Multiplexed Proximity Ligation Assays to Profile Putative Plasma Biomarkers Relevant to Pancreatic and Ovarian Cancer

Simon Fredriksson,1,5* Joe Horecka,1 Odd Terje Brustugun,1,6 Joerg Schlingemann,4 Albert C. Koong,2 Rob Tibshirani,3 and Ronald W. Davis1

BACKGROUND: Sensitive methods are needed for biomarker discovery and validation. We tested one promising technology, multiplex proximity ligation assay (PLA), in a pilot study profiling plasma biomarkers in pancreatic and ovarian cancer.

METHODS: We used 4 panels of 6- and 7-plex PLAs to detect biomarkers, with each assay consuming 1 μL plasma and using either matched monoclonal antibody pairs or single batches of polyclonal antibody. Protein analytes were converted to unique DNA amplicons by proximity ligation and subsequently detected by quantitative PCR. We profiled 18 pancreatic cancer cases and 19 controls and 19 ovarian cancer cases and 20 controls for the following proteins: a disintegrin and metalloprotease 8, CA-125, CA 19-9, carboxypeptidase A1, carcinoembryonic antigen, connective tissue growth factor, epidermal growth factor receptor, epithelial cell adhesion molecule, Her2, galectin-1, insulin-like growth factor 2, interleukin-1α, interleukin-7, mesothelin, macrophage migration inhibitory factor, osteopontin, secretory leukocyte peptidase inhibitor, tumor necrosis factor α, vascular endothelial growth factor, and chitinase 3–like 1. Probes for CA-125 were present in 3 of the multiplex panels. We measured plasma concentrations of the CA-125–mesothelin complex by use of a triple-specific PLA with 2 ligation events among 3 probes.

RESULTS: The assays displayed consistent measurements of CA-125 independent of which other markers were simultaneously detected and showed good correlation with Luminex data. In comparison to literature reports, we achieved expected results for other putative markers.

CONCLUSION: Multiplex PLA using either matched monoclonal antibodies or single batches of polyclonal antibody should prove useful for identifying and validating sets of putative disease biomarkers and finding multimarker panels.

© 2008 American Association for Clinical Chemistry

Large numbers of putative biomarkers for many diseases are being identified through methods that include gene expression arrays, mass spectrometry, bioinformatics, and cell biology in general. Newly identified biomarkers require performance screening using precious biobanked plasma samples for early-phase evaluation before clinical validation. Thus, there is a need for sensitive and multiplex protein detection methods with low sample consumption.

Two currently used multiplex protein detection procedures, antibody arrays and fluorescent bead-based assays, are hampered by slow assay development time for new markers, limited multiplex, antibody cross-reactivity, and sometimes lack of sensitivity. Proximity ligation was developed to conveniently and sensitively detect proteins by converting them to DNA molecules for subsequent quantification (1, 2). The technique uses pairs of antibodies coupled to DNA oligonucleotides that, upon solution phase binding to target proteins, gain an increase in local concentration, enabling hybridization of a third oligonucleotide. A DNA ligase then seals the nick, uniting the 2 DNA strands. The ligation product is then detected by PCR using primers flanking the ligation site.

We recently established a multiplex version of proximity ligation assay (PLA)7 with improved sensitivity and dynamic range (3). Multiple proteins are thereby converted into unique amplicons for separate detection by real-time quantitative PCR (qPCR).
contrast to solid-phase and bead-based approaches for multiplexed protein detection, PLA detects only the cognate pairs from dual antibody binding events. This is achieved by using either unique ligation sites or matched primer pair combinations for distinct detection. As an aspect of the probe sequence design, any possible cross-reactive binding to a related antigen by 2 noncognate probes does not result in a detected signal. General performance specifications of PLA and comparisons to conventional immunoassays have been published (3).

Here, we configured 4 panels of 6- and 7-plex assays to evaluate multiplex PLA for biomarker profiling. We chose both low (fmol/L) and high (nmol/L) abundance markers. Because proximity ligation can use either a matched monoclonal antibody pair or a single batch of affinity purified polyclonal antibody for antigen recognition (2), we devised assays using both types of probes. We analyzed plasma samples from cases and controls.

