Multiplexed protein detection by proximity ligation for cancer biomarker validation

Simon Fredriksson[1], William Dixon[1], Hanlee Ji[1], Albert C Koong[2], Michael Mindrinos[1] & Ronald W Davis[1]

We present a proximity ligation–based multiplexed protein detection procedure in which several selected proteins can be detected via unique nucleic-acid identifiers and subsequently quantified by real-time PCR. The assay requires a 1-µl sample, has low-femtomolar sensitivity as well as five-log linear range and allows for modular multiplexing without cross-reactivity. The procedure can use a single polyclonal antibody batch for each target protein, simplifying affinity-reagent creation for new biomarker candidates.

Protein-based biomarkers in blood hold a great promise as diagnostic markers indicative of disease states and outcomes in clinical cancer management. To validate large sets of candidate markers in bio-banked sample collections, multiplexed and sensitive detection technologies with low sample consumption are required. Such technologies will not only be critical for the discovery of biomarker panels potentially leading to increased predictive power in diagnostics but may eventually permit early stage disease detection[1,2].

Most conventional immunoassays rely on a solid support for capture of the target protein and for the removal of excess secondary reporter antibody by washing. The sensitivity of these sandwich assays is limited by the nonspecific binding of the secondary reporter to the surface of the solid support. Also, any single noncognate binding event by a detection antibody will give rise to a false positive signal. This is especially challenging when performing multiplexed reactions with many detection antibodies, requiring extensive and careful optimization through selection of particular antibody combinations to minimize nonspecific cross-reactivity[3,4]. To reduce this scalability bottleneck and lower sample consumption, we configured a protein detection technology, the proximity ligation assay, to remove potential cross-reactive signals when performed in multiplex. In multiplex, we also show the assay can operate with simple affinity reagents such as one single batch of polyclonal antibody per target protein, and we apply the assay in clinical blood samples for biomarker validation.

The technical ability to analyze nucleic acids is far more advanced than that for proteins. Proximity ligation has been developed in an attempt to bridge this gap and thereby improve the limit of detection, specificity, dynamic range and multiplex capability[5,6]. The technology is a unique approach for protein quantification (often seen in nucleic acid detection) wherein the analyte promotes the creation of a new and distinct reporter molecule. This is in contrast to heterogenous methods, which require washes to remove unbound secondary reporters after solid-phase capture. The homogenous, or liquid phase version of the technology uses a pair of proximity probes each composed of an antibody linked to an oligonucleotide. As the two antibodies bind the protein analyte in solution, the local concentration of the two corresponding oligonucleotides increases. This allows the hybridization of one connecting oligonucleotide to both probes, thus enabling an enzymatic ligation joining the 3′ end of the first probe with the 5′-end of the second probe. This leads to the formation of a unique target reporter amplicon containing specific molecular bar codes, a frequently used procedure for multiplexed detection of nucleic acids[7]. These molecular bar codes serve as primer sites, of which some are universal for all protein analytes, whereas others are target-specific for quantification by real-time PCR. The assay reporter signal is dependent on a proximal and dual recognition of each target analyte providing high specificity. We outline the multiplex detection procedure in **Figure 1,** and detailed protocols are available in the **Supplementary Methods** online.

To suppress the level of background ligations, we incubated the proximity probes at subsaturating concentrations during the initial incubation with the sample. Therefore, the sensitivity and dynamic range of the assay is dependent on the affinity of the antibody[6].

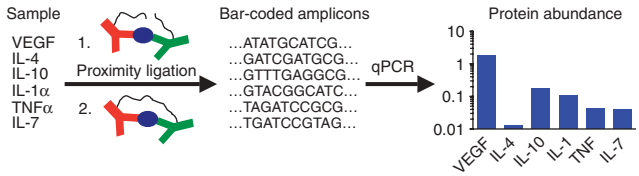

Figure 1 | Schematic outline of multiplexed proximity ligation. The sample is first incubated with sets of proximity probes, antibodies equipped with oligonucleotides in step 1. A connector oligonucleotide is then added in step 2 uniting proximal sequences forming unique bar-coded amplicons representing each target protein. This is followed by preamplification of all ligation products. Finally, the surrogate markers for the proteins comprised within the pool are analyzed by real-time PCR in a miniaturized format, generating quantitative protein-abundance data.

