST1R
PTK2
TYK2
DDR2
FLT1
PTK6
FGR
MERTK
PTK2B
YES1
EPHA2
SYK
TNK2
MET
FER
IGF1R
AXL
ERBB3
PDGFRA
JAK2
DDR1_YES1
EGFR_LYN
EPHA4_FRK
HCK_HCK
HCK_SYK
LYN_PTK2
PTK2_PTK2
PTK2_TYK2
ABL2_DDR1
DDR1_HCK
FGR_YES1
MET_PTK2
PTK2_TNK2
SYK_YES1
AXL_JAK2
PDGFRA_TNK2
AXL_TNK2
EGFR_ERBB3
JAK2_PTK2
PDGFRA_PTK2
JAK2_TNK2
2.52
Enrichment of hyper-Phospho
2.04
1.56
1.08
0.6
0.12
−0.36
−0.84
−1.32
−1.8
Enrichment of hypo-Phospho
−2.28