Genome-wide DNA copy number predictors of lapatinib sensitivity in tumor-derived cell lines

Joel Greshock,[1] Jie Cheng,[2] David Rusnak,[1] Anne Marie Martin,[3] Richard Wooster,[1] Tona Gilmer,[1] Kwan Lee,[2] Barbara L. Weber,[1] and Tal Zaks[3]

[1]Translational Medicine Oncology, GlaxoSmithKline, King of Prussia, Pennsylvania and [2]Department of Biostatistics and Data Management and [3]Medicines Development Center, Oncology, GlaxoSmithKline, Collegeville, Pennsylvania

Abstract

A common aim of pharmacogenomic studies that use genome-wide assays on panels of cancers is the unbiased discovery of genomic alterations that are associated with clinical outcome and drug response. Previous studies of lapatinib, a selective dual-kinase inhibitor of epidermal growth factor receptor (EGFR) and HER2 tyrosine kinases, have shown predictable relationships between the activity of these target genes and response. Under the hypothesis that additional genes may play a role in drug sensitivity, a predictive model for lapatinib response was constructed from genome-wide DNA copy number data from 24 cancer cell lines. An optimal predictive model which consists of aberrations at nine distinct genetic loci, includes gains of HER2, EGFR, and loss of CDKN2A. This model achieved an area under the receiver operating characteristic curve of ~0.85 (80% confidence interval, 0.70–0.98; $P < 0.01$), and correctly classified the sensitivity status of 8 of 10 head and neck cancer cell lines. This study shows that biomarkers predictive for lapatinib sensitivity, including the previously described copy number gains of EGFR and HER2, can be discovered using novel genomic assays in an unbiased manner. Furthermore, these results show the utility of DNA copy number profiles in pharmacogenomic studies. [Mol Cancer Ther 2008;7(4):935–43]

Introduction

Lapatinib is a potent, orally active dual kinase inhibitor of the epidermal growth factor receptor (EGFR) and HER2 tyrosine kinase receptors. In tumor cells that overexpress either EGFR or HER2, lapatinib treatment results in growth arrest and cell death (1), suggesting that it may be possible to predict clinical response based on a tumor's genetic profile. Although larger *in vitro* screens confirm strong associations between EGFR and HER2 amplification with lapatinib responsiveness, some tumor models seem to respond in the absence of these markers (2, 3). This observation indicates that additional proteins may cooperate with HER2 in promoting lapatinib responsiveness, thereby forming the basis of a more inclusive means of predicting clinical response. Such proteins may include the linear cascades that constitute the mitogen-activated cascade, the stress-activated protein kinase cascade, protein kinase C, and the Akt pathway (4).

Altered DNA copy number is a hallmark feature of the cancer genome. A number of tumor suppressor genes have been mapped to homozygous deletions, including RB1 and SMAD4/DPC4 (5, 6). Similarly, oncogenic copy number amplifications have been described in such loci as *PDGFRA, EGFR*, and *HER2* (7–9). Until recent technological advances, the assembly of more complete lists of genes whose copy number is altered in cancer has been limited by nonsystematic techniques (10). However, the development of high-resolution microarray-based genome-wide DNA copy number has remedied this limitation. Although earlier BAC clone–based platforms provided a resolution of ~1 Mb (11), more recent oligonucleotide-based platforms such as Representational Oligonucleotide Microarray Analysis decreased probe spacing considerably (12). Single-nucleotide polymorphism arrays (SNP chips) represent a type of high-density oligonucleotide-based microarrays that have been adapted for DNA copy number analyses (13) and perform comparably with traditional two-channel arrays (14).

The increasing role of pharmacogenomics in oncology compound development and treatment has closely paralleled the advent of high-throughput genomics platforms. A major attraction of using genome-wide assays is the increased ability to more comprehensively characterize genomic mechanisms of drug sensitivity. For example, gross patterns in gene expression effectively show that *in vitro* response to a farnesyltransferase inhibitor of the Ras pathway is directly correlated with Ras pathway activity (15). Translational research focuses on the discovery of rational predictive markers derived from such genomic assays that could then be used for response enrichment in a clinical setting. A growing number of these studies include an effective predictor of bortezomib response in myelomas that was retrospectively calculated and validated from gene expression microarrays run on patient tissue collected from multiple independent clinical sites (16). Although genome-wide expression profiling is currently the most widespread platform for pharmacogenomic studies, the mapping of

Received 9/18/07; revised 12/19/07; accepted 1/10/08.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Tal Z. Zaks, Translational Medicine Oncology, GlaxoSmithKline, 1250 South Collegeville Road, UP 4W-4230, Collegeville, PA 19426. Phone: 610-917-5124; Fax: 610-917-4830. E-mail: Tal.Z.Zaks@gsk.com

doi:10.1158/1535-7163.MCT-07-2072

specific DNA copy number alterations to multiple pathways is also effective in delineating markers associated with the responsiveness to targeted therapy (reviewed in ref. 17). Fixed genetic alterations such as DNA copy number gains and losses are appealing for clinical study because (*a*) they are relatively stable *in vivo* over time, and (*b*) they are faithfully modeled in cell lines *in vitro* (18, 19). For example, copy number studies have associated specific DNA copy number alterations in ovarian carcinomas with carboplatin response (20), whereas amplifications of the multidrug resistance gene *MDR1* are correlated with therapeutic resistance (21, 22).

The growing use of targeted therapies in cancer treatment reflects an increasing knowledge of key oncogenic pathways. Difficulties in predicting response to targeted therapies are likely a consequence of the limited global knowledge of altered genetic pathways in each cancer. Also, unique genomic profiles observed in each cancer suggests the existence of many more contributing pathways than previously thought. For example, contrasting sets of recurring alterations found to differentiate between certain types of lung cancers indicate that common oncogenic pathways may be subject to unique alterations (23). Collectively, these results suggest that developing genome-wide strategies to discover predictive models of therapeutic response can be more sensitive than traditional single-gene approaches. In this study, we present a model that takes advantage of global characterizations of DNA copy number to accurately predict *in vitro* responses to lapatinib. The model, which emphasizes high-level copy number alterations such as homozygous deletions and oncogenic amplifications, rediscovers the *HER2* locus, which is already a known predictor of clinical activity. Furthermore, this model discovers novel markers of response that are capable of classifying a panel of non–HER2-amplified head and neck cancers. Eventually, these DNA alterations may provide clinically relevant markers that allow the selection of patients most likely to benefit from therapy.

