Translational Oncology 13 (2020) 100830

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Original Research

Candidate biomarker assessment for pharmacological response

William C. Reinhold [a,*], Fathi Elloumi [a,c], Sudhir Varma [a,b], Jacques Robert [d], Gordon B. Mills [e], Yves Pommier [a]

[a] Developmental Therapeutic Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, United States of America
[b] HiThru Analytics LLC, Laurel, MD, USA
[c] General Dynamics Information Technology, Falls Church, VA 22042, United States of America
[d] Inserm unité 1218, Université de Bordeaux, France
[e] Knight Cancer Institute, Oregon Health & Science University, Portland, OR 97239, United States of America

ARTICLE INFO

Article history:
Received 26 March 2020
Received in revised form 23 June 2020
Accepted 26 June 2020
Available online xxxx

ABSTRACT

Using the information from our CellMiner (https://discover.nci.nih.gov/cellminer/) and CellMinerCDB (https://discover.nci.nih.gov/cellminercdb/) web-based applications, we identified 3978 molecular events with significant links to pharmacological response for genes that are either targets, biomarkers, or have established causal linkage to drugs. Molecular events included DNA copy number, methylation and mutation; and transcript; and whole or phospho-protein expression for the NCI-60 human cancer cell lines. While all forms of molecular data were informative in some (gene-drug) pairings, the type of significantly linked molecular events was found to vary widely by drug. Some forms of molecular data were found to have more frequent significant correlation than others. Leading were phosphoproteins as measured by antibody (31%), followed by transcript as measured by microarray (16%), and total protein levels as measured by mass spectrometry or antibody (14%). All other measurements ranged between 5 and 11%. Data reliability was underscored by concordant results when using differing drugs with the same targets, as well as different measurements of the same molecular parameter. The significance of correlations of the various molecular parameters to the pharmacological responses provides functional indication of those parameters that are biologically relevant for each gene-drug pairing, as well as comparisons between measurement types.

Introduction

The recognition of predictive biomarkers for pharmacological response is of great interest in clinical applications. Biomarkers have allowed the development of new treatment strategies, including those based on the molecular profiles of tumors allowing matching of targeted therapies with patients [1]. However, for any target gene of interest, there are many molecular features that might be predictive. Currently, DNA mutations are the favored biomarker due to their relative ease of access, purification, stability and assessment. In order to assess biomarker potential for specific drugs across multiple molecular features of their known biomarkers (such as BRAF for vemurafenib), targets (such as EGFR for afatinib, or TOP1 for topotecan) or causally linked genes (such as SLFN11 for topotecan, etoposide, cisplatin, gemcitabine, cytarabine or mitoxantrone), we systematically reviewed the drugs for significant association to molecular features of those genes using the deep NCI-60 cell line screen information.

Materials and methods

CellMiner datasets and comparisons

The National Cancer Institute 60 (NCI-60) cell line datasets used for this study were accessed in CellMiner and downloaded from the Download Data Sets tab\Download Processed Data set (https://discover.nci.nih.gov/cellminer). The drug activities (growth inhibition 50%) were accessed from Download Processed Data Set\Compound activity: DTP NCI-60. All drug data were generated by the Developmental Therapeutics Program (https://dtp.nci.nih.gov/). The activities of 129 drugs were evaluated. The numbers of molecular markers measured by each molecular marker platform were 23,232 for aCGH, 17,553 for DNA methylation, 12,706 amino acid changing variants, 9143 protein function-affecting variants, 25,040 transcripts as measured by microarray, 23,826 transcripts as measured by RNAseq, 3162 proteins as measured by mass spectrometry, 93 proteins as measured by antibodies by NCI, and 347 proteins as measured by

* Corresponding author at: National Cancer Institute, 9000 Rockville Pike, Building 37, room 5041, Bethesda, MD 20892, United States of America.
E-mail address: wcr@mail.nih.gov. (W.C. Reinhold).

http://dx.doi.org/10.1016/j.tranon.2020.100830

antibodies by MD Anderson (MDA). The exact CellMiner selections made to download each form of molecular data is detailed in Table 1 footnotes. p values were determined from Pearson's correlation coefficients and were calculated using R computing (https://www.r-project.org).

Results

Relationships between specific pharmacological response and molecular alterations

We assessed molecular data for DNA copy number, methylation, amino acid-affecting mutations, protein-function-affecting mutations, transcript levels as assessed by microarray and RNAseq, total protein levels as assessed by mass spectrometry and antibody, and phosphoprotein levels as assessed by antibody for 167 genes compared to 128 FDA-approved or clinical trial drug activities (Supplemental Table 1). This yielded 3978 drug-gene comparisons with direct linkage, i.e. the same gene that is a biomarker for, targeted by, or causally linked to a given drug is assessed for change at the molecular level. Table 1 presents a subset of that data that exhibit significant correlations between molecular and activity data.

Among well-documented connections, robust and significant mutation-drug positive correlations were observed for BRAF V600E-vemurafenib and dabrafenib. An additional small contribution to these drugs activity profiles may be made by (BRAF) DNA copy number alteration. Also, SLFN11 transcript level was significantly positively correlated with carboplatin, mitoxantrone, topotecan, gemcitabine, and melphalan activities (DNA-damaging drugs of different types).

Multiple less known or novel relationships are detailed in Table 1 and/or Supplemental Table 1, including the relative importance of the various molecular features for predicting response to each drug. Transcript expression of ALK has significant positive correlations to multiple ALK inhibiting drugs, including alectinib, AP-26113, crizotinib and LDK-378. Transcript and/or protein expression of EGFR has significant positive correlations to multiple EGFR inhibiting drugs, including afatinib, erlotinib, gefitinib and lapatinib. Lapatinib, which inhibits ERBB2 as well, is also positively correlated to ERBB2 transcript and protein expression. ESR1 protein and/or transcript expression are significantly positively correlated to the ESR1 inhibitors fulvestrant, raloxifene and tamoxifen. MAP2K1 (MEK1) phosphoprotein expression has significant positive correlations to several multi-target MEK inhibitors, including cobimetinib, PD-98059, selumetinib and trametinib. PRKCA (protein kinase C alpha) transcript and protein expression both have significant positive correlation to the PRKCA inhibitors midostaurin and staurosporine. TOP2A DNA copy number and transcript expression are correlated positively to the TOP2 inhibitors amonafide, daunorubicin, doxorubicin, etoposide, idarubicin, mitoxantrone and teniposide. TUBB6 has significant negative correlations to multiple tubulin affecting drugs, including docetaxel, dolstatin 10, eribulin mesilate, ixabepilone, paclitaxel, vinblastine, and vinorelbine, for transcript and/or protein expression.

