Karyotypic “state” as a potential determinant for anticancer drug discovery


*Genetics Branch and ‡Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892

Edited by Albert de la Chapelle, Ohio State University, Columbus, OH, and approved January 7, 2005 (received for review July 30, 2004)

Cancer is a genetic disease caused by genomic instability. In many cancers, this instability is manifested by chromosomal reconfigurations and karyotypic complexity. These features are particular hallmarks of the epithelial cancers that are some of the malignancies most resistant to long-term control by current chemotherapeutic agents. We have asked whether we could use karyotypic complexity and instability as determinants for the screening of potential anticancer compounds. Using a panel of well-characterized cancer cell lines, we have been able to identify specific groups of chemical compounds that are more cytotoxic toward the relatively more karyotypically complex and unstable panel members. Thus, we delineate an approach for the identification of “lead compounds” for anticancer drug discovery complementary to those that are focused at the outset on a given gene or pathway.

Materials and Methods

NCI-60 Drug Discovery Panel. The NCI-60 panel of cancer cell lines was developed by the National Cancer Institute for in vitro anticancer drug screening and includes cell lineages derived from different tissues (lung, renal, colorectal, ovarian, breast, prostate, central nervous system, melanoma, and hematological malignancies). The details of the growth inhibitory assay have been described elsewhere (13–16). Cells are initially inoculated into 96-well microtiter plates in amounts (5,000–40,000 cells per well) based on the doubling times of the individual cell lines. After an initial 24-hour growth period, five concentrations of a test agent as well as a diluent control are added to sets of wells followed by a 48-h growth period. The cells are then sulfurhodamine B stained and subjected to automated read-out of density as measured by absorbance. By using the seven absorbance measurements (time 0 (Tz), control growth (C), and test growth in the presence of drug at the five concentration levels (Ti)), the percentage growth was calculated at each of the drug concentrations levels. Percentage growth inhibition was calculated as [[(Ti – Tz)/(C – Tz)] × 100] for concentrations in which Ti ≥ Tz and [(Ti – Tz)/Tz] × 100 for concentrations in which Ti < Tz.

G1s0 was calculated from [(Ti – Tz)/(C – Tz)] × 100 = 50, which is the drug concentration resulting in a 50% reduction in the net protein increase (as measured by sulfurhodamine B staining) in control cells during the drug incubation. For this analysis, sensitivity was expressed as the negative logarithm of the G1s0.

Since 1990, these cell lines have been exposed to >100,000 compounds. Screening results and chemical structural data on

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: SC, structural complexity; SH, structural heterogeneity; NC, numerical complexity; NH, numerical heterogeneity; MN, modal number; G1s0, 50% growth inhibitory concentration.

†A.V.R., S.L., and G.T. contributed equally to this work.

§To whom correspondence should be addressed. E-mail: kirschil@exchange.nih.gov.
compounds that are not covered by a confidentiality agreement are publicly available (http://dtp.nci.nih.gov). We selected growth inhibition data on 1,429 compounds that had been tested at least four times on all or most of the NCI-60 panel.

Cytogetic Analysis. We performed spectral karyotypic analysis on 59 cancer cell lines from the NCI-60 panel (the MDA-N cell line was unavailable because of restricted access). Details and results of the karyotypic analysis have been described (16). From this panel, 58 independent cell lines were selected for further analysis.

Every cell line in the NCI-60 panel presented karyotypic abnormalities, with notable individual variations among the cell lines at the level of karyotypic complexity and heterogeneity. We defined the complexity of the karyotype on the basis of three factors: ploidy changes, numerical abnormalities, and structural rearrangements. Evaluation of ploidy for each cell line was based on the determination of the MN and the range of the total number of chromosomes among metaphases. Numerical chromosomal complexity (NC) was expressed in relation to the cell line ploidy level, according to the International Conventional System for Human Cytogenic Nomenclature (ICSN) convention (18), and was calculated as a sum of the number of deviations of each specific chromosome from the designated ploidy level.

SC was expressed as the number of different structurally rearranged chromosomes present in two or more metaphases. Chromosomes were counted as structurally abnormal if they contained translocations, deletions, duplications, insertions, inversions, or homogeneously staining regions. Identical rearrangements present in two or more metaphases were designated as clonal; rearrangements present in only one metaphase were designated as nonclonal, according to the ICSN convention (18).