Materials and Methods

Cell Lines, Cell Culture, and Reagents

A panel of 34 cell lines of mixed histology were obtained from American Type Culture Collection and were cultured in RPMI 1640 supplemented with 10% heat-inactivated fetal bovine serum, 2 mmol/L of glutamine. Similarly, for the purpose of characterizing a diploid reference in the copy number analysis of the tumor-derived cell lines, DNA from a panel of 20 cell lines composing a human genome diversity panel were obtained (Coriell Institute, Camden, NJ; ref. 24). For tumor-derived cell lines, DNA was extracted before any compound treatment (i.e., untreated cells).

Proliferation Assays

Cells were plated into 24-well plates at a density of 2×10^5 to 5×10^5 and grown in cell line–specific medium without or with increasing concentrations of lapatinib (ranging between 0.008 and 10 μmol/L). Cells were harvested by trypsinization on day 7 and counted using a particle counter (Z1, Beckman Coulter, Inc.). Growth inhibition was calculated as a percentage of the untreated controls. Experiments were carried out twice or thrice in duplicate for each cell line. The log of the fractional growth inhibition was plotted against the log of the drug concentration, and the IC_{50}s were interpolated from the resulting linear regression curve fit (Calcusyn, Biosoft). IC_{50} values <1.0 μmol/L were considered as sensitive, 1.0 to 2.5 μmol/L as intermediate, and values >2.5 μmol/L were considered resistant. Proliferation data for several of these lines have been previously presented in ref. (2).

SNP Chip/Transcript Analysis

DNA was extracted from each line using Mini DNeasy kit (Qiagen, Inc.) and subsequently purified. For the 24 cell lines of mixed tumor type composing the gtrainingh panel, two aliquots (250 ng each) were digested with the restriction enzyme *Xba*I or *Hin*dIII (New England Biolabs), whereas a similar quantity of material was digested with *Sty* and *Nsp* enzymes for the 10 head and neck lines in the validation set. Digested DNA was subsequently ligated to an adaptor and amplified by PCR using Platinum Pfx DNA Polymerase (Invitrogen), yielding a product of ~250 to 2,000 bp. For each enzyme digest, PCR was carried out in four 100-μL aliquots, pooled, purified, quantified, normalized to 40 μg/45 μL, and fragmented with DNase I to yield a size range of ~25 to 200 bp. The fragmented products of the training set and test set were then labeled, denatured, and hybridized to the Affymetrix 100K and 500K chip, respectively (Affymetrix Inc.). Upon completion of hybridization, each assay was washed and stained using Affymetrix fluidics stations. Image data were acquired using the GeneChip Scanner 3000. For all 34 cell lines, transcript abundance was quantified in triplicate by using the Affymetrix U133 Plus2 GeneChips. First, cell lines were plated in triplicate and lysed in TRIzol. Lysates were captured with chloroform and purified using QIAGEN RNeasy Mini Kit (QIAgen, Inc.). cDNA was prepared from 5 μg of total RNA using the Invitrogen SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen, Inc.) and amplified using the ENZO BioArray High-Yield RNA Transcript Labeling Kit (Enzo Biochem, Inc.). Finally, the samples were fragmented and hybridized to the HG-U133Plus2 GeneChips, stained, and scanned according to the manufacturer's protocols.

Data Analysis

All SNP chip images (''CEL files''), were extracted using the Affymetrix Genotype software, and analyzed using a modification of the RLMM implementation in the R programming language (25). Specifically, a quantile intensity normalization was carried out for the CEL file–extracted intensity scores. Next, SNP-wise copy number ratios ($\log_2$ scale) were calculated for all cancer cell lines by dividing the SNP intensity score by the respective median intensity score for the lymphoblastic reference panel. Subsequently, data were adjusted under the assumption that the median copy number for all samples was diploid.

Finally, copy number inferences were made by circular binary segmentation to reduce noise (e.g., from unmasked complex sequences in the target) and provide a consensus score for all regions of the genome based on at least two underlying SNPs (26). For all subsequent analysis, the genome was divided into 100 kb units in which each was assigned the circular binary segmentation–extracted value for every cancer cell line. Transcript abundance was estimated by normalizing all probe signal intensities to a value of 150 using the mas5 algorithm in the Affymetrix Microarray Analysis Suite 5.0. For subsequent analyses, the average intensity of each probe was used.

Model Building and Classification

An externally cross-validated prediction model, using methods described in ref. (27), was constructed using 24 tumor-derived cell lines of mixed histologies. The predictive model framework is solely determined by the three variables in the feature selection step. The feature selection step included (*a*) converting $\log_2$ segment scores to zero if they were less than the data threshold t; (*b*) calculating the mean difference between phenotypic groups for each feature (100 kb in this case) and filtering out features with a small mean difference by applying a mean difference threshold d, and also removing neighboring redundant features of the included features; and (*c*) estimating the P values of the mean difference of selected features using 10,000 random permutations and subsequently filtering out insignificant features by applying a P value threshold. The remaining features were then used to build a weighted-voting model as described in ref. (28). To find the optimal model, a grid search was carried out in the following search space: data threshold t between 0.2 and 0.5; mean difference threshold d between 0.2 and 0.5; P value threshold between 0.01 and 0.05.