CellMinerCDB visualization and exploration of pharmacological-molecular relationships

Our CellMinerCDB web-application provides a rapid, on-the-fly ability to visualize and explore relationships between drugs, genomic and proteomic parameters [2,3] (https://discover.nci.nih.gov/cellminercdb). Fig. 1 provides several examples taken as snapshots. Because each cell line is represented on the plots and data can be readily downloaded to Microsoft Excel-compatible files, CellMinerCDB provides granularity and insight, which is better displayed by scatter plots than simple correlation analyses. It includes examples of the ability to visualize drug responses of cell lines with molecular alterations of interest (Fig. 1A with the labeled cell lines having the BRAF V600E mutation); clinically relevant response profiles to a drug by cancer tissue(s) of origin (Fig. 1B and C); the plateauing of activity values (Fig. 1D); the ability to compare drug response across variable tissues of origin (Fig. 1E); and high correlation due to a single widely variant cell line (Fig. 1F).

Fig. 1A visualizes the relationship between MAP2K1 phosphoprotein (MEK) expression as measured by reverse-phase protein array (RPPA) and the MAP2K1-inhibitor PD-98059. The labeled cell lines have BRAF mutated at V600E and are the most responsive. The plateauing of the activity values at the low end indicates that the sensitivity of those cell lines was outside the assay range. Fig. 1B visualizes the significant positive correlation between ERBB2 protein expression and the activity of the ERBB2-inhibitor lapatinib for the combined breast and lung cancer subset (chosen due to their clinical relevance). Fig. 1C extends the Table 1 (positive correlation) relationship between estrogen receptor ESR1 protein expression and the ESR1-inhibitor fulvestrant for the breast cancer subset. The three non-expressor/poor responders at the bottom left are all triple negative cell lines. Both ESR1 transcript and protein phosphorylation at S118 were also found to be predictive (by positive correlation) for fulvestrant activity in the breast cancer cell lines. Fig. 1D visualizes the significant positive correlation between TOP2A transcript expression and the TOP2 inhibitor mitoxantrone. The plateauing of activity values at the high end indicates those cell responses were outside of the assay range. Fig. 1E visualizes of the significant positive correlation between PRKCA S657 phosphoprotein (RPPA) expression and the PRKCA inhibitor midostaurin, revealing the renal cancers as most responsive. Fig. 1F visualizes the positive correlation between ALK and one of the four ALK-inhibitors for which it has significant correlation. The high correlation is shown to be due to a single leukemic cell line, SR.

Relationships between pharmacological response and classes of molecular alterations

The forms of molecular information presented in Supplemental Table 1 were each assessed across all drug-gene pairs for which there was information. Table 2 presents the synopsis of this analysis, including the percent that were predictive with significant correlations for each molecular measurement. DNA copy number and methylation data were predictive of pharmacological response 10 and 9% of the time, respectively. Two approaches of compiling the DNA mutation data demonstrated significant correlation in 6 and 5% of cases tested. RNA transcript, either measured by microarray or RNAseq performed better, at 16 and 11%, respectively. Total protein was predictive for 14% of cases for both mass spectrometry and total protein as measured by RPPA by either the NCI Genomics and Pharmacology Facility or the M.D. Anderson Systems Biology Department [4,5] (https://tcpaportal.org/mclp/#/). Strikingly, phospho-protein levels were the most predictive at 31%.

Discussion

While there have been notable successes using DNA mutations as biomarkers for pharmacological intervention, such as BRAF V600E for vemurafenib, and EGFR L858R and G719X (primary, sensitizing), and T790M (secondary, desensitizing) for EGFR inhibitors, it remains in doubt that mutational data will be able to supply the broader indications required for a personalized medicine approach for most patients [6–8]. A key goal of personalized medicine is to take full advantage of the many drugs that have robust response in some proportion of patients or model systems of a particular cancer type or subset [9]. Our studies indicate that assessing expression levels of pharmacologically-linked genes, such as the SLFN11 transcript [10], will provide broader pharmacological insight than DNA mutation alone. The detailed and systematic approach adopted in this study provides both verification of known relationships, and novel insights regarding the relation between specific drugs and their biomarkers, targets, or causally linked genes.

Among previously known examples, the BRAF V600E mutation is a well- established biomarker for vemurafenib and dabrafenib [6]. The SLFN11 transcript-DNA damaging drug relationships are also well established, although not yet employed as biomarkers [11–13]. ESR1 is a clinical, diagnostic biomarker for fulvestrant, raloxifene, and tamoxifen

Table 1
Pharmacological versus molecular relationships of directly or causally linked gene-drug pairs.

