UPPSALA
UNIVERSITET

UPTEC X 14 009
Examensarbete 30 hp
April 2014

Optimization of the multiplexed Proximity Ligation Assay for detection of blood-based biomarkers

Martin Lundberg

UPPSALA UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 14 009	**Date of issue 2014-04**

Author

Martin Lundberg

Title (English)

Optimization of the multiplexed Proximity Ligation Assay for detection of blood-based biomarkers

Title (Swedish)

Abstract

The Proximity Ligation Assay (PLA) is a relatively new method which utilizes the strength of both immunoassays and DNA detection. PLA has the capacity of high multiplexing due to the high specificity achieved with both dual protein-binding and dual primer binding during detection with Real-Time PCR. We developed a multiplexed PLA protocol that can measure 28 biomarkers in human EDTA plasma. The method was tested on 46 individuals diagnosed with colorectal cancer and 48 age matched healthy controls. The results are very promising as we re-discover the most well-known biomarkers for colorectal cancer and also find some potential new markers (significance tested with students T-test with $p<0.05$). Further improvements of the protocol are needed to decrease the variation.

Keywords

PLA, colorectal cancer, biomarker, multiplex, qPCR, Olink

Supervisors

Dr. Erika Assarsson
Olink AB

Scientific reviewer

Patrik Lindstedt
Olink AB

Project name	Sponsors
PROACTIVE	**EU FP7**
Language: **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages: **45**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 471 4687

Optimization of the multiplexed Proximity Ligation Assay for detection of blood-based biomarkers

Martin Lundberg

Populärvetenskaplig sammanfattning

Idag så är kolorektalcancer en av världens vanligast förekommande cancersorter, med över 200 000 dödsfall i Europa varje år. Skälet till den höga andel människor som inte överlever länge efter diagnos av sjukdomen beror på att man hittar cancern i ett sent skede, där behandlingsalternativen inte är tillräckligt effektiva. För att tidigare upptäcka kolorektalcancer behövs nya metoder som effektivt kan användas för att rutinmässigt leta efter riskfaktorer. Detta är en stor utmaning, då det finns tiotusentals potentiella biologiska markörer att undersöka.

Detta arbete beskriver hur en relativt ny metod kallad Proximity Ligation assay, eller PLA, utvecklas för detta ändamål. Metoden är unik i det att man i teorin kan mäta flera olika biologiska markörer samtidigt, utan att det påverkar mätresultaten, något som andra metoder har problem med. Detta gör tekniken lämplig för att leta efter många markörer snabbt, och därmed identifiera nya markörer som kan avslöja sjukdomar i ett tidigt skede.

Examensarbete 30 hp
Civilingenjörsprogrammet Molekylär Bioteknik

Upppsala Universitet, april 2014

Table of contents

1. Abbrevations

Amplicon	complete DNA-sequence that is measured
Ag	Antigen
Ab	Antibody
Biomarker	Biological marker
Connector	Oligonucleotide complementary to both ends of 3'- and 5'-PLA probes
CRC	Colorectal cancer
Ct	Cycle threshold
CU	University of Copenhagen
CV	Coefficient of variation
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
FBB	Fish gelatine Blocking Buffer
fM	femto Molar
GFP	Green Fluorescent Protein
GP	Goat Probe
HAMA	Human anti-mouse antibody
Hook effect	When too much antigen is present
IgG	Immunoglubulin G
In silico	Performed on computer
In solution	Experiment performed in solution in tubes rather than on solid surface
mAb	Monoclonal antibody
Multiplex	Measuring several analytes simultaneously
Oligo	Oligonucleotide
pAb	Polyclonal antibody
PCR	Polymerase Chain Reaction
PE	Phycoerythrin
PLA	Proximity Ligation Assay
pM	pico Molar
qPCR	Quantitative Real-Time PCR
RT-PCR	Real-Time PCR
Spike-in	Add a known amount of antigen
Taqman probe	Oligonucleotide labeled with a fluorescent quencher and emitter

2. Introduction

2.1. Cancer

One of modern society's biggest health problem is cancer. With high mortality rates and expensive treatments it is a huge burden not only for the patient, but for the whole society. Current treatments available for several types of cancer are not sufficient for survival if initiated at a late stage of the disease. If the treatment instead could start in an early stage, the survival rate would increase drastically for most types of cancer[1]. By recognising this situation, it is easy to understand that early diagnostic tools are very much needed in order to decrease cancer mortality rates.

Colorectal cancer (CRC) is the second most deadly cancer in the western world, with over 50 000 deaths in the US[2,3] and more than 200 000 deaths in Europe each year[4]. With less than 12% five year survival rate in case of late stage diagnosis and more than 90% five year survival rate if diagnosis occurred in an early stage (see Figure 1), it is indeed important to develop a better diagnostic tool for early detection[3].

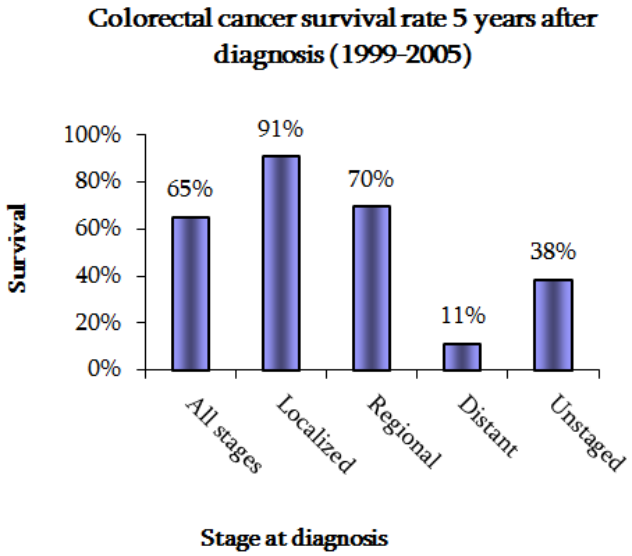

Figure 1. Survival rate for different stages of colorectal cancer when diagnosis occurred. With almost 80% difference between distant and localized tumours it is easily understood that early detection is crucial to increase the survival rate.

Localized stage of colorectal cancer comprises tumours located in the colon or rectum (see Figure 2). The regional tumours have then spread to the regional lymph nodes and in case of distant stage the cancer has spread to other organs. The localized tumours can be surgically removed, but when reaching the late stages of cancer, chemotherapy will be the only remaining treatment option.

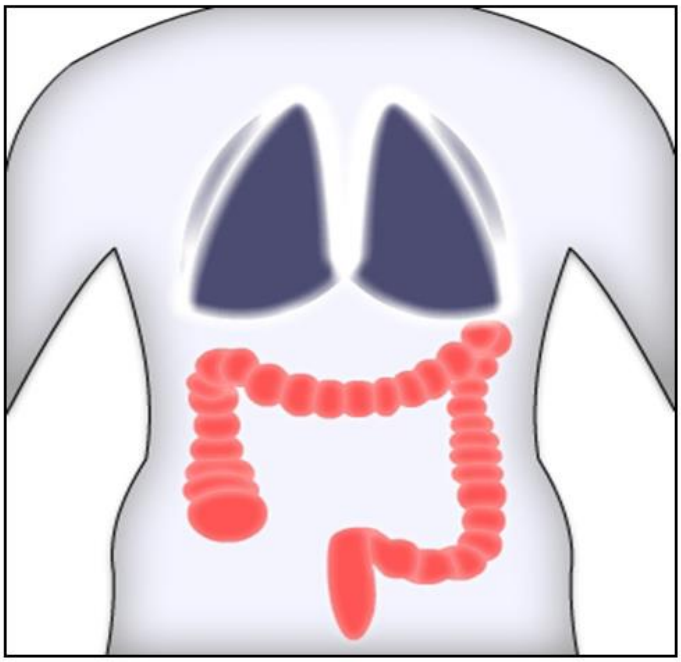

Figure 2. The localized tumour stage of colorectal cancer comprises tumours situated in the colon and rectum (red). Regional tumours have spread to lymph nodes in the area and when reaching the distant stage the tumours have spread to other organs as well.

2.2. Methodology

Today there is a lack of diagnostic methods for early detection of several types of cancer. One reason might be that cancer research up until today mainly has focused on finding biomarkers that single-handedly can detect cancer in different body fluids, mainly blood plasma. Plasma has the advantages of being relatively easy to preserve, having the largest fluid proteome and is also easy to obtain from patients[5]. However, it has been proven hard to find cancer specific biomarkers due to the complex plasma composition; with an estimate of more than 100 000 proteins with concentrations spanning over 12 orders of magnitude[5]. Methods used for diagnostic purposes need to have high sensitivity and specificity when discriminating between cases and controls, which is often hard to accomplish when measuring only one biomarker. Although there are many markers significantly up- or down-regulated for most diseases, overlaps in protein levels between groups are very common[1]. By detecting multiple biomarkers simultaneously and perform multivariate data analysis, it could be easier to discriminate cancer patients from controls. It could be compared with the separation of two points in a coordinate system; for one dimension (or one marker) the two points (case and control levels) can only move on a straight line, but if you increase the number of variables to two you suddenly can vary both these variables and get a better separation.

This is why it is desirable to have a robust method for detection of biomarkers in multiplex[1]. For a biomarker discovery platform the multiplexing capacity is very important for the high throughput. Achieving high throughput with multiplexing is not only a fast way to find new markers, but also makes it easier to find panels of markers. When screening for candidate biomarkers for diagnostic tools, a method with high sensitivity is more likely to find new, important and low abundant proteins in plasma.

Many proteins in plasma require a method with low femtomolar sensitivity to be detected[5].

A rather new technology suited for this kind of detection is *in solution* Proximity Ligation Assay (PLA®) which is based on proximity ligation. PLA® has previously been used for detection of biomarkers in blood plasma without cross-reactivity and with sensitivities down to low femtomolar[6]. The method consumes very small sample volumes; down to 1 µL - which also makes it optimal for studies on biological samples that are often hard to obtain in large volumes. Due to the theoretically high multiplexing capacity of the method it is well suited for high throughput analysis for screening of new biomarkers[6].

In solution PLA® is an immuno based assay. The technology uses two antibodies that recognize the same antigen, on these antibodies there are oligonucleotides (oligos) attached. When both antibodies bind to the same antigen these oligos are in proximity and can then ligate and form one single strand. When this ligation product is formed it represents one antigen. The reason for using DNA detection is partly due to the high sensitivity that can be reached with Real-time quantitative PCR (RT-qPCR), but the most important property is the high specificity which makes it suitable for multiplexing; by using different oligo sequences it is possible to form barcode like sequences, each one with its own specific PCR primers. Due to the very specific binding of DNA primers in PCR it is possible to achieve high multiplexing without any mis-binding of different primers. Even if one primer (out of two in one pair) misbind, it does not yield an exponential amplification of the template unless the other primer also misbind to the same template, hence this has extremely low probability to occur.

Let us assume that the forward primer for analyte X has a low affinity for analyte Y's forward primer site, and that 1% of the total Y-amplicons are bound by forward X and also extended. Now after the first cycle there is 1% of the start-material that has a forward X incorporated. Now let us also assume that the reverse X primer has affinity for the reverse Y primer site and binds 1% of all the amplicons. When this primer has extended (after cycle #2) there is four types of amplicons (with corresponding complementary strands);

- ~96% - Forward Y and Reverse Y (not amplified with X-primers, start material).
- ~2% - Forward X and Reverse Y (only forward X has extended two cycles).
- ~2% - Forward Y and Reverse X (only reverse X has extended two cycles).
- ~0.02% - Forward X and Reverse X (1% of reverse X bound the first round of forward X amplicons (0.01%) and 1% of the forward X bound the first round of reverse X amplicons (0.01%).

The 0.02% of amplicons with both forward and reverse X will then be exponentially amplified, but will lag behind from the beginning with 2 cycles + ~12 .3 (2^12.3 = 5042 times equals 0.02%). This means that if the actual Ct-value for analyte Y is 15 (read with Y-primers), the value read with X-primers is ~29 – 20,000 times lower than the real value. Even such a poor design of primers does not yield a noticeable signal.

Two conjugates targeting different proteins can also be ligated, forming an amplicon with forward X and reverse Y sites. This event occurs all the time in the "background ligation"; probes interact now and then as they float around, and are then

accidentally joined. This is also what sets the background for matched conjugates. These random events depend mainly on the conjugate concentration and ligation time. But to detect an amplicon that has non-matched sites it requires the same non-matched primers in the RT-qPCR detection. This also means that if one conjugate (X) cross-reacts with another analyte (Y) in the same panel, the amount of non-matched amplicons increases with the Y concentration. But in order for this antibody cross-reactivity to affect the readout (X or Y) it requires misbinding of the X or Y primers as well.

2.3. Biomarkers for colorectal cancer

Most markers selected for this screening are either known to be up- or downregulated in CRC patients or in other types of cancer for the pilot round in the project. Many interesting CRC biomarkers have been found by other researchers, but are not yet confirmed to be appropriate for clinical diagnostic purposes (see table 5 in section 5.10 for some examples). By including some less well-studied markers and not only focus on the already known, we hope to be able to find new important markers for CRC. However, this is most likely to pay off at a later stage when more than a hundred biomarkers will be screened.

2.4. Aim and challenges

The aim with this project is to optimize the *in solution* PLA® technique for multiplex detection of biomarkers in blood plasma. The overall purpose is to develop a cheap and robust high throughput method to find and measure new biomarkers with clinical utility for CRC, which if successful could be adapted for screening of markers relevant for other diseases as well. Funding for developing this technique comes from the European Union in the Proactive project, which is a collaboration between Olink AB (Sweden), Uppsala University (Sweden), University of Copenhagen (Denmark), Innova Biosciences (United Kingdom) and Integromics (Spain).

There are many parameters to optimize when using multiplexed *in solution* PLA®, and it is a great challenge to develop a high throughput platform with this technology. This study is one of the largest ever to be performed with this amount of biomarkers on such a large number of samples of cases and controls. Since the method is relatively new to the scene it will most certainly raise new questions and obstacles along the way which needs to be addressed and solved.