Receiver Operating Curve Analysis, Bootstrapping, and External Validation

To evaluate the performance of a model, leave-one-out cross-validation was carried out (with a new feature selection for each cross-validation step) and area under the receiver operating curve (AUROC) was calculated. Cross-validation AUROC showed that model performance was generally good when the variables were in the following ranges: t between 0.3 and 0.5, d between 0.35 and 0.45, and P between 0.03 and 0.05. The optimal model had the variables $t = 0.45$, $d = 0.4$, and $P = 0.03$. Leave-one-out cross-validation showed that the performance of the optimal model had an AUROC score of 0.85. To establish confidence intervals, 200 runs of bootstrapping were conducted and an AUROC of each run was calculated using leave-one-out cross-validation. The mean of the AUROC was 0.85, with an 80% confidence interval (0.70–0.98). To estimate the statistical significance of the optimal model, the phenotypic grouping information of the 24 cell lines was randomly permuted 500 times. Less than 1% of the random data sets had an AUROC score of 0.85 or above. This showed that the optimal model was significant at $P < 0.01$. Finally, an independent panel of 10 head and neck cancer cell lines were classified using the loci and variables in the optimal model derived from the training data. All lines in this test set lack amplification of the *HER2* locus (the most significant predictor in the training set), which addresses whether additional predictors of sensitivity were identified.

Results

Detection of Lapatinib Sensitivity and Copy Number Alterations in a Mixed Panel of Cancer Cell Lines

The effects of lapatinib on cell growth were studied in 24 human cancer cell lines of mixed histology. These cell lines exhibited a wide range of sensitivities based on their IC_{50} values measured in the proliferation assays (range, 0.025–7.71 μmol/L; Table 1). Several showed hypersensitive phenotypes, having IC_{50} scores of <0.1 μmol/L.

Many cancer cell lines in this panel are subject to complex rearrangements. For example, 1q in the breast cancer cell

Table 1. Sensitivities of 34 tumor-derived cell lines

	Cell line	Tissue of origin	Lapatinib IC_{50} (μmol/L) ± SE
Training set	BT474	Breast	0.025 ± 0.004
	NCI-N87	Gastric	0.028 ± 0.002
	HN5	Head/neck	0.029 ± 0.005
	SKBr3	Breast	0.032 ± 0.005
	MDA-MB-453	Breast	0.079 ± 0.025
	A431	Skin, epidermoid	0.23 ± 0.06
	NCI-H322	Lung	0.92 ± 0.02
	BxPC3	Pancreatic	1.41 ± 0.13
	SKOV3 ATCC	Ovarian	1.25 ± 0.11
	786O	Renal	1.82 ± 0.17
	MDA-MB-468	Breast	2.32 ± 0.21
	DLD-1	Colon	3.63 ± 0.86
	LNCaP	Prostate	4.69 ± 0.52
	MCF7	Breast	4.82 ± 1.14
	T47D	Breast	4.83 ± 1.12
	A549	Lung	4.98 ± 0.49
	HT-29	Colon	5.24 ± 1.01
	RKO	Colon	5.48 ± 0.52
	MX1	Breast	5.55 ± 1.55
	HCT 116	Colon	5.87 ± 2.04
	SKMEL28	Melanoma	5.90 ± 0.37
	MDA-MB-231	Breast	7.01 ± 0.51
	PC3	Prostate	7.15 ± 2.24
	Colo 205	Colon	7.70 ± 1.03
Validation set	SCC-15	Head/neck	0.171 ± 0.003
	FaDu	Head/neck	0.216 ± 0.025
	SCC-25	Head/neck	0.222 ± 0.065
	CAL-27	Head/neck	0.308 ± 0.107
	SCC-9	Head/neck	0.331 ± 0.084
	SCC-12	Head/neck	0.674 ± 0.059
	SCC-13	Head/neck	1.063 ± 0.196
	SCC-4	Head/neck	1.851 ± 0.387
	RPMI2650	Head/neck	2.931
	SW579	Head/neck	3.074 ± 0.277

NOTE: This includes both predictive model training and validation cell line panels.

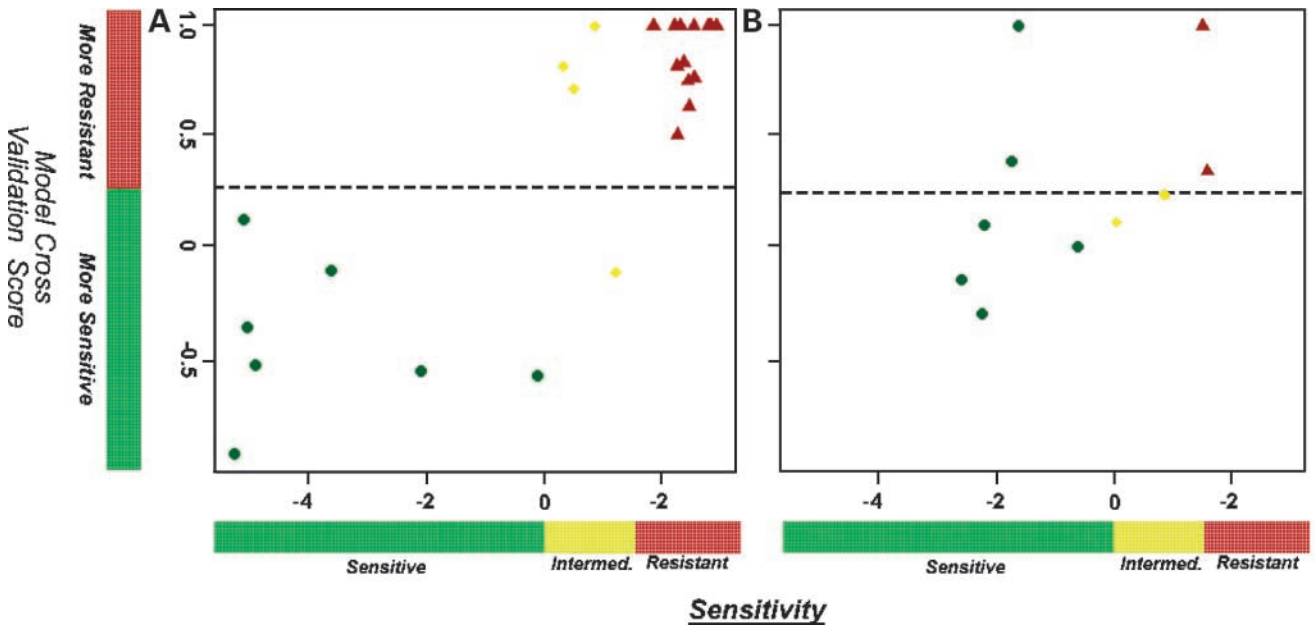

Figure 1. A panel of cancer cell lines show correlations between model scores and measured sensitivity using IC_{50} as a metric. **A,** cross-validation scores for the training set. **B,** classification scores for 10 HER2-negative head and neck cell lines.