Drug[b]			Drug and molecular measurement intersection		p values for gene's molecular measurement versus drug activity correlations[a]									
					DNA				RNA		Protein			
Names	NSC[c]	MOA[d]	Gene[e]	Function[f]	Copy #[g]	Methylation[h]	Mut_AA[i,j]	Mut_PFA[i,k]	Microarray[l]	RNAseq[m]	Mass spec.[n]	Reverse phase protein array		
												DTP[o]	Antibody target[p]	MDA[q]
Crizotinib	756645	PK: YK, MET, ALK	ALK	Bio., Tar.	0.862	**0.023**	0.287	0.487	**0.001**	0.145	na	NA	NA	NA
AP-26113	761191	PK: YK, ALK, EGFR	ALK	Target	0.538	**0.029**	**0.049**	0.192	**4.4E−05**	0.260	na	NA	NA	NA
Alectinib	764040	PK: YK, PIK3, ALK	ALK	Bio., Tar.	0.285	**0.016**	0.131	0.639	**1.7E−09**	0.375	na	NA	NA	NA
LDK-378	777193	PK: ALK	ALK	Target	0.842	**0.002**	0.256	0.894	**1.5E−07**	0.128	na	NA	NA	NA
Vemurafenib	761431	PK: YK, BRAF	BRAF	Bio., Tar.	**0.019**	0.438	**2.2E−21**	**6.1E−23**	0.110	0.903	na	NA	BRAF_pS445	0.123
Dabrafenib	764134	PK: BRAF	BRAF	Bio., Tar.	**0.049**	0.569	**6.4E−17**	**1.5E−17**	0.101	0.646	na	NA	BRAF_pS445	0.220
Erlotinib	718781	PK: YK, EGFR	EGFR	Target	0.723	0.101	**0.044**	0.087	**0.001**	**1.5E−04**	0.252	NA	EGFR	**0.002**
Erlotinib	718781	PK: YK, EGFR	EGFR	Target	-	-	-	-	-	-	-	-	EGFR_pY1068	**0.004**
Erlotinib	718781	PK: YK, EGFR	EGFR	Target	-	-	-	-	-	-	-	-	EGFR_pY1173	**0.049**
Lapatinib	727989	PK: YK, EGFR, ERBB2	EGFR	Target	0.233	0.609	0.870	0.964	0.960	0.629	0.727	NA	EGFR_pY1068	**0.021**
Afatinib	750691	PK: YK, EGFR	EGFR	Bio., Tar.	0.947	0.265	0.276	0.338	**0.012**	**0.005**	0.555	NA	EGFR	**0.027**
Afatinib	750691	PK: YK, EGFR	EGFR	Bio., Tar.	-	-	-	-	-	-	-	-	EGFR_pY1068	**0.001**
Gefitinib	759856	PK: YK, EGFR	EGFR	Bio., Tar.	0.806	0.356	0.673	0.263	**0.023**	**0.014**	0.138	NA	EGFR	**0.022**
Lapatinib	745750	PK: YK, EGFR, ERBB2	ERBB2	Bio., Tar.	**0.013**	0.884	0.411	0.471	**0.010**	**0.007**	**0.028**	NA	HER2:ERBB2	**0.004**
Fulvestrant	719276	Ho| SERM	ESR1	Bio., Tar.	0.421	**0.009**	0.587	na	**2.2E−14**	**7.1E−15**	na	0.064	ESR1	**1.7E−08**
Fulvestrant	719276	Ho| SERM	ESR1	Bio., Tar.	-	-	-	-	-	-	-	-	ESR1_pS118	**6.0E−07**
Raloxifene	747974	Ho| SERM	ESR1	Bio., Tar.	0.386	**0.015**	0.984	na	**9.4E−05**	**3.9E−05**	na	0.527	ESR1	**0.002**
Raloxifene	747974	Ho| SERM	ESR1	Bio., Tar.	-	-	-	-	-	-	-	-	ESR1_pS118	**8.2E−06**
Tamoxifen	180973	Ho| SERM	ESR1	Bio., Tar.	0.627	0.361	0.769	na	0.145	0.146	na	0.113	ESR1_pS118	**0.010**
PD-98059	679828	PK: STK, MAP2K1	MAP2K1	Target	0.611	0.797	0.350	0.350	0.513	0.520	0.604	0.881	MAP2K1_pS217S221	**4.4E−08**
Selumetinib	741078	PK: STK, MAP2K1	MAP2K1	Target	0.383	0.954	0.348	0.348	0.951	0.978	0.770	0.802	MAP2K1_pS217S221	**3.0E−07**
Trametinib	758246	PK: STK, MAP2K1	MAP2K1	Target	0.198	0.881	0.223	0.223	0.577	0.824	0.472	0.754	MAP2K1_pS217S221	**0.001**
Cobimetinib	768068	PK: STK, MAP2K1	MAP2K1	Target	0.265	0.595	0.303	0.303	0.597	0.758	0.602	0.879	MAP2K1_pS217S221	**1.1E−05**
Staurosporine	618487	PK: PRKCA	PRKCA	Target	0.702	0.950	0.118	0.118	**1.9E−05**	**5.7E−06**	na	0.702	PRKCA	**0.004**
Staurosporine	618487	PK: PRKCA	PRKCA	Target	-	-	-	-	-	-	-	-	PRKCA_pS657	**1.9E−05**
Midostaurin	656576	PK: PRKCA, STK	PRKCA	Target	0.277	0.603	0.692	0.692	**1.9E−04**	**1.4E−04**	NA	0.296	PRKCA_pS657	**1.2E−04**
Midostaurin	656576	PK: PRKCA, STK	PRKCA	Target	-	-	-	-	-	-	-	NA	PRKCA	**0.009**
Carboplatin	241240	A7| AlkAg	SLFN11	Causal	0.382	**6.0E−06**	0.717	NA	**9.0E−06**	**0.001**	NA	NA	NA	NA
Mitoxantrone	301739	TOP2	SLFN11	Causal	0.164	**0.010**	0.141	na	**4.4E−08**	**5.6E−05**	na	NA	NA	NA
Topotecan	609699	TOP1	SLFN11	Causal	0.133	**2.6E−04**	0.684	na	**3.3E−13**	**1.9E−08**	na	NA	NA	NA
Gemcitabine	613327	Ds	SLFN11	Causal	0.329	**0.001**	0.239	na	**7.8E−10**	**4.7E−08**	na	NA	NA	NA
Melphalan	757098	A7| AlkAg	SLFN11	Causal	0.550	**4.7E−04**	0.205	NA	**1.2E−09**	**1.3E−05**	NA	NA	NA	NA
Amonafide	308847	TOP2	TOP2A	Target	**0.023**	0.478	NA	NA	0.112	**0.015**	NA	NA	NA	NA

(continued on next page)

Table 1 (*continued*)