Different metaphases of the same cell line sometimes present different numbers of like chromosomes. We described this variation as NH. It corresponds to the number of cell-to-cell variations of similar centromeres, assessed for each chromosome. Loss of a centromere in only one or two cells or gain in only one cell was not considered in the calculation of NH because of the possibility of mechanical loss or gain during preparation of the metaphase spreads. Any specific centromere displaying a higher number of gains or losses was considered to show variability, and tallied as "one point" in the NH index.

SH was estimated as the number of nonclonal (i.e., present in only one metaphase, according to the ICSN nomenclature) structurally abnormal chromosomes per metaphase. For example, if, among 10 analyzed metaphases, 5 different nonclonal structurally rearranged chromosomes were found, the SH would be calculated as \( S/H = 0.5 \) nonclonal structurally rearranged chromosomes per metaphase.

Statistical Analysis. For statistical analyses we used the software package SAS 8.2 (SAS Institute, Cary, NC). With the five karyotypic parameters (SC, NC, SH, NH, and MN) and the susceptibility to 1,429 drug measured in the panel of NCI-60 cell lines (n = 58 as the effective sample size), we calculated Pearson correlation coefficients for all of the possible relationships (a total of 7,145 = 5 × 1,429 correlation coefficients). The range of values of correlation coefficients was from \(-0.623\) to \(0.524\). Among these correlations, 5,634 (78.9%) were negative, and 1,511 (21.1%) were positive, where the overall mean of correlations was \(-0.14\) with standard deviation 0.17. A positive correlation between sensitivity to a given compound and an increased level of a given karyotypic parameter means that cell lines presenting higher values for that specific parameter would be more sensitive to the growth inhibitory action of that chemical agent. The 95% bootstrap confidence intervals of Pearson’s correlation coefficient were calculated by using the empirical percentiles method with balanced resampling of 10,000 iterations (19). By bootstrap resampling, we avoided parametric assumptions about the distributions of the variables and incorporated possible nonnormal distributional characteristics. With these calculations, we declared a correlation between a karyotypic parameter and a drug significant at the 0.05 α level if the entire 95% bootstrap confidence interval was above (or below) the hypothesized zero when the observed correlation coefficient was positive (or negative). For 10,000 bootstrap iterations with 95% confidence interval, the component of resampling error has a standard error of no more than 0.002. To control the overall family-wise type 1 error rate for testing 7,145 statistical tests applied to the same experiment, we then used a \( P \) value adjustment procedure by choosing a threshold cut-off \( P \) value < 0.001, allowing, in expectation, a maximum of 7 false positives among 483 significant correlations identified by this procedure.

To further validate the significant correlations (bootstrap \( P < 0.05 \)), we performed the bootstrap analysis for the significant Pearson’s correlation coefficients between the karyotypic parameters of the 58 cell lines and their drug sensitivity after leaving out, one at a time, each cell lineage of the nine lineages present. Similar analyses were performed leaving out the six mismatch repair defective cell lines or the 18 p53 wild-type cell lines. An additional control for the latter analysis was performed in which a random set of 18 cell lines was removed (performed with 200 iterations).

Results and Discussion

Distribution of Correlations. An analysis was performed comparing sensitivity data of the 1,429 drug subset (expressed as the negative logarithm of GI_{50}) and each of the following karyotypic parameters: SC, SH, NC, NH, and MN. A positive correlation between sensitivity to a given compound and an increased level of a given karyotypic parameter means that cell lines presenting higher values for that specific parameter would be more sensitive to the growth inhibitory action of that agent.

The overall number of correlation coefficients between karyotypic parameters (MN, SC, NC, SH, and NH) and drug sensitivity values across the NCI-60 is 1,429 × 5 = 7,145. Among those correlations, 5,634 (78.9%) are negative and 1,511 (21.1%) are positive. After Bootstrap testing (two-sided \( P \) value < 0.05), the number of statistically significant correlation coefficients was 2,068: 1,946 negative (94.1%) and 122 positive (5.9%).

NH was the parameter showing the highest number of significant correlations (626 drugs), followed by NC and SH (541 and 487 drugs, respectively). SC and MN present in the different cell lines showed considerably fewer significant correlations than the other parameters (285 and 129, respectively), suggesting that these two factors had relatively less of a global impact on cell line resistance or sensitivity to these 1,429 agents.

The distributions of the significant positive and negative correlations for the different karyotypic parameters are shown in Fig. 1. Structural heterogeneity of karyotypes positively correlated with 52 drugs, whereas SC correlated with only 6. NH and NC correlated positively with 30 and 24 drugs, respectively, and MN with 10.