line BT474 contains more than 20 calculated copy number transitions, and includes an amplification of the *AKT3* gene. Other recurring high-level changes are apparent, including homozygous losses of the *CDKN2A* locus, as well as high-level amplifications of the oncogene c-MYC. Furthermore, copy number aberrations seen in this study seem to be concordant with those noted in earlier studies. For instance, the breast cancer cell line SK-BR-3 shows a balanced loss of 8p and an amplification of the *HER2* locus, which are alterations noted in previous studies of this cell line (11, 12, 33). Two additional cell lines harbored DNA amplifications of the *HER2* locus (NCI-N87 and BT-474), whereas a total of six lines were EGFR-amplified (HN5, A431, MDA-MB-468, SCC-15, Calu-27, and SCC-4). In all these cases, elevated protein abundance was confirmed by Western blot analysis (ref. 3; Supplemental Tables S1 and S3).[4] All raw data microarray are available for download in the National Center for Biotechnology Information's Gene Expression Omnibus repository (series ID GSE9585; ''reference samples'' associated with Gene Expression Omnibus series ID GSE7822).

Model Building

A weighted voting predictive model for lapatinib sensitivity was constructed using genome-wide DNA copy number measurements derived from 100K Affymetrix SNP Chips from 24 tumor-derived cell lines of mixed histology. Model performance, evaluated using leave-one-out cross-validation, was found to be significant ($P < 0.01$). This significance is reflected in the strong relationship between cross-validation scores for each cell line and its relative sensitivity (Fig. 1A; ROC cross-validation analysis shown in Supplemental Fig. S1; Supplemental Table S1).[4] Notably, cell lines with intermediate sensitivity to lapatinib spanned the boundary between resistant and sensitive.

Nine distinct genomic regions were identified as contributing to the optimal model (Fig. 2). Regions encompassing the EGFR and HER2 genes (7p11.2 and 17q12, respectively) were among these loci (Fig. 2). Several other cancer-related genes map to the loci in the optimal model. Included in the model are deletions of the familial melanoma locus CDKN2A (9p21.3), DNA gains of the candidate oncogene ZNF217(20q13.2) as well as losses of the candidate tumor suppressor gene PTPRD (9p23; Table 2; additional details for all genes available in Supplemental Table S2).[4] There seemed to be multiple contributing rearrangements in each cell line, in which those with the highest sensitivity had rearrangements in all nine genomic regions of the optimal model, whereas less sensitive lines harbored fewer alterations.

Classification of Head and Neck Lines

To independently assess the optimal model, an additional data set consisting of 10 head and neck cell lines was analyzed for sensitivity to lapatinib and genome-wide copy number alterations. The head and neck cell lines had a range of sensitivities to lapatinib, including six sensitive cell lines ($IC_{50} < 0.10$ μmol/L) and two resistant cell lines ($IC_{50} > 2.5$ μmol/L; Fig. 3; Supplemental Table S3).[4] Although, high EGFR amplification is prevalent in head and neck tumors and is a poor prognostic indicator (29),

[4] Supplementary material for this article is available at Molecular Cancer Therapeutics Online (http://mct.aacrjournals.org/).

HER2 amplifications have not been previously reported. As expected, several lines in this panel (3 of 10; 30%) harbored EGFR amplification (more than five copies), whereas none seemed to have amplification of HER2. Predictive model scores that were generated for each head and neck cell line seemed to be generally related to its sensitivity to lapatinib (Fig. 1B; Supplemental Table S3).[4] Given score boundaries derived from cross-validation, the optimal model successfully predicted two of two of the lapatinib-resistant cell lines, and four of six of the lapatinib-sensitive cell lines, whereas both of the intermediate-sensitivity cell lines had scores close to the boundary between resistant and sensitive (all intermediate to those correctly classified as most sensitive or resistant). Specifically, sensitive lines scored <0.15 whereas resistant cell lines scored >0.3. Only two cell lines, CAL27 and SCC9, had scores discordant with their measured sensitivities; this may be due to the presence of additional alterations that confer resistance.

In order to determine the effects that DNA copy number alterations may have on transcript levels, a comparison was made between sensitive cell lines versus the resistant cell lines (including both training and test panels) for the genes mapping to the regions composing the optimal performing predictive model. Statistical significance was assessed by simple t test, where a P value threshold of 0.05 was applied. For those probes expressed at a minimum of 50 units in at least one sensitivity group, a total of 55 probes (67.1%; $n = 82$) representing 31 unique genes (70.5%; $n = 44$) were differentially expressed (Table 2). The fold change (mean of sensitive/mean of resistant) of 54 of 55 (98.2%) differentially expressed probe sets was concordant with the DNA copy number changes exhibited by the regions to which they map (30 of 31 genes; 96.8%; Supplemental Table S4).[4] This high concordance between DNA copy number and transcript abundance is highly significant for the genes mapping to 17q12, in which 20 of 22 (90.9%) probe sets showed differential expression and fold changes concordant with the DNA copy number gains seen in the sensitive group.

Discussion

By using full genome copy number profiles and associated sensitivities to lapatinib in a panel of 24 cell lines of mixed histologic origins, a predictive model was constructed for the purpose of identifying the likelihood of response in untested cell lines. Both the sensitivity categorizations and model variables were subsequently applied to an independent panel of non–HER2-amplified head and neck cancer cell lines ($n = 10$). Both leave-one-out cross-validation and the independent validation set suggest that copy number profiles can efficiently predict cell line sensitivity to lapatinib even in the absence of HER2 amplification.

