Evaluating HER2 amplification status and acquired drug resistance in breast cancer cells using Raman spectroscopy

Xiaohong Bi
Brent Rexer
Carlos L. Arteaga
Mingsheng Guo
Anita Mahadevan-Jansen
Evaluating HER2 amplification status and acquired drug resistance in breast cancer cells using Raman spectroscopy

Xiaohong Bi,a,b Brent Rexer,c Carlos L. Arteaga,c Mingsheng Guo,d and Anita Mahadevan-Jansen*a

aVanderbilt University, Department of Biomedical Engineering, VU Station B #351631, 2301 Vanderbilt Place, Nashville, Tennessee 37235
bUniversity of Texas Health Science Center at Houston, Department of Nanomedicine and Biomedical Engineering, CABIR 3SCR.4678, 1881 East Road, Houston, Texas 77054
cVanderbilt University, Vanderbilt-Ingram Cancer Center, Nashville, Tennessee 37232
dVanderbilt University, Vanderbilt Center for Quantitative Sciences, Nashville, Tennessee 37232

Abstract. The overexpression of human epidermal growth factor receptor 2 (HER2) is associated with increased breast cancer recurrence and worse prognosis. Effective treatments such as trastuzumab and lapatinib for patients with HER2 overexpression target the blockade of HER2 signaling activities but are often limited by the emergence of acquired drug resistance. This study applied Raman spectroscopy to differentially identify the amplification status of HER2 in cells and to characterize the biochemical composition of lapatinib resistant and sensitive HER2+ breast cancer cells in response to the drug. Raman spectra from BT474 (HER2+ breast cancer cell), MCF-10A (HER2− control), and HER2+ MCF-10A (HER2+ control) were analyzed using lasso and elastic-net regularized generalized linear models (glmnet) for multivariate statistical analysis and were discriminated to groups of different HER2 expression status with an overall 99% sensitivity and specificity. Enhanced lipid content and decreased proteome were observed in HER2+ cells. With lapatinib treatment, lapatinib-resistant breast cancer cells demonstrated sustained lipogenesis compared with the sensitive cells. © The Authors.

Keywords: Raman spectroscopy; breast cancer; human epidermal growth factor receptor 2; lapatinib; drug resistance.

1 Introduction

The human epidermal growth factor receptor 2 (HER2) protein and the HER2/neu oncogene make an important contribution to the regulation of cell proliferation and survival. Amplification of HER2/neu gene occurs in approximately 20% to 25% of human breast cancers1 and is associated with a more aggressive disease course and poor prognosis.2 Besides increased disease recurrence and short survival, HER2 amplification status is predictive for resistance to certain endocrine and chemotherapeutic agents.3 Patients with HER2+ breast cancer require specific treatments targeting blockade of HER2 activity with monoclonal antibodies [e.g., trastuzumab (Herceptin)] and small molecule tyrosine kinase inhibitors (e.g., lapatinib). Therefore, assessment of HER2 status is pivotal in therapeutic decision making for breast cancer patients.

Current testing for amplification of HER2/neu gene relies on two validated techniques: immunohistochemical (IHC) analysis and fluorescence in situ hybridization (FISH). The IHC analysis stains HER2 protein on the cytoplasmic membrane through specific antibody and scores protein expression status based on the color of the stains. This qualitative test is subjected to interlaboratory variation and is relatively less sensitive than FISH. The FISH targets the encoding DNA and thus directly probes HER2 gene amplification. The main disadvantages of this method include the expensive cost and time-consuming analysis that can take days. Although both assays are extensively used in routine clinical diagnosis, the discrepancy between IHC and FISH varies from 3% to 50% (Refs. 4, 5). A combined FISH and IHC approach was suggested to achieve more efficient outcome. Development of new methodologies that can provide accurate and objective determinations on HER2 status is desirable.