2.5. Strategy

In order to optimize the technology to work as required, several factors needs to be considered and improved. The methodology must be robust, in other words have a high grade of technical repeatability to retrieve similar results over several runs. In the future it also needs to be reproducible in order to set up the assay in other labs. It is also important to investigate possible cross-reactivity of the antibodies and mis-binding of the PCR primers used. Also since all plasma samples are unique and contains varying amounts of inhibitory factors that affects the efficiency of the ligase (details below), some exogenous standards (or spike-ins) needs to be added to the samples. The normalization with these spike-ins are important to monitor to be sure that they work as thought. By including at

least two spike-ins it is possible to use one as a normalizer and the second one as a control. If the normalized control values still vary it is an indication that one (or both) of the spikes is unsuitable for normalization purposes. Since many different evaluations need to be performed during a specific time period (milestone for this EU project), we are prepared to put some problems aside for fixing later, depending on the severity.

3. Materials and Methods

3.1. Samples

EDTA (Ethylenediaminetetraacetic acid, used to deactivate metal-dependant enzymes) blood plasma samples from controls and CRC diagnosed patients were obtained from University of Copenhagen (CU). A set of 20 cases (10 female, 10 male, median age 70) and 20 controls (14 female, 6 male, median age 66) were used for initial testing and optimizations. For later studies an additional set of 77 cases (37 female, 40 male, median age 73) and 77 controls (37 female, 40 male, median age 72) were used.

3.2. Conjugates

In this PLA® protocol we used two monoclonal antibodies or one polyclonal batch (containing several antibodies specific to the same antigen), which recognized the same antigen but different epitopes, to bind antigens with dual recognition. Attached to these "two" antibodies were two different oligos, one with a free 3' end and one with the 5' end free, see Figure 3. These molecules are called PLA® conjugates. The oligos were designed in silico, meaning that hairpin structures, melting temperatures and GC content were analysed with computer algorithms. These oligos had unique "barcode sequences" in order to work together in multiplex.

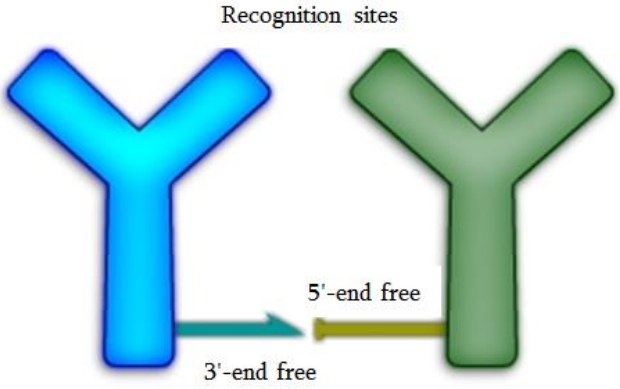

Figure 3. PLA® conjugates. Each conjugate comprise one antibody (two recognition sites each) and one attached oligo; either with the 5' end or the 3' end free. Together these work as a pair and provide dual recognition by two antibodies and also dual recognition in the amplification of the signal (PCR).

In the early stage of the project, conjugates made by Olink Bioscience were used. These were used for small scale experiments and initial testing.

With a new technology developed by Innova Biosciences Ltd (Innova) the antibodies were conjugated with the oligos quicker, easier and with higher success rate than before. The oligo sequences were designed by Olink Bioscience and antibodies were commercially purchased. A list of conjugates used can be seen in Table 1.

As mentioned, the PLA conjugates consisted of one antibody and one oligo. This is a truth with modification; it is difficult to achieve 100% conjugation of exactly one oligo to each antibody. If free antibodies were present in the reaction they would bind antigens without yielding a signal. To avoid this, excess of oligos were added during conjugation. Due to the excess of oligos there were always at least one oligo attached to each antibody. There were no problem with having more than one oligo per antibody, in fact it only increased the signal which were for the better most of the times. However if too many were attached the risk of damaging the active site of the antibody increased.

Table 1. List of conjugates generated with their corresponding oligo sequence number. Each oligo sequence number has a matched couple of 3'- and 5'-oligos. Antigens available are indicated with a cross.

Sequence ID	Antibody specificity	Antibody type	Antigen	Sequence ID	Antibody specificity	Antibody type	Antigen
7	CEACAM 5	polyclonal	X	45	CEACAM 5	monoclonal	X
7	TIMP1	polyclonal	X	45	GDNF	polyclonal	X
8	CEACAM 1	polyclonal	X	45	MIF	polyclonal	X
8	CEACAM 8	polyclonal		45	TGF beta1	polyclonal	X
8	EGFR	polyclonal	X	51	Fractalkine (CX3CL1)	polyclonal	X
8	VEGF	polyclonal	X	51	Her3	polyclonal	X
32	**GFP**	polyclonal	X	51	IL4	polyclonal	X
33	Calgranulin A (S100A8)	polyclonal		51	Mesothelin	polyclonal	
33	EGF	polyclonal	X	53	Her4	polyclonal	X
33	Her2	polyclonal	X	53	MIP-1	polyclonal	X
33	IL6	polyclonal	X	53	SLPI	polyclonal	X
35	CTGF	polyclonal		53	TNFa	polyclonal	X
35	CXCL12 (SDF-1)	polyclonal	X	57	CA 242	monoclonal	X
35	IL8	polyclonal	X	57	IL-1a	polyclonal	X
35	OPN	polyclonal	X	57	Calgranulin A (S100A8)	polyclonal	
38	**PE**	polyclonal	X	57	YKL-40	polyclonal	

3.3. Antigens

Antigens were purchased for most of the assays in order to evaluate the performance of the conjugates. The list of antigens used can be seen in Table 1.

3.4. Panel setup

The first milestone in the Proactive project were to generate four biomarker panels with ten assays in each, in other words we needed to run 10-plex reactions four times, which made it possible to measure 40 biomarkers per sample. Out of the ten assays per panel, two were used for measuring the spike-ins. The panels used the same set of ten primers for qPCR detection, and also the same conjugates for spike-ins (Green Fluorescent Protein (GFP) and Phycoerythrin (PE)). By limiting the amount of assays (and hence sequences) to ten per panel, the risk were lower to get PCR primer mis-binding compared to running all assays in one panel (four times the amount of primer pairs). Also, by dividing the assays into several panels it was possible to separate the assays with regard to plasma abundance (this was needed due to limited range of all immunoassays); two panels (A and B) were used to measure low abundant biomarkers and were run on undiluted plasma samples. The last two panels (C and D) contained assays for mid- and high abundant biomarkers and required the plasma to be diluted 50 times. If the samples were not diluted for the high abundant markers, these assays would have reached the hook effect (see Figure 7 and section 3.7) and yielded unreliable or useless data. The assays were divided according to reported plasma concentration of their respective biomarker, see Table 2.

Table 2. PLA panels setup. The assays are divided in different plasma dilutions according to the estimated abundance of the biomarker in plasma. Panel A and B are used on undiluted plasma and panel C and D on plasma samples diluted 50 times. The reference assays (GFP and PE) are included in all panels.

Panel A		**Panel B**		**Panel C**		**Panel D**	
Plasma dilution 1:1		*Plasma dilution 1:1*		*Plasma dilution 1:50*		*Plasma dilution 1:50*	
Seq ID	**Antigen**	**Seq ID**	**Antigen**	**Seq ID**	**Antigen**	**Seq ID**	**Antigen**
7	CEACAM5 pAb	7	CEACAM5 pAb	7	TIMP1	7	TIMP1
8	VEGF	8	CEACAM1	8	EGFR	8	CEACAM8
32	**GFP**	**32**	**GFP**	**32**	**GFP**	**32**	**GFP**
33	IL6	33	EGF	33	Her2	33	S100A8
35	IL8	35	CTGF	35	OPN	35	CXCL12 (SDF-1)
38	**PE**	**38**	**PE**	**38**	**PE**	**38**	**PE**
45	CEACAM5 mAb	45	GDNF	45	TGF beta1	45	MIF
51	IL4	51	Fractalkine (CX3CL1)	51	Her3	51	Mesothelin
53	TNFa	53	MIP-1	53	Her4	53	SLPI
57	IL-1a	57	CA 242	57	S100A8	57	YKL-40

3.5. Polymerase Chain Reaction (PCR)

PCR is the most widely used method for amplifying nucleic acids *in vitro*. By using a DNA polymerase with thermal cycling one can generate millions of identical copies of a single DNA strand. Specificity is obtained by using two sequence specific primers for dual recognition[7]. In the thermal cycle there are three steps. First denaturation of the double stranded DNA and activation of the polymerase. Next step is annealing of the primers to the template, and the third step is elongation of the primers along the template. Each double stranded DNA molecule is then used as two templates to create two double-stranded DNA strands, resulting in a duplication of the desired DNA sequence each cycle, see Figure 4.

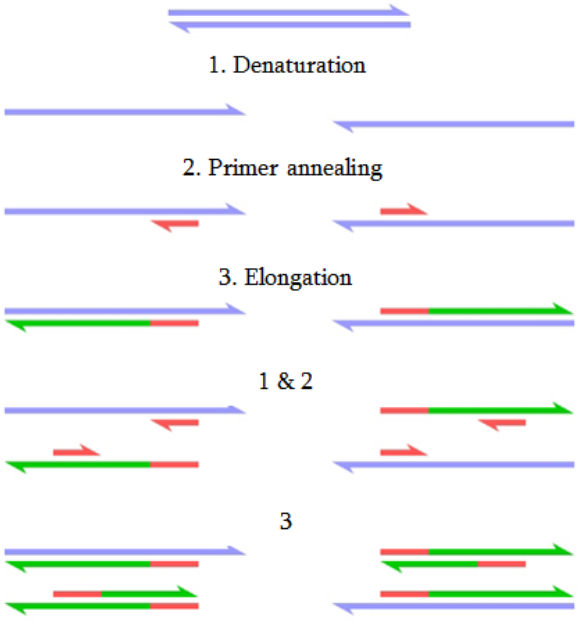

Figure 4. Polymerase chain reaction (PCR). The method amplifies DNA strands by thermal cycling. There are three steps in the thermal cycle; the first is denaturating (1) of the dsDNA original template (blue), the second is the annealing step (2) where specific primers (red) bind the different ssDNA and the third is the elongation part (3) where a DNA polymerase synthesises a new DNA strand (green).

3.6. Real-Time Quantitative PCR (RT-qPCR)

The idea of qPCR is to measure the quantity of DNA molecules in a sample, and by measuring the quantity after each cycle, and not just at the end point, it is called "real time". By doing so, one quantifies the amount of a specific DNA sequence present in a sample by comparing to a standard curve. There are several alternative approaches for quantifying DNA with PCR, in this project we used a Taqman® probe (Applied Biosystems)[8].

The Taqman® probe is an oligonucleotide with an attached fluorophore in one end of the oligo and a quencher molecule on the other end. When exposing the fluorophore to photons of a certain wavelength it absorbs them and emit photons at another wavelength (fluorescence) at which the quencher molecule absorbs the photons. This probe binds a specific DNA sequence and is cleaved off when a new strand is synthesized (requires a polymerase with 5' exonuclease activity). When the probe is cleaved the fluorophore and quencher are separated and the emitted light from the fluorophore is no longer quenched, and can then be measured[8].

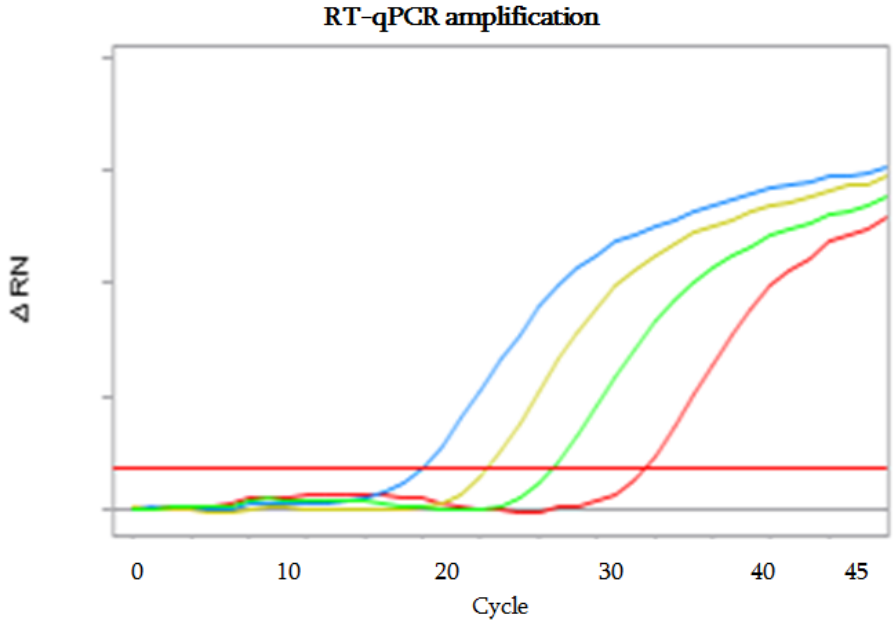

Figure 5. Shown here is a typical qPCR run for four different samples. The sample with most starting material will reach detectable fluorescent intensity after the fewest cycles and will be located to the left of the other samples curves. A threshold (red horizontal line) is set when all curves has reached the exponential phase. The cycle number at which this threshold is reached will be that curve's Ct value.

By measuring the intensity of the emission after each cycle and plot it against the amount of cycles performed, you see a curve that starts with background fluorescence, followed by an exponential phase and a plateau. The DNA is duplicated each cycle, but is present in too low concentrations to be detected before the exponential phase. The plateau is reached when the PCR reaction is saturated. By comparing different concentrations of initial DNA, the curve shifts with the number of cycles required to reach a certain intensity level. To compare different samples an intensity threshold is set where all samples have reached the exponential phase, see Figure 5. The cycle number it took for the sample to reach that intensity is called the Ct value. Since a sample with high concentration reaches the threshold faster, it will get a low Ct value.