The most significant finding of this study is the unbiased rediscovery, from genome-wide copy number data, of the EGFR and HER2 amplicons as predictors of lapatinib sensitivity. This is consistent with previous studies demonstrating that lapatinib treatment leads to the inhibition of EGFR and HER2, and consequently results in growth arrest and/or apoptosis in a variety of EGFR and HER2-dependent tumor cell lines (1, 30). This is also concordant with the correlation between HER2 expression and the

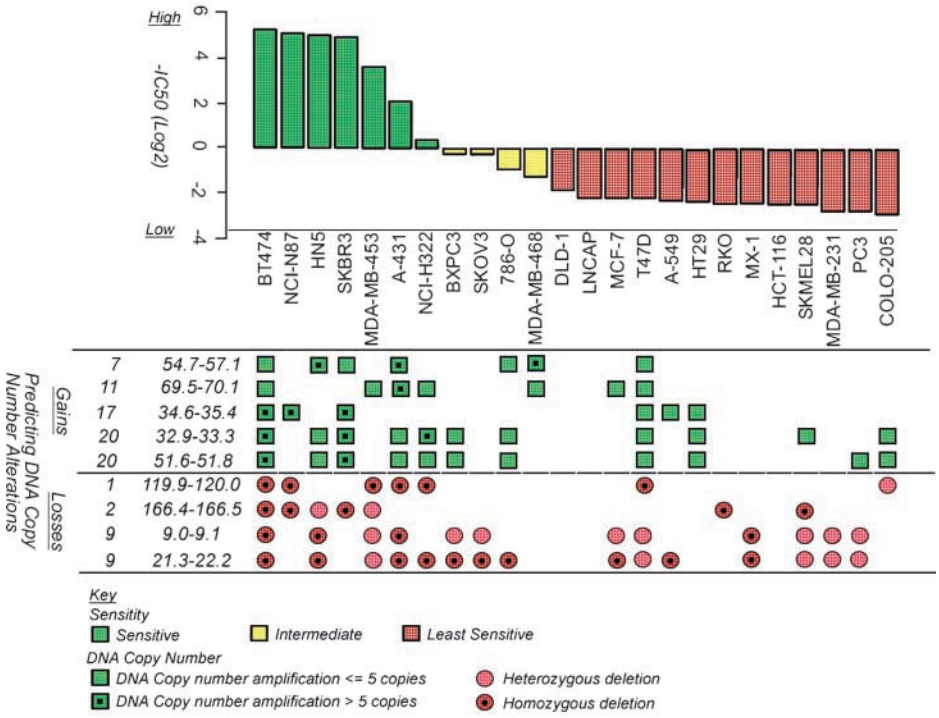

Figure 2. DNA alterations to nine distinct genomic regions that are associated with lapatinib sensitivity compose the optimal calculated model. These regions show high rates of alteration in the most sensitive cell lines and much lower rates of alteration in the most resistant lines.

Table 2. Features used for prediction in the optimal model and genes mapped to their respective genomic regions.

	Region	Start (Mb)	End (Mb)	Genes
Chromosomal gain	7p11.2	54.7	57.1	***EGFR, ECOP, LANCL2***, *MRPS17*, ***GBAS, PSPH, CCT6A,*** *SUMF2, PHKG1,* ***CHCHD2***
	11q13.3	69.5	70.1	*TMEM16A,* ***FADD, PPFIAI, CTTN, SHANK2***
	17q12	34.7	35.5	***FBXL20***, *CAB2,* ***PPARBP***, *PPP1R1B,* ***STARD3, TCAP, HER2 GRB7***, *ZPBP2,* ***GSDML, PSMD3***, *CRK7, C17orf37*
	20q11.22	32.9	33.3	*GGTL3, GSS,* ***ACAS2***, *MYH7B,* ***TRPC4AP***, *PROCR,* ***C20orf44***
	20q13.2	51.6	51.8	*ZNF217*
Chromosomal loss	1p12	119.9	120	*TBX15, WARS2, HAO2, HSD3B1, HSD3B2,* ***PHGDH***
	2q24.3	166.4	166.5	*GALNT3*
	9p24.1	9	9.1	***PTPRD***
	9p21.3	21.3	22.2	*IFNB1, IFNA7, IFNA5,* ***KLHL9***, *IFNA13, IFNA1, CDKN2A,* ***MTAP***

NOTE: Genes in boldface were differentially expressed ($P < 0.05$), with fold change between sensitive and resistant cell lines concordant with that of DNA copy number.

ability of lapatinib to inhibit HER2, Raf, Akt, and Erk phosphorylation in a panel of breast cancer cell lines (2). Notably, other known members of the HER2 signaling pathway, such as PTEN and AKT, were not identified as useful predictors.

Although it can be expected that copy number increases of target genes are competent predictors of drug sensitivity in cell lines, their recognition from an unbiased genome–wide association study has several important implications. First, it suggests that molecular profiling using genome-wide assays can be a powerful and valid means of identifying biomarkers predictive of clinical response. This has been shown even for less selective agents, such as molecular profiles that predict for taxane activity in the treatment of breast cancer (31). Second, genome-wide profiles could point to additional genetic aberrations that further influence the efficacy of an agent in the absence of (or in conjunction with) primary markers; the fact that clinical responses may occur in only ~25% of an already-enriched population (32) highlights the potential existence of additional predictive markers. Finally, given that genome-wide copy number aberrations in cell lines have been shown to correlate with primary tumor tissues in individual tumor types (e.g., breast cancer; refs. 33, 34) as well as across histologies (18), finding such loci *in vitro* can yield clinically relevant testable hypotheses. In the current study, discerning both target genes of lapatinib (EGFR and HER2) from a relatively limited number of cell lines of mixed histologies infers the potential to delineate molecular markers that span tumor types. The clinical activity of lapatinib against HER2+ breast cancer has recently been proven (35), and there are ongoing studies in patients with HER2+ tumors of other histologies (i.e., gastric, bladder, and ovarian cancers). Although substantial activity has not been seen in unselected patients in phase II studies of lung and colorectal cancers (36), a retrospective analysis of a phase III study in renal cell cancer suggests benefit in patients whose tumors markedly overexpress EGFR (amplification of HER2 is extremely rare in renal cell cancer; ref. 37).