Drug[b]			Drug and molecular measurement intersection		p values for gene's molecular measurement versus drug activity correlations[a]									
Names	NSC[c]	MOA[d]	Gene[e]	Function[f]	DNA				RNA		Protein			
					Copy #[g]	Methylation[h]	Mut_AA[i,j]	Mut_PFA[i,k]	Microarray[l]	RNAseq[m]	Mass spec.[n]	Reverse phase protein array		
												DTP[o]	Antibody target[p]	MDA[q]
Daunorubicin	82151	TOP2	TOP2A	Target	**0.035**	0.233	NA	NA	**0.002**	0.229	NA	NA	NA	NA
Doxorubicin	123127	TOP2	TOP2A	Target	**0.040**	0.183	NA	NA	**0.004**	0.767	NA	NA	NA	NA
Etoposide	141540	TOP2	TOP2A	Target	**0.018**	0.204	NA	NA	**0.001**	0.283	NA	NA	NA	NA
Idarubicin	256439	TOP2	TOP2A	Target	**0.001**	0.128	NA	NA	**0.004**	**0.050**	NA	NA	NA	NA
Mitoxantrone	279836	TOP2	TOP2A	Target	**0.003**	0.157	NA	NA	**8.3E−05**	0.150	NA	NA	NA	NA
Teniposide	122819	TOP2	TOP2A	Target	**0.021**	0.442	NA	NA	**2.7E−04**	0.334	NA	NA	NA	NA
Docetaxel	628503	Tu| Tu-stab	TUBB6	Target	0.118	0.413	0.912	0.912	**0.050**	0.140	0.554	NA	NA	NA
Dolastatin 10	376128	Tu	TUBB6	Target	0.528	0.742	0.575	0.575	**0.009**	**0.012**	**0.013**	NA	NA	NA
Eribulin mesilate	707389	Tu	TUBB6	Target	0.947	0.252	0.064	0.064	**0.003**	**0.001**	**0.005**	NA	NA	NA
Ixabepilone	747973	Tu| Tu-stab	TUBB6	Target	0.691	0.884	0.890	0.890	**0.018**	0.183	0.145	NA	NA	NA
Paclitaxel	125973	Tu| Tu-stab	TUBB6	Target	0.809	0.598	**0.039**	**0.039**	**0.003**	**0.004**	**0.018**	NA	NA	NA
Vinblastine	90636	Tu| Tu-frag	TUBB6	Target	0.767	0.244	0.517	0.517	**0.003**	**0.002**	**0.026**	NA	NA	NA
Vinorelbine	608210	Tu| Tu-frag	TUBB6	Target	0.476	0.548	0.074	0.074	**0.029**	**0.008**	**0.021**	NA	NA	NA

PK (protein kinase), STK (Serine threonine kinase), or YK (tyrosine kinase), followed by a gene name indicates an inhibitor of that gene. For full MOA designation, see Supplemental Table 1.

[a] p values from Pearson's correlations calculated for drug activity versus the specified molecular alteration, with significant values (≤ 0.05) in bold. DNA copy number is from array comparative genomic hybridization (aCGH); Illumina DNA methylation is from Infinium HumanMethylation450; Mut_AA and Mut_PFA are from whole exome sequencing (WES); RNA microarray is a 5 Platform Gene Transcript/Average; RNAseq is RNA sequencing from Illumina HiSeq 2000; Protein Mass Spec is mass spectrometry; reverse phase protein array as measured either at MD Anderson or the NCI. All CellMiner downloads are available at (https://discover.nci.nih.gov/cellminer/loadDownload.do). "NA" is no data available. "-" is used in rows for which multiple antibodies exist, to avoid duplication of data in those rows.

[b] All drugs are either Food and Drug Administration approved or in clinical trial.

[c] NSC is National Service Center number.

[d] MOA is mechanism of action. AlkAg is alkylating agent. AM is anti-metabolite. Apo is apoptosis inducer. BCR-ABL is BCR-ABL inhibitor. Df is antifols. Ds is DNA synthesis inhibitor. Ho is hormone. Hg-SMO is Hedgehog-Smoothen - B catenin. Mito is mitochondrial affecting. NonCan is non-cancer, SERM is selective estrogen receptor modulator. TOP1 is topoisomerase 1 inhibitor. TOP2 is topoisomerase 2 inhibitor. Tu is tubulin affecting. Tu-stab is tubulin stabilizing. Tu-frag is tubulin fragmenting.

[e] Genes are both related to the drugs described in the first three columns, and are measured as described for columns six through fifteen.

[f] The genes relationship to the drug, as a target, biomarker as defined at https://www.fda.gov/drugs/science-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling, or has causal linkage. Bio., Tar. is biomarker and target.

[g] DNA copy number data downloaded from CellMiner\Download Data Sets\DNA:combined aCGH\gene summary.

[h] DNA methylation downloaded from CellMiner\Download Data Sets\Download Processed Data Sets\DNA:Illumina 450K methylation\gene_average.

[i] Mut_AA is amino acid changing mutations, including variants defined as missense, nonsense, splice-sense, frameshift, read-through, non-frameshift insertions or deletions. Downloadable at CellMiner\Download Data Sets\Download Processed Data Sets\DNA:Exome Seq\AA changing.

[j] The relevant mutations in EGFR are at L858R and G719 (sensitizing), and T790M (desensitizing) for the EGFR targeting drugs. These mutations are not present, and so the mutational results are presented as "na".

[k] Mut_PFA are protein function affecting mutations. The variant criteria include those from the amino acid changing mutants (footnote i), absent from both the 1000 Genomes and ESP5400, and a sift score less than or equal to 0.05 or a polyphen score greater than or equal to 0.85. May be downloaded from CellMiner\Download Data Sets\Download Processed Data Sets\DNA:Exome Seq\Protein function affecting.

[l] Microarray is the 5 platform transcript z-scores. May be downloaded from CellMiner\Download Data Sets\Download Processed Data Sets\RNA:5 Platform Gene Transcript\Average z scores.

[m] RNAseq is RNA sequencing data. Downloaded from CellMiner. May be downloaded from Data Sets\Download Processed Data Sets\RNA:RNA-seq\composite expression.

[n] Mass spec is mass spectrometry as measured by SWATH. May be downloaded from CellMiner\Download Data Sets\Download Processed Data Sets\Protein:SWATH (Mass spectrometry\Protein.

[o] DTP antibody arrays. May be downloaded from CellMiner\Download Data Sets\Download Processed Data Sets\Protein:Lysate Array\log2.