3.7. *In solution* PLA®

The PLA® method is based on proximity ligation of two oligos attached to two antibodies (PLA® conjugates) recognizing different epitopes on the same antigen. The antibodies can be polyclonal or monoclonal, as long as the pair recognizes separate epitopes on the antigen. If the two antibodies bind the same antigen their attached oligos is in proximity and can then be ligated using a connector oligo together with a DNA ligase (proximity ligation). The connector contains a sequence partly complementary to both oligos. Once the oligos are ligated, PCR is performed on the ligated sequence to increase the signal before detection. To detect the amount of antigens (or now amount of ligated oligos) qPCR is performed[9]. See Figure 6 for an overview, and the identity of the protocol (section 3.7.1) for details.

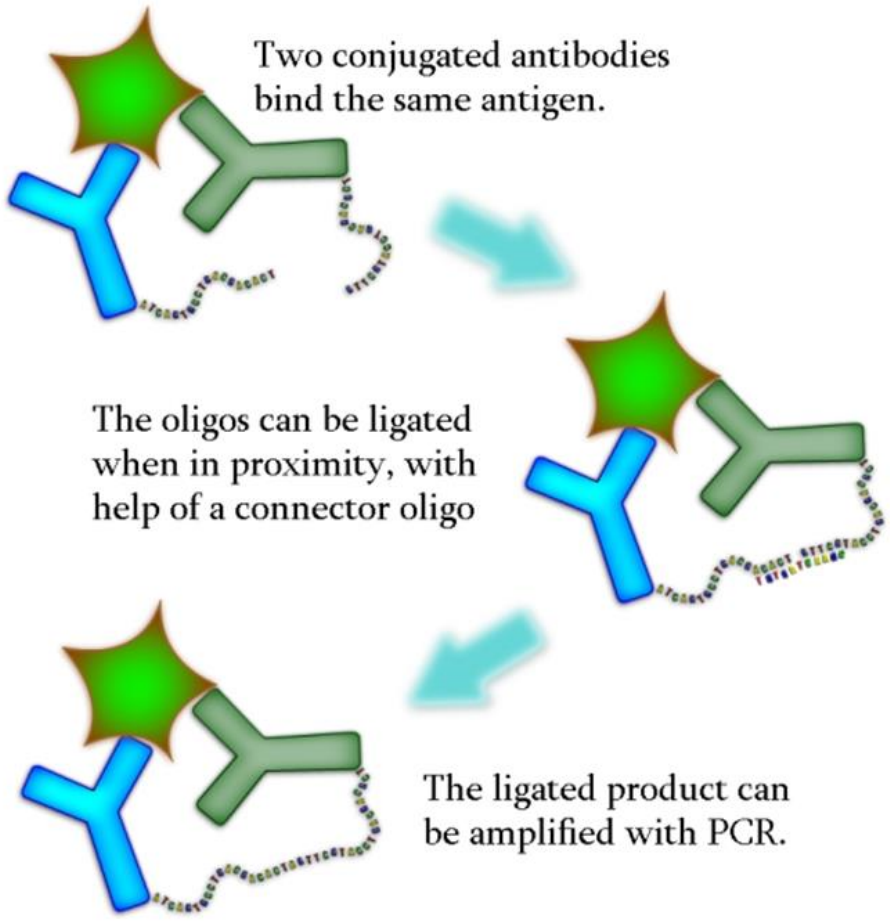

Figure 6. "Protein to DNA conversion" with in solution PLA®. First two conjugated antibodies bind the same antigen (upper left). The two conjugate-oligos are then in proximity (middle) and with the help of a connector oligo they can be ligated (lower left). These ligated products can then be amplified by PCR and quantified by qPCR.

There is always some background ligation which sets the assay background. The background ligation is depending on the concentration of conjugates, this due to the stochastic distribution of PLA® conjugates in the solution which means that some conjugates is always in proximity by random and will then ligate to form background.

The *in solution* PLA® method only works within a certain concentration span of analytes (picomolar range for most of our assays), the range can be adjusted with several parameters, the most important one being conjugate concentration. Although low conjugate concentration decrease the background level, too low concentration makes the low signals very variable. If too much antigen is present the conjugates will have a higher chance to bind an antigen by their own which will not yield any proximity for two conjugates. This is defined as the "hook effect" and is described further in Figure 7. The easiest way to work around the hook effect is to dilute the sample plasma until the antigen concentration is within the dynamic range of the assay. Another way is to artificially lower the antigen concentration by adding free antibodies to the sample, which then will occupy the binding sites on the antigens.

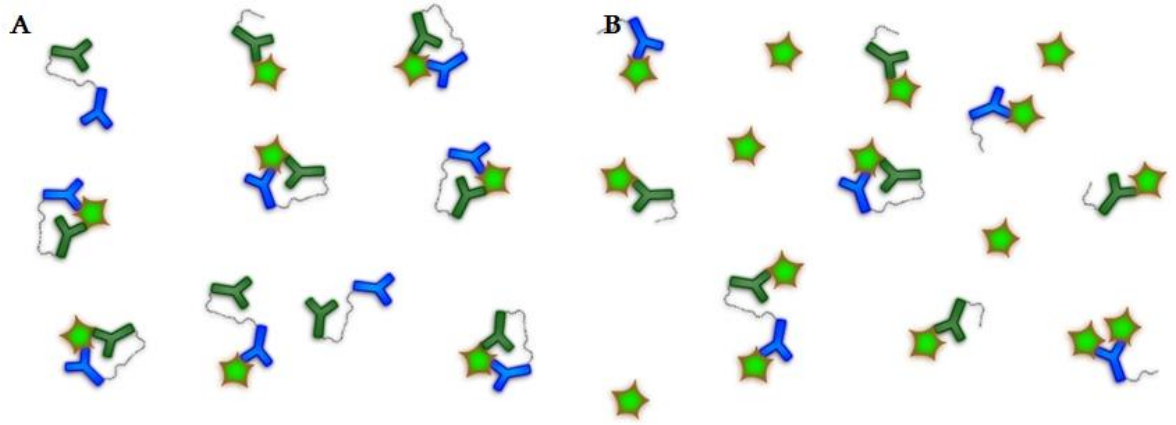

Figure 7. In A) a medium concentration of antigen is incubated with PLA® conjugates. B) shows the "hook effect" which is seen when there is too much antigen available. At high concentrations of antigen, each conjugate will have a higher probability to bind one antigen alone - and thus decreasing the signal with increasing antigen concentration.

The ligation step of PLA® is the most important step, where a bound antigen gives rise to a complete ligation product that can be detected. When detecting antigens in blood plasma it has been established by Fredriksson *et al.* that the T4 DNA ligase is inhibited, and does not ligate at full efficacy[6]. What causes this inhibition is yet not fully known, but it has been seen that this inhibition level varies between individuals and will need to be normalized, though this inhibitory factor can be quite high.

Multiplexing with *in solution* PLA® is achieved by using different antibodies coupled to specific “barcode oligos”. By using these barcode oligos it is possible to use unique PCR primers for each assay, and thus choose which ligated product to amplify and detect. One part of the oligo is the site for PCR primer binding (unique) and another site the site for ligation and detection (identical for all assays). Primers and other sequences is easiest designed *in silico* to fit together. This is what makes the multiplexing capacity of PLA® higher than for other immuno assays, so a thorough design is crucial. If one of the antibodies is cross-reactive this will not be a great problem; PLA® will not amplify and detect this kind of cross-reactivity (due to the need of both antibodies to be cross-reactive) and hence not give a false positive signal. Let us say that we use a polyclonal antibody

(pAb) with 5% cross-reactivity against another protein. We then have 5% cross-reactivity for both conjugates, and they will have a $5^2\%$ = 0.25% chance to bind the same "wrong antigen" which then will yield a false positive ligation product. But even if we have a very high concentration of this crossreactive antigen it will result in hooking before we reach a high false positive signal. What this can result in though, is that the conjugates may be occupied by wrong antigen, making them unavailable for the correct antigen.

If one of the conjugates bind the antigen for another assay in the same panel this may result in proximity between two mismatched conjugates. This product will however not be detected when using primer pairs specific to a matched conjugate pair.

3.7.1. In solution PLA® protocol for blood plasma samples:

1. Some panels requireed a 50 fold dilution of the plasma samples as the first step.
2. Blocking - the plasma was mixed 1:2 with stabilizing agents, blocking agents and reference antigens. Incubation at 25°C for 20 min allowed the blocking agents to be bound before adding the real conjugates.
3. Conjugate incubation - the PLA® conjugates were added in multiplex, also mixing the sample 1:2 once again. Incubation was done at 4°C for 16 h.
4. Ligation - DNA ligase and the connector oligo were added to the sample and the reaction were then diluted 1:25 to reduce background ligation by random proximity. Incubation at 37°C during 10 min for ligation was followed by 65°C inactivation of the enzyme for 10 min.
5. UNG treatment - Uracil-DNA glycosylase were added to degrade the connector (containing uracil). Incubation was done at 25°C for 5 min.
6. Pre-amplification - The samples were amplified in multiplex with PCR for 15 cycles before they were analysed with qPCR. This step was necessary to increase the signal and thereby reduce variance in the qPCR analysis.
7. Dilution - The samples were diluted 1:20 to reduce the amount of transferred primers from the pre-amplification step which could have interfered in the qPCR.
8. qPCR - The samples were analysed with qPCR, using X different primer pairs for each sample, where X is the grade of multiplexing.
9. Data analysis - The Ct values were determined and normalized with the spike-ins, giving a relative plasma level of each biomarker.

(A more detailed protocol can be seen in Appendix 1)

3.7.2. Conjugate quality control - negative control

It was important to ensure that the conjugates did not give any false positive background signals. Any pair of given conjugates always have a chance to be ligated even if no antigen is present – and hence always yield a background level. By testing each assay with only one of the conjugates we were not expecting a signal, and when using both conjugates a normal background level were expected. If one conjugate alone would yield a signal it might have been contaminated with another oligo, which could affect assay performance. Another negative control that we used were to perform the ligation step without ligase, and hence not having any ligation products, this was more of a qPCR detection control –

ensuring that we did not have any primer dimers forming, or wrong sequences amplified by our primers.

3.7.3. Panel evaluation

For the assays for which antigens were available, the sensitivity, linearity of dilution, assay range and recovery were determined. By measuring antigens of different concentrations and in plasma, we estimated these parameters roughly. We could also tell if the plasma concentration were within the dynamic range for the assays in these experiments.

3.7.4. Normalization

In plasma there were some molecules that inhibit the T4 DNA ligase and hence decreased the ligation efficiency. The level of inhibition differed in each plasma sample and needed to be normalized. There were also experimental variation in different steps, for example pipetted volumes and temperatures which affected the end results for that sample. To be able to normalize all variation, two internal standards, or spike-in antigens, were added to all samples.

We used GFP and PE as our standards. These antigens were chosen because they are exogenous for humans, which was the most important property for the references. The standards had to be exogenous so that the added amount of antigen was equal in each sample. When comparing reference signals from different plasma samples we were able to use the variation in these spikes to adjust the signal for other assays. Since everything is done in multiplex, all technical and inhibitory variation were assumed to be equal for all assays – and hence contained in the spike-ins signals. By using two standards it was possible to determine if they correlated well with each other from one plasma sample to another, which was crucial for the standards. A good correlation would be a positive sign that normalization would work to reduce variation.

There are numerous ways to normalize data. The statistical analysis of this project was mainly performed by another research group in the Proactive project (Uppsala Academic Hospital (UAH), Sweden). Therefore, no detailed algorithms used are presented here. However, it was important for us to evaluate if normalization would be helpful. Mainly we looked at the internal standards, but we also compared some assay data to the literature.

Two normalization methods that were tested was: one or two spikes at Ct value level, where we used the same threshold value for one entire qPCR run and then compensated the different assays for one sample based on its GFP and/or PE value. For the final data normalization, linear scale normalization was used. The Ct level method was the easiest and fastest to perform, and also the one we used for all our evaluations. The methods give the same relative values, but at different type of scales.

3.7.5. Cross-reactivity

There are several types of cross-reactivity that can occur in multiplex immuno based assays; the antibodies can have affinity for more than one antigen and there can also be human antibodies targeting animal IgG present in plasma which could yield false signals. Another phenomenon that can appear is mis-binding of PCR primers (a primer pair can bind and amplify the wrong template if the primer design is poorly done). It is not easy to detect antibody cross-reactivity; there are thousands of proteins that are potential binders, and it is difficult and time consuming to test for binding of each and every one of these.

We relied on the antibody producers quality controls regarding antibody specificity and only used affinity purified antibodies, and did not test for cross-reactivity further. We did however test for unspecific binding of our PLA conjugates, and the easiest way was to use two mismatched PLA conjugates in singleplex and look at the differences in signal with and without plasma present. Testing for primer cross-reactivity was done by using mismatched primers on samples derived from a PLA reaction with matched couple of conjugates (singleplex); there should be no signal at all if using the mismatched primers in the qPCR.

3.7.6. Blocking agent

Some individuals have developed antibodies against other species antibodies (for example Human Anti Mouse Antibodies (HAMA)), and if these are present in the plasma they will bind our conjugates to yield a false positive signal.

By adding free antibodies in excess we found that it was possible to reduce this false signal. But we have also seen that free IgG is not sufficient in all cases. There appeared to be some molecules that bound our conjugates in some plasma samples. The extent of this effect differed between individuals and needed to be eliminated to reduce false positive signals. Olink Bioscience recently filed a patent for a blocking agent to be used with *in solution* PLA for this specific reason. By adding random sequence conjugated antibodies generated from the same species as our PLA conjugates, we were able to decrease the false positive signal. These random conjugates were incubated with the plasma before adding our specific assay conjugates. For this project Innova Biosciences conjugated new blocking agents for us, which needed to be evaluated before use. When we investigated the efficiency of the blocking agent it was crucial to look at unspecific binding. The unspecific binding was detected by using mismatched primers in the qPCR detection; these should yield a background level unless an unspecific conjugate binding were present in the plasma. So by looking at the specific and non-specific signals with different concentrations of the blocking agent we assessed the blocking efficiency.