The novel *in vitro* predictors observed in this study are partially composed of genes with established roles in oncogenesis, apoptosis, and tumor suppression. The *CDKN2A* locus (9p21.3) contains three related genes (ARF, p15INK4b, and p16INK4a) that encode distinct tumor suppressor proteins, all of which have well-established roles in tumor suppression (38). Both the loss of heterozygosity and complete DNA deletion of CDKN2A

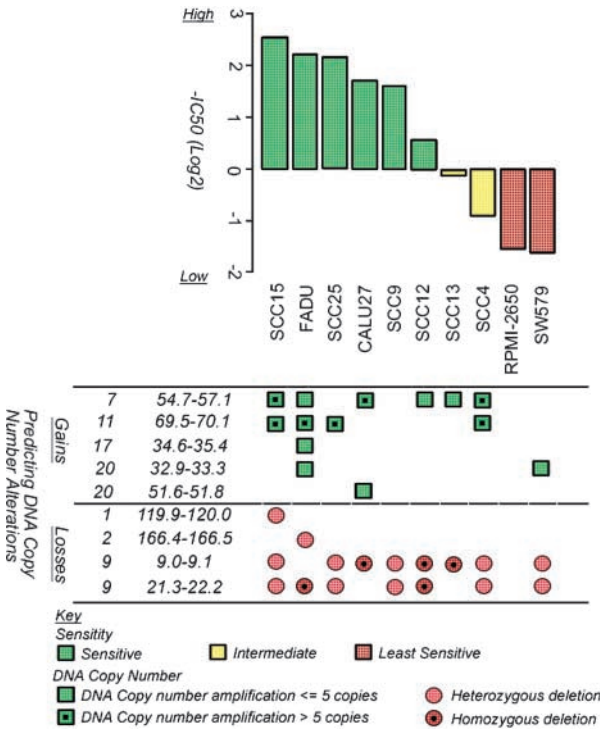

Figure 3. Alterations observed in 10 HER2-negative head and neck cell lines seem largely concordant with those in the training set. Sensitive lines seem to have an accumulation of alterations to these regions, whereas resistant lines harbor fewer alterations.

are common occurrences in a wide range of tumors including melanoma (39), leukemias (40), and non–small cell lung cancer (41). The recurring amplification of ZNF217 (20q13.2) has also been identified in studies of breast cancer and ovarian carcinomas. Specifically, ZNF217, a Kruppel-like zinc-finger protein, which acts as an oncogene targeting E-cadherin, Akt phosphorylation, and cell survival (42–45). The FADD gene (11q13.3), the amplification of which was associated with lapatinib sensitivity in this study, is frequently mutated in human malignancies. FADD seems to play a central role in the control of programmed cell death, as loss of expression in tumor cells confers survival and growth advantages (46). PTPRD (9p23), another candidate tumor-suppressor gene, has been identified by high-resolution genomic analysis in both cell lines and primary tumors for neuroblastomas (47) and non–small cell lung cancer (41, 48), suggesting an important contribution to carcinogenesis in these tumor types.

A high degree of concordance (98.2%) was found between DNA copy number and transcript abundance of probes mapping to DNA model regions. Although the direct effects of DNA copy number have been well documented (e.g., ref. 49), this result suggests that computational models derived from transcript data may be reverse-engineered in a way that allows clinical testing by DNA aberrations (and vice versa). Furthermore, this has functional implications for the genes that reside in the copy number regions, suggesting that the proteins encoded by these genes, as is the case with EGFR and HER2, may affect the clinical responsiveness to lapatinib.

Recent studies have shown the *in vitro* sensitivity of EGFR-amplified head and neck cancers to lapatinib (3). As expected from previous studies, HER2 amplification was absent in this panel of head and neck cell lines, and provided no prognostic significance (3). Thus, the nine-locus predictive model highlights the utility of additional genomic alterations in predicting lapatinib response in head and neck cancer cell lines. Limited array-CGH studies of head and neck cancers corroborate many of these loci as candidate genes for driving the development and progression of head and neck cancer (50, 51). Collectively, the sensitivity predicted from a limited set of fixed genetic alterations in a subset of head and neck cell lines suggests the potential for clinical benefit. A recent phase II study of the combination of lapatinib with chemotherapy and radiation in head and neck cancer showed promising results (52), and phase III studies are ongoing.

Patient enrichment strategies have guided the development of several cancer therapies. In the cases of trastuzumab with HER2 and of imatinib with BCR-ABL, PDGFRA, or KIT, clinical activity has correlated best with the presence of specific fixed genetic aberrations (e.g., ref. 53). An immunohistochemical assay for detecting EGFR positivity has been included in the development of EGFR inhibitors, at least in CRC; however, to date, this assay has not been an adequate predictor of clinical activity (54). Furthermore, more recent data suggest that the best predictor is the absence of a downstream mutation in KRAS. Whether this is due to a biological reason (i.e., the presence of a genetic aberration encodes for oncogenic addiction; ref. 55) or a technical one (whereby quantifying the expression of a specific protein is less reliable than quantifying the relative amounts of a given gene) is unclear. Given this and the technical limitations in proteomic profiling, the strategy of many recent pharmacogenomic studies has been to quantify relationships between transcript abundance and clinical response to cancer therapy (15, 16, 56). To a lesser extent, DNA alterations have been used for this purpose (e.g., ref. 20). As with other genome-wide profiling platforms, microarray-based DNA copy number studies provide an unbiased, parallel, high-throughput interrogation of many genes; however, pharmacogenomic studies of DNA copy number may be advantageous for several reasons. First, genetic aberrations seem to be relatively stable *in vivo* over time and place (i.e., between primary tumors and metastasis), likely more so than RNA or protein changes. Second, genetic aberrations are uniquely derived from the tumor cells (whereas gene transcripts likely represent stromal cells as well). Finally, as noted above, genetic aberrations are faithfully represented in cell lines, allowing for hypotheses to be generated and tested *in vitro*. Limitations of array-based copy number analysis include the requirement of sufficient tissue material, the relative expense of the technique, and the lack of availability of array equipment in the clinical setting (17, 57, 58). In addition, copy number analyses will obviously not discover specific mutations (e.g., *EGFR* and *KRAS*) relevant to drug responsiveness.