[1]Stanford Genome Technology Center, Bio-X, 318 Campus Dr., Stanford, California 94305, USA. [2]Department of Radiation Oncology, Stanford, California 94305, USA. Correspondence should be addressed to S.F. (simon.fredriksson@stanford.edu) and R.W.D. (dbowe@stanford.edu).

RECEIVED 20 OCTOBER 2006; ACCEPTED 23 JANUARY 2007; PUBLISHED ONLINE 18 MARCH 2007; DOI:10.1038/NMETH1020

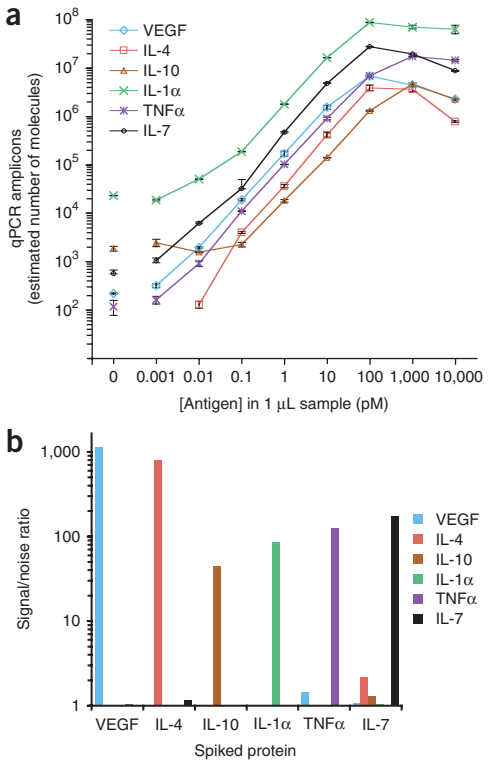

Figure 2 | Multiplex proximity ligation assays. (**a**) Standard curves of diluted analytes detected in six-plex assays. Each data point represents duplicate detection at the qPCR stage (error bars, 1 s.d.). (**b**) qPCR analysis of individual analytes spiked into chicken plasma samples. Colored bars represent single qPCR measurements. Each *x*-axis category represents one added antigen.

Optimally, the assays should use high concentrations of proximity probes to promote target binding and ensure a wide dynamic range while maintaining low levels of background ligation events. The background noise in proximity ligation is derived from two main sources: first, proximity probes nonspecifically binding to each other, and second, the connector oligonucleotide binding to two freely diffusing probes, enabling ligation. By probe sequence design *in silico*, empirical experimentation and connector reconfiguration, we considerably improved the performance, permitting the detection of proteins in low-femtomolar concentrations in 1-μl samples over a five-log linear range in multiplex (**Fig. 2a**). A large linear range is critical for multiplexed biomarker assays as proteins vary over large concentration ranges in the blood[8]. Assays using medium- to low-affinity antibodies are enhanced over 100-fold in sensitivity by permitting the use of tenfold-higher probe concentrations with still lower background compared to previous assay configurations[6]. For optimization details and a comparison to previous proximity ligation configurations and also enzyme-linked immunosorbent assay (ELISA), see **Supplementary Figures 1** and **2** online.