Most of our antibodies used were affinity purified from goat serum, so the most suitable blocking agent for us was a random goat IgG conjugated with random oligos ("goat probe" or GP). Previously we have used a GP conjugated by Olink, but in the PROACTIVE project we used one that is conjugated by Innova (when up scaling it was important to have a reproducible method for conjugating the GP).

3.7.7. Storing buffers evaluation

Some conjugates were going to be used during the entire 3-year project, and hence the conjugate stocks needed to be stored at a stable environment (including storing buffer). Previously an Olink buffer called Fish gelatine Blocking Buffer (FBB) had been used for dilution and storing of the conjugates (the buffer components are confidential). And since the new Innova conjugates were made with a different chemistry we investigated if our in-house buffer maintained the activity of the conjugates or if we would need to switch to another storing buffer.

3.8. Data analysis

3.8.1. Amplicons

A Ct value is defined as the number of PCR cycles it takes to reach a fluorescence threshold; so a high concentration in a sample will reach the fluorescence threshold early, giving a low Ct *value*. When converting Ct values to a linear scale we named the new values "amplicons" - or number of ligation products. And since the Ct values are presented in an inverted log2-scale we needed a maximum value to set as one amplicon (or background), we used 38 or 30 depending on which instrument used, when comparing values it did not matter which Ct to choose as background – the relation between samples will be equal either way. The formula used to estimate the amplicons from Ct values was:

$$\boldsymbol{Amplicons\ (amp) = 2^{(38-Ct)}}$$

3.8.2. Limit of detection and sensitivity

By making standard curves for the assays it was possible to determine the limit of detection (LOD) and sensitivity for a certain assay. To do this it was necessary to have the specific antigen for that assay in a known concentration. By looking at the linear phase of the standard curve (Ct values) and taking the concentration divided by 2^(delta Ct) (or dCt, which is the Ct value compared to the buffer Ct) we estimated a sensitivity for this specific assay by performing a simple extrapolation from the lowest measurable concentration down to background. To determine the background level, two times the standard deviation of the buffer measurements was subtracted from the buffer Ct value to get the lowest practical measurement possible.

$$\boldsymbol{Sensitivity = \frac{C}{2^{dCt}}}$$

$$\boldsymbol{Lower\ LOD = Ct(buffer) - 2 * STD(Ct(buffer))}$$

3.8.3. Recovery by addition

Recovery is a measure of how well the assay works in plasma samples compared to buffer. To calculate the recovery we needed to add (or spike in) a known amount of antigen into buffer and plasma and compare with unspiked buffer and plasma. By comparing the signal to noise ratio for buffer and plasma we were able to calculate to which extent the signal to

noise ratio is reduced in plasma, and of course it was important to take the endogenous level into consideration.

$$\boldsymbol{Recovery} = \frac{\boldsymbol{amp(plasma + antigen) - amp(plasma)}}{\boldsymbol{amp(buffer + antigen) - amp(buffer)}}$$

3.8.4. Recovery by dilution

Also known as linearity of dilution. This is another important assay property which tells us about how well our assay performs in the media used (plasma in our case). By diluting the plasma to a certain concentration we could either see an increase or decrease in signal. Increase in signal was usually seen when the assay had hooked, but could also occur when there were no antigen present and we were diluting plasma inhibition. A third possibility could be diluting agents interfering in the antigen-binding. In contrast, a decrease in signal indicates that we were diluting the antigen measured, where an optimal scenario were a loss of one Ct value for each two fold dilution.
In general the linearity of dilution was much better at low plasma concentration for PLA (low plasma inhibition) but it was also much dependent on the assay quality.

3.8.5. Normalization

The idea with the internal standards we added was to normalize all the other assays with these values. The normalization method was obviously different depending on what scale that was used; Ct (log scale) or amplicons (linear scale). When working with log scales we normalized by subtracting the internal standard value to compensate for all variation (interference) in the assay. When using linear scales we divided with the internal standard instead to achieve the same normalization effect.

$$\boldsymbol{Ct_{Normalized}^{Assay\ 1,Sample\ 1}} = \boldsymbol{Ct_{Raw}^{Assay\ 1,Sample\ 1}} - \boldsymbol{Ct_{Raw}^{Standard,Sample\ 1}}$$

$$\boldsymbol{Signal_{Normalized}^{Assay\ 1,Sample\ 1}} = \boldsymbol{Amp_{Raw}^{Assay\ 1,Sample\ 1}} / \boldsymbol{Amp_{Raw}^{Standard,Sample\ 1}}$$

3.9. Comparison of different qPCR instruments

For the optimization and evaluation of all assays we either used 96 well or 384 well plates on an Applied Biosystems 7900HT Fast Real-Time PCR System. But as the project required higher throughput with time, it was decided to advance to a high capacity microfluidic qPCR system (Biomark™ from Fluidigm®) which has a capacity of analysing 48 samples with 48 assays on one chip. However, it was important to know whether or not the PLA® method adapted well to this instrument or if further optimization of the protocol was needed. Replicate samples run on the 7900HT system were also run on the Biomark™ system and a standard deviation and Coefficient of Variation (CV = stdev(replicates)/average(replicates)) was calculated for each sample and assay.

4. Results

4.1. *In solution* PLA®

Below, all results from the different PLA evaluation experiments are presented.

4.1.1. Conjugates from Innova compared with conjugates from Olink

Two new conjugate pairs, with different oligo/antibody ratios, produced with Innovas new conjugation chemistry were compared with conjugates made from standard Olink in-house protocol. This to see if the signal and signal-to-noise ratio were sufficient to use or if more optimizations were needed from Innova. The assay used for evaluation was detection of ICAM (also known as CD-54). Three conjugate concentrations were tested; 50 pM (standard), 150 pM and 500 pM at the incubation step. To determine the dynamic range of the assays, a wide concentration span of ICAM was used, ranging from 0.1 pM to 100 nM. The results are presented in Figure 8.

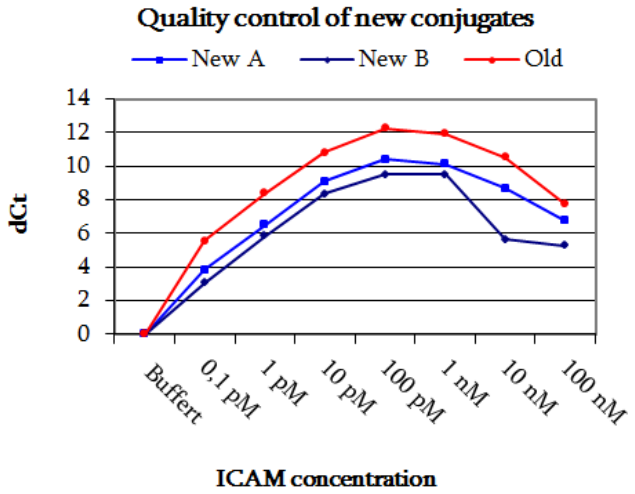

Figure 8. New Innova conjugates for an ICAM assay were compared with in-house conjugates. 150 pM conjugate concentration were used for this experiment to determine the dynamic range, sensitivity and signal strength. Antigen concentrations range from 0.1 pM to 100 nM. The in-house conjugates had a higher signal-to-noise ratio but a lower signal (data not shown).

4.1.2. Preliminary PLA analysis of test plasma samples

As a test set of plasma samples we had received 20 control and 20 CRC plasma samples from CU. These samples where used for evaluation of the different assays when developing and optimizing the protocol. When the first shipment of conjugates arrived from Innova they were evaluated on these 20+20 samples with the initial protocol (see Appendix 1 for details). Results are presented in two columns, grouped by case and controls, see Figure 9. Note that there are some assays missing due to a delayed production of these conjugates. The conjugates for the two references (GFP and PE) were yet to be received, so normalization was not possible in this initial experiment on plasma samples.

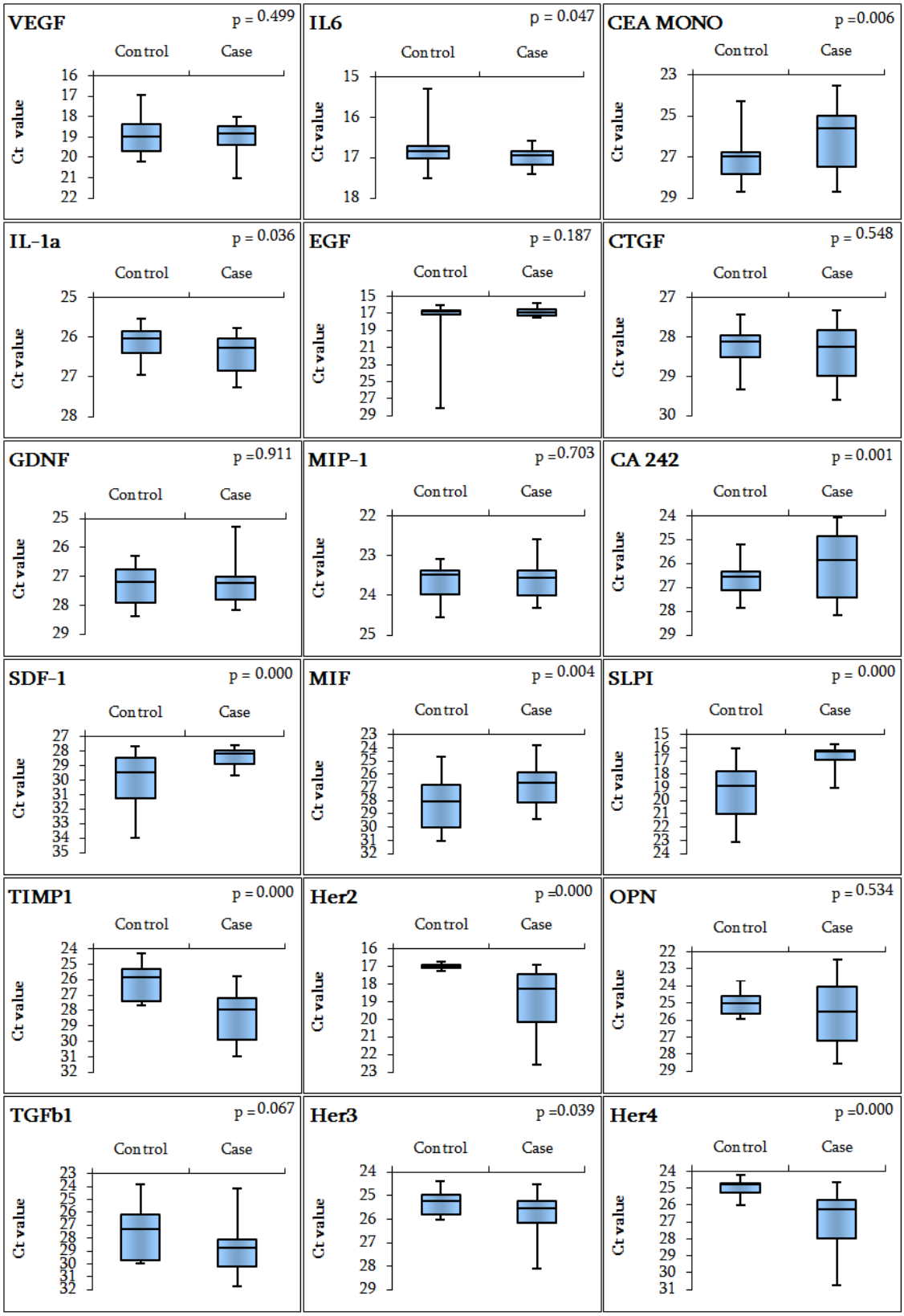

Figure 9. Preliminary data from 20 control and 20 case samples. Data for all assays available are shown in two columns (controls vs. cases). Boxes represent the interquartile range and whiskers the max and minimum value. The bold line in the box indicates the median. Values in the top right corner indicate the p-value of the students T-test. A cut-off value of $p<0.05$ was chosen as significant difference between the groups.

4.1.3. Panel quality control - negative control

All four full panels were tested at the same time at a conjugate concentration of 300 pM. Ct values above 32 that have been pre-amplified are quite high for this technology. The number of amplicons transferred from the ligation was then very few. If even fewer amplicons were transferred it was very random how many that happens to be transferred. So a Ct value above ~ 32 were quite unreliable and was considered negative, and as we can see in Figure 10 all the negative control samples were either not detected at all, or detected with a very high Ct value. Ct values at 32 compared to 22 will be 2^10 (=1024) times lower if comparing amplicons, which would not affect the real background value.

