A Search for Novel Cancer/Testis Antigens in Lung Cancer Identifies VCX/Y Genes, Expanding the Repertoire of Potential Immunotherapeutic Targets

Ayumu Taguchi1, Allen D. Taylor2, Jaime Rodriguez1, Müge Çeliktaş3, Hui Liu1, Xiaotu Ma4, Qing Zhang5, Chee-Hong Wong2, Alice Chin2, Luc Girard6,6, Carmen Behrens7, Wan L. Lam6, Stephen Lam8, John D. Minna8,6,9, Ignacio I. Wistuba1, Adi F. Gazdar5,10, and Samir M. Hanash3

Abstract

Cancer/testis (CT) antigens are potential immunotherapeutic targets in cancer. However, the expression of particular antigens is limited to a subset of tumors of a given type. Thus, there is a need to identify antigens with complementary expression patterns for effective therapeutic intervention. In this study, we searched for genes that were distinctively expressed at a higher level in lung tumor tissue and the testes compared with other nontumor tissues and identified members of the VCX/Y gene family as novel CT antigens. VCX3A, a member of the VCX/Y gene family, was expressed at the protein level in approximately 20% of lung adenocarcinomas and 35% of squamous cell carcinomas, but not expressed in normal lung tissues. Among CT antigens with concordant mRNA and protein expression levels, four CT antigens, XAGE1, VCX, IL13RA2, and SYCE1, were expressed, alone or in combination, in about 80% of lung adenocarcinoma tumors. The CT antigen VCX/Y gene family broadens the spectrum of CT antigens expressed in lung adenocarcinomas for clinical applications. Cancer Res; 74(17); 4694–705. ©2014 AACR.

Introduction

Many aberrantly expressed antigens are currently being targeted as therapeutic cancer vaccines (1, 2). Cancer/testis (CT) antigens are specifically expressed in immune-privileged tissues, notably the testis and placenta, and exhibit aberrant expression in various types of cancer (3). Thus, they are highly immunogenic, without self-tolerance, and are considered promising targets for cancer immunotherapy with fewer adverse effects. However, the expression of particular CT antigens is limited to only a subset of patients with a particular tumor type. This is the case for members of the MAGE family and NY-ESO-1, which are expressed in 10% to 30% and 10% to 30% of non–small cell lung cancers (NSCLC), respectively (4, 5), and are currently being targeted in clinical trials of vaccine immunotherapy (6). Moreover, the effectiveness of immunotherapy targeting of MAGEA3 or NY-ESO-1 is so far limited (7, 8).

In this study, we identified VCX/Y genes as novel CT antigens. VCX/Y genes, mostly located on the X chromosome, may have arisen from a common ancestor to known CT antigens SPANX and CSAG2 (9). VCX may be involved in mRNA stability (10). The expression levels of VCX/Y mRNAs were reduced in normal, immortalized lung cell lines compared with lung cancer cell lines. Reactivity against the protein product of the VCX3A gene was observed in approximately 20% of lung adenocarcinoma and 35% of squamous cell carcinoma (SCC) tumors using tissue microarrays, whereas no expression was detected in normal lung tissues. We further examined the expression patterns of multiple CT antigens in lung adenocarcinomas. We found that one or more of four CT antigens (XAGE1, IL13RA2, SYCE1, and VCX) were expressed in about 80% of lung adenocarcinoma tumors.

Materials and Methods

Cell culture and transfection

All lung cancer cell lines used in this study were cultured in RPMI1640 containing 10% FBS and 1% penicillin/streptomycin cocktail (Gibco). Human bronchial epithelial cells (HBEC30-KT) were cultured with keratinocyte serum-free medium (Life
Technologies, Inc.) that contained 50 μg/mL bovine pituitary extract (Life Technologies) and 5 ng/mL EGF (Life Technologies). Human small airway epithelial (HSAEC1-KT) cells were cultured with Small Airway Epithelial Cell Growth Medium (SAGM, Lonza) according to the manufacturer’s protocol. For the siRNA transfection experiments, two siRNAs that targeted VCX (#1: Hs_VCX_8 FlexiTube siRNA, cat. #SI04173568, and #2: Hs_VCX_10 FlexiTube siRNA, cat. #SI04187295) were obtained from Qiagen. PC-9 cells were transfected at a final concentration of 50 nmol/L siRNA using Lipofectamine RNAiMAX (Life Technologies), according to the manufacturer’s instructions. Seventy-two hours after transfection, protein was harvested using RIPA buffer for Western blotting, and an MTS assay (CellTiter 96 Aqueous One Solution Cell Proliferation Assay, Promega) was performed according to the manufacturer’s protocol. Three independent experiments for the MTS assay were performed in triplicate. For treatment with 5-aza-2’-deoxycytidine (5-Aza-dC; Sigma), cells were incubated with 2 μmol/L 5-Aza-dC. Media were changed every 24 hours. After 5 days, total RNA was extracted from cell lines using the RNeasy Plus Mini Kit (Qiagen).