**Materials and Methods**

**ASSAY PROCEDURE**

Multiplex PLA was performed as described (3) with the following modifications. To 10 μL human plasma sample, we added 10 μL sample dilution buffer (Olink AB) along with recombinant platelet-derived growth factor (PDGF)-BB added at a concentration of 300 pmol/L to assess reaction quality. The amount of PDGF-BB added is far above the endogenous PDGF-BB present in human plasma. We incubated the mixtures at 25 °C for 30 min before incubation with proximity probes and also further diluted them 50-fold to enable quantification of high-abundance proteins. The 4 μL proximity probe incubations were supplemented with 1 mL/L Triton X-100, 0.1 g/L aprotinin, and 10 nmol/L phenylmethylsulfonyl fluoride (Sigma), and ligation reactions were supplemented with 1 mL/L Triton X-100 and 0.1 g/L BSA. Preamplification was performed with 10 nmol/L each pool-PCR primer for 13 cycles with a 4-min extension at 60 °C, and preamplification products were diluted 10-fold in 1× TE. We prepared 384-well plates with target protein–specific qPCR detection primers lyophilized at the bottom of each well, to which we added sample and qPCR master mix. Antibody sources and oligonucleotide probe and primer sequences are listed in Supplemental Data Table 1. We studied the following proteins: a disintegrin and metalloprotease 8 (Adam 8), CA-125, CA 19-9, carboxypeptidase A1 (CPA1), carcinoembryonic antigen (CEA), connective tissue growth factor (CTGF), epidermal growth factor receptor (EGFR), epithelial cell adhesion molecule (EpCAM), Her2/neu (ErbB2), galectin-1, insulin-like growth factor 2 (IGF-2), interleukin (IL)-1α, IL-7, mesothelin, macrophage migration inhibitory factor (MIF), osteopontin, secretory leukocyte peptidase inhibitor (SLPI), tumor necrosis factor α (TNF-α), vascular endothelial growth factor (VEGF), and chitinase 3–like 1 (YKL-40).

**DATA ANALYSIS**

We converted cycle threshold (Ct) values derived from qPCR to estimated number of starting amplicons, or PLA units (calculated as $10^{(-0.301 \times \text{Ct} + 11.45)}$). To control for potential assay failures, we analyzed signal from the added PDGF-BB, which showed no evidence of statistically significant outliers. For predictive analysis of microarrays (PAM) and P value calculations, all data were transformed from linear to log10 scale to reduce nonnormality in distribution. We used 2-tailed, non-equal variance, Student’s t test to analyze significance of difference between case and control groups for each candidate biomarker (Excel). To determine which subset of markers may best differentiate between the case and control samples, we used the nearest shrunken centroids technique, implemented in the pamr package in R (4). The misclassification rate was determined by 10-fold cross-validation.

**HUMAN PLASMA SAMPLES**

Nineteen human EDTA plasma samples for the ovarian cancer study were provided by the Pacific Ovarian Cancer Research Consortium. All samples were collected presurgery. Four specimens were stage I, 10 stage III, 5 stage IV, and 1 not determined; histological subtypes were 13 serous, 2 clear cell, 2 mucinous, 1 endometroid, and 1 not determined. The ovarian cancer–matched control samples were obtained from healthy women during their regular screening mammograms and collected and stored with identical procedures (Fred Hutchinson Cancer Research Center, Institutional Review Board #5317). Eighteen EDTA plasma
samples for the pancreatic study were collected from 13 men and 5 women in accordance with an institutional review board–approved protocol. Median age was 60 years (range 41–81 years). All had adenocarcinoma and were either previously untreated (50%) or had recurrent/progressive disease after treatment (50%) at the time of specimen collection. Untreated patients were locally advanced/unresectable, and recurrent patients were all locally recurrent/progressive. The age-matched controls for the pancreatic cancer samples were provided by healthy volunteers.

Results

We combined individual proximity probe pairs to configure 4 multiplex panels, each for detecting 6 or 7 markers. Individual proximity probes were made by covalent attachment of DNA oligonucleotides to antibodies (3) (see Supplemental Table 1). We used these panels to assay protein concentrations in plasma samples from pancreatic and ovarian cancer cases and age-matched controls (Figs. 1 and 2, respectively). As expected, the best performing marker for pancreatic cancer was found to be the clinically used CA 19-9 (Fig. 1). Likewise, CA-125 was the best-performing marker for ovarian cancer (Fig. 2). All samples were analyzed as undiluted and 50-fold diluted to ensure that high abundance markers (more than approximately 1 nmol/L) were also within the assay’s linear range. IGF-2, SLPI, YKL-40, EGFR, and osteopontin behaved as high-abundance markers, and data presented for them are from assays of diluted samples (Figs. 1 and 2). Median values for the 2 sample groups, cases and controls, with corresponding P values are presented in Supplemental Data Table 2.

Table 1 shows comparisons between literature reports of the expected concentrations for the markers assayed here and results obtained with multiplex PLA. When published information was available, concordant results were obtained for CA-125, CA 19-9, CEA, YKL-40, EGFR, MIF, ErbB2 (conflicting reports for ErbB2 in ovarian cancer), mesothelin, osteopontin, TNF-α, and VEGF. There were only a few instances of disagreement between the marker concentrations re-
ported in the literature and those measured by PLA, as indicated by having a $P$ value $< 0.05$. We believe these discordances to be attributable to differences in measurement technologies and/or the particular antibodies used.