In each assay, we used either a single batch of an affinity-purified polyclonal antibody or a matched monoclonal antibody pair. To reduce antibody cross-reactivity, we took advantage of features available for specific nucleic acid ligation and detection by PCR. We designed the six proximity probe pairs with either unique ligation sites on the connector sequence or unique primer combinations for quantitative PCR (qPCR) detection. In the presence of a complex mixture containing chicken plasma, we spiked 10 pM of one of the six target proteins into the incubation together with all the six probe pairs and recorded the signal over background for each reaction (**Fig. 2b**). We observed no antibody cross-reactivity even without optimization of conditions or requirement of selection of particular antibody combinations. This demonstrates potential for scalability beyond that of solid phase–based arrays, presently limited to ~10–20-plex. In comparison, multiplexing with multicolor bead–based assays has in some cases been reported to result in a loss of sensitivity, resulting in certain proteins becoming undetectable in plasma[9]. As cross-reactive events are not detected in proximity ligation, the assay is uniquely modular in the sense that new analytes can be added to an existing panel of prevalidated sequence systems without the need for reconfirming assay specificity at the antibody level (S.F. and R.W.D., unpublished data). The characteristics, reagents and performance of all six proximity ligation assays are summarized in **Supplementary Table 1** online and oligonucleotide sequences are available in **Supplementary Table 2** online. Multiplexing will become increasingly important as biomarker panels of multiple proteins can improve diagnostic performance[10].

To validate the technology in biological samples, we analyzed human plasma of 14 healthy age-matched control patients and 20 patients with untreated locally advanced pancreatic carcinoma for these six analytes in one multiplex panel and also for three additional markers in a separate higher-abundance multiplex panel. These first six proteins targeted for detection were not initially chosen for their relevance in pancreatic cancer but taken to represent a variety of low-abundance plasma proteins. We quantified each protein using standards of diluted analytes (**Fig. 3**). The expected concentrations of these proteins in plasma for healthy subjects spanned from undetectable amounts of IL-4 up to low-picomolar amounts of VEGF (**Supplementary Table 1**). Because of each assay's limit of detection and, more importantly, the expected abundance of the particular analyte, we did not detect

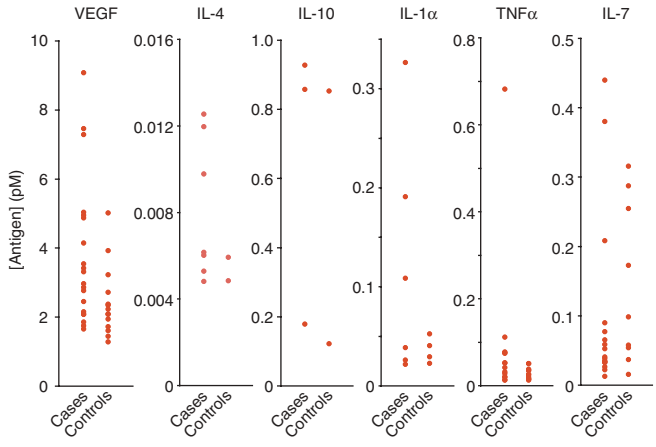

Figure 3 | Multiplex proximity ligation of six proteins and subsequent detection by qPCR in plasma samples from 20 patients with pancreatic cancer and 14 matched healthy controls. Omitted data points were below the limit of detection. The mean VEGF level for controls and pancreatic cancer cases were significantly different, 2.42 pM and 3.80 pM, respectively ($P = 0.03$). The differences for the other proteins were not statistically significant.

all analytes in all samples. We detected VEGF in all samples and, in agreement with some literature reports, several pancreatic cancer samples had significantly elevated amounts of VEGF compared to healthy controls ($P = 0.03$)[11]. However, conflicting reports on VEGF levels and also those of other cytokines in plasma samples from pancreatic cancer demonstrate the need for improved tests with higher predictive value, possibly through multiplexing[12]. By reusing three of the same probe sequence pairs of the six-plex assay, we ran a separate three-plex panel for the high-abundance proteins IGF-2, CA-125 and mesothelin on the same sample set, which showed elevated amounts of these proteins compared to controls in some cases (**Supplementary Fig. 3** online). We performed a colorimetric sandwich ELISA for VEGF and TNFα, which showed correlation between the two methods (**Supplementary Fig. 4** online). Intra-assay coefficient of variance for the entire multiplexed detection procedure was 12%, whereas samples analyzed on different occasions could benefit from the inclusion of internal reference proteins, either spiked in or of endogenous 'housekeeping'-type (S.F. and R.W.D., unpublished data; **Supplementary Fig. 5** online). This pilot study, which previously had been complicated by the limited sample availability, was made possible by the low sample consumption per data point of six-plex proximity ligation at 0.16 µl compared to 100 µl per data point for six individual ELISAs. Additionally, some of the target proteins are below the ELISA limit of detection.