**Quantitative real-time reverse transcription-PCR**

Complementary DNA samples were prepared using the High-Capacity cDNA Reverse Transcription Kit (Life Technologies). A TaqMan PCR assay was performed with a 7500 Fast Real-Time PCR System, TaqMan PCR master mix, commercially available primers, FAM-labeled probes for the VCX3A gene (Hs01491277_gH), and VIC-labeled probes for 18S (Hs99999901_S1), according to the manufacturer’s instructions (Life Technologies). Each sample was run in triplicate. Ct values for the VCX3A gene were calculated and normalized to Ct values for 18S (ΔCt). The ΔΔCt values were then calculated by being normalized to the ΔCt value for the control.

**The Cancer Genome Atlas dataset**

The Cancer Genome Atlas (TCGA) cancer methylomes, measured using Illumina Human Methylation 450 BeadChips, were downloaded from the TCGA website, where the methylation levels of 485,577 CpG sites are represented as a β value of 0 to 1 (β = 1 for fully methylated sites and β = 0 for fully unmethylated sites). RNA-seq data of TCGA cancer samples were also downloaded. The normalized count from RSEM (11) for each gene was extracted from files with the extension ".rsem.genes.normalized_results." Between-sample quantile normalization was performed using the R software package preprocessCore. The samples were classified as tumors (primary solid tumor, code: 01) or nontumors (solid tissue normal, code: 11) according to the sample naming conventions specified in the code tables of the TCGA and corresponding barcodes.

**Tissue samples**

Eighty-three pairs of treatment-naïve lung adenocarcinoma tissues and adjacent nontumor lung tissues had been obtained with informed consent (12, 13). EGFR and KRAS mutations were screened as previously described (13). Tumor tissues were considered TITF1/NKX2-1 negative when mRNA expression levels were < 0.5-fold lower than those in adjacent normal tissues.

**Gene expression analysis**

Total RNA was extracted from cell lines and tissues using the RNeasy Plus Mini Kit (Qiagen). RNA quality and concentration were checked using the Experion Automated Electrophoresis System (Bio-Rad) according to the manufacturer’s protocol. Expression data were obtained using Illumina Human WG-6 v3.0 Expression BeadChips (Illumina) in the Genomics Core at The University of Texas Southwestern Medical Center at Dallas (Dallas, TX). Bead-summarized data were obtained using Illumina BeadStudio software and preprocessed using the R software package MBCB (Model-based Background Correction for Beadarray) for background correction and probe summarization. Preprocessed data were quantile normalized and log transformed.

**Mass spectrometry analysis**

A mass spectrometry analysis of cell lines was performed as previously described (14). In brief, all cell lines were grown in RPMI1640 (Pierce) that contained 10% of dialyzed FBS (Invitrogen), 1% penicillin and streptomycin cocktail, and 13C-l-lysine instead of regular lysine for seven passages, according to the standard SILAC (stable isotope labeling by amino acids in cell culture) protocol (15). The same batch of cells was used to analyze the whole cell extract (WCE), conditioned media, and cell surface proteins. WCEs, conditioned media, and cell surface proteins were fractionated by reverse-phase chromatography using 2 mg, 1 mg, and 0.5 mg of total protein, respectively. Cell extracts were reduced and alkylated with iodoacetamide before chromatography. Protein digestion and identification by liquid chromatography/tandem mass spectrometry was performed as described previously (16, 17). The tandem mass spectra were searched against a composite database of IPI human (v3.57) and IPI bovine (v3.43) proteins. A fixed modification of 6.020129 mass units was added to lysine residues for database searching to account for incorporation of the 13C-l-lysine isotope. To estimate the significance of the peptide and protein matches, we used the tools PeptideProphet and ProteinProphet. We selected all peptides that had been initially identified with a PeptideProphet probability minimum of 0.05 and not shared between humans and bovines and submitted them to ProteinProphet. Protein Propheth infers a minimal set of proteins that can explain the peptide evidence, assigning a probability to each protein on the basis of the combined peptide probabilities. The derived protein identifications were filtered at a maximum 5% error. The total number of spectral counts for each protein group in ProteinProphet was used in the semiquantitative analysis. Each dataset was normalized to the total number of spectral counts of each compartment.

**Tissue microarrays**

The tissue microarrays used in this study comprised 463 surgically resected NSCLC tumor specimens (305 adenocarcinomas and 158 SCCs) collected under an institutional review
board–approved protocol and archived as formalin-fixed, paraffin-embedded specimens in The University of Texas Specialized Program of Research Excellence thoracic tissue bank at The University of Texas MD Anderson Cancer Center (Houston, TX). Tissue microarrays were prepared with a manual tissue arrayer (Advanced Tissue Arrayer ATA100, Chemicon International) using 1 mm diameter cores in triplicate for tumors, as described previously (18). Histologic sections that were 4 µm in thickness were then prepared for the subsequent immunohistochemical analysis.