As for MIF levels in pancreatic cancer, we observed an average increase of 30% ($P = 0.15$), close to the reported 50% increase (5). Regarding CPA1, a protein expressed only in pancreatic tissue, the levels reported in the literature were measured with an enzymatic assay and showed only a slight increase of 50% in activity and a large variation between samples (6, 7). In the ovarian cancer sample set, only a few cases had increased concentrations of either CA 19-9 or CEA, which could be explained by these markers being increased mostly in the mucinous histological subtypes represented by only 2 stage I cases in this sample set (8). IGF-2 is reported to be downregulated in ovarian cancer, and we observed a 20% reduction of the median value ($P = 0.073$), nearly reflecting the expected results. IL-1$\alpha$ when assayed for ovarian cancer was the only marker with results clearly conflicting with literature reports (9).

Some markers have known binding proteins that could affect their detection by antibodies, depending on epitope accessibility when the marker is bound to another protein. This is particularly important to note for IGF-2 and the soluble receptors EGFR and ErbB2. Several explanations have been proposed for the reduced concentrations measured for serum markers such as EGFR and ErbB2 in cancer patients, including blocked detection by ligand binding, formation of autoantibodies, and enhanced clearance rates from serum (10).

To demonstrate the modularity and reproducibility of PLA, probes for CA-125 were included in 3 of the 4 panels. The measured CA-125 concentrations for each ovarian cancer case and matched control sample are compared in Fig. 3. Irrespective of which other proteins were simultaneously detected with CA-125 in the
multiplex panels, the same measured concentrations of CA-125 were obtained.

We also evaluated the correlation between multiplex PLA and a bead-based Luminex assay (11) for the ovarian cancer sample set. PLA reached a detection plateau for samples with the highest CA-125 concentrations, which is due to the inherent hook effect encountered in all dual-binding homogeneous assays performed with binders in subsaturating conditions. This also accounts for the reduced correlation of the 2 assays for samples of high CA-125 concentration. For this reason, we routinely assay both undiluted and 50-fold diluted samples.

Diagnostic accuracy can often be improved by the use of multiple markers complementing each other for increased specificity and sensitivity. To illustrate how such panels may be found using PLA data, we performed PAM (4) investigating which subset of markers might improve classification accuracy compared to CA 19-9 alone for pancreatic cancer, and compared to CA-125 alone for ovarian cancer. By assessing the cross-validation error rate for pancreatic cancer, CA 19-9, YKL-40, CA-125, CEA, osteopontin, ErbB2, Adam 8, SLPI, VEGF, IGF-2, and CTGF showed promise combined into a panel assay. Additional analysis of more samples in an independent sample set is required to confirm the potential utility of these assays (work in progress). We were unable to identify any potential cross-validated improvements for the ovarian cancer set using a multimarker approach. Future studies will show if other putative markers can better complement CA-125, besides the chosen few selected in this study.

Because CA-125 and mesothelin form a complex and together mediate cell adhesion and possibly influence metastasis (12, 13), we attempted to measure levels of this complex in plasma. PLA-based interaction assays can be configured in many ways, including those

Table 1. Comparison of biomarker levels reported in the literature with levels measured by PLA in this study.

<table>
<thead>
<tr>
<th>Protein marker</th>
<th>Pancreatic cancer</th>
<th>Ovarian cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PLA</td>
<td>Literature</td>
</tr>
<tr>
<td>Adam 8</td>
<td>—</td>
<td>NI</td>
</tr>
<tr>
<td>CA-125</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>CA 19-9</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>CPA1</td>
<td>—</td>
<td>↑</td>
</tr>
<tr>
<td>CEA</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>YKL-40</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>CTGF</td>
<td>—</td>
<td>NI</td>
</tr>
<tr>
<td>EGFR</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>EpCAM</td>
<td>—</td>
<td>NI</td>
</tr>
<tr>
<td>ErbB2</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Galectin-1</td>
<td>—</td>
<td>NI</td>
</tr>
<tr>
<td>IGF-2</td>
<td>↑</td>
<td>↑ (tissue)</td>
</tr>
<tr>
<td>IL-1α</td>
<td>—</td>
<td>NI</td>
</tr>
<tr>
<td>IL-7</td>
<td>—</td>
<td>NI</td>
</tr>
<tr>
<td>Mesothelin</td>
<td>—</td>
<td>↑ (tissue)</td>
</tr>
<tr>
<td>MIF</td>
<td>—</td>
<td>↑</td>
</tr>
<tr>
<td>Osteopontin</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>SLPI</td>
<td>↑</td>
<td>NI</td>
</tr>
<tr>
<td>TNF-α</td>
<td>—</td>
<td>NI</td>
</tr>
<tr>
<td>VEGF</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>CA-125–mesothelin</td>
<td>ND</td>
<td>NI</td>
</tr>
</tbody>
</table>