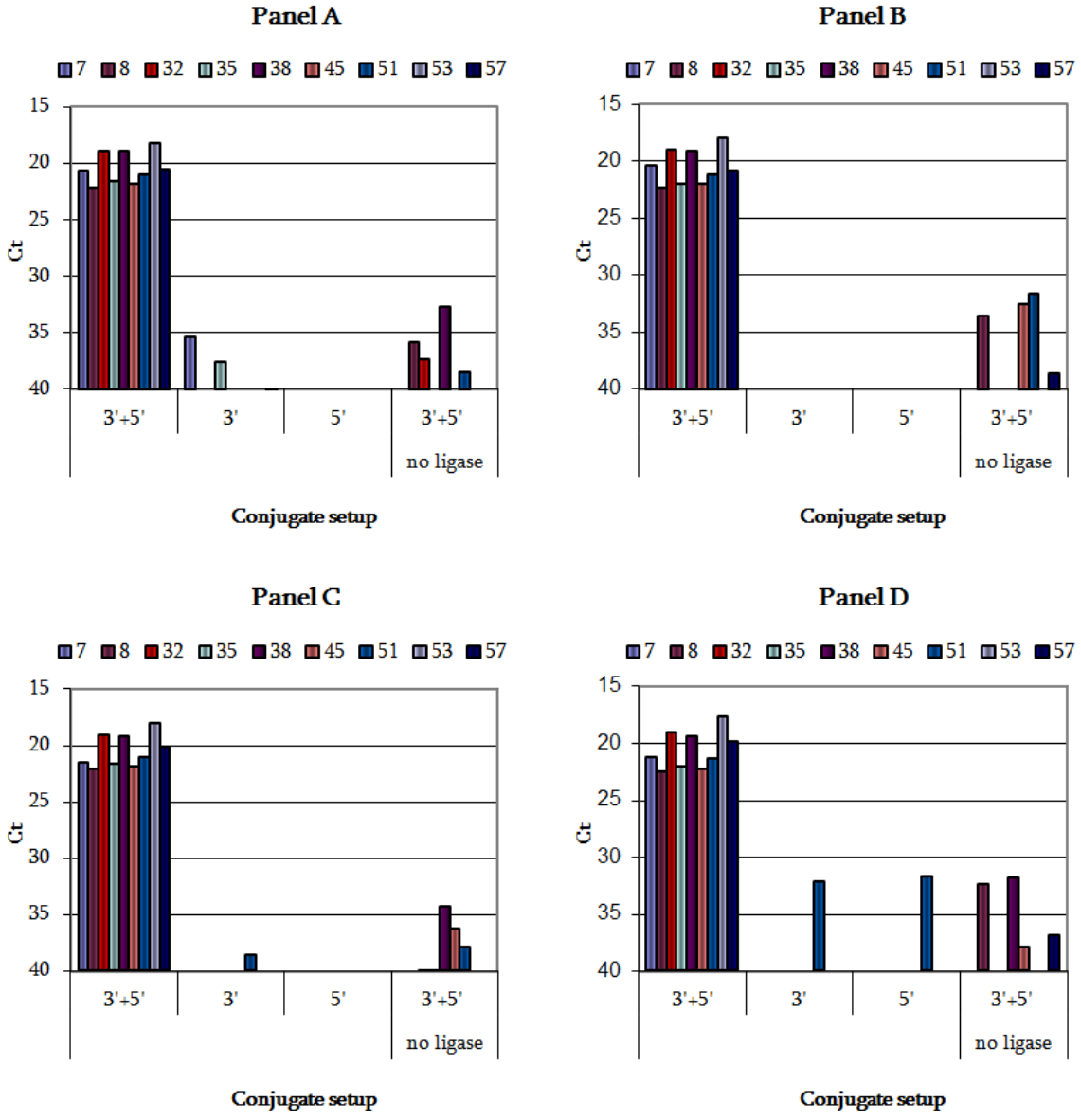

Figure 10. Quality control for the background level of each assay, three negative and one positive control for each assay were included. All the negative controls have either a very high Ct value or were not detected in the qPCR. The coloured bars represent the different PLA conjugates (see Table 2).

4.1.4. Panel evaluation

These tests were performed at 50 pM conjugate concentration in multiplex (except GFP and PE which were tested in singleplex). Four panels with nine assays each were used, where two assays are present in two panels (test for inter panel correlation). Antigens were available for 23 assays and the results are presented in Figure 11.

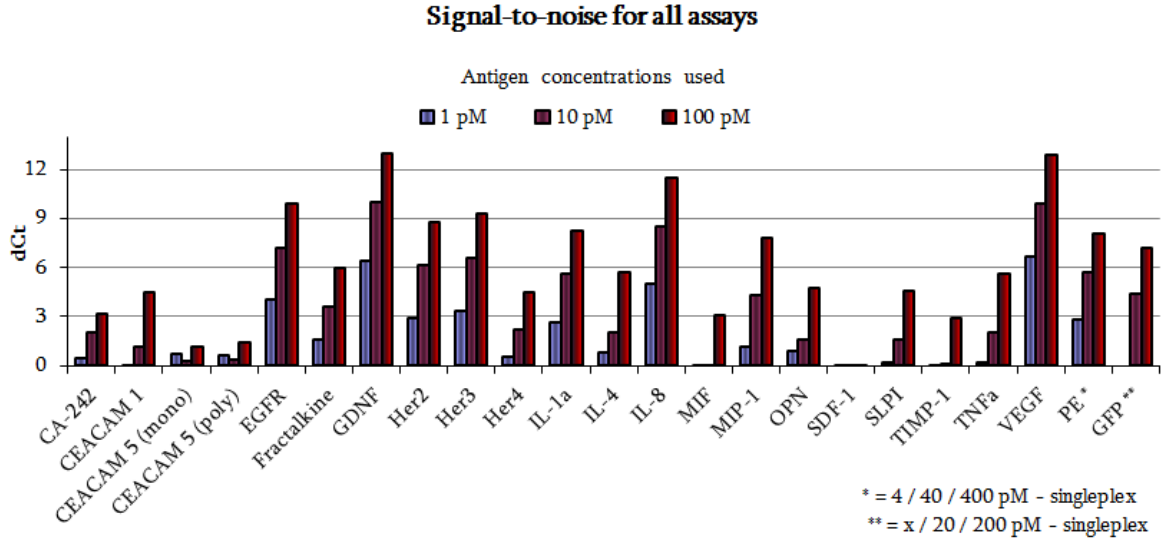

Figure 11. Signal-to-noise for the 23 assays where antigens were available. Three antigen concentrations were spiked and measured in multiplex (except for GFP and PE). The higher dCt value the better sensitivity for the assay.

The sensitivity varied significantly between assays. The assay with the highest sensitivity was VEGF with a sensitivity of ~10 fM at 50 pM conjugate concentration. Several assays had relatively poor sensitivity (>1 pM), and by using 300 pM conjugate concentration we got worse sensitivity (~70 fM for VEGF). Over all, the good assays got worse sensitivity with higher conjugate concentration due to higher background, but some poor assays appeared to improve – this was almost certainly due to the more stable background levels at higher conjugate concentrations, not due to higher sensitivity.

Recovery also varies greatly between assays, and as expected the recovery was much better at low plasma concentrations (compared to undiluted) where there was less inhibition of the ligation (see Table 3). Recovery was determined for some assays where a level above background could be determined. 10 pM antigen were spiked into plasma in multiplex and measured with 50 pM conjugate concentration. Recovery was calculated before and after normalization with GFP (see 3.8.5) – and the normalization showed very much improvement in recovery, see Table 3.

Table 3. Recovery by addition before and after normalization. The lower plasma concentrations have better recovery in general. The recovery was determined by adding 10 pM Ag to plasma and calculate the relative signal compared to antigen in buffer. When normalising data the risk is that the spike-in signal varies more than the measured analyte signal, then the normalization will change the analyte more than it should - overcompensating. IL-4 (italic) has probably been overcompensated in the normalization and hence gives such high recovery.

Recovery by addition in plasma for 14 assays

	Undiluted plasma			1:50 Diluted plasma	
Assay	**Raw data**	**Normalized**	**Assay**	**Raw data**	**Normalized**
VEGF	25%	98%	EGFR	77%	98%
IL8	27%	109%	OPN	66%	109%
IL4	*71%*	*288%*	Her2	61%	84%
TNFa	21%	77%	Her3	88%	122%
IL-1a	23%	94%	Her4	65%	87%
GDNF	14%	70%			
Fractalkine (CX3CL1)	23%	116%			
MIP-1	9%	49%			
CA 242	5%	18%			
Average	**18%**	**79%**	**Average**	**71%**	**100%**

4.1.5. Cross-reactivity tests

When we looked at the data from the panel evaluations it was observed that some mismatched primers (not all combinations were tested since that would be 20x20 different combinations, but at least one primer from each sequence were cross-tested with another primer) gave a higher signal with increasing antigen concentration. Overall it seemed to be specifically two mismatched primer combinations that yields noticeable signal in several panels (8 forward with 32 reverse and 7 forward with 45 reverse). When the same mismatched primers yield signal in more than one panel the most likely reason for this was mis-binding of one of the primers. Figure 12 a) shows an example of VEGF/GFP (8 forward with 32 reverse) cross-reactivity. To determine whether this ”false positive” (no reported VEGF/GFP interaction found in literature) signal came from antibody unspecificity or primer mis-binding we tested these possibly problematic conjugates in singleplex. Matched and mismatched conjugates were tested along with matched and mismatched primers to find the source of the cross-reactivity. The “VEGF + GFP (8+32)” data of this experiment is displayed in Figure 12 c). We did notice a signal when using matched VEGF conjugates with both matched and mismatched primers if VEGF antigen were present, see Figure 12 b). However, as mismatched conjugates were used in the PLA reaction, no increase in signal with antigen present was observed. This suggested that the false signal arose from unspecific primer binding.

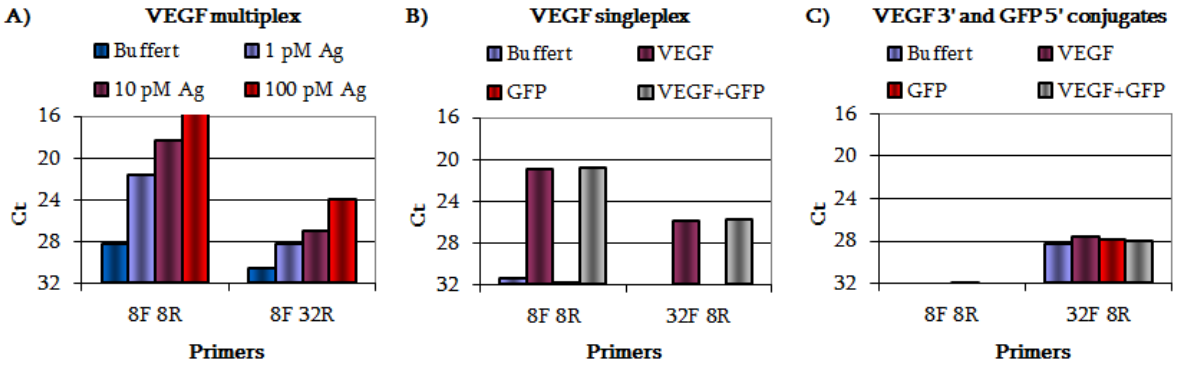

Figure 12. A) Higher signal with increased antigen concentration was observed with some of the mismatched primers; VEGF (8) and GFP (32). To determine the source of the cross-reactivity seen in (A), mismatched primers were tested with matched VEGF conjugates (B) on samples spiked with VEGF, GFP and VEGF+GFP. C) Mismatched conjugates with the corresponding primers were also tested with the same antigen spikes as in (B).

4.1.6. Blocking agent evaluation

Different concentrations of the new GP were tested along with the old GP to see the blocking efficiency; this could be estimated by using mismatched primers for the detection – which should be at a background level. And if this mismatched detection was high, there was probably unspecific binding in the system. The mean Ct values obtained for a number of matched and mismatched primers can be seen in Figure 13.

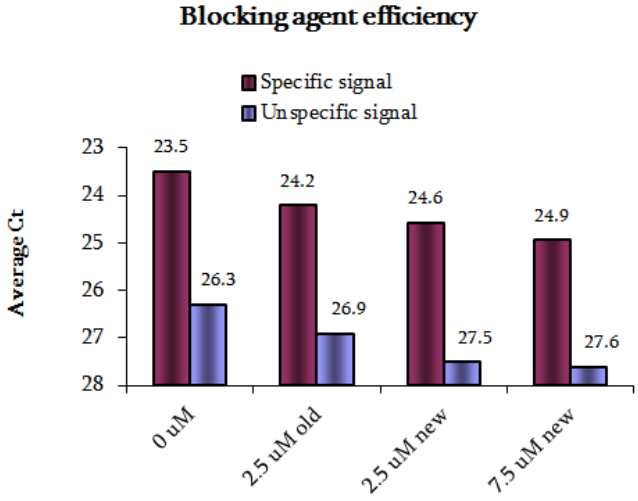

Figure 13. Goat probe efficiency, mean Ct values for the old GP and two concentrations of the new batch from Innova. Two series are shown, mean Ct values for six matched primer couples and nine mismatched couples.

4.1.7. Storing buffers evaluation

Three kinds of conjugates (VEGF, GFP and CA-242) were diluted in four buffers; StabilZyme SELECT® and StabilZyme® NOBLE from SurModics® and two in house buffers, FBB and Duolink® storing buffer. A PLA® reaction was performed at the time of dilution to use as a reference point, after which the conjugates were stored at three different temperatures (4°C, 37°C and 45°C) during 20 days before the next PLA® reaction. In Figure 14 the VEGF and GFP data are shown for the different conditions.