**Western blot analysis and immunohistochemical analysis**

Monoclonal anti-mouse antibody raised against human VCX3A (Abnova) was used for the immunohistochemical and Western blot analyses. β-Tubulin (Cell Signaling Technology) was used as a loading control for the Western blot analysis. Tissue microarray slides and lung tissue sections were immunostained in a Dako Autostainer Plus automated machine. The slides were deparaffinized and hydrated, and antigen retrieval was performed using a decoiler with a Dako pH 6.0 target retrieval solution (Dako North America, Inc.). Intrinsic peroxidase activity was blocked using 3% methanol and hydrogen peroxide for 10 minutes, and serum-free protein block (Dako) was used for 5 minutes to block nonspecific antibody binding. Slides were then incubated with antibodies against human VCX3A (1:1,000 dilution) for 30 minutes at ambient temperature. After being washed three times in Tris-buffered saline, slides were incubated for 30 minutes with Dako Envision + Dual Link at ambient temperature. They were then washed three more times, incubated with Dako chromogen substrate for 5 minutes, and counterstained with hematoxylin. Formalin-fixed, paraffin-embedded whole-section specimens, except for the omission of the primary antibodies, were used as negative controls. For the immunostaining scoring, two board-certified surgical pathologists participated in the analysis of this TMA (H. Liu and J. Rodriguez). The scoring system evaluates the intensity of staining (0: negative, 1+: weakly positive, 2+: moderately positive, and 3+: strongly positive) and percentage of staining distribution in the tumor cells (from 0 to 100% of the cells) for each TMA core. Three TMA cores from the same patient were evaluated and the percentage of staining distribution in the tumor cells (from 0 to 100% of the cells) for each TMA core. The TMA cores from the same patient were evaluated and the percentage of staining distribution was averaged to provide the final scoring. In this study, tumors with 20% or more of the cells having 2+ or 3+ staining intensity were considered as positive. Before the TMA scoring, the two pathologists evaluated together at a double head microscope the immunostaining pattern for VCX3A in control slides and in the TMA slides to agree on the criteria for intensity and distribution. Then one of the pathologists (H. Liu) performed the scoring of the TMA slides. Finally, the two pathologists (H. Liu and J. Rodriguez) met again to review the TMA scoring together in the double head microscope and agreed on any question or discrepancy about the evaluation.

**Serological immunoreactivity against arrayed VCX3A protein**

A recombinant protein microarray (Version 5 Protoarrays, Invitrogen) containing VCX3A protein was utilized to determine Ig-based immunoreactivity in plasmas from 96 female subjects that included 12 and 17 subjects subsequently diagnosed with lung and colon adenocarcinoma, respectively. One control was matched to colon adenocarcinoma cases and 2 to 5 controls were matched to individual lung adenocarcinoma cases. The array study was performed as part of discovery studies of biomarker candidates for the early detection of epithelial tumor types (19, 20). Plasma samples were randomly allocated to the arrays within batches to eliminate potential bias, and were each hybridized with an individual protein microarray at a dilution of 1:500 in probe buffer (Synthetic Block in 150 mmol/L PBS, 0.1% Tween-20) for 1.5 hours at 4°C. IgG reactivity was assessed quantitatively using an indirect immunofluorescence protocol, and local background-subtracted median spot intensities for downstream statistical analysis were generated using GenePix Pro 6.1. Background-subtracted median spot intensities from individual slides were quantile normalized, and individual clones were normalized to the mean and SD of the normal samples before analysis.

**Results**

**Identification of potential CT antigens in lung cancer**

To discover potential CT antigens in lung cancer, we identified genes that were overexpressed in the testis compared with normal tissues using BioGPS (http://biogps.org/; GSE1133; ref. 21; Fig. 1A for the workflow). We focused on CT antigens on the X chromosome (CT-X antigens) because their gene expression is more testis-restricted than that of non-X CT antigens (22). We compared the mRNA expression profiles of 1,013 probes for 632 genes on the X chromosome between 5 testis and 47 other normal tissues, excluding the placenta, ovary and neuronal tissue, as CT antigens may be expressed in those tissues (22). Forty-five probes that corresponded to 37 genes had more than 2-fold higher mRNA expression levels in testis than in nontestis tissues (P < 0.05, Mann–Whitney U test; Supplementary Table S1). A subset of 23 of these genes had not been previously recognized as CT antigens (CTDatabase; http://www.cta.lncc.br/; ref. 23).