[p] MD Anderson antibody arrays Identifiers, followed by colon, followed by the Gene Name. _p indicates a phosphorylation. For full names, see Supplementary Table 1.

[q] MD Anderson antibody arrays. May be downloaded from CellMiner\Download Data Sets\Download Processed Data Sets\Protein:Antibody Array\Protein.

treatment of breast cancer [14]. The current study results are consistent with all of these, providing support for the approach implemented herein.

Multiple examples from this study demonstrate novel, or a mix of novel and known relationships with clinical context. For example, ALK (frequently activated by gene rearrangement) positive metastatic non-small cell lung cancer (NSCLC) patients are treated with alectinib. In the current study, it is demonstrated that a lymphoma cell line (SR) overexpressing ALK and harboring an activating ALK-NPM1 fusion has increased sensitivity to the ALK-inhibiting alectinib (Fig. 1F). SR also has significant increased sensitivity to the ALK-inhibiting NSCLC drugs AP-26113, crizotinib, and LDK-378 (Supplemental Table 1) indicating a potential novel use for these drugs [15,16].

The EGFR inhibitors afatinib, erlotinib, gefitinib, and lapatinib are used clinically for lung cancer. Unfortunately, the clinically relevant EGFR mutations used as biomarkers for EGFR inhibitors are not present in these cell lines. However, in their absence, EGFR expression as measured by transcript, total protein, and/or phosphoprotein are significantly (positively) correlated to the four drug responses (Table 1). In the NSCLC subset,

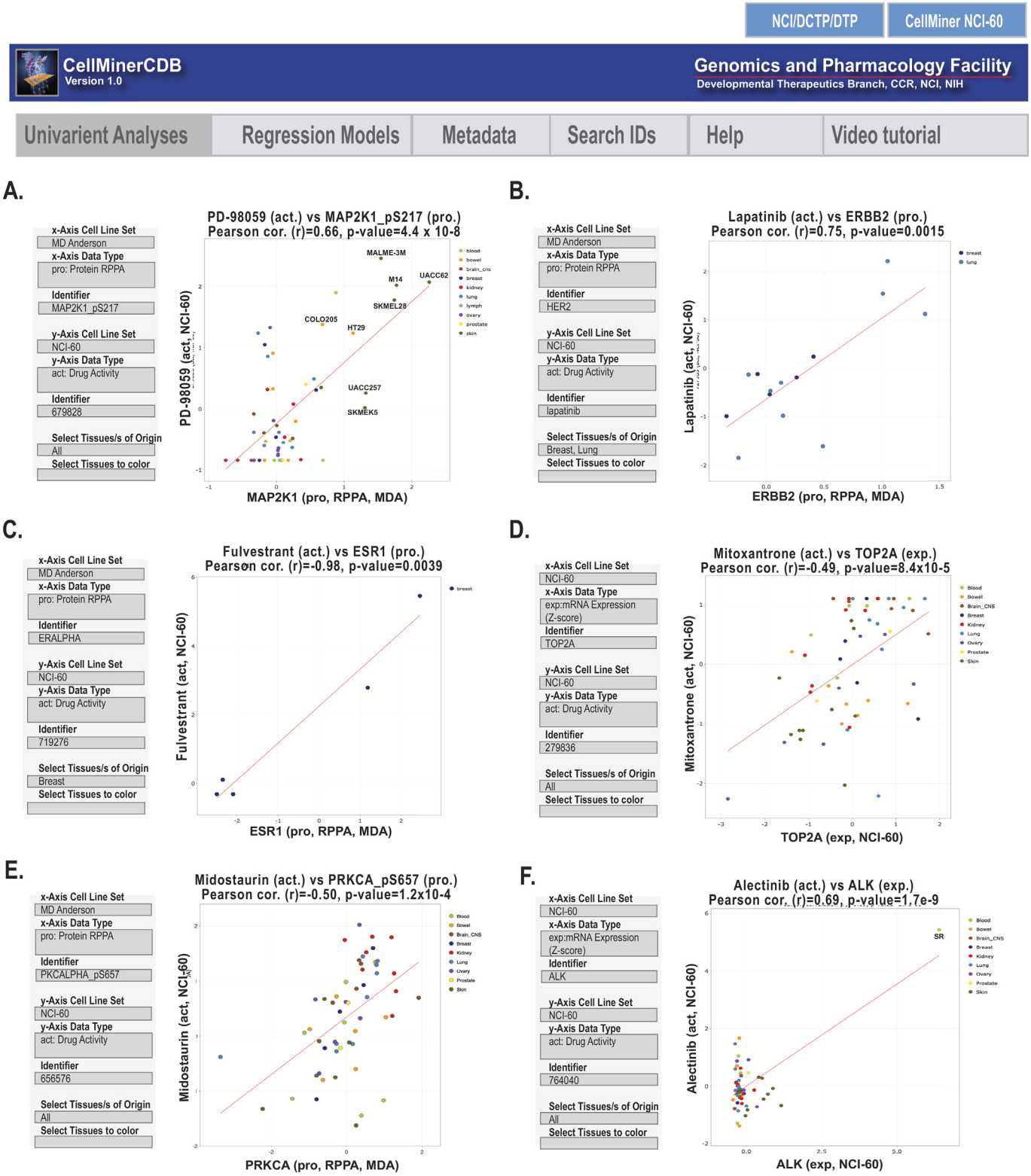

Fig. 1. Examples of CellMinerCDB univariate analyses. A. MAP2K1 phosphoprotein expression levels as measured by RPPA comparison to PD-98059 activity in the NCI-60. The labeled cell lines all have the BRAF V600E mutation. B. ERBB2 (HER2) protein expression as measured by RPPA comparison to lapatinib activity in the combined NCI-60 breast and ovarian cell lines. C. ESR1 protein expression as measured by RPPA comparison to fulvestrant activity in the breast cell lines. ERALPHA is an identifier used by MDA for ESR1. D. TOP2A transcript expression from microarray comparison to mitoxantrone drug activity in the NCI-60. E. PRKCA phosphoprotein expression levels as measured by RPPA comparison to midostaurin activity in the NCI-60. F. ALK transcript expression from microarray comparison to alectinib activity in the NCI-60. In all panels, the input parameters are shown on the left. Each dot is a NCI-60 cell line, with the color code defined by the legend on the right. Regression lines are included in red. X- and y-axes, correlations (r) and p values are as defined within each panel. Exp. is microarray expression using z score, and act is drug activity using z scores from the NCI/DTP. Pro is protein, RPPA is reverse phase protein array as done at M.D. Anderson. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Comparisons of molecular predictors of drugs to pharmacological response.[a]