Plasma or serum level (or cancer tissue staining, if so noted) changes as reported in the literature: up regulated (↑), down regulated (↓), conflicting information reported (↑ / ↓), no change (—), no information found (NI), not determined (ND). PLA data with P values <0.05 were classified as either up- or downregulated; otherwise, no change (—).
that generate a signal from more than 2 recognition events by the detection of multiple proximity-dependent nucleic acid interactions (14–16). Of the several options available for detecting multiple binding events, we used one originally proposed by Landegren and Fredriksson (17), whereby 1 detection amplicon is generated by 2 connectors mediating the ligation of oligonucleotides coupled to 3 proximity probes. In terms of the oligonucleotides coupled to the proximity probes, the assay is configured such that the 3′ end of 1 CA-125 probe is ligated to the 5′ end of a partially double-stranded mesothelin probe, which in turn is ligated at its 3′ end to the 5′ end of the other CA-125 probe. This triple-specific assay detected increased concentrations of the CA-125–mesothelin complex in several of the ovarian cancer cases (Fig. 2). Further analysis with additional samples will be required to assess the diagnostic utility of the CA-125–mesothelin complex as a marker for ovarian cancer detection and disease progression.

**Discussion**

Multiplex PLA is a promising technology for scalable and sensitive analysis of candidate plasma biomarkers. We have shown here that PLA has good precision and potential for higher scalability than achievable by current solid-phase approaches. Signals from antibody cross-reactive events are not detected in the dual-specific nucleic acid reporter system used. This lack of assay cross-reactivity in PLA enables a modular use of the specific assays within a panel for reconfiguration to fit particular multiplex needs. The need to carefully assess antibody cross-reactivity is thereby greatly reduced, as seen by the consistent results for CA-125 in 3 of the multiplex panels used.

In our previous report on multiplexed PLA, a calibration curve was used for quantification of each protein (3). An impracticality of that approach was the variable storage stability of the individual antigen standards. On further multiplex expansion, this...
may become an increasing difficulty, whereas the relative profiling procedure presented here alleviates this challenge. This also stresses the important point of uniform sample handling in biomarker discovery studies.

In the present study, we used assays based on both matched monoclonal antibodies and affinity-purified polyclonal antibodies raised against native recombinant proteins. The correlation between our PLA data and biomarker levels reported in the literature demonstrates that the dual and proximal recognition characteristics of PLA provide the required specificity for analysis of human plasma samples even with polyclonal antibodies. When configuring new assays for novel markers in an initial screening setting, it is worth noting that polyclonal antibodies require less time and cost to generate than matched monoclonal antibodies. In downstream diagnostic settings, however, monoclonal antibodies may prove advantageous because of their batch-to-batch consistency. Multiplexed assays are in general more cost efficient than conventional assays by the reduction in assay time and reagent consumption per data point. This is also the case for PLA.

As seen by the detection of the CA-125–mesothelin complex, proximity probes are highly suitable for assaying protein–protein interactions by reconfiguring the ligation and detection oligonucleotides used.

With further developments of the assay and with additional expansion of the multiplexing, PLA-derived amplicons could be analyzed on a number of quantitative detection platforms such as DNA microarrays, mass spectrometry, microfluidic real-time PCR, and others. Biomarker research targeting large sets of putative markers and novel panel combinations in precious and limited biobanked blood samples can be rapidly enhanced with this profiling procedure.

Grant/funding Support: This study was sponsored by the Swedish Research Council, NIH grant 2P01HG00025, Glue grant U54GM062119, the Canary Foundation, and the Liu Bie Ja Cha and Family Fellowship in Cancer. J.S. was supported by a Marie Curie inter-European fellowship.

Financial Disclosures: S.F. holds shares in Olink AB (Uppsala, Sweden).

Acknowledgments: Ovarian cancer samples and their accompanying CA-125 Luminedx data were provided by the Pacific Ovarian Cancer Research Consortium (P50 CA83636; PI: Nicole Urban). We thank Angela Chu for help with qPCR data processing, Heidi Schlingemann for technical support, Stephanie Chang for collecting and organizing sample information, and David Schwartz and Jim Williams at Solulink for the synthesis of the oligonucleotide-antibody conjugates.

References