Four genes, BEX1, NXF3, TCEAL3, and VCX2, were considered plausible candidates for encoding CT antigens because their family members were predominantly localized to the X chromosome, which is a typical characteristic of CT-X antigens (Supplementary Table S1; ref. 3). To determine whether these genes are expressed in lung cancer, we profiled the mRNA expression of four gene families (BEX, NXF, TCEAL, and VCX) in 113 NSCLC (78 adenocarcinoma, 3 adenosquamous cell carcinoma, 10 SCC, 12 large cell carcinoma, and 10 other NSCLCs), 29 SCLC, 2 carcinoid, 4 mesothelioma, and 59 normal lung immortalized cell lines (30 normal bronchial epithelial cells and 29 normal small airway epithelial cells). VCX3 family genes were coexpressed at a higher level in a subset of NSCLC and SCLC cell lines than in normal bronchial epithelial cell lines or normal small airway epithelial cell lines (Fig. 1B). VCX, VCX3A, and VCX3B were predominantly expressed compared with VCX2 and VC1, BEX1, BEX2, and BEX4, and BEX3 were highly expressed.
in some NSCLC cell lines and most SCLC cell lines but were also expressed in normal cell lines, except for BEX5 (Supplementary Fig. S1). These data suggested that BEX5 is a potential CT antigen in lung cancer, particularly SCLC. Two members of the NXF gene family, NXF2 and NXF2B, are known CT antigens (CT39), and mRNA expression levels of the other NXF family genes were not significantly different between cancer and nontumor cell lines. Of the TCEAL family genes, TCEAL2 was strongly expressed in most SCLC cell lines, compared with nontumor cell lines indicating that it is a potential cancer testis antigen in SCLC, together with BEX5.

The expression of CT-X genes is associated with promoter demethylation and more broadly with genomically hypomethylated (3, 6). Thus, we determined whether mRNA expression levels of VCX/Y genes were associated with hypomethylation. The methylation status of the 5 CpG sites in VCX, 3 in VCX2, 4 in VCX3A, and 5 in VCX3B was analyzed in TCGA data (Supplementary Fig. S2A). Methylation data were available for 4, 1, 3, and 2 CpG sites in VCX, VCX2, VCX3A, and VCX3B respectively, for lung adenocarcinomas (406 tumors and 32 nontumor tissues) and lung squamous cell carcinomas (252 tumor and 43 nontumor tissues). The methylation levels of all CpG sites were significantly lower in cancer tissues compared with nontumor tissues (Fig. 2A and Supplementary Fig. S2B). We further evaluated the correlation between mRNA and methylation levels of VCX/Y family genes in cancer tissues. mRNA levels of VCX/Y genes were also obtained from the TCGA RNA-seq data (326 adenocarcinomas and 217 SCCs for VCX; 377 adenocarcinomas and 248 SCCs for VCX2; and 378 adenocarcinomas and 248 SCCs for VCX3A and VCX3B). mRNA levels were significantly and inversely correlated with methylation for all VCX/Y genes in lung adenocarcinomas (VCX, Spearman correlation coefficient = −0.22, P < 0.0001; VCX2, Spearman correlation coefficient = −0.16, P = 0.0025; VCX3A, Spearman
correlation coefficient = −0.21, P < 0.0001; and VCX3B, Spearman correlation coefficient = −0.18, P = 0.0004), whereas the correlation was not significant in SCC, except for VCX3A (VCX, Spearman correlation coefficient = −0.03, P = 0.6983; VCX2, Spearman correlation coefficient = 0.02, P = 0.8127; VCX3A, Spearman correlation coefficient = −0.16, P = 0.0099; and VCX3B, Spearman correlation coefficient = −0.07, P = 0.2598; Fig. 2B), suggesting that VCX/Y gene expression is epigenetically regulated in lung adenocarcinoma and that additional regulatory mechanisms for VCX/Y gene expression may occur in SCC that are not associated with the CpG sites tested. In addition, VCX3A gene expression levels were determined in 8 NSCLC cell lines after treatment with the DNA demethylating agent 5-Aza-dC. VCX3A gene expression was clearly induced with 5-Aza-dC treatment in all 8 NSCLC cell lines, further supporting the epigenetic regulation of VCX/Y gene expression (Fig. 2C). VCX3A protein expression was confirmed by Western blotting in PC-9 cells, in which VCX/Y genes were highly overexpressed (Figs. 1B and 2D). Of note, cell proliferation was inhibited in PC-9 cells by treatment with VCX siRNA (Fig. 2E and F), indicating an oncogenic role for VCX/Y family genes in NSCLC.