Molecular measurement[c]	Platform[d]	Gene drug pairs[b]		
		Total[e]	# significant[f]	%
DNA copy number	aCGH	654	65	9.9
DNA methylation	Microarray	598	54	9.0
DNA mutation, AA changing	WES	503	28	5.6
DNA mutation, Prot Fun Aff	WES	416	21	5.0
RNA transcript	Microarray	646	102	15.8
RNA transcript	RNAseq	658	73	11.1
Protein	Mass. spec.	222	31	14.0
Protein, total DTP	RPPA	70	10	14.3
Protein, total (MDA)	RPPA	133	18	13.5
Phosphoprotein (MDA)	RPPA	78	24	30.8

RNAseq is RNA sequencing. Mass spect. is mass spectrometry.; RPPA is reverse phase protein array.
See Table 1 footnotes [a] and [f] through [o] for details and source of data.
[a] See Supplemental Table 1 for detailed drug, target, and gene molecular alteration information.
[b] Drug-gene pairs with direct linkage, that is genes targeted by or causally linked to those drugs.
[c] AA is amino acids. Prot Fun Aff is protein function affecting. DTP is Developmental Therapeutics Program.
[d] aCGH is array comparative genomic hybridization. WES is whole exome sequencing.
[e] Total number of gene-drug pairs for that molecular measurement.
[f] Number of significant Pearson's correlations $p \leq 0.05$, for drug activity versus molecular alteration.

EGFR phosphorylation at Y1068 is significantly positively correlated to response of the same four drugs ($p = 0.00029$, 0.0047, 0.039 and 0.0059 respectively). Lapatinib is also an ERBB2 inhibitor used to treat lung and HER2 over-expressing breast cancer [17]. In the current study, ERBB2 total protein expression was found to be significantly positively correlated to the combined breast and NSCLC cell lines, consistent with clinical application (Fig. 1B). Of potential interest clinically, ovarian cell lines also demonstrate significant positive correlations between lapatinib response and HER2 expression as measured by transcript, total protein, or phosphoprotein ($p = 0.035$, 0.034, and 0.016 respectively), providing candidate biomarkers for ovarian cancers. As prior ovarian clinical trials have failed, improved recognition of responsive sub-populations is desirable [18]. MEK inhibitors are being used or tried clinically for the treatment of melanoma, NSCLC and thyroid cancer, oftentimes as a complement to vemurafenib when BRAF is mutated.

In the current study, MAP2K1 (MEK1) phosphoprotein is significantly (positively) correlated to the pharmacological response of four MEK inhibitors, cobimetinib, PD-98059, selumetinib, and trametinib (Table 1), primarily driven by the presence of the BRAF V600E mutation. BRAF V600E, an FDA-approved biomarker for trametinib, is reported here to serve the same purpose for PD-98059 (Fig. 1A, $p = 4.4 \times 10^{-8}$) selumetinib and trametinib ($p = 3.0 \times 10^{-7}$ and 6.2×10^{-4}, respectively) [19]. Selumetinib and trametinib activity are also both found to be significantly positively correlated to the MAP2K1 pS217S221 phosphoprotein in breast ($p = 5.6 \times 10^{-4}$ and 6.6×10^{-3}, respectively), with the unusual feature of the triple negative cell lines being more responsive than the estrogen positive cell lines, providing a rationale for testing the clinical utility for these drugs.

Midostaurin is a clinical trial drug for acute myeloid leukemia that inhibits PRKCA (among other kinases). PRKCA transcript and protein phosphorylation expression are shown to be significantly (positively) correlated to response (Table 1). The renal cancers are the best responders to midostaurin (Fig. 1E), with significant positive correlations to transcript levels also found for breast ($p = 0.045$), identifying additional potential applications.

TOP2 inhibitors are currently used clinically for multiple cancers. In the current study, TOP2A DNA copy number and transcript expression are shown to have significant (positive) correlation to seven TOP2 inhibitors, amonafide, daunorubicin, doxorubicin, etoposide, idarubicin, mitoxantrone, and teniposide (Table 1). TOP2A is not currently a biomarker for these drugs, but its implementation as such may improve patient response predictability. Tubulin-affecting drugs are currently used across a broad spectrum of cancers. In the current study, TUBB6 expression as measured by microarray, RNAseq, and/or mass spectrometry have significant negative correlation to the tubulin-affecting drugs docetaxel, dolstatin 10, epothilone B, eribulin mesilate, ixabepilone, paclitaxel, vinblastine, vincristine, and vinorelbine (Table 1). Negative correlations show that as the tubulin expression increases, the drug activities decrease. TUBB6 is not currently a biomarker for any of these tubulin-affecting drugs but may be a candidate.

Table 2 provides an overview comparison of predictors of pharmacological response. As the clinical community currently uses DNA mutation as their primary form of biomarker, it is of note that in this comparison, DNA mutation was the worst performer. DNA copy number and methylation both were somewhat better, suggesting a potentially expanded role for their measurements. RNA transcript and total protein expression were both more predictive (than DNA copy number and methylation), arguing for their increased application as well. SLFN11, with its broad applicability for the many workhorse DNA-damaging chemotherapies, is at the top of the list as a candidate biomarker among RNA transcripts. The mass spectrometry measurements appear to also have potential, especially as more genes become assessable due to technical advancement. However, the challenges in implementing mass spectrometry into clinical usage may indicate that its optimal role may currently be the identification of potential biomarkers to be assessed by alternative approaches. The clear best performer, however, was activating or inactivating protein phosphorylations.