The larger number of negative correlations between a karyotypic parameter and drug sensitivity is consistent with the general concept that a primary basis for treatment resistance in cancer relates to the genetic diversity and genomic instability of cancer cell populations (20, 21). The fact that the most significant successes in cancer treatment have been achieved for hematologic malignancies that possess a relatively lower degree of karyotypic complexity and heterogeneity may be relevant to this point. Among the drugs reflecting this pattern were thiopurine, mechlorethamine, busulfan, geldanamycin, methotrexate, carbustine, 5-FU, and rifamycin. Many chemotherapeutic drugs, however, did not show any statistically significant correlation with any of the karyotypic parameters, suggesting that sensitivity
to these compounds might not be related to the karyotypic state of the tumor cells. Examples are gemcitabine, mitoxantrone, fludarabine, actinomycin, cisplatin, etoposide, camptothecin, cytarabine (araC), or daunorubicin.

The positive correlations between drug sensitivity and karyotypic complexity and heterogeneity found in this analysis (122 statistically significant positive correlations) provide a distinct opportunity to identify agents that are more active against karyotypically complex and chromosomally unstable cancer cells. Such cells would typically be found in the epithelial cancers, which cause so much therapeutic concern and frustration.

Criteria Applied to Select the Most Reliable Associations. Our strategy was to apply additional filters on the results of this analysis to come up with, as a first pass, conservative, reliable, and biologically meaningful associations between drug activity and karyotypic state. We did this filtering so as to minimize potential false-positive correlations. The following criteria were applied for selection of potential lead compounds among chemical agents represented by 122 statistically significant ($P < 0.05$) positive correlations. First, we identified drugs and compounds significantly associated with more than one karyotypic parameter. We had previously demonstrated that certain of these parameters had some correlation with each other (16). Therefore, if the same compound emerged as positive against two or more distinct and previously related parameters, it passed this initial screen. Second, we identified similar chemical compounds positively and significantly correlated with the same karyotypic parameter. Here, we made the assumption that, if multiple agents of similar chemical structure or mechanism of action independently emerged as active in the presence of a particular parameter, additional study of the class was warranted.

Based on fulfilling one or both of these criteria, we designated 96 of 122 positive correlations as our focus group. Fifty-three chemical compounds were involved in these correlations; 29 of them were associated with more than one karyotypic parameter.

Furthermore, in consideration of the multiple comparisons of correlation coefficients, we looked at a more stringent cutoff of $P < 0.001$ (see Statistical Analysis). Six positively correlated compounds were so identified; all 6 of them had already been included in the focus set of 53 on the basis of the first two screening criteria.

Groups of Compounds. Grouping of 53 selected compounds based on their functional classification or chemical structure allowed us to identify seven groups of compounds with more growth inhibitory activity against karyotypically complex and/or heterogeneous cancer cells than against karyotypically simple and homogenous ones. These groups are listed in Table 1. Compounds that could not be functionally or structurally grouped were categorized in Table 1 as “Others.”

Table 1 lists a broad spectrum of compounds; for the most part, one group of compounds is not related to another by basic chemical structures nor by known mechanism of action. They are a rich resource for identification of “lead compounds” and further studies. These groups of compounds are described in Table 2 based on their representation in the panel of 1,429 drugs and significance of correlations. The list of all 53 selected compounds and the specifically associated karyotypic parameter is available in Table 3, which is published as supporting information on the PNAS web site.

In some cases, differential cytotoxic activity of compounds within our groupings can be linked to known differences in their mechanisms of action. In others, the observed dissimilarity points to the existence of unappreciated differences in the mechanisms of action of closely related compounds. For example, among five cytochalasins included in the panel, three (cytochalasins D, E, and H) had significant correlations of their growth-inhibitory activity with NH of karyotypes in cancer cell lines, but two others (cytochalasins A and B) did not. All cytochalasins bind actin and alter its polymerization, but cytochalasin D is ten times more active than cytochalasin B (22). Additionally, cytochalasins A and B bind to glucose transporters, but cytochalasins D, E, and H do not (23), implying that there are, in fact, distinct subgroups of cytochalasins. Their cytotoxic effects are consistent with this distinction.

Ellipticine analogs in the panel of 1,429 drugs have been divided into four subgroups: derivatives of ellipticine, ellipticinium, olivacin, and olivacinum. Ellipticines showed very different associations with karyotypic parameters than did ellipticiniums. Among 17 ellipticine derivatives, only one had a positive correlation with karyotypic heterogeneity, and the other 16 had negative correlations (Table 2). In contrast, every one of the 16 ellipticinium derivatives was positively correlated with karyotypic heterogeneity. Although the structural backbone of these two classes is identical, ellipticines are neutral whereas ellipticiniums are charged (Fig. 2).