Tissue microarray-based analysis of VCX3A protein expression in NSCLC

Next, we determined VCX3A protein expression in NSCLC tumors using a tissue microarray. VCX3A protein was strongly expressed in normal testis (data not shown) but was negative or very weakly positive in normal lung tissue (Fig. 3A). VCX3A expression was predominantly found in the nucleus (Fig. 3B). A tumor was considered positive if 20% or more of tumor cells were moderately or strongly stained. While non-tumor adjacent tissues from 20 adenocarcinomas and 8 SCCs in the tissue microarray were not stained with VCX3A (data not shown), 118 (25.5%) of 463 tumors were positively stained, unrelated to sex, age (> or ≤ 65 years), smoking status, stage, or mutational status for EGFR and KRAS. Reactivity was significantly higher in SCC than in adenocarcinoma (37.3% and 19.3%, respectively; P < 0.0001, Fisher exact test; Table 1).
CT antigen expression in NSCLC cell lines

Given that no single CT antigen is widely expressed in NSCLC, we assessed whether a combination of CT antigens exhibited complementary expression and would thus in principle, provide a more effective approach to immunotherapy than would single antigens. We first determined CT antigen expression at both the mRNA and protein levels in 38 NSCLC cell lines, given the potential for discordance between RNA and protein expression (24). Of 255 known CT antigens in the CTDatabase (23), 182 genes were profiled in 38 NSCLC cell lines using Illumina WG-6 v3.0 gene expression microarrays (Supplementary Table S2), including 87 (47.5%) CT-X antigens and 96 (52.5%) non-X CT antigens. Next, we determined the protein expression levels of CT antigens in the same NSCLC cell line set with quantitative in-depth proteomic profiling of three compartments: WCE, conditioned media, and cell surface proteins (14, 17). One hundred fifteen (45.1%) of 255 CT antigens were identified in at least one cell line, and 83, 69, and 83 CT antigens were identified in WCE, media, and the cell surface, respectively (Fig. 4A and Supplementary Table S3), consisting of 47 (40.9%) CT-X antigens and 68 (59.1%) non-X CT antigens. One hundred nine CT antigens were profiled at both the mRNA and protein levels. The mRNA expression levels of 10 CT antigens (CT45A5, GAGE7, IL13RA2, MAGEA10, MAGEA4, MAGEB2, NY-ESO-1, PAGE5, SYCE1, and XAGE1) were significantly and positively correlated with protein expression levels in at least one compartment (P < 0.05, Spearman correlation; Fig. 4B). In addition, mRNA levels were significantly higher in cell lines with protein expression of these CT antigens compared with cell lines without protein expression of these CT antigens (P < 0.05, Mann–Whitney U test). Protein expression levels were more widely distributed than were mRNA levels, suggesting translational regulation of CT antigens. All CT antigens, except for SYCE1, with significant correlation between mRNA and protein levels.
levels were CT-X antigens. Peptides encoded in the VCX gene family (VCX, VCX2, VCX3A, and VCX3B) are largely indistinguishable and were identified by mass spectrometry in the WCE of PC-9 cells (Fig. 4C). VCX/Y proteins are highly homologous and have a paucity of arginine and lysine residues in their sequences, resulting in few detectable tryptic peptides for identification by mass spectrometry. A search for VCX peptides among other cell lines only yielded identification of VCX3A in SW620 colon adenocarcinoma cell line.

mRNA expression profiles of CT antigens in lung cancer tissues

We determined the mRNA expression profiles of 10 CT antigens with concordant mRNA and protein expression levels, together with VCX as a representative gene for the VCX/Y family genes and MAGEA3, which is a current target for vaccine immunotherapy. We evaluated 83 lung adenocarcinomas and adjacent nontumor tissues (Supplementary Table S4). Unsupervised hierarchical clustering of gene expression in tumors revealed distinct CT antigen expression patterns. Interestingly, lung adenocarcinomas yielded three clusters on the basis of XAGE1 mRNA levels: high, intermediate, and low (Fig. 5A, top heatmap). In the intermediate and low clusters, VCX, IL13RA2, or SYCE1 was highly expressed in most tumors. Thus, we selected XAGE1, VCX, IL13RA2, and SYCE1 to form a CT antigen panel and compared the mRNA expression levels of these four genes in tumor and adjacent nontumor tissues. The number of tumors with CT antigen overexpression (more than 2-fold higher mRNA expression levels in cases vs. controls) was 55 (66.3%) for XAGE1, 26 (31.3%) for VCX, 10 (12.0%) for IL13RA2, and 9 (10.8%) for SYCE1 (Fig. 5A, bottom heatmap; Table 2). One or more of the CT antigens in the panel were overexpressed in 66 (79.5%) lung adenocarcinomas. XAGE1 overexpression was significantly associated with TITF1/NKX2-1–positive lung adenocarcinomas (P = 0.0054, Fisher exact test; Table 2 and Supplementary Table S4). While overexpression of VCX and SYCE1 was not associated with the adenocarcinoma patients’ clinical characteristics, IL13RA2 overexpression was significantly less common in current smokers (P = 0.0436, χ² test).