Since protein and particularly phosphoprotein data are highly predictive of responses to therapeutic drugs, it is logical to develop Clinical Laboratory Improvement Amendments (CLIA) compliant approaches to quantitatively measure phosphoprotein levels [20]. The Knight Diagnostic Laboratories have recently established a CLIA assay that provides a quantitative readout of oligonucleotide-tagged antibodies using the Nanostring platform [21]. This GeneTrails©Intracellular Protein Signaling panel comprehensively measures cell surface receptors as well as proteins and phosphoprotein markers in the RAS/ERK and PI3K/AKT pathways, and requires only a single formalin-fixed, paraffin-embedded (FFPE) slide. Assays for additional signaling and immune biomarkers are in development that will aid in the precise application of therapies likely to benefit patients.

Different platform measurements of the same gene's transcript or protein levels are not always identical, as seen for ALK RNA as measured by microarray or RNAseq, or for EGFR protein as measured by mass spectrometry or RPPA (Table 1). These differences are attributable to a combination of experimental and platform variation and are not unexpected. Variations in results from different drugs with the same targets result from a combination of experimental reproducibility and drug targeting differences. Pharmacologically, there are always multiple molecular parameters that affect outcome, a point to keep in mind when considering biomarker selection (s). Consequently, it is generally the case that true effects of single molecular changes on pharmacological response do not survive multiple testing corrections, thus the use of uncorrected p values is preferable.

There are, of course, other cell line sets that might have been used. These include, but are not limited to, the Cancer Cell Line Encyclopedia (CCLE, https://portals.broadinstitute.org/ccle), Cancer Therapeutics Response Portal (CTRP, https://portals.broadinstitute.org/ctrp/), and Genomics of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org/). Each of these have the advantage of having more cell lines (~1000), and thus represent more tissue of origin types and subsets, as well as combinations of molecular features. The NCI's cell line set has the more complete set of molecular and phenotypic data, including some 22,000 compound and drug activities, as well as 5355 two-drug combinations. Going forward, users will be well advised to choose those subsets of data that are of greatest relevance to them and to mine drugs responses and genomic parameters across all database using the CellMiner tools (http://discover.nci.nih.gov/cellminercdb) [2,3].

Cancer cell lines used as two-dimensional monocultures suffer from all of the limitations inherent in being simplistic as compared to multiple other systems. Using two-dimensional co-cultures, three-dimensional monocultures and three-dimensional co-cultures each provide increasingly complex systems designed to provide information on cells interactions or response to more complex structure. When jumping to animal models or patients, multiple forms of cellular differentiation, interactions between tissues, monitoring by immune responses, and exposure to hormonal influence (among many other considerations) are added. Nonetheless, 2-dimensional monocultures provide a basis for the increased understanding of human physiology at the cellular and biochemical levels. Despite their relatively simplified design, the combinatorial influences and interactions for their component genes and functional pathways remain largely not understood.

Conclusion

For clinically relevant insight, cell lines remain the cornerstone for hypothesis generation and testing, both for the effects of molecular alterations as well as pharmacological responses. Examples of clinical relevance of cell lines from our group include the introduction of therapeutic strategies based on the novel structure TOP1 inhibitors, LMP776 and LMP744, and the recognition of causal linkage between SLFN11 expression and multiple classes of DNA-damaging drugs, both discovered using NCI-60 cell line data [13,22].

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tranon.2020.100830.

Authors' contributions

WCR provided the concept, design, and execution. FE and SV provided database and programming support. JR provided clinical input and kinase inhibitor expertise. GBM provided protein assay expertise. YP provided clinical input and pharmacological expertise. All provided editorial review and comments.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Our studies are supported by the Intra-mural Program of the National Cancer Institute, Center for Cancer Research (ZIC BC 011509).