In summary, we show here that an unbiased model derived from high-resolution genome-wide DNA copy number profiles estimates the *in vitro* efficacy of lapatinib. This model correctly identified increases in copy number of the two known target receptors, EGFR and HER2, as well as additional regions harboring several cancer-related genes. Collectively, these alterations provide more robust estimations of cell line sensitivity to lapatinib compared with target genes alone. These analyses illustrate for the first time that genome-wide copy number analyses can be used to discover clinically relevant predictors of sensitivity to anticancer therapeutics.

Acknowledgments

We thank the Discovery Technology Group at GlaxoSmithKline for all gene expression assay work.

References

1. Rusnak DW, Lackey K, Affleck K, et al. The effects of the novel, reversible epidermal growth factor receptor/ErbB-2 tyrosine kinase inhibitor, GW2016, on the growth of human normal and tumor-derived cell lines *in vitro* and *in vivo*. Mol Cancer Ther 2001;1:85–94.

2. Konecny GE, Pegram MD, Venkatesan N, et al. Activity of the dual kinase inhibitor lapatinib (GW572016) against HER-2-overexpressing and trastuzumab-treated breast cancer cells. Cancer Res 2006;66:1630–9.

3. Rusnak DW, Alligood KJ, Mullin RJ, et al. Assessment of epidermal growth factor receptor (EGFR, ErbB1) and HER2 (ErbB2) protein expression levels and response to lapatinib (Tykerb, GW572016) in an expanded panel of human normal and tumour cell lines. Cell Prolif 2007; 40:580–94.

4. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. Nat Rev Mol Cell Biol 2001;2:127–37.

5. Hahn SA, Schutte M, Hoque AT, et al. DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 1996;271: 350 – 3.

6. Lee WH, Bookstein R, Hong F, Young LJ, Shew JY, Lee EY. Human retinoblastoma susceptibility gene: cloning, identification, and sequence. Science 1987;235:1394 – 9.

7. Fleming TP, Saxena A, Clark WC, et al. Amplification and/or overexpression of platelet-derived growth factor receptors and epidermal growth factor receptor in human glial tumors. Cancer Res 1992;52: 4550 – 3.

8. Lin FK, Suggs S, Lin CH, et al. Cloning and expression of the human erythropoietin gene. Proc Natl Acad Sci U S A 1985;82:7580 – 4.

9. Merlino GT, Xu YH, Ishii S, et al. Amplification and enhanced expression of the epidermal growth factor receptor gene in A431 human carcinoma cells. Science 1984;224:417 – 9.

10. Weber BL. Cancer genomics. Cancer Cell 2002;1:37 – 47.

11. Snijders AM, Nowak N, Segraves R, et al. Assembly of microarrays for genome-wide measurement of DNA copy number. Nat Genet 2001;29: 263 – 4.

12. Lucito R, Healy J, Alexander J, et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. Genome Res 2003;13:2291 – 305.

13. Bignell GR, Huang J, Greshock J, et al. High-resolution analysis of DNA copy number using oligonucleotide microarrays. Genome Res 2004; 14:287 – 95.

14. Greshock J, Feng B, Nogueira C, et al. A comparison of DNA copy number profiling platforms. Cancer Res 2007;67:10173 – 80.

15. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 2006;439: 353 – 7.

16. Mulligan G, Mitsiades C, Bryant B, et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. Blood 2007;109:3177 – 88.

17. Anderson JE, Hansen LL, Mooren FC, et al. Methods and biomarkers for the diagnosis and prognosis of cancer and other diseases: towards personalized medicine. Drug Resist Updat 2006;9: 198 – 210.

18. Greshock J, Nathanson K, Martin AM, et al. Cancer cell lines as genetic models of their parent histology: analyses based on array comparative genomic hybridization. Cancer Res 2007;67:3594 – 600.

19. Neve RM, Chin K, Fridlyand J, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 2006;10:515 – 27.

20. Osterberg L, Levan K, Partheen K, Helou K, Horvath G. Cytogenetic analysis of carboplatin resistance in early-stage epithelial ovarian carcinoma. Cancer Genet Cytogenet 2005;163:144 – 50.

21. McDonald SL, Stevenson DA, Moir SE, et al. Genomic changes identified by comparative genomic hybridisation in docetaxel-resistant breast cancer cell lines. Eur J Cancer 2005;41:1086 – 94.

22. Pang E, Hu Y, Chan KY, et al. Karyotypic imbalances and differential gene expressions in the acquired doxorubicin resistance of hepatocellular carcinoma cells. Lab Invest 2005;85:664 – 74.

23. Zhao X, Weir BA, LaFramboise T, et al. Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. Cancer Res 2005;65: 5561 – 70.

24. Cann HM, de Toma C, Cazes L, et al. A human genome diversity cell line panel. Science 2002;296:261 – 2.

25. Rabbee N, Speed TP. A genotype calling algorithm for affymetrix SNP arrays. Bioinformatics 2006;22:7 – 12.

26. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 2004;5:557 – 72.

27. Ambroise C, McLachlan GJ. Selection bias in gene extraction on the basis of microarray gene-expression data. Proc Natl Acad Sci U S A 2002; 99:6562 – 6.

28. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531 – 7.

29. Chung CH, Ely K, McGavran L, et al. Increased epidermal growth factor receptor gene copy number is associated with poor prognosis in head and neck squamous cell carcinomas. J Clin Oncol 2006;24: 4170 – 6.