The CT antigen panel was more broadly expressed than were NY-ESO-1 and MAGEA3 (Fig. 5A, bottom heatmap), which are currently being studied in clinical trials as vaccine immunotherapy. Therefore, we determined the immunogenicity of the four CT antigens. We used two epitope prediction algorithms, BIMAS (http://www-bimas.cit.nih.gov/; Table 1. VCX3A expression in NSCLC tissue microarray

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total</th>
<th>Positive, n (%)</th>
<th>Negative, n (%)</th>
<th>P²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive, n (%)</td>
<td>Negative, n (%)</td>
<td>P²</td>
</tr>
<tr>
<td>Sexa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>210</td>
<td>56 (26.7)</td>
<td>154 (73.3)</td>
<td>0.4335</td>
</tr>
<tr>
<td>Female</td>
<td>216</td>
<td>50 (23.1)</td>
<td>166 (76.9)</td>
<td></td>
</tr>
<tr>
<td>Ageb</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;65 years</td>
<td>246</td>
<td>61 (24.8)</td>
<td>185 (75.2)</td>
<td>1.0000</td>
</tr>
<tr>
<td>≤65 years</td>
<td>180</td>
<td>45 (25.0)</td>
<td>135 (75.0)</td>
<td></td>
</tr>
<tr>
<td>Smoking statusa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>176</td>
<td>41 (23.3)</td>
<td>135 (76.7)</td>
<td>0.5784</td>
</tr>
<tr>
<td>Former</td>
<td>203</td>
<td>55 (27.1)</td>
<td>148 (72.9)</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>47</td>
<td>10 (21.3)</td>
<td>37 (78.7)</td>
<td></td>
</tr>
<tr>
<td>Histologic type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>305</td>
<td>59 (19.3)</td>
<td>246 (80.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>SCC</td>
<td>158</td>
<td>59 (37.3)</td>
<td>99 (62.7)</td>
<td></td>
</tr>
<tr>
<td>Stagea</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>219</td>
<td>48 (21.9)</td>
<td>171 (78.1)</td>
<td>0.2125</td>
</tr>
<tr>
<td>II</td>
<td>101</td>
<td>33 (32.7)</td>
<td>68 (67.3)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>87</td>
<td>20 (23.0)</td>
<td>67 (77.0)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>19</td>
<td>5 (26.3)</td>
<td>14 (73.7)</td>
<td></td>
</tr>
<tr>
<td>Mutationb</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KRAS</td>
<td>51</td>
<td>13 (25.5)</td>
<td>38 (74.5)</td>
<td>0.4901</td>
</tr>
<tr>
<td>EGFR</td>
<td>22</td>
<td>5 (22.7)</td>
<td>17 (77.3)</td>
<td></td>
</tr>
<tr>
<td>KRAS and EGFR wild</td>
<td>103</td>
<td>18 (17.5)</td>
<td>85 (82.5)</td>
<td></td>
</tr>
</tbody>
</table>

*Information was not available for 37 samples.
*Mutations were not investigated in 287 samples.
*P values were calculated by Fisher exact test or χ² test.
Five peptides that consisted of nine amino acid residues for XAGE1, 22 for VCX, 31 for IL13RA2, and 20 for SYCE1 were predicted to bind HLA molecules, with high scores, in both algorithms (100 in BIMAS and 20 in SYFPEITHI; Fig. 5B; Supplementary Table S5). Interestingly, VCX contains identical immunogenic 9-mer peptides in different locations: EPLSQESEV at starting amino acid positions 112, 132, and 152; QESEVEEPL at 116, 136, and 156; EPLSQESQV at 122, 142, 162, and 182; and QESQVEEPL at 126, 146, and 166. HLA binding of EPLSQESEV and EPLSQESQV was predicted to be HLA-B*51:01 restricted, and HLA binding of QESEVEEPL and QESQVEEPL was predicted to be HLA-B*40:01 restricted (Supplementary Table S5). Thus, these repeating amino acid sequences may amplify the immune response against VCX in patients with HLA-B*51:01 or HLA-B*40:01. As part of prior

ref. 25) and SYFPEITHI (http://www.syfpeithi.de/; ref. 26). Five peptides that consisted of nine amino acid residues for XAGE1, 22 for VCX, 31 for IL13RA2, and 20 for SYCE1 were predicted to bind HLA molecules, with high scores, in both algorithms (≥ 100 in BIMAS and ≥ 20 in SYFPEITHI; Fig. 5B; Supplementary Table S5). Interestingly, VCX contains identical immunogenic 9-mer peptides in different locations: EPLSQESEV at starting amino acid positions 112, 132, and 152; QESVEEPL at 116, 136, and 156; EPLSQESQV at 122, 142, 162, and 182; and QESQVEEPL at 126, 146, and 166. HLA binding of EPLSQESEV and EPLSQESQV was predicted to be HLA-B*51:01 restricted, and HLA binding of QESVEEPL and QESQVEEPL was predicted to be HLA-B*40:01 restricted (Supplementary Table S5). Thus, these repeating amino acid sequences may amplify the immune response against VCX in patients with HLA-B*51:01 or HLA-B*40:01. As part of prior