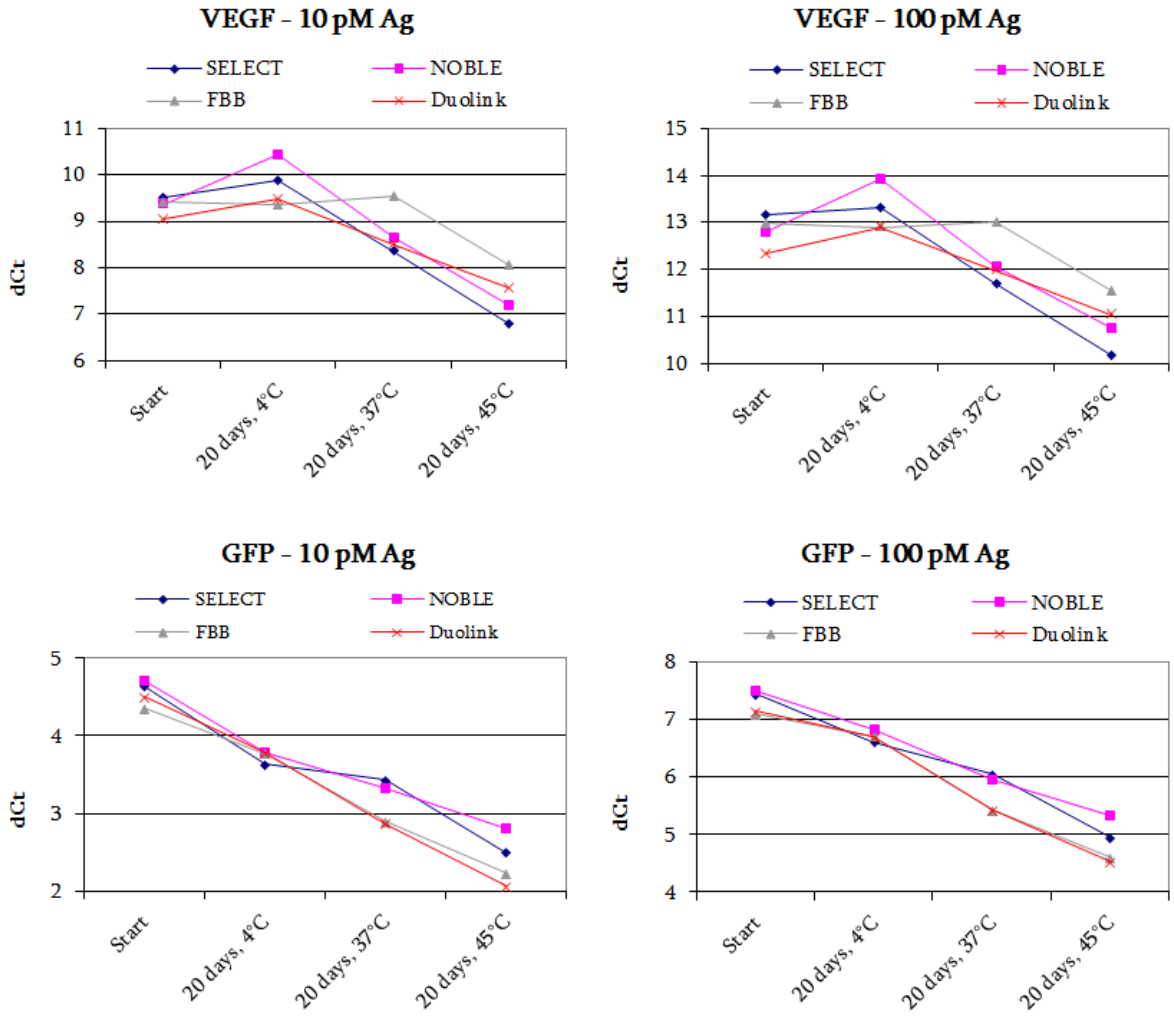

Figure 14. Test of different storing buffers for the conjugates. The conjugates were stored in four different buffers; StabilZyme SELECT®, StabilZyme® NOBLE, FBB and Duolink® storing buffer for 20 days at three different temperatures.

4.2. qPCR instrument comparison

Only a small scale comparison of the two qPCR instruments was performed. The sample set was a panel evaluation with different concentrations of plasma and antigens which made the signal range very wide (>15 Ct values). A scatter plot of all matched Ct values from a 384 well plate and one 48x48 chip can be seen in Figure 15 a. The signals correlated well overall between the instruments, but had quite high variation when reaching high Ct values. This pattern was also observed when plotting the standard deviation (Std) from the replicates in the 48x48 chip (Figure 15 b). The reason for the high variation on the low signals for the micro fluidic instrument was that there were very small reaction volumes, and that there were not enough templates present.

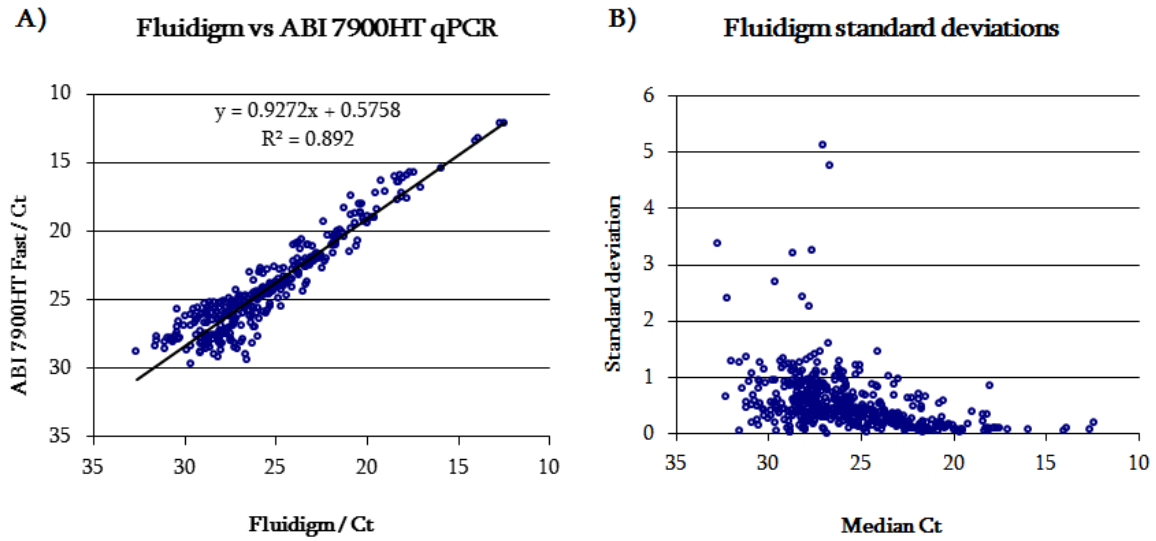

Figure 15. qPCR instrument comparison. All Ct values compared from the 384 wells qPCR plate system with the 48x48 microfluidic qPCR chip are shown in a). In b) the median Ct for replicates are plotted against the standard deviation for the replicates analysed in the 48x48 chip, demonstrating more variation at higher Ct values.

4.3. Normalization

4.3.1. Evaluation of normalization

We performed a normalization test by adding three antigens to a small set of plasma samples; two exogenous and one endogenous marker. Two of the "spike-ins" were the same as for the final runs; GFP (5 pM), PE (50 pM) and the third analyte measured was IL-4 (5 pM) which was very low abundant in plasma samples (crucial when we did not want the endogenous levels to affect the spiked level). For this normalization a mean Ct value for GFP was calculated, then the GFP, PE and IL-4 Ct values for each sample were shifted (addition or subtraction) so that the GFP value were the same for all samples (subtracting Ct values corresponds to division of linear values, see section 3.8.5). The normalization appeared to be working quite well between the two exogenous spikes, in other words they correlated well with each other, with the exception for one outlier. This did not apply for the IL-4, which still varied between samples after normalization, see Figure 16.

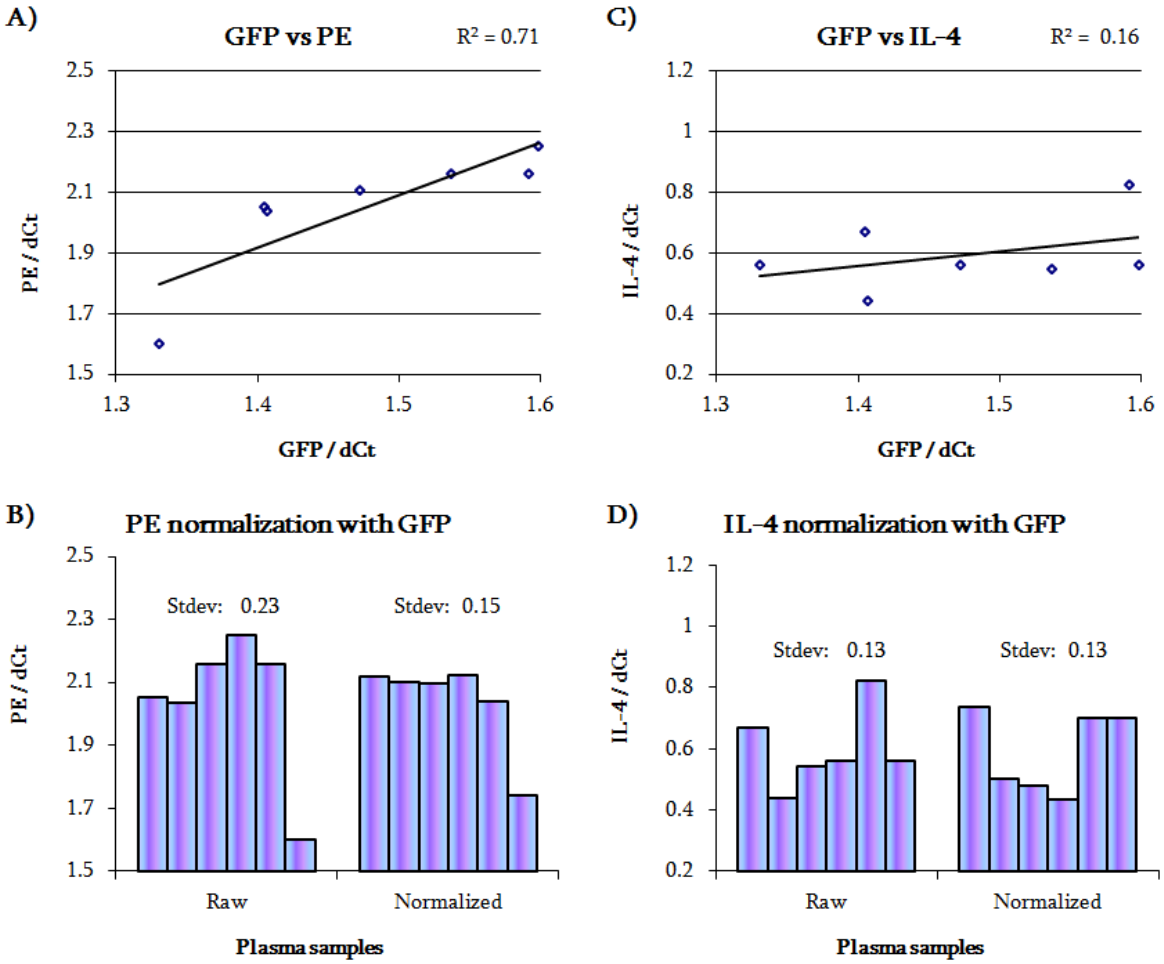

Figure 16. A) GFP and PE correlated well before normalization (R^2=0,71 including one outlier). B) The normalisation of PE with GFP worked fine but was not sufficient for the outlier sample (standard deviation decreased from 0.23 to 0.15). C) GFP and IL-4 did not correlate and it can also be seen that the normalization of this assay is poor (D); no decrease in standard deviation.

4.3.2. Normalization of the entire data set

There was some unexpected interference within our data that seemed to appear in all panels and to be independent of which sample it was. GFP and PE correlated quite well for all panels with the exception for some outliers (data not shown). Due to this interference in some of the signals, the normalization of PE with GFP did not work as wanted; there seemed to be some assay interference depending on the PLA plate position (96 well plates with samples). When we compared normalized PE values from different plates there seemed to be a pattern which also appeared in many of our other assays after normalization (see Figure 17 for two examples). The samples at the "edges" of the PLA plate seemed to differ in Ct values compared to the samples located in the middle of the plate. Both spike-ins have been used to try and normalize the data, but both seemed to fail. This interference probably arose during the preamplification, where small temperature differences has influenced different PCR reactions (read assays) to various extent, and made it next to impossible to normalize with our standards.

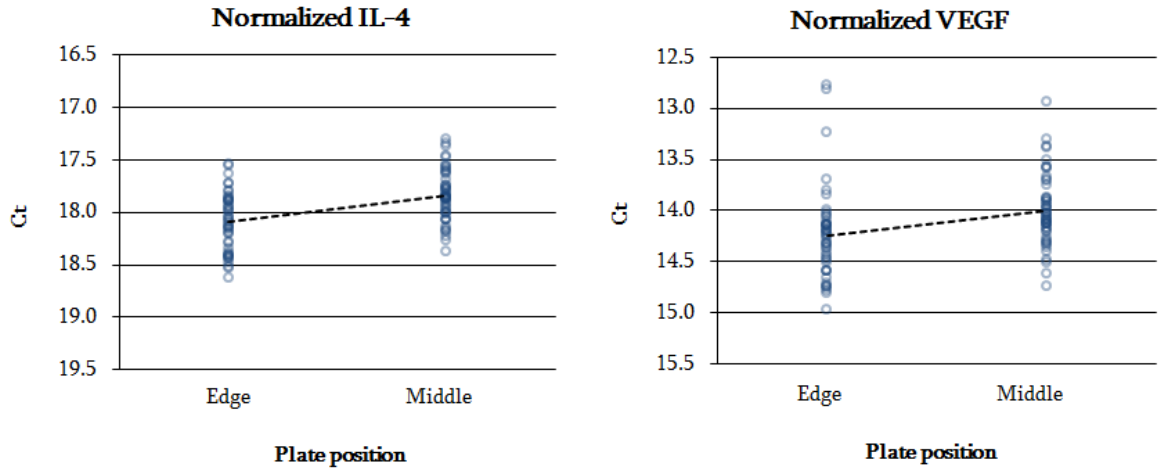

Figure 17. The "plate effect" for normalized values of VEGF and IL-4. Signals varied significantly along the samples position on the PLA reaction plates. This pattern was observed for all of the assays. The columns indicate the "edge" and the "middle" parts of the PLA reaction plate. These both regions contain the same amount of case and control samples. Dashed lines represent the average value of both groups.

To try and read out something from the data several normalization methods were used – without real success. Finally we assumed equal distribution of the two case groups and two control groups from different positions on the plates, and shifted the entire population so the edge and middle had the same median values (performed on all assays). So the first step we did was to normalise with GFP and then shift the median. This normalization method could of course not be used to diagnose a patient, but it gave a better hint about how the levels of different markers were distributed between the populations.