Despite significant response rate, a significant proportion of HER2+ breast cancer patients still relapse and develop therapeutic resistance to HER2-targeted regimens such as lapatinib and trastuzumab. Lapatinib is a dual small molecule tyrosine kinase inhibitor that inhibits both epidermal growth factor receptor and HER2. Treatment with lapatinib has been shown to inhibit downstream signaling pathways of HER2 including PI3K-Akt.6 It has shown antitumor activity both in vitro and in vivo by inhibiting cell proliferation and inducing apoptosis.6 Because lapatinib was approved only recently, patient cohorts in which to interrogate mechanisms of acquired resistance have yet to be established.

Raman spectroscopy (RS) is an optical technique that can detect molecular components or biophysical microenvironment in the tissue. Distinctive Raman signatures have been identified to be arising from biomolecules such as nucleic acids, proteins, and lipids. Diagnostic information can thus be achieved by detecting subtle biochemical changes in multiple biomolecules in diseased tissue. In breast cancer research, RS has been applied to distinguish between malignant, normal, and benign breast tissues with significant sensitivity and specificity.7 Several
groups reported HER2 amplification assessment using surface-enhanced RS (SERS). However, SERS requires involvement of nanoparticles and antibodies and as such, does not take full advantage of the intrinsic nature of RS. Meanwhile, conventional RS has been successful in collecting Raman spectra from cells and revealed distinct spectral features from the cytoplasm and chromosome of a live single cell. Hartsuiker et al. reported significant variance in lipid content among breast carcinoma cell lines with varying expression levels of HER2. However, the aforementioned study only provided qualitative analysis on the biochemical properties of single cells and did not explore the potential of RS in differentiating cells based on HER2 status. The goal of the present article is to utilize RS along with multivariate statistical analysis to evaluate the feasibility of this technique in determining HER2 amplification status. Given the sensitivity of RS in detecting molecular phenotypes of cells, it is further applied to characterize the acquired drug resistance of HER2+ breast cancer cells to lapatinib. Instead of interrogating single cells at subcellular level, the ultimate goal of the current study is to characterize HER2-associated biochemical variances at the tissue level, and thus a collection of cells in pellets was investigated in this article. The HER2 overexpressing breast cancer cell line BT474 and breast epithelial cell line MCF-10A with and without HER2 amplification (as positive and negative controls) were compared. Breast cancer cells with acquired lapatinib resistance were also characterized by RS and compared with lapatinib-sensitive breast cancer cells BT474.

2 Materials and Methods

2.1 Cell Lines

Totally four different cell lines were investigated in this study (Table 1) including human breast cancer cell line BT474 that overexpresses HER2 protein, human breast epithelial cell MCF-10A with HER2 amplification (HER2+ control), MCF-10A without HER2 amplification (HER2− control), and BT474 with acquired lapatinib resistance. The controls were obtained by transducing MCF-10A cells with a retroviral vector-encoding human HER2 or with vector alone, as previously described. To generate drug-resistant cells, HER2+ breast cancer cells BT474 were cultured in increasing concentrations of lapatinib, following published methods. BT474 cells exhibit an IC50 to lapatinib of ≤0.1 μM (Ref. 3). At this time, BT474 lapatinib-resistant cells are growing in the laboratory in the presence of 2-μM lapatinib, whereas the parental BT474 serves as drug-sensitive cells. Both lapatinib resistant and sensitive cells were treated with lapatinib overnight (12 to 16 h) before being processed for RS measurements.

All the cell lines were seeded in Petri dishes at a density of 1 × 10^5 cells/ml and grown in a Dulbecco’s modified Eagle medium in a humidified 5% CO2 atmosphere at 37°C. The cells were scrapped free from the dish using a disposable cell scraper, transferred to a 15-ml centrifuge tube, and centrifuged at 1000 rpm for 3 min at room temperature. The supernatant was aspirated, and the cells were washed by three series of saline, followed by centrifugation at 1000 rpm for 3 min after each wash. After the last wash, the cell pellets were transferred to a quartz microscope slide for Raman measurement.

2.2 Raman Spectra Collection

Raman spectra were acquired using a confocal Raman microspectroscopy (Ramanscope Mark III, Renishaw Inc., Gloucestershire, United Kingdom) with a 785-nm diode laser (Innovative Photonic Solutions, Monmouth Junction, New Jersey). Each spectrum was acquired through a 50 × 0.75 NA objective (Leica Microsystems Inc., Buffalo Grove, Illinois, 50x N Plan) using 30-s exposure time and two accumulations with a laser power of 30 mW on the sample. Raman shift of the collected spectra is 300 to 1800 cm−1, and the spectral resolution is 3 cm−1. Multiple spectra were acquired from each pellet. A summary of cell lines and the total number of spectra from each group is listed in Table 1. Background spectrum of the quartz slide was subtracted from each sample spectrum using the “automatic spectral subtraction function” in GRAMS/Al spectroscopy software (Thermo Fisher Scientific Inc., Waltham, Massachusetts). The resulting spectrum equals to the source sample spectrum minus a factored background. The final background factor is automatically determined by iteratively subtracting the background multiplied by various factors in the range of 0 to 2 in a step of 0.0001 until the residual reaches the tolerance of 1e-6. Fluorescence background was subtracted using an automated modified polynomial fitting method written in MATLAB (MathWorks, Natick, Massachusetts). To account for inherent variation in intra- and intersample absolute signal intensities, the spectra were normalized to their respective mean intensity in the fingerprint range of Raman shift 700 to 1750 cm−1 (Ref. 3).

2.3 Data Analysis and Discrimination Algorithm Development

Raman spectra were pooled and analyzed using high-dimension multivariate statistical method in R package called glmnet. A fast and probability-based algorithm for high-dimension discrimination with lasso and elastic-net regularized generalized linear models (glmnet) was applied to classify the cell lines using k-fold cross-validation by k = 5. Glmnet method includes both ridge-regression and lasso penalties for optimally reducing the dimensions through elastic-net. The glmnet package in R allows the data to be fit to follow the probability distributions with multivariate statistical models such as linear, logistic and multinomial, Poisson, and Cox regression models. Here, we used logistic and multinomial model for the analysis purpose.
by the parameter family “multinomial.” The misclassification rate was set up as the optimal destination to find out the minimum misclassification rate for the best discrimination status under a specific selection of the parameter lambda value. The advantages of glmnet method are its flexibility to include a wide range of common situations and to work on large high-dimension dataset.

3 Results and Discussion

3.1 Comparing HER2+ BT474 Breast Cancer Cells with HER2 Positive and Negative Controls

Raman spectra were collected from HER2+ breast cancer cells BT474, HER2+ control (MCF-10A/HER2), and HER2− control (MCF-10A). The mean spectra with 1× standard deviation (shaded area) are shown in Fig. 1 with offset for clarity in display. Spectral characteristics at various wavenumbers were observed to be different between the cell lines. The spectral regions that showed statistically significant differences across groups are marked with peak positions, which mainly appear in the region of lipid (1304, 1338, and 1660 cm⁻¹), protein (852, 872, 1003, and 1260 cm⁻¹), and nucleic acids (721, 758, 781, and 1088 cm⁻¹) bands. The Raman bands centered in the region of 1440 to 1450 cm⁻¹ and at 1304 cm⁻¹ arise from CH₂ deformation modes of protein and lipids and thus will be assigned in this study along with other accompanying bands for both molecules.

To further investigate the biochemical differences between the cell lines, spectra were subtracted to reveal the difference between HER2+ BT474 cancer cells and controls. The difference spectra were calculated by subtracting the normalized mean spectrum from HER2+ control from HER2− control [Fig. 2(a)] as well as BT474 from HER2− [Fig. 2(b)] and HER2+ [Fig. 2(c)] controls. Increased nucleic acid content in cancer cells is revealed by the positive difference at 752 and 1088 cm⁻¹ that were assigned to the phosphodiester groups in nucleic acids (p < 0.05, Figs. 2(b) and 2(c)). Slightly enhanced nucleic acid content was also observed in the difference spectra between HER2+ and HER2− controls [p < 0.01, Fig. 2(g)], indicating that HER2 overexpression leads to active cell proliferation. The intense positive difference at 1594 cm⁻¹ [Figs. 2(f) and 2(c)] can be attributed to the ring C=C stretch vibrations of nucleic acids in agreement with the elevation of nucleic acid content in HER2+ cells indicated by 752 and 1088 cm⁻¹.

The difference spectra of BT474 to the controls [Figs. 2(b) and 2(c)] suggest decreased proteome with HER2 amplification by the strong negative spectral features at the amide III (1247 and 1270/1278) and 1438 cm⁻¹ (p < 0.01). The latter has previously been assigned to CH₂ and CH₃ deformations in the side chain of collagen in normal breast tissue. The protein amide I band at 1650 cm⁻¹ did not show consistent variation as amide III or the band at 1438 cm⁻¹, which confounded the interpretation on proteome content in the current study. The Raman bands at 1003 and 1625 cm⁻¹ arise from the ring structures of phenylalanine and tryptophan/tyrosine, respectively. Both bands are more intense in cancer cells [Fig. 2(d) and 2(e), p < 0.01], suggesting possible elevation in the expression of such aromatic amino acids rich proteins.

The positive difference at 1344 cm⁻¹ in Figs. 2(b) and 2(c) arises from the wagging mode of CH₂ and CH₃ in lipids or proteins. This enhancement along with the concurrent increase (p < 0.01) at 1666 cm⁻¹, which was assigned to the C=C stretching in unsaturated lipids, indicates elevated lipid content in HER2+ BT474 cancer cells. HER2+ control also demonstrates a slight increase in lipid composition than HER2− control evidenced by the positive difference at 1457 and 1663 cm⁻¹ [Fig. 2(f), p < 0.05].

The enhanced lipid content in breast cancer cells comes from exacerbated synthesis of fatty acid and phospholipids. The lipogenesis regulated by fatty acid synthase is associated with upregulation of HER1/HER2 tyrosine kinase receptors in the
Our study indicates augmented lipogenic process with HER2 overexpression in both cancer and control epithelial cells, suggesting that the excessive HER2 expression might play an important role in promoting the upregulation of fatty acid synthase. This is in agreement with a previous report that the HER2 overexpression stimulates the fatty acid synthase promoter.

The probability-based multivariate statistical analysis in glmnet was used to discriminate the cell lines and to identify specific spectral features that are important for classification. The confusion matrix for classification from the glmnet analysis is shown in Table 2. The rows of the table show the numbers of spectra measured for each cell line (true), whereas the columns list the number of spectra predicted to each group. The sensitivity and specificity of discrimination for each cell line are calculated by the ratio of true prediction (bold) to the sum of all true values in each row and to the sum of all predicted values in each column, respectively. The spectra from BT474 (131 out of 132) were discriminated from controls with 99.3% sensitivity and 100% specificity. HER2+ and HER2− controls were classified into the correct groups with 100% sensitivity and 98.8% and 100% specificities, respectively, indicating the potential of RS in identifying HER2+ breast cancer cells.

### 3.2 Characterizing Lapatinib-Sensitive HER2+ BT474 Cancer Cells and Lapatinib-Resistant HER2+ BT474 Cancer Cells with Treatment

Mean Raman spectra from lapatinib-treated breast cancer cells ($N = 104$) and lapatinib-resistant cells ($N = 104$) consist of similar band contours with low variations within each group (Fig. 3). Statistical analysis of the spectra showed significant difference ($p < 0.05$) in Raman bands, which are labeled with peak positions on Fig. 3. Phospholipid (1307, 1333, and 1657 cm$^{-1}$), protein (855, 1156, and 1173 cm$^{-1}$), and nucleic acids (721 and 781 cm$^{-1}$) contribute to the major differences between the cell lines.

Figure 3 shows the difference of spectra from lapatinib-resistant and lapatinib-sensitive cells. With lapatinib treatment, drug-resistant cancer cells have a higher phospholipid content with respect to the drug-sensitive cancer cells, as suggested by the positive differences at 778, 1364 and 1384 cm$^{-1}$ (Fig. 3, $p < 0.01$). Content of unsaturated lipids appears decreased as indicated by the negative difference at 1651 cm$^{-1}$ ($p < 0.01$). The prominent negative differences at 1282, 1310, 1341, and 1680 cm$^{-1}$ reflect decreased proteome in lapatinib-treated drug resistant cells ($p < 0.05$). Significant positive differences are observed at 730 and 778 cm$^{-1}$ that arise from the ring breathing of nucleotides indicating enhanced nucleic acid abundance in the drug-resistant cancer cells.

Treatment with lapatinib in drug-sensitive cells blocks HER2 phosphorylation and downstream PI3K-Akt signaling. Conversely, in resistant cells, lapatinib inhibits HER2 phosphorylation but the PI3K-Akt activity is recovered, which involves in the lipogenic pathway in breast cancer cells. The greater lipid content indicates active endogenous lipogenesis in drug-resistant cancer cells despite the treatment.

### 4 Conclusions

This study reports a sensitive identification of HER2 amplified cell lines, which proves the potential of RS in determining HER2 status in breast cancer cells. Spectral differences reveal variability in HER2-associated biochemical composition in the cell lines. Variation in lipid content is closely related to HER2 status in this study, consistent with previously reported positive regulation between fatty acid synthase and HER2 signaling activities by Menendez et al. Interestingly, a recent report suggests that targeting fatty acid synthase may be a strategy to overcome resistance to HER2-targeted therapies.

The biochemical differences detected in the current study are the basis of spectral assessment for HER2 amplification status. Along with multivariate analysis, breast cancer cells can
potentially be distinguished in situ by directly importing a blinded test spectrum into the established dataset. Future studies using fiber optics-coupled RS on excised tumor samples can bring this method closer to clinical application in HER2 testing. Our findings indicate that RS has the potential to provide a new avenue for studying molecular mechanisms of acquired drug resistance as well as evaluating the response of cancer cells to therapeutic treatments.

Acknowledgments

This work was performed in the Biomedical Photonics Laboratory at Vanderbilt University. The authors would like to acknowledge the grant support from National Cancer Institute 2P50 CA098131-06 (C.A.), T32 CA119910 (B.R.), and K25CA149194-01 (X.B.).

References

42. Q. Wu et al., “Intensities of E. coli nucleic acid Raman spectra excited selectively from whole cells with 251-nm light,” Appl. Chem. 72(13), 2981–2986 (2000).

Bi et al.: Evaluating HER2 amplification status and acquired drug resistance...


**Xiaohong Bi** is an assistant professor of nanomedicine and biomedical engineering at the University of Texas Health Science Center at Houston. She received her BS degree in biochemistry from Nanjing University, China, and her MS and PhD degrees in chemistry from Rutgers University, Newark. Her research focuses on the development of optical techniques for pathogenesis characterization and cancer theragnosis.

**Brent Rexer** is an assistant professor of medicine at Vanderbilt University. He completed fellowship and residency training in hematology/oncology at Vanderbilt University Medical Center, after attending medical school at Vanderbilt and earning a PhD degree in biochemistry. His research aims to understand mechanisms of resistance to targeted therapies in cancer and to identify potential therapies to overcome that resistance. In particular, he is focused on studying resistance to lapatinib in breast cancer.

**Carlos L. Arteaga** obtained his MD at the University of Guayaquil in Ecuador. He trained in internal medicine and medical oncology at Emory University and the University of Texas Health Sciences Center in San Antonio, respectively. He holds the Donna S. Hall Chair in Breast Cancer Research and serves as a professor of medicine and cancer biology, an associate director for translational/clinical research, and a director of the Breast Cancer Program of the Vanderbilt-Ingram Cancer Center at Vanderbilt University.

**Mingsheng Guo** is a bioinformatician at the Center for Quantitative Sciences at Vanderbilt University. He received his PhD degree from Nanjing University in China. Previously, he worked on the application of Raman spectroscopy for cornea laser surgery and characterization of antifreeze glycoprotein at Fisk University. His current research focuses on the feature discrimination of Raman spectra and the analysis of RNA/DNA sequence.

**Anita Mahadevan-Jansen** received her BS and MS degrees in physics from the University of Bombay, India, and an MS and PhD degrees in biomedical engineering from the University of Texas at Austin. She joined the Vanderbilt Engineering Faculty in 1996. She is currently the Orrin H. Ingram Professor of Biomedical Engineering at Vanderbilt University and holds a secondary appointment in the Department of Neurological Surgery.