30. Xia W, Mullin RJ, Keith BR, et al. Anti-tumor activity of GW572016: a dual tyrosine kinase inhibitor blocks EGF activation of EGFR/erbB2 and downstream Erk1/2 and AKT pathways. Oncogene 2002;21:6255 – 63.

31. Potti A, Dressman HK, Bild A, et al. Genomic signatures to guide the use of chemotherapeutics. Nat Med 2006;12:1294 – 300.

32. Baselga J, Tripathy D, Mendelsohn J, et al. Phase II study of weekly intravenous trastuzumab (Herceptin) in patients with HER2/neu-overexpressing metastatic breast cancer. Semin Oncol 1999;26:78 – 83.

33. Naylor TL, Greshock J, Wang Y, et al. High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization. Breast Cancer Res 2005;7:R1186 – 98.

34. Chin K, DeVries S, Fridlyand J, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell 2006; 10:529 – 41.

35. Muss HB. Targeted therapy for metastatic breast cancer. N Engl J Med 2006;355:2783 – 5.

36. Smylie M, Blumenschein GR, Dowlati A, et al. A phase II multicenter trial comparing two schedules of lapatinib (LAP) as first or second line monotherapy in subjects with advanced or metastatic non-small cell lung cancer (NSCLC) with either bronchioloalveolar carcinoma (BAC) or no smoking history. In: American Society of Clinical Oncology Annual Meeting, McCormick Place, Chicago, IL, June 1 – 5, 2007. p. 7611.

37. Latif Z, Watters AD, Bartlett JM, Underwood MA, Aitchison M. Gene amplification and overexpression of HER2 in renal cell carcinoma. BJU Int 2002;89:5 – 9.

38. Kim WY, Sharpless NE. The regulation of INK4/ARF in cancer and aging. Cell 2006;127:265 – 75.

39. Curtin JA, Fridlyand J, Kageshita T, et al. Distinct sets of genetic alterations in melanoma. N Engl J Med 2005;353:2135 – 47.

40. Strefford JC, Worley H, Barber K, et al. Genome complexity in acute lymphoblastic leukemia is revealed by array-based comparative genomic hybridization. Oncogene 2007;26:4306 – 18.

41. Weir BA, Woo MS, Getz G, et al. Characterizing the cancer genome in lung adenocarcinoma. Nature 2007;450:893 – 8.

42. Cowger JJ, Zhao Q, Isovic M, Torchia J. Biochemical characterization of the zinc-finger protein 217 transcriptional repressor complex: identification of a ZNF217 consensus recognition sequence. Oncogene 2007;26: 3378 – 86.

43. Huang G, Krig S, Kowbel D, et al. ZNF217 suppresses cell death associated with chemotherapy and telomere dysfunction. Hum Mol Genet 2005;14:3219 – 25.

44. Li P, Maines-Bandiera S, Kuo WL, et al. Multiple roles of the candidate oncogene ZNF217 in ovarian epithelial neoplastic progression. Int J Cancer 2007;120:1863 – 73.

45. Shimada M, Imura J, Kozaki T, et al. Detection of Her2/neu, c-MYC and ZNF217 gene amplification during breast cancer progression using fluorescence *in situ* hybridization. Oncol Rep 2005;13:633 – 41.

46. Tourneur L, Buzyn A, Chiocchia G. FADD adaptor in cancer. Med Immunol 2005;4:1.

47. Stallings RL, Nair P, Maris JM, et al. High-resolution analysis of chromosomal breakpoints and genomic instability identifies PTPRD as a candidate tumor suppressor gene in neuroblastoma. Cancer Res 2006;66: 3673 – 80.

48. Sato M, Takahashi K, Nagayama K, et al. Identification of chromosome arm 9p as the most frequent target of homozygous deletions in lung cancer. Genes Chromosomes Cancer 2005;44:405 – 14.

49. Carrasco DR, Tonon G, Huang Y, et al. High-resolution genomic profiles define distinct clinico-pathogenetic subgroups of multiple myeloma patients. Cancer Cell 2006;9:313 – 25.

50. Hunter KD, Parkinson EK, Harrison PR. Profiling early head and neck cancer. Nat Rev Cancer 2005;5:127 – 35.

51. Sparano A, Quesnelle KM, Kumar MS, et al. Genome-wide profiling of oral squamous cell carcinoma by array-based comparative genomic hybridization. Laryngoscope 2006;116:735 – 41.

52. Harrington KJ, Bourhis J, Nutting CM, et al. A phase I, open-label study of lapatinib plus chemoradiation in patients with locally advanced squamous cell carcinoma of the head and neck (SCCHN). In: Atlanta (GA): American Society of Clinical Oncology Annual Meeting; 2006. p. 5553.

53. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 2001;344:783 – 92.

54. Scartozzi M, Bearzi I, Berardi R, Mandolesi A, Fabris G, Cascinu S. Epidermal growth factor receptor (EGFR) status in primary colorectal tumors does not correlate with EGFR expression in related metastatic sites: implications for treatment with EGFR-targeted monoclonal antibodies. J Clin Oncol 2004;22:4772 – 8.

55. Weinstein IB. Cancer. Addiction to oncogenes—the Achilles heal of cancer. Science 2002;297:63 – 4.

56. Huang F, Reeves K, Han X, et al. Identification of candidate molecular markers predicting sensitivity in solid tumors to dasatinib: rationale for patient selection. Cancer Res 2007;67:2226 – 38.

57. Anderson JE, Kopecky KJ, Willman CL, et al. Outcome after induction chemotherapy for older patients with acute myeloid leukemia is not improved with mitoxantrone and etoposide compared to cytarabine and daunorubicin: a Southwest Oncology Group study. Blood 2002;100: 3869 – 76.

58. Roses AD. Pharmacogenetics and drug development: the path to safer and more effective drugs. Nat Rev Genet 2004;5:645 – 56.

Molecular Cancer Therapeutics

Genome-wide DNA copy number predictors of lapatinib sensitivity in tumor-derived cell lines

Joel Greshock, Jie Cheng, David Rusnak, et al.

Mol Cancer Ther 2008;7:935-943.