4.4. Complete run results

After many optimizations of the normalization method we were quite satisfied with it. At first the outliers were removed from the data set; the ratio of the two spike-in controls (PE / GFP on linear scale) were assumed to have a normal distribution and values within 95% confidence interval (CI) were used for the analysis. All Ct values were then converted to amplicons (Ct(30)=1, see methods) and divided by corresponding GFP amplicons to normalize. The normalized PE-signal appeared to lower at the edges of all the plates, indicating uneven pre-amplification in the PCR block. To compensate for this, the median of all samples at the edges were shifted to match the median of the rest (assuming equal assay interference for the shifted samples). The results presented should be treated with caution however, since the normalization method used may introduce false variation of samples depending on the location on the plate. Data for all assays except for GFP (used as single reference and is equal in all samples) and PE (no endogenous marker) are shown in Figure 18.

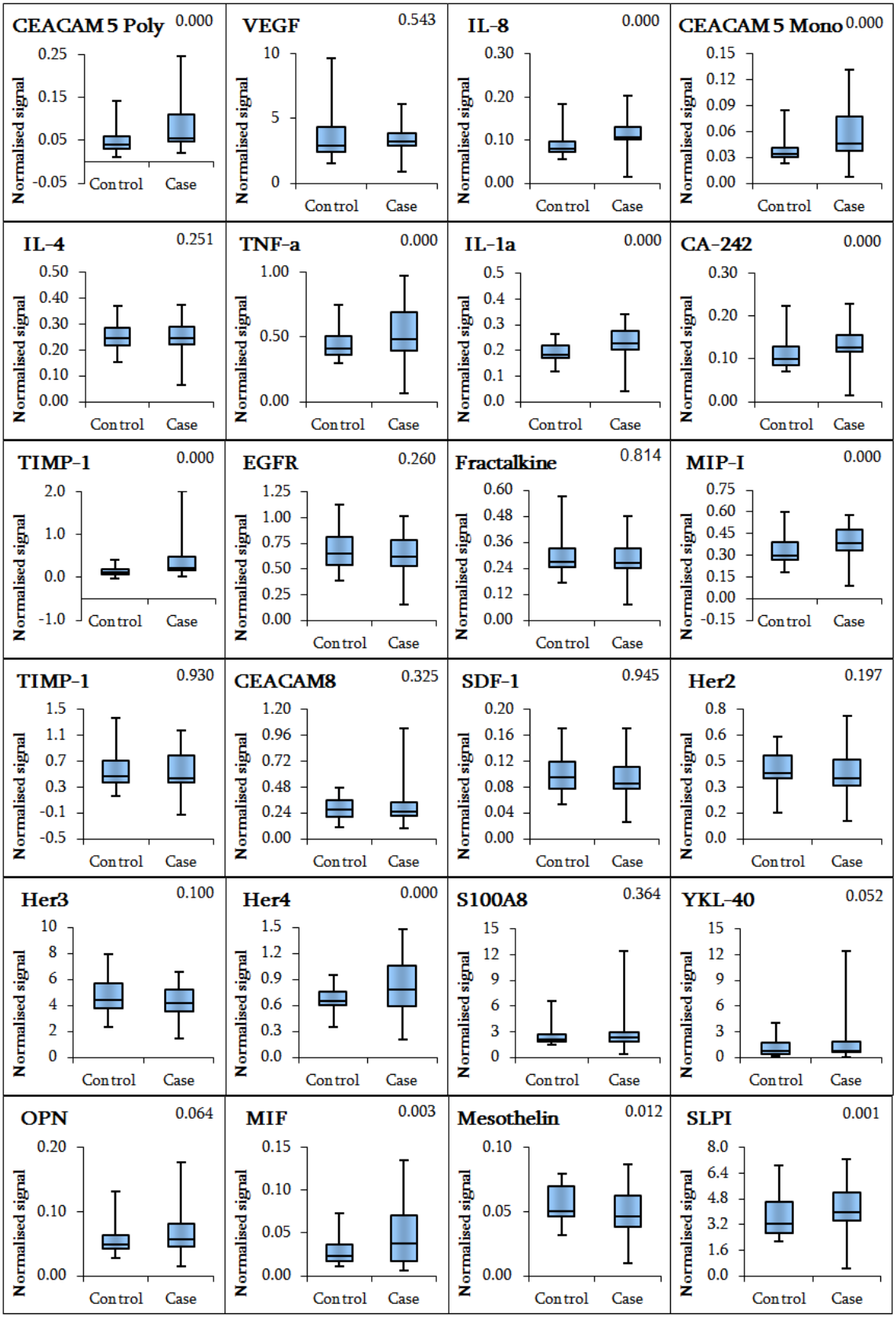

Figure 18. Results from 48 controls and 46 cases analyzed with four panels. Boxes indicate the interquartile range and whiskers the max and minimum value. The horizontal line inside the boxes indicates the median value. Values in the top right corner display the p-value of the students T-test. The Y-axis show assay/GFP ratio values and does not have a unit.

5. Discussion

5.1. Test of new conjugates

The new conjugates received from Innova proved to have a higher signal than the old ones (data not shown), however they did seem to have a slightly lower signal/noise ratio, see Figure 8. Although we lost some signal to noise it was good to have a strong signal in the assays, so we decided to use the new conjugates for all our large scale tests from then on. The new conjugation technique was still under development and optimization, so if the performance would change it would be for the better.

5.2. Preliminary data results

The quick overview of the distribution between control and case patients for the assays received were really interesting. Note that there was no spike-ins within these samples so no normalized data was available. But if the plasma inhibition factors are independent of the diagnosis (assumed) then the overall distribution should be statistically correct between the two groups. The PLA® experiments were performed in parallel and the qPCR analysis one plate for each panel (384 well plates). When only having 40 plasma samples the 384 well system was suitable for this kind of experiment - every sample and assay for one panel fits in one plate so that no comparisons between two qPCR runs were needed. The results shown in Figure 9 look really promising; several assays indicated big differences between the case and control groups. The results were translated into up/down-regulated or no difference between groups by looking at the p-value (threshold = 0.05), data can be seen in Table 5. One thing that was bothering with these results was that the sequence 33 always showed very low Ct values (including buffer) compared to the overall data (see IL6, EGF and Her2 in Figure 9). It also appeared that some of the other assays followed this sequence's signal. To confirm that there was a problem with sequence 33, a negative control experiment in multiplex was performed. As suspected the negative controls (one conjugate only) yielded signals, which indicated a problematic sequence. The decision to remove sequence 33 and to reconstruct the panels was made to get more reliable results in the end.

Innova also told us that they had some problems conjugating some of the antibodies (TIMP-1 and TGF-β) which could explain the low signals in these assays. To be able to use the promising TIMP-1 marker, another antibody couple were conjugated for this assay.

5.3. Quality control - negative controls

The purity of the conjugates was of great importance - a contaminated or erroneously conjugated PLA conjugate could give a false positive signal or no signal at all. The expected results for our negative controls was that the 3' + 5' mixed conjugates should have a background level that is specific to that of the oligo sequence and of course dependent on the conjugate concentration. The 3' solo, 5' solo and 3' + 5' without ligase should not yield a real signal at all. Judging by Figure 10, everything seemed to be in order here. The background level varied a little bit with the oligo sequence (or PCR primer) and most of the negative controls did not yield a signal. However we could see some really low

signals coming up in some of the negative controls, this sometimes happens if some primers form a "primer-dimer" and yields a low false positive signal.

5.4. Assay evaluations

All assay evaluations were performed in multiplex with a mix of all antigens in different concentrations. The assays not shown in Figure 11 had no antigen available. The chart clearly shows that there was a huge difference between all the assays. Some of them did not work at all within this concentration span, while some yielded very high signal/noise ratios. We ranked the assays as poor (dCt < 3), average (3 <= dCt < 8) and good (dCt >= 8) for 100 pM antigen detection (see Table 4). There were several reasons for why the poor assays (CEACAM5 mono and poly, SDF-1 and TIMP-1) behaved as they did, some explanations are: the antigen were degraded, wrong concentration span for the assay (too low abundant or hook effect), the conjugation did not work or that the conjugation damaged the active site of the antibody.

Table 4. The evaluated assays were ranked as poor (dCt < 3), average (3 <= dCt < 8) and good (dCt >= 8) for 100 pM antigen detection. For the assays where no antigen were available there is no data available (N/A).

Good	Average	Poor	N/A
EGFR	CA-242	CEACAM 5 (mono)	CEACAM 8
GDNF	CEACAM 1	CEACAM 5 (poly)	CTGF
Her2	Fractalkine	SDF-1	Mesothelin
Her3	Her4	TIMP-1	S100A8
IL-1a	IL-4		YKL-40
IL-8	MIF		
VEGF	MIP-1		
PE	OPN		
	SLPI		
	TNFa		
	GFP		

5.5. Unspecificity and primer mis-binding evaluation

When we tried to find the source of the false positive signal seen in the panel evaluation different conjugate mixes and primer mixes in singleplex were tested. The data received clearly suggested primer mis-binding. The mismatched conjugates did not yield a difference in signal when we added corresponding antigens (both antibodies did not bind the same antigen), but if mismatched primers were used on a matched couple of conjugates they gave a signal difference when adding antigen (one of the primers binding where it should not). And since PCR requires double recognition of primers we were not that worried about **two** primers mis-binding to give us a false signal, so in our qPCR step when we used a single primer pair in each well we would detect only the correctly ligated product.

5.6. Blocking agent evaluation

To detect some unspecific binding of our conjugates we used both matched and mismatched primers. When adding the blocking agent the unspecific antibody bindings, detected by using mismatched primers, should decrease to background level and the specific bindings should decrease to a certain level (true signal). The unspecific binding mainly consists of factors binding the antibodies (for example human anti goat antibodies), and by adding random IgG molecules from goat (most of our conjugates are based on goat IgG) we eliminated a great deal of these antibodies interacting with our conjugates. There appeared to be no great difference between the 2.5 and 7.5 pM Innova blocking agent, so it was decided to use 2.5 pM.

5.7. Test of storing buffers

A temperature stress test was performed for three weeks at three different temperatures. It appeared to be no buffer that really stood out (neither good nor bad) in general. The 37°C condition was probably more reliable than 45°C due to some evaporation in the tubes (and not equal volumes evaporated from all tubes). The commercial buffers appeared to be slightly better when storing the GFP conjugates, but not with VEGF. We decided to continue using FBB as standard buffer for the Innova conjugates as well.

5.8. qPCR instrument comparison

In a normal 10µL qPCR reaction Ct values above 35 are unreliable due to small amounts of templates. In the Biomark™ instrument the reaction volume is 1000 times lower, so when reaching Ct values above 25 we had the same effect for this instrument. To reach a maximum Ct value of 25 we looked at the lowest signal to know how much the signal needs to increase. From the panel evaluations performed with the preliminary protocol the highest Ct value were ~30, and to compensate for differences in Ct values between the qPCR instruments (~2 Ct) we needed to gain around 7 Ct values to be on the safe side.

To achieve this boost in Ct values several things needed to be changed in the PLA® protocol. All the different changes were evaluated separately and concluded to not disturb the assays too much. The first thing we changed were the conjugate concentration, going from 50 pM to 300 pM, giving a theoretical increase of 2.6 Ct values (resulted in better precision but poorer sensitivity). The other changes were in the pre-amplification step; first we added twice the sample amount (1 Ct), then we added two cycles to the PCR protocol, from 15 to 17 (where we started to lose stability), and the last thing we did to increase the Ct values were to dilute the product less, going from 20 times to five times (2 Ct). The last step is normally not such a good idea since remaining primers from the pre-amplification will disturb the qPCR detection, but we also decreased the multiplex primer-concentration five times during the pre-amplification. For further details on the protocol changes see Appendix 1.

5.9. Normalization

5.9.1. Evaluation of normalization

The early normalisation tests with GFP, PE and IL-4 worked well from our point of view. GFP seemed to normalize PE really well (except for some outliers). Samples with high PE signal also had high GFP signal, in other words their signals correlated with each other. When assessing the efficiency of normalization we looked at inter sample variation of PE before and after normalization with GFP, and although variation varieed between experiments the normalization improved the variation in the vast majority of experiments. When observing spiked IL-4 in plasma samples it did not correlate well with GFP (or PE). This could be explained by the endogenous expression of IL-4, which we had hoped to be negligible in comparison. In other words – the poor IL-4 normalization did not concern us because the endogenous levels of IL-4 were too close to the spiked level.

5.9.2. Complete run normalization

To normalize the data from the full run proved to be much harder however. It appeared that we had some robustness problems during the PLA® experiment and that different assays were affected to different degrees. We believe that this occured during the pre-amplification when we used all the wells in the PCR machine – and that there was some temperature differences between wells. After a confirmation of this we noticed that our pre-amplification was sensitive to temperature differences. And since every assay had different primers with different melting temperatures they were biased at different degrees. The GFP normalization did not normalize for this kind of variation between assays, and might be insufficient. In total we had 12x8 samples in one plate, and it seemed that the first and last three strips (edges) were affected similarly and that the six middle strips are quite similar. So we decided to shift the median for the three first strips (controls) to the median of the first three strips in the middle (also controls). In the same way we shifted the last three strips (cases) to a group of similar distribution of cancer stages within the last three middle strips.

Since we took into consideration the known diagnosis of the patients this is no method that can be used for diagnostic purpose. When trying to estimate a diagnosis all the data need to be treated exactly the same, and that was not the case here. Here we just screened for new markers that were up or down-regulated in CRC patients that perhaps later on will be used in a diagnostic tool. One could argue that normalization would not be needed at all, but in our experience we reduce the variation and case – control group overlap if we normalized the data.

5.10. Full screening run

The reported literature data, the preliminary PLA® results and the full run results are shown in Table 5. The arrows indicate up or down-regulation in CRC cases tested with a student's T-test and the threshold set at $p<0.05$.

Table 5. Reported data, PLA® preliminary data and the PLA® full run results. Arrows (↑ and ↓) indicates up- and down-regulation respectively in CRC blood plasma. (-) indicates no significant difference between populations. Significance was determined with the Students T-test with the threshold value set to p=0.05.