Figure 4. Transcriptomic and proteomic profiling of CT antigens in 38 NSCLC cell lines. A, Venn diagrams of CT antigens identified by mass spectrometric analysis in WCE, cell surface (surface), and conditioned media (media) of 38 NSCLC cell lines. B, CT antigens with significant correlation of expression levels between mRNA and protein. P values for comparison of mRNA expression levels in cell lines with or without CT antigen protein expression were calculated by Mann-Whitney U test. C, schema of human VCX/Y family proteins. Black bars indicate peptides identified in the PC-9 cell line. Presented numbers are based on NP_038480.2 (VCX), NP_057462.2 (VCX2), NP_057463.2 (VCX3A), and NP_001001888.2 (VCX3B).
studies that were aimed at identifying tumor antigens associated with autoantibodies in epithelial tumors types, we conducted discovery studies using recombinant and natural protein microarrays (19, 20). Given our discovery of novel CT antigens, we examined their association with an autoantibody response using protein microarrays. VCX3A, which we have found to be overexpressed in both lung and colon cancer cell lines, was represented in recombinant protein microarrays. We assessed its Ig-based reactivity in prediagnostic sera from 29 plasmas from subjects diagnosed with lung or colon adenocarcinoma and 29 pools of matched controls. Serum Ig autoantibody reactivity against VCX3A was significantly higher among cases compared with controls \( (P = 0.020, \text{Mann–Whitney} \ U \text{test}) \) and a receiver operating characteristic (ROC) analysis yielded an area under the curve (AUC) of 0.678, suggesting the occurrence of an immune response against VCX3A in a subset of subjects with lung and colon adenocarcinoma and potential usefulness of VCX3A autoantibodies as a part of a diagnostic biomarker panel (Fig. 5C).

Because antigen presentation occurs in an HLA-restricted manner, we determined the allelic distribution of HLA molecules with potential binding to our CT antigen panel, which yielded 17 HLA class I molecules. The allelic frequencies of 13 HLA class I molecules were available for two U.S. populations in the Allele*Frequencies database (http://www.allelefrequencies.net/; ref. 27; Supplementary Table S6). For African-Americans, the sum of the allele frequency of HLA-A molecules was 36.1% and that of HLA-B was 20.1%; and for Caucasians, it was 65.8% for HLA-A and 37.6% for HLA-B.

**Discussion**

We identified the VCX/Y gene family as encoding novel CT antigens initially by mining gene expression data. These genes were initially identified as testis specific and were associated with X-linked nonspecific mental retardation (28, 29). They exhibited in our study several characteristics of CT-X antigens: testis-specific expression, a multigene family on the X chromosome, overexpression in a subset of NSCLC cell lines and tumor tissues, as confirmed by a tissue microarray analysis, no overexpression in nontumor lung cell lines or tissues, and coexpression with family genes (3, 6). Our findings also suggest epigenetic regulation of VCX/Y gene expression, which is another typical characteristic of CT-X antigens. There is also independent evidence that VCX gene expression is epigenetically regulated (30). Moreover, the results of a recent genomic study suggested that VCX/Y and the known CT antigens SPANX and CSAG2 are derived from a common ancestor and that VCX/Y and SPANX are paralogs with a similar protein structure that consists of C-terminal acidic repeats of variable lengths (9). To our knowledge, the relevance of VCX/Y to cancer has not been previously investigated. VCX/Y proteins

---

**Table 2.** Association of CT antigen overexpression status and clinical characteristics of 83 lung adenocarcinoma patients

<table>
<thead>
<tr>
<th>Clinical characteristics</th>
<th>All</th>
<th>XAGE1</th>
<th>VCX</th>
<th>IL13RA2</th>
<th>SYCE1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Yes</td>
<td>No</td>
<td>P&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Yes</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>24</td>
<td>20</td>
<td>4</td>
<td>0.7665</td>
<td>18</td>
</tr>
<tr>
<td>Female</td>
<td>59</td>
<td>46</td>
<td>13</td>
<td>0.0943</td>
<td>37</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;65</td>
<td>54</td>
<td>46</td>
<td>8</td>
<td>0.0943</td>
<td>38</td>
</tr>
<tr>
<td>&lt;65</td>
<td>29</td>
<td>20</td>
<td>9</td>
<td>17.12</td>
<td>17</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>39</td>
<td>30</td>
<td>9</td>
<td>0.7805</td>
<td>30</td>
</tr>
<tr>
<td>Former</td>
<td>14</td>
<td>12</td>
<td>2</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Never</td>
<td>30</td>
<td>24</td>
<td>6</td>
<td>17.12</td>
<td>17</td>
</tr>
<tr>
<td>Mutation&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KRAS</td>
<td>33</td>
<td>26</td>
<td>7</td>
<td>0.6638</td>
<td>22</td>
</tr>
<tr>
<td>EGFR</td>
<td>20</td>
<td>17</td>
<td>3</td>
<td>13.7</td>
<td>7</td>
</tr>
<tr>
<td>KRAS, EGFR wild</td>
<td>27</td>
<td>20</td>
<td>7</td>
<td>18.9</td>
<td>8</td>
</tr>
<tr>
<td>TITF1/NKX2-1 expression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>68</td>
<td>57</td>
<td>11</td>
<td>0.0706</td>
<td>50</td>
</tr>
<tr>
<td>Negative</td>
<td>15</td>
<td>9</td>
<td>6</td>
<td>5.10</td>
<td>4</td>
</tr>
<tr>
<td>Stage&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>36</td>
<td>30</td>
<td>6</td>
<td>0.7154</td>
<td>25</td>
</tr>
<tr>
<td>II–IV</td>
<td>18</td>
<td>14</td>
<td>4</td>
<td>13.5</td>
<td>5</td>
</tr>
</tbody>
</table>