Separation of cancer cell lines on the basis of their responses to ellipticinium and ellipticine members of this class of compounds has been reported by Shi et al. (24). The ellipticiniums, but not ellipticines, were on average more potent against p53 mutant cells than against p53 wild-type ones. Recently Peng et al. (25) reported that a number of ellipticine-like compounds were able to restore mutant p53 transcription function. However, ellipticine and ellipticinium derivatives shared this property. It is unlikely, therefore,
that the specific action of ellipticine derivatives on cell lines with more heterogeneous karyotypes can be explained solely by this targeting of p53 function. Studies on the mechanisms of cytotoxicity of the ellipitcines and analogs indicate a variety of mechanisms of action, including DNA intercalation, inhibition of topoisomerase II, covalent alkylation of macromolecules, and generation of cytotoxic free radicals (26, 27). Because essentially only the charged analogs (ellipticine and olivacinium derivatives) had growth-inhibitory action, including DNA intercalation, inhibition of topoisomerase II, targeting of p53 function. Studies on the mechanisms of cytotoxicity of the ellipitcines and analogs indicate a variety of mechanisms of action, including DNA intercalation, inhibition of topoisomerase II, covalent alkylation of macromolecules, and generation of cytotoxic free radicals (26, 27). Because essentially only the charged analogs (ellipticine and olivacinium derivatives) had growth-inhibitory activity positively correlated with the parameter of karyotypic heterogeneity (Table 2), this correlation suggests that mechanisms specific to these charged derivatives may be involved in their cytotoxic activity against karyotypically unstable cancer cells.

We are not yet saying that these findings necessarily imply directly causal relationships. One alternative explanation would be that an agent targeted a particular cell lineage that just happened to be more karyotypically complex. To explore this possibility, we reanalyzed the data for compounds from each group, sequentially leaving out one and then another of each of the nine lineages in the panel. The essential features of the correlations that we describe and the groups of compounds that we identified were not changed by these additional tests. We also considered the possibility that other cellular “states” like mismatch repair status or p53 gene status might be the critical factors acting as determinants of sensitivity or resistance to these compounds. We therefore performed additional bootstrap analyses similar to the one described for determination of the influence of cell lineage, but now leaving out the six mismatch repair defective cell lines present in the panel or the 18 p53 wild-type cell lines present in the panel. In neither case did our fundamental results change. At least 85% of the correlations that we have described remained significant. In many cases, the value of the correlation actually increased. For the mismatch recognition and repair (MMR) analysis, only 1 correlation out of 26 representative compounds became insignificant; for the p53 study, 4 of 26 became insignificant. The loss of significance for 2 of these 4 compounds seemed to be a function of the markedly reduced number of cell lines (40 vs. 58) that were available for the analysis, because, for these 2 compounds, the random removal of any 18 cell lines (repeated 200 times) resulted in loss of significance >60% of the time.

The NCI-60 panel of cancer cell lines represents a rather distinctive resource. There is no other cell line cohort at this time that has been so systematically subjected to cytotoxicity assay from such a wide-ranging group of compounds. However, independent data supporting our findings have emerged from a comparison of drug sensitivity in two closely related cell lines derived from such a wide-ranging group of compounds. Nevertheless, GI₅₀ data for many chemical agents available for the analysis, because, for these 2 compounds, the random removal of any 18 cell lines (repeated 200 times) resulted in loss of significance >60% of the time.

The NCI-60 panel of cancer cell lines represents a rather distinctive resource. There is no other cell line cohort at this time that has been so systematically subjected to cytotoxicity assay from such a wide-ranging group of compounds. However, independent data supporting our findings have emerged from a comparison of drug sensitivity in two closely related cell lines derived from such a wide-ranging group of compounds. Nevertheless, GI₅₀ data for many chemical agents available for the analysis, because, for these 2 compounds, the random removal of any 18 cell lines (repeated 200 times) resulted in loss of significance >60% of the time.

The NCI-60 panel of cancer cell lines represents a rather distinctive resource. There is no other cell line cohort at this time that has been so systematically subjected to cytotoxicity assay from such a wide-ranging group of compounds. However, independent data supporting our findings have emerged from a comparison of drug sensitivity in two closely related cell lines derived from such a wide-ranging group of compounds. Nevertheless, GI₅₀ data for many chemical agents available for the analysis, because, for these 2 compounds, the random removal of any 18 cell lines (repeated 200 times) resulted in loss of significance >60% of the time.