The workflow of multiplexed proximity ligation assays conceptually resembles that of cDNA synthesis, but for targeted proteins only. The subsequent quantitative detection should be performed on the appropriate platform depending on multiplexing level, for example, DNA microarrays or high-throughput microfluidic real-time PCR. As suggested by the performance of multiplexed nucleic acid detection procedures using molecular bar codes of minimal similarity[7] and the ease with which this six-plex was set up, we anticipate the assay to scale well without cross-talk between the sequences designed and linked to the antibodies. Such potential interference would then require redesign of individual sequences. Proximity ligation has previously been applied to the detection of single proteins[5], protein-protein interactions[13], protein modifications[14], bacteria and virus particles[15] and now also in multiplex for biomarker validation in a pilot study using clinical blood samples.

The presented assay can use either affinity-purified polyclonal antibodies or matched monoclonal pairs. The generation of a single batch of polyclonal antibody is both more rapid and less expensive compared to generation of matched monoclonal antibody pairs. The presented procedure may provide a detection platform facilitating a broad generation of binders and scalable assays for validation of new biomarker candidates of the plasma proteome.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

This work was supported by the Swedish Research Council, the Canary Foundation, the Liu Bie Ju Cha and Family Fellowship in Cancer and the US National Institutes of Health (Center Grant 2P01HG000205 and Glue Grant U54GM062119).

AUTHOR CONTRIBUTIONS

S.F., M.M. and R.W.D. initiated the project; S.F. and R.W.D. designed experiments and methods; S.F. and W.D. performed experiments; A.C.K. supplied clinical samples; S.F., H.J., A.C.K. and R.W.D. analyzed data; S.F. drafted the initial manuscript and all authors contributed to subsequent revisions.

COMPETING INTERESTS STATEMENT

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/naturemethods.

Published online at http://www.nature.com/naturemethods
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Aebersold, R. *et al. J. Proteome Res.* **4**, 1104–1109 (2005).
2. Etzioni, R. *et al. Nat. Rev. Cancer* **3**, 243–252 (2003).
3. Elshal, M.F. & McCoy, J.P. *Methods* **38**, 317–323 (2006).
4. de Jager, W. & Rijkers, G.T. *Methods* **38**, 294–303 (2006).
5. Fredriksson, S. *et al. Nat. Biotechnol.* **20**, 473–477 (2002).
6. Gullberg, M. *et al. Proc. Natl. Acad. Sci. USA* **101**, 8420–8424 (2004).
7. Hardenbol, P. *et al. Nat. Biotechnol.* **21**, 673–678 (2003).
8. Anderson, N.L. & Anderson, N.G. *Mol. Cell. Proteomics* **1**, 845–867 (2002).
9. Liu, M.Y. *et al. Clin. Chem.* **51**, 1102–1109 (2005).
10. Scholler, N. *et al. Clin. Cancer Res.* **12**, 2117–2124 (2006).
11. Karayiannakis, A.J. *et al. Cancer Lett.* **194**, 119–124 (2003).
12. Garcea, G., Neal, C.P., Pattenden, C.J., Steward, W.P. & Berry, D.P. *Eur. J. Cancer* **41**, 2213–2236 (2005).
13. Söderberg, O. *et al. Nat. Methods* **3**, 995–1000 (2006).
14. Zhu, L. *et al. Biol. Chem.* **387**, 769–772 (2006).
15. Gustafsdottir, S.M. *et al. Clin. Chem.* **52**, 1152–1160 (2006).