DISSERTATION

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

Syafrizayanti

born in Padang, Indonesia

Oral examination: March 11, 2014

Production of Personalized Protein Microarrays

Optimized production of protein microarrays and the establishment of processes for the representation of protein conformations that occur in individual patients

Referees: Prof. Dr. Stefan Wiemann

Prof. Dr. Stefan Wölfl

Thesis declaration

I hereby declare that this thesis has been written only by the undersigned and without any assistance from third parties. Furthermore, I confirm that no sources have been used in the preparation of this thesis other than those indicated in the thesis itself.

Heidelberg, 2014

_____

Syafrizayanti

Part of this work has been accepted for publication:

Syafrizayanti, Betzen C, Hoheisel JD, Kastelic D. Methods for analysing and quantifying protein-protein interactions. *Exp. Rev. Prot.* (2014) in press.

Another manuscript will be submitted while this thesis is evaluated:

Syafrizayanti, Lueong SS, Hoheisel JD. Production of personalized protein microarrays (2014).

Lueong SS, Syafrizayanti, Hoheisel JD. Production of functional protein microarrays from cDNA products and functional applications (2014).

Poster presentations

Syafrizayanti, Lueong S, Hoheisel JD, Personalised proteomics by means of individualised protein microarrays, *HUPO 12th Annual World Congress*, Yokohama, September 2013

Bilen S, Syafrizayanti, Hoheisel JD, Antibody specific isoform protein interactions using *in situ* protein microarray platform, *Functional Genomics and Proteomics - Applications, Molecular Diagnostics & Next Generation Sequencing*, DECHEMA, February 2012

Syafrizayanti, Hoheisel JD, Solid phase amplification of covalently bound of cDNA for in situ cell free protein expression towards personalized proteome, *PhD DKFZ PhD Retreat*, Weil der Stadt, July 2011

Syafrizayanti, Fugazza M, Hoheisel J, On-chip PCR products for *in situ* protein expression, *Scientific Retreat Research Program B: Structural and Functional Genomics*, May 2011

Syafrizayanti, Di C, Fugazza M, Hoheisel JD, On-chip PCR products for *in situ* protein expression, *Functional Genomics - Next Generation Applications and Technologies* , DECHEMA, February 2011

....taruntuak anak nagari

dibaliak pandakian ado panurunan

dibaliak panurunan ado pandakian

Acknowledgments

Many people have professionally and personally supported me during the time of my Ph.D. and have given generously of their time and expertise. My wholehearted gratitude is expressed to:

Dr. Jörg Hoheisel, head of division Functional Genome Analysis (B070), DKFZ, who provided valuable and constructive suggestions throughout the planning and development of this research. His willingness to so generously share his knowledge and time is truly appreciated.

Prof. Dr. Stefan Wiemann (DKFZ) and Prof. Dr. Stefan Wölfl (IPMB), the referees on my thesis advisory committee, are sincerely thanked for their guidance, enthusiastic encouragement, and useful comments and discussions on this research work.

Special thanks go to Dr. Anette Jacob (B070, DKFZ and Peps 4LS GmbH) for her guidance and valuable support. Advice given by Dr. Andrea Bauer (B070, DKFZ) has been a great help in my research. Dr. Christoph Schröder (B070, DKFZ) kindly and professionally helped in data analysis.

Anke Mahler and Marie-Christine Leroy-Schell kindly helped in all sorts of administrative issues. Technical assistance was skillfully provided by Melanie Bier, Stefanie Kutschmann, and Sandra Widder – all of whose professional devotion is greatly appreciated.

My special thanks for useful help and discussion go to Smiths Lueong, Harish Srinivasan, Pedro Simonini, Mohanachary Amaravadi and to all other current members of B070, DKFZ and also to its former members Reham Helwa, Cuixia Di, and Manuel Fugazza.

Theodor C. H. Cole provided precious advice and I thank him for language editing and proofreading my drafts and manuscripts as well as Erika Siebert-Cole for her kind hospitality.

The Directorate General of Higher Education, Republic of Indonesia, granted financial support through a DIKTI scholarship.

My deepest gratitude goes to my entire family: especially to my parents for their support and encouragement throughout my studies. Your prayers have sustained me thus far. I am, and always will be thankful and indebted to you for all the sacrifices you made on my behalf. Words cannot express my love and gratefulness to my brother Adrial and my sister Syafrimayani.

My beloved husband Ikhwan Resmala Sudji was always there at all times to support me in all ways, and especially in moments when no one else would answer my queries.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

APID	Agile Protein Interaction Data Analyzer
APTES	Aminopropyltriethoxysilane
BIND	Biomolecular Interaction Network Database
BioGRID	Biological General Repository for Interaction Datasets
BSA	Bovine serum albumin
cDNA	Complementary DNA
CV	Coefficients of variations
DAPA	DNA array to protein array
DARPin	Designed ankyrin repeat protein
dH_2O	Distilled water
DIP	Database of Interacting Proteins
DNA	Deoxyribonucleic acid
ELISA	Enzyme linked immunosorbent assay
ENCODE	Encyclopedia of DNA Elements
GENECODE	Encyclopedia of Genes and Gene Variants
GFPs	Green fluorescent proteins
GMR	Giant magnetoresistance
GST	Glutathione *S*-transferase
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
His tag	Polyhistidine
HMHM	1-(5-Hydroxy-5-methylhexyl)-methylxanthine
HPLC	High performance liquid chromatography
HPRD	Human Protein Reference Database
IntAct	Molecular Interaction Database
LIMMA	Linear Models for Microarray Data
LP	Laser Power
MAPK	Mitogen-activated protein kinase
MAPPIT	Mammalian Reverse Two-Hybrid Technology
MBP	Maltose binding protein
MINT	Molecular Interaction Database

MIST	Multiple spotting technique
mRNA	Messenger RNA
mtDNA	Mitochondrial DNA
NAPPAs	Nucleic acid programmable protein arrays
Ni-NTA	Nickel-nitrilotriacetic acid
ORFs	Open reading frames
PBS	Phosphate buffer saline
PBST	Phosphate buffer saline with Tween-20
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PMT	Photomultiplier tube
PPI	Protein–protein interaction
ProCAT	Protein Chip Analysis Tool
PTB	Phosphotyrosine binding
RBS	Ribosomal binding site
RefSeq	NCBI Reference Sequence Database
RNA	Ribonucleic acid
RT-PCR	Reverse transcription-polymerase chain reaction
SARS	Severe Acute Respiratory Syndrome Diagnostics
SNPs	Single nucleotide polymorphisms
SPR	Surface plasmon resonance
SSC	Saline–sodium citrate buffer
Y2H	Yeast to hybrid

SUMMARY

Despite remarkable progress in understanding biology and disease at the level of nucleic acids, insights into the relevant biochemical processes frequently remain preliminary, since much regulation and activity occurs at the protein level through control of gene expression and variations of protein conformation. In particular, the effect of such variations on protein interactions is critical for a better description of biology and disease. Protein microarray technology provides a means to such ends and is a growing field of proteomics, with a high potential for analytical and functional applications in biology and medicine. On the basis of sequence information from individuals, it is possible to characterize disease-specific protein isoforms that result from mutations, polymorphisms, and splice variants with **personalized protein microarrays**. During my thesis, I developed such a technique.

As a first step, **solid-phase PCR** is applied to copy a particular tissue's RNA/cDNA onto the microarray surface, using for each gene a specific primer pair that is attached to the chip surface. The generated DNA templates are firmly attached to and specifically oriented on the array surface. The solid-phase PCR successfully amplified DNA