The NCI-60 panel of cancer cell lines represents a rather distinctive resource. There is no other cell line cohort at this time that has been so systematically subjected to cytotoxicity assay from such a wide-ranging group of compounds. However, independent data supporting our findings have emerged from a comparison of drug sensitivity in two closely related cell lines derived from such a wide-ranging group of compounds. Nevertheless, GI₅₀ data for many chemical agents available for the analysis, because, for these 2 compounds, the random removal of any 18 cell lines (repeated 200 times) resulted in loss of significance >60% of the time.
more sensitive to only 6 compounds (sensitivity to 3 compounds was equal for both cell lines). Moreover, when we compared the sensitivity of these two cell lines to 22 drugs identified in our study as those showing significant negative correlations with NH, the cell line showing less NH (HT-29) now was more sensitive to 16 of them, compared with only 3 for KM20L2 (binomial test; H0: \( P = 0.5 \); 29/44, \( P = 0.013 \)). Given that the genomes of these two cell lines, although closely related, are still distinct, the fact that NH emerges as a determinant of drug sensitivity supports this concept of anticancer drug discovery.

Relationships between karyotypic parameters and sensitivity of cancer cells to certain classes of agents are diagrammed in Fig. 3. We identified several groups of compounds associated mostly with karyotypic instability. For example, derivatives of benzodithiophene-4,8-dione and combretastatin had growth inhibitory activity associated with structural chromosomal instability only. Combretastatins are new promising anticancer agents with a variety of effects, including targeting tumor blood vessels and inhibition of tubulin assembly (28–30). Cytochalasins have been associated with NH, and ellipticinimines have been associated with both structural and numerical heterogeneity. Antibiotics with anticancer activities formed the largest group of drugs identified in this study. Their growth inhibitory activity showed a dichotomous pattern associated either with numerical karyotypic changes (heterogeneity, complexity, and MN) or with SH alone (Fig. 3). Different members of the 2-propenamide group showed a dichotomous pattern of action as well, being associated either with SC or NH. The fuchsin compounds showed correlations of their cytotoxicity with both NC and SH (Fig. 3).

A search for clinical data involving the drugs and compounds identified in this study as being more active against karyotypically complex cancer cells reveals that a number of them have already been assessed in Phase I and II clinical trials (elliptinium, ellipticinum acetate, hydroxymethyllellipticinum, combretastatin A4, chromomycin A3, rapamycin, mitramycin, and depsipeptide). Several trials were concluded without further advance into Phase III studies because of limited response and/or toxicity. Other clinical trials are ongoing (for example, combretastatin A4, rapamycin, and depsipeptide) (31–33). Review of these clinical studies with consideration of possible stratification based on tumor karyotypic parameters might now be warranted.

Conclusions

These data support the thesis that it is possible to discover potential anticancer agents based on association of their activity with a determinant of genomic state rather than with an already identified gene product. In this study, using a correlation analysis and subsequent application of conservative and biologically relevant selection procedures, we have identified chemical compounds that express more growth-inhibitory activity toward cancer cell lines with more complex and/or unstable karyotypes. The successful identification of potential lead compounds based on the analysis of 1,429 well documented agents opens up a quite extensive field of inquiry. These findings suggest that it would now be worthwhile to search drug databases (including the remaining \( \approx \)100,000 compounds tested in the NCI drug discovery program) for additional derivatives or relatives of compounds that we have identified to assess their efficacy on tumors or cell lines manifesting the features of karyotypic complexity and heterogeneity that seem to confer sensitivity. Similarly, a search should be conducted for unique compounds (outside of those included in the panel of 1,429) whose cytotoxic or growth inhibitory action is consistent with an activity correlated with karyotypic complexity and/or instability.

The karyotypic parameters associated with the activities of these compounds may well be markers for underlying genes or pathways that are the true targets of these agents. Extended analysis of these compounds in model systems (e.g., yeast) and study of these cell lines in gene expression assays may provide some leads in this regard. It is equally important, however, to recognize that certain agents may be active against the “state” of complexity or instability itself rather than against any contributing gene product or pathway. It is a plausible hope that the assessment of chromosomal state of a cancer cell population may provide a future guide for the selection of drugs active against aggressive and intractable cancers.

We thank Susan Holbeck for information concerning possible relevant mechanisms of drug action, W. Michael Kuehl and Zhen Wang for helpful discussions, and Ernest Hamel (Developmental Therapeutics Program, National Cancer Institute, Frederick, MD) for providing the KM20L2 cell line.