Marker	PLA preliminary data	PLA final data	Literature	Literature reference
CA-242	↑	↑		NF
CEACAM 1	NA	↑	↓	15
CEACAM 5 (mono)	↑	↑	↑	10
CEACAM 5 (poly)	NA	↑	↑	10
CEACAM 8	NA	-		NF
CTGF	-	↑		NF
EGF	-	NA		NF
EGFR	↓	-	↑	21
Fractalkine	NA	-		NF
GDNF	-	↑		NF
Her2	↓	-	↑	17,18
Her3	↓	-		NF
Her4	↓	↑		NF
IL-1a	↓	↑	↑	13
IL-4	NA	-		NF
IL-6	↓	NA		NF
IL-8	↑	↑	↑	11
Mesothelin	NA	↓		NF
MIF	↑	↑		23
MIP-1	-	↑		NF
OPN	-	-	↑	22
S100A8	NA	-	↑	16
SDF-1	↑	-	↓	19
SLPI	↑	↑	↑	24
TGFb1	-	NA		NF
TIMP-1	↓	↑	↑	20
TNFa	-	↑	↑	11,12
VEGF	-	-	↑	14
YKL-40	↑	-	↑	25

* NF: No literature data found

** NA: Not applicable

5.10.1. Panel A

In this panel we had the most commonly used biomarker for colorectal cancer; CEACAM 5. We detected it both with a polyclonal set of antibodies and two matched monoclonal. This marker was expected to be up-regulated in CRC and that was also seen with both our PLA® assays[10]. The markers IL-8, TNF-α and IL1-α have also been reported to be elevated in CRC patients (IL-1 α in tissue) of CRC patients[11,12,13], and our assays showed the same pattern.

VEGF was documented to be elevated in blood plasma of CRC patients[14]. This was also one of the best assays available for PLA® with a very high signal and good sensitivity. We could however not see any elevated levels in the cases compared to controls. IL-4 appeared to have similar expression in cases as in controls as well, but here no literature data was found.

5.10.2. Panel B

Most of the assays (CEACAM5, CEACAM1, CTGF, GDNF, MIP-1 and CA-242) appeared to be up-regulated in cases. The only literature data found was for CEACAM 1 where it was observed to be down-regulated in adenomas and carcinomas[15]. Fractalkine was the only assay in this panel with no significant difference between the populations.

5.10.3. Panel C

We did not expect TIMP-1 or S100A8 to yield a signal in this panel due to the hook effect (both were high abundant in blood plasma and reported to be elevated in CRC blood plasma[20,16]). The rearrangement in the panels made us sacrifice these markers to rescue the human epidermal growth factor receptors (Her family). The only marker that seemed to be significantly different in cases compared to controls was Her4. No literature data could be found for this marker though. SDF-1 and CEACAM 8 were poor assays during the evaluation and did not show any significance here either. Some of the markers in this panel were however documented to be up- or down-regulated in CRC patients[17, 18, 19], but it was hard to know if we lacked the precision, the sensitivity or if the normalization was affecting the dataset.

5.10.4. Panel D

Most of the markers in this panel have been reported to be up-regulated (although some in tissue) in CRC patients[20, 21, 22, 23, 24, 25]. What appeared to be correct for us was then TIMP-1, MIF and SLPI. Mesothelin showed a lower concentration in our cases than controls, but no literature data could be found.

5.11. Summary

The multiplexed proximity ligation assay is a very promising tool for biomarker research. The sensitivity is comparable to other immuno based assays, but the advantage of double recognition in the detection step makes it perfect for multiplexing in theory. The results presented are most in accordance to the literature, and one could see this study as a biomarker screening with the correct answers. We re-discovered many potential biomarkers and did only see one assay showing conflicting data compared to the literature. This is very promising data for a PLA design that has existed for only a few months.

Assay development might seem straight forward, but there are many parameters to take into consideration when optimizing these kinds of assays. This particular method is special due to the fact that we work with both protein binding events as well as DNA hybridizations, which means that we need to be very careful when pulling in the different parameters; everything is connected.

The protocol used in this study needs to be further improved, especially if the assay should be transferred to other laboratories. As mentioned above, our biggest problem was the temperature differences in the PCR instrument and this will have to be very robust in order to achieve better repeatability and reproducibility. This is not as trivial as it might seem; increasing binding strength could yield more unspecific primer binding among other things. One possible way around this problem is to use a universal primer sequence for all of the assays instead of 9 different in the pre amplification step. By doing this all the assays would be affected the same way, and the normalization would work (in theory!). But this would require new design of all the sequences along with new conjugations of antibodies, altogether a quite expensive story.

6. References

[1] R. Etzioni, N. Urban, S. Ramsey, M. McIntosh, S. Schwartz, B. Reid, J. Radich, G. Anderson and L. Hartwell. The Case For Early Detection. *Nature Reviews.* **2003**. Vol. 3.

[2] DG. Ward, N. Sugget, Y. Cheng, W. Wei, H. Johnson, LJ. Billingham, T. Ismail, MJO. Wakelam, PJ. Johnson and A. Martin. Identification of serum biomarkers for colon cancer by proteomic analysis. *British Journal of Cancer*. **2006**. Vol 94: 1898-1905.

[3] National Cancer Institure, Surveillance Epidemiology and End Results 1975-2006. http://seer.cancer.gov/csr/1975_2006/browse_csr.php?section=6&page=sect_06_table.11.htm (downloaded October 2009)

[4] P. Boyle and J. Ferlay. Cancer incidence and mortality in Europe, 2004. *Annals of Oncology.* **2005.** Vol 16: 481-488.

[5] N. L. Anderson and N. G. Anderson. The Human Plasma Proteome: History, Character, and Diagnostic Prospects. *Mol. Cell. Proteomics.* **2002**. Vol 1: 845-867.

[6] S. Fredriksson, W. Dixon, H. Ji, A.C. Koong, M. Mindrinos and R.W. Davis. Multiplexed protein detection by proximity ligation for cancer biomarker validation. *Nature Methods*. **2007**.

[7] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, H. Erlich. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring harb Symp Quant Biol*. **1986**. Vol. 51, Pt 1:263-73.

[8] C. A. Heid, J. Stevens, K. J. Livak and P. M. Williams. Real Time Quantitative PCR. *Genome Res*. **1996**. Vol. 6: 986-994.

[9] M. Gullberg, S. M. Gústafsdóttis, E. Schallmeiner, J. Jarvius, M. Bjarnegård, C. Betsholtz, U. Landegren and S. Fredriksson. Cytokine detection by antibody-based proximity ligation. PNAS. **2004**. Vol. 101, no. 22: 8420-8424.

[10] J. Schneider and G. Schulze. Comparison of tumor M2-pyruvate kinase (tumor-M2-PK), carcinoembryonic antigen (CEA), carbohydrate antigens CA 19-9 and CA 72-4 in the diagnosis of gastrointestinal cancer. *Anticancer Res.* **2003**. Vol 23: 5089-5093.

[11] J. Hartwich, SJ. Konturek, P. Pierzchalski, M. Zuchowicz, PC. Konturek, W. Bielański, K. Marlicz, T. Starzyńska and M. Ławniczak. Molecular basis of colorectal cancer – role of gastrin and cyclooxygenase-2. *Med Sci Monit.* **2001**. Vol 7: 1171-1181.

[12] L. Tessitore, B. Vizio, O. Jenkins, I. De Stefano, C. Ritossa, JM. Argiles, C. Benedetto and A. Mussa. Leptin expression in colorectal and breast cancer patients. *Int J Mol Med.* **2000**. Vol 5: 421-426.

[13] Y. Matsuo, H. Sawai, J. Ma, D. Xu, N. Ochi, A. Yasuda, H. Takahashi, H. Funahashi and H. Takeyama. IL-1alpha secreted by colon cancer cells enhances angiogenesis: the relationship between IL-1alpha release and tumor cells' potential for liver metastasis. *J Surg Oncol.* **2009**. Vol 99: 361-367.

[14] P. Myśliwiec, Z. Piotrowski, B. Zalewski, A. Kukliński and K. Pawlak. Plasma VEGF-A and its soluble receptor R1 correlate with the clinical stage of colorectal cancer. *Rocz Akad Med Bialymst.* **2004**. Vol 49, suppl 1: 85-7.

[15] MJ. Bermúdez-Soto, M. Larrosa, JM. Garcia-Cantalejo, JC. Espín, FA. Tomás-Barberan and MT. García-Conesa. Up-regulation of tumor suppressor carcinoembryonic antigen-related cell adhesion molecule 1 in human colon cancer Caco-2 cells following repetitive exposure to dietary levels of a polyphenol-rich chokeberry juice. *J Nutr Biochem.* **2007**. Vol 18:259-271.

[16] HJ. Kim, HJ. Kang, H. Lee, ST. Lee, MH. Yu, H. Kim and C. Lee. Identification of S100A8 and S100A9 as serological markers for colorectal cancer. *J Proteome Res.* **2009**. Vol 8:1368-1379

[17] M. Pohl, I. Stricker, A. Schoeneck, K. Schulmann, S. Klein-Scory, I. Schwarte-Waldhoff, M. Hasmann, A. Tannapfel, W. Schmiegel and A. Reinacher-Schick. Antitumor activity of the HER2 dimerization inhibitor pertuzumab on human colon cancer cells in vitro and in vivo. *J. Cancer Res Clin Oncol.* **2009**. Vol 135:1377-1386

[18] S. Essapen, H. Thomas, M. Green, C. De Vries, MG. Cook, C. Marks, C. Topham and H. Modjtahedi. The expression and prognostic significance of HER-2 in colorectal cancer and its relationship with clinicopathological parameters. *Int J Oncol.* **2004.** Vol 24: 241-248

[19] J. Dimberg, A. Hugander, S. Löfgren and D. Wågsäter. Polymorphism and circulating levels of the chemokine CXCL12 in colorectal cancer patients. *Int J Mol Med.* **2007**. Vol 19: 11-15

[20] C. Frederiksen, AF. Lomholt, GJ. Davis, BL. Dowell, MA. Blankenstein, IJ. Christensen, N. Brünner and HJ. Nielsen. Changes in plasma TIMP-1 levels after resection for primary colorectal cancer. *Anticancer Res.* **2009**. Vol 29:75-81

[21] S. Goodin. Development of monoclonal antibodies for the treatment of colorectal cancer. *Am J Health Syst Pharm*. **2008**. Vol 65(11 Suppl 4):S3-7

[22] D. Agrawal, T. Chen, R. Irby, J. Quackenbush, AF. Chambers, M. Szabo, A. Cantor A, D. Coppola and TJ. Yeatman. Osteopontin identified as lead marker of colon cancer progression, using pooled sample expression profiling. *J Natl Cancer Inst*. **2002**. Vol 94:513-521

[23] JM. Wilson, PL. Coletta, RJ. Cuthbert, N. Scott, K. MacLennan, G. Hawcroft, L. Leng, JB. Lubetsky, KK. Jin, E. Lolis, F. Medina, JA. Brieva, R. Poulsom, AF. Markham, R. Bucala and MA. Hull. Macrophage migration inhibitory factor promotes intestinal tumorigenesis. *Gastroenterology.* **2005**. Vol 129: 1485-1503

[24] RI Jr. Garver, KT. Goldsmith, B. Rodu, PC. Hu, EJ. Sorscher and DT. Curiel. Strategy for achieving selective killing of carcinomas. *Gene Ther.* **1994**. Vol 1: 46-50

[25] EV. Høgdall, JS. Johansen, SK. Kjaer, PA. Price, L. Christensen, J. Blaakaer, JE. Bock, E. Glud and CK. Høgdall. High plasma YKL-40 level in patients with ovarian cancer stage III is related to shorter survival. *Oncol Rep.* **2003**. Vol 10: 1535-1538

Appendix 1

Initial Protocol **Final Protocol**

Sample dilution:

Plasma samples diluted 1:50 in buffert for use on panel C and D

Plasma samples diluted 1:50 in buffert for use on panel D

Sample blocking:

Diluent mix	Conc	1x
Plasma Diluent	-	75
Goat Probe (Olink)	2,67 uM	25
Total		100

Mix 4 ul sample (diluted or not) with 4 ul diluent mix
Incubate RT 20 min

	Diluent mix	Conc	1x
	Plasma Diluent	-	84,7
	Goat probe (Innova)	19 uM	13,2
5 pM final	GFP	10 nM	0,2
50 pM final	PE	10 nM	2
	Total		100

Mix 10 ul diluent mix with 10 ul undiluted plasma samples
Incubate 20 min at RT

Probe incubation:

Probe mix	Conc	1x
Probe 3´	10 nM	0,4
Probe 5´	10 nM	0,4
BSA	10%	4
Triton X-100	1%	4
FBB		31,2
Total		40

(One for each panel)

Mix 2 ul diluted plasma and 2 ul probe mix
Incubate o/n 4?C

	Probe mix	Conc	1x
	Probe 3´	10 nM	2,4
	Probe 5´	10 nM	2,4
2 nM	*anti-YKL-40	1 uM	0,08
2 nM	*anti-TIMP-1	1 uM	0,08
	BSA	10%	4
	Triton X-100	1%	4
	FBB		27,04
	Total		40

(One for each panel)
* Only in panel D

Mix 2 ul diluted plasma and 2 ul probe mix
Incubate o/n 4?C

Ligation

Ligation mix	Conc	1x
Sample		
10x buffert		10
H2O		85
T4 ligase	1:150	1
Total		96

Incubate in PCR machine: 37C 10 min
65C 10 min

Ligation mix	Conc	1x
Sample		
10x buffert		10
H2O		85
T4 ligase	1:150	1
Total		96

Incubate in PCR machine: 37C 10 min
65C 10 min

UNG

UNG mix	1x
U-excition mix	2,75
PBS	8,75
Total	11,5

Add 1 µl to each sample
Incubate 5 min at RT

UNG mix	1x
U-excition mix	2
PBS	9,5
Total	11,5

Add 1 µl to each sample
Incubate 5 min at RT