References

[1] M. Kalia, Biomarkers for personalized oncology: recent advances and future challenges, Metabolism 64 (2015) S16–S21.
[2] V.N. Rajapakse, A. Luna, M. Yamade, L. Loman, S. Varma, M. Sunshine, F. Iorio, F.G. Sousa, F. Elloumi, M.I. Aladjem, A. Thomas, C. Sander, K.W. Kohn, C.H. Benes, M. Garnett, W.C. Reinhold, Y. Pommier, CellMinerCDB for integrative cross-database genomics and pharmacogenomics analyses of cancer cell lines, iScience 10 (2018) 247–264.
[3] W.C. Reinhold, S. Varma, M. Sunshine, F. Elloumi, K. Ofori-Atta, S. Lee, J.B. Trepel, P.S. Meltzer, J.H. Doroshow, Y. Pommier, RNA sequencing of the NCI-60: integration into CellMiner and CellMiner CDB, Cancer Res. 79 (13) (2019) 3514–3524.
[4] S. Nishizuka, L. Charboneau, L. Young, S. Major, W.C. Reinhold, M. Waltham, H. Kouros-Mehr, K.J. Bussey, J.K. Lee, V. Espina, P.J. Munson, E. Petricoin 3rd, L.A. Liotta, J.N. Weinstein, Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 14229–14234.
[5] T. Guo, A. Luna, V.N. Rajapakse, C.C. Koh, Z. Wu, W. Liu, Y. Sun, H. Gao, M.P. Menden, C. Xu, L. Calzone, L. Martignetti, C. Auwerx, M. Buljan, A. Banaei-Esfahani, A. Ori, M. Iskar, L. Gillet, R. Bi, J. Zhang, H. Zhang, C. Yu, Q. Zhong, S. Varma, U. Schmitt, P. Qiu, Q. Zhang, Y. Zhu, P.J. Wild, M.J. Garnett, P. Bork, M. Beck, K. Liu, J. Saez-Rodriguez, F. Elloumi, W.C. Reinhold, C. Sander, Y. Pommier, R. Aebersold, Quantitative proteome landscape of the NCI-60 cancer cell lines, iScience 21 (2019) 664–680.
[6] S. Patrawala, I. Puzanov, Vemurafenib (RG67204, PLX4032): a potent, selective BRAF kinase inhibitor, Future Oncol. 8 (2012) 509–523.
[7] M.H. Hong, H.R. Kim, B.C. Ahn, S.J. Heo, J.H. Kim, B.C. Cho, Real-world analysis of the efficacy of rebiopsy and EGFR mutation test of tissue and plasma samples in drug-resistant non-small cell lung cancer, Yonsei Med. J. 60 (2019) 525–534.
[8] L.R. Cardon, T. Harris, Precision medicine, genomics and drug discovery, Hum. Mol. Genet. 25 (2016) R166–R172.
[9] P. Jonsson, B.S. Taylor, Transforming biomarker development with exceptional responders, Trends Cancer 4 (2018) 3–6.
[10] W.C. Reinhold, A. Thomas, Y. Pommier, DNA-targeted precision medicine; have we been caught sleeping? Trends Cancer 3 (2017) 2–6.
[11] J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan, A.A. Margolin, S. Kim, C.J. Wilson, J. Lehar, G.V. Kryukov, D. Sonkin, A. Reddy, M. Liu, L. Murray, M.F. Berger, J.E. Monahan, P. Morais, J. Meltzer, A. Korejwa, J. Jane-Valbuena, F.A. Mapa, J. Thibault, E. Bric-Furlong, P. Raman, A. Shipway, I.H. Engels, J. Cheng, G.K. Yu, J. Yu, P. Aspesi, M. de Silva, K. Jagtap, M.D. Jones, L. Wang, C. Hatton, E. Palescandolo, S. Gupta, S. Mahan, C. Sougnez, R.C. Onofrio, T. Liefeld, L. MacConaill, W. Winckler, M. Reich, N. Li, J.P. Mesirov, S.B. Gabriel, G. Getz, K. Ardlie, V. Chan, V.E. Myer, B.L. Weber, J. Porter, M. Warmuth, P. Finan, J.L. Harris, M. Meyerson, T.R. Golub, M.P. Morrissey, W.R. Sellers, R. Schlegel, L.A. Garraway, The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity, Nature 483 (2012) 603–607.
[12] J. Murai, A. Thomas, M. Miettinen, Y. Pommier, Schlafen 11 (SLFN11), a restriction factor for replicative stress induced by DNA-targeting anti-cancer therapies, Pharmacol. Ther. 201 (2019) 94–102.
[13] G. Zoppoli, M. Regairaz, E. Leo, W.C. Reinhold, S. Varma, A. Ballestrero, J.H. Doroshow, Y. Pommier, Putative DNA/RNA helicase Schlafen-11 (SLFN11) sensitizes cancer cells to DNA-damaging agents, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 15030–15035.
[14] Table of pharmacogenomic biomarkers in drug labeling, https://www.fda.gov/drugs/science-research-drugs/table-pharmacogenomic-biomarkers-drug-labeling.
[15] H.G. Drexler, S.M. Gignac, R. von Wasielewski, M. Werner, W.G. Dirks, Pathobiology of NPM-ALK and variant fusion genes in anaplastic large cell lymphoma and other lymphomas, Leukemia 14 (2000) 1533–1559.
[16] M. Ghandi, F.W. Huang, J. Jane-Valbuena, G.V. Kryukov, C.C. Lo, E.R. McDonald 3rd, J. Barretina, E.T. Gelfand, C.M. Bielski, H. Li, K. Hu, A.Y. Andreev-Drakhlin, J. Kim, J.M. Hess, B.J. Haas, F. Aguet, B.A. Weir, M.V. Rothberg, B.R. Paolella, M.S. Lawrence, R. Akbani, Y. Lu, H.L. Tiv, P.C. Gokhale, A. de Weck, A.A. Mansour, C. Oh, J. Shih, K. Hadi, Y. Rosen, J. Bistline, K. Venkatesan, A. Reddy, D. Sonkin, M. Liu, J. Lehar, J.M. Korn, D.A. Porter, M.D. Jones, J. Golji, G. Caponigro, J.E. Taylor, C.M. Dunning, A.L. Creech, A.C. Warren, J.M. McFarland, M. Zamanighomi, A. Kauffmann, N. Stransky, M. Imielinski, Y.E. Maruvka, A.D. Cherniack, A. Tsherniak, F. Vazquez, J.D. Jaffe, A.A. Lane, D.M. Weinstock, C.M. Johannessen, M.P. Morrissey, F. Stegmeier, R. Schlegel, W.C. Hahn, G. Getz, G.B. Mills, J.S. Boehm, T.R. Golub, L.A. Garraway, W.R. Sellers, Next-generation characterization of the Cancer Cell Line Encyclopedia, Nature 569 (2019) 503–508.
[17] Y. Ito, N. Tokudome, T. Sugihara, S. Takahashi, K. Hatake, Does lapatinib, a small-molecule tyrosine kinase inhibitor, constitute a breakthrough in the treatment of breast cancer? Breast Cancer 14 (2007) 156–162.
[18] A.A. Garcia, M.W. Sill, H.A. Lankes, A.K. Godwin, R.S. Mannel, D.K. Armstrong, R.L. Carolla, M.K. Liepman, N.M. Spirtos, E.G. Fischer, K.K. Leslie, A phase II evaluation of lapatinib in the treatment of persistent or recurrent epithelial ovarian or primary peritoneal carcinoma: a gynecologic oncology group study, Gynecol. Oncol. 124 (2012) 569–574.
[19] A. Greco, D. Safi, U. Swami, T. Ginader, M. Milhem, Y. Zakharia, Efficacy and adverse events in metastatic melanoma patients treated with combination BRAF plus MEK inhibitors versus BRAF inhibitors: a systematic review, Cancers (Basel) 11 (2019).
[20] C.L.I.A. (CLIA), https://www.fda.gov/medical-devices/ivd-regulatory-assistance/clinical-laboratory-improvement-amendments-clia.
[21] J. Lee, G.K. Geiss, G. Demirkan, C.P. Vellano, B. Filanoski, Y. Lu, Z. Ju, S. Yu, H. Guo, L.Y. Bogatzki, W. Carter, R.K. Meredith, S. Krishnamurthy, Z. Ding, J.M. Beechem, G.B. Mills, Implementation of a multiplex and quantitative proteomics platform for assessing protein lysates using DNA-barcoded antibodies, Mol. Cell. Proteomics 17 (2018) 1245–1258.
[22] A. Thomas, Y. Pommier, Targeting topoisomerase I in the era of precision medicine, Clin. Cancer Res. 25 (2019) 6581–6589.