Multiplexed analysis of lung cancer for distinguishing adenocarcinoma from squamous cell carcinoma

Sireesha Kaunumalle1, Denise Hallman2, Vidya Kamath1, Yousef Al-Kofahi1, Kathleen Bove2, Rod Beck2, Fiona Ginty1, Edward Kim1, Brian D Sarachan1, Mike Lazare2,Christina Lowes1, Brian Ring2, Teng Jin Ong3, Qing Li1, Alberto Santamaria-Pang1

1GE Global Research, Niskayuna, NY ; 2Clarient Diagnostic Services, GEHC, Aliso Viejo, CA; 3Medical Diagnostics, GEHC, Princeton, NJ ; 4Levine Cancer Institute, Charlotte, North Carolina

BACKGROUND
Lung cancer is a leading cause of cancer related deaths, 80% of which are classified as non-small cell carcinomas (NSCLC). Differentiating the two main subtypes of NSCLC, adenocarcinoma and squamous cell carcinoma (SCC), is crucial for therapeutic decision-making. Current methods for characterizing lung cancer subtypes may involve DAB staining on up to 7 tissue sections, depending on complexity of diagnosis. A recent survey of 94 US pathology practices at the ASCO 2013 meeting, concluded insufficient sample availability in 6% of all NSCLC samples recently handled by these pathologists. We have shown with multiplexed analysis of FFPE tissue that 11 proteins can be measured on a single section. We also evaluated the performance of the analytical workflows in automatic biomarker scoring and in AD and SCC discrimination, with reference to the Pulmotype test of five markers with an overall sensitivity and specificity of 95% and 87%.

METHODS
The protein markers included in the study were comprised of the five antibodies from the Pulmotype test: MUC1, CK5/6, TRIM29, CEACAM5 and SLC7A5. In addition, markers TTF-1, p40, CK7, p63, NapsA were selected based on literature reports. The entire set of fluorescently conjugated antibodies against these 11 markers was multiplexed on two cohorts comprising of a 378 core tissue microarray (TMA) with 213 cases of AD or SCC diagnosed (cohort II). A second 74 core TMA with 50 cases of AD or SCC was stained with the Pulmotype markers (cohort II). Grayscale fluorescent images were algorithmically transformed into images that closely resemble traditional diamino benzidine (DAB) stain. This allows the pathologist to view each fluorescent biomarker image as a virtually created DAB stain.

RESULTS
Serial sections were DAB stained and manually scored and then compared with the algorithmically generated mDAB scores with the manual scores. The entire set of fluorescently conjugated antibodies against these 11 markers was multiplexed on two cohorts comprising of a 378 core tissue microarray (TMA) with 213 cases of AD or SCC diagnosed (cohort II). A second 74 core TMA with 50 cases of AD or SCC was stained with the Pulmotype markers (cohort II). Grayscale fluorescent images were algorithmically transformed into images that closely resemble traditional diamino benzidine (DAB) stain. This allows the pathologist to view each fluorescent biomarker image as a virtually created DAB stain.

CONCLUSIONS
We have shown that differential diagnosis of AD and SCC may be achieved using a multiplexed panel of markers in a single tissue section, when compared to the Pulmotype test panel. Concordance between fluorescence and mDAB images and DAB shows transferability of the two detection methods. Furthermore, we have demonstrated that image and data analysis tools can be applied for consistent automatic biomarker scoring.

REFERENCES
1. Richter PD et. al., Correlation between biopsy type and insufficient tissue availability: Is any biopsy good in five solid cancer types, c.129.133 (abstract, suppl) abstr e22136
2. Gerdes MJ et. al., Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. PNAS 2013 v.110(26), pp.11862-11867

Manuscript scoring of 11 targets demonstrated excellent concordance between fluorescence and DAB (Table 2). Concordance was also demonstrated between manual DAB scores and automated IF metrics for the five Pulmotype markers with an overall sensitivity and specificity of 95% and 87% (Figure3). Statistically significant validation in the 10 cases of AD, the 20 cases of SCC, the 50 cases of NSCLC and the 97 cases of normal lung showed a sensitivity of 95%, specificity of 95%, and a Cohen’s Kappa of 0.85. A new marker, Pulmotype type and Pulmotype were

Figure 1: Sample analysis workflow

Figure 2: Example stains from cohort 2, staining on the left represents traditional DAB and staining on the right represents mDAB and middle represents fluorescence image.

Figure 3. In the above confusion matrix rows represent the automatic scores and columns the manual scores. The diagonal elements represent the number of images where the manual and automated scores match. The non-diagonal elements represent incorrectly scored images.

MANUAL SCORING METHOD – Table 1

<table>
<thead>
<tr>
<th>Marker Criteria for manually scoring image as positive</th>
<th>CK5/6</th>
<th>MUC1</th>
<th>CK7</th>
<th>CK14</th>
<th>P63</th>
</tr>
</thead>
<tbody>
<tr>
<td>When cytoplasm staining was present on greater than 1% of invasive tumor cells</td>
<td>when staining was present on greater than 10% of invasive tumor cells including those cases where staining was predominantly basally located</td>
<td>when staining was present on greater than 1% of the nuclei of invasive tumor cells</td>
<td>when staining was present on greater than 10% of invasive tumor cells including those cases where staining was predominantly basally located</td>
<td>when staining was present on greater than 10% of invasive tumor cells including those cases where staining was predominantly basally located</td>
<td>when staining was present on greater than 10% of invasive tumor cells including those cases where staining was predominantly basally located</td>
</tr>
</tbody>
</table>

Comparison of Algorithmically Generated Molecular DAB images to Traditional Brightfield IHC Images

Cohort I: Subtype Classification

Table 2

Cohort II: Automated Scoring

A linear model was trained on a set of 25 samples to automatically generate binary scores for each image. The resulting model was used to predict scores on 50 test images. Performance of the linear model was tested by comparing the automatic scores with the manual scores for each biomarker.