<sup>a</sup>Mutations were not studied in three patients.

<sup>b</sup>Staging information was not available for 29 patients.

<sup>c</sup>P values were calculated by Fisher exact test or \( \chi^2 \) test.
are localized in the nucleus and may be involved in regulating ribosome assembly during spermatogenesis (31), inhibiting mRNA decapping and regulating mRNA stability (10). In our study, VCX siRNA inhibited cell proliferation in PC-9 cells, suggesting that VCX/Y genes have oncogenic roles. Therefore, further functional studies are warranted.

An integrative mRNA analysis and proteomic profiling of 38 NSCLC cell lines revealed that mRNA and protein expression levels were significantly correlated with each other for 10 CT antigens. We analyzed 83 lung adenocarcinomas samples to determine the gene expression profiles of these 10 CT antigens, VCX, and MAGEA3, which is currently being targeted in clinical trials. Our analysis resulted in a panel of four CT antigens (XAGE1, VCX, IL13RA2, and SYCE1); one or more of these was overexpressed in 80% of lung adenocarcinomas, highlighting the potential of this combination for immunotherapy in lung cancer.

XAGE1 expression and an associated autoantibody response have been observed in NSCLCs (32), and several XAGE1 transcript variants have been identified. Most XAGE1 peptides identified by mass spectrometry in our study were common in various XAGE1 isoforms. However, consistent with the findings of a previous study (33), XAGE1b-unique peptides were most commonly identified among XAGE1 isoform-unique peptides in 38 NSCLC cell lines (data not shown). Five 9-mer XAGE1 peptides were identified as highly immunogenic on the basis of two distinct epitope prediction algorithms, BIMAS and SYFPEITHI, and several additional immunogenic peptides have been identified (34).

IL13RA2 is known to be overexpressed in solid tumors, particularly glioblastoma multiforme. Although IL13RA2 has been associated with inflammation in lung epithelial cells, its role in lung cancer has not been investigated (35). Okada and colleagues recently reported the results of a phase I/II trial using α-type 1 polarized dendritic cells loaded with synthetic peptides for glioma-associated antigen epitopes, including EphA2, IL13 receptor-α2, YKL-40, gp100, and polynosinic-polycytidylic acid for HLA-A2(+) patients with recurrent malignant glioma. Of 22 patients in the study, two experienced objective clinical tumor regression. One patient
with recurrent glioblastoma experienced a complete response (36).

SYCE1 is a constituent of the synaptonemal complex during meiosis (37); however, its role in cancer has not been well studied. Approximately 20% of lung adenocarcinomas did not exhibit expression of the CT antigens in our panel (Fig. 5A). XAGE1 (38) and IL13RA2 (39) are known to be induced by 5-Aza-dC. Our data indicate that similar epigenetic regulation occurs in VCX/Y genes. Thus, for patients whose tumors do not express CT antigens, a combination of CT antigen-targeted therapy and DNA demethylating agents may be promising (6).

In conclusion, we found that VCX/Y family genes are novel CT antigens in lung cancer on the basis of an analysis of gene expression data; we further confirmed VCXA expression at the protein level in a tissue microarray analysis. Importantly, we identified a panel of CT antigens, consisting of XAGE1, VCX, IL13RA2, and SYCE1, which were expressed in most lung adenocarcinomas. This finding suggests that targeting this four-antigen panel will be a promising immunotherapy approach for lung adenocarcinoma.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

References
A Search for Novel Cancer/Testis Antigens in Lung Cancer Identifies VCX/Y Genes, Expanding the Repertoire of Potential Immunotherapeutic Targets


Updated version

Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-13-3725

Supplementary Material

Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2014/06/27/0008-5472.CAN-13-3725.DC1

Cited articles

This article cites 39 articles, 12 of which you can access for free at:
http://cancerres.aacrjournals.org/content/74/17/4694.full#ref-list-1

Citing articles

This article has been cited by 3 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/74/17/4694.full#related-urls

E-mail alerts

Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions

To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/74/17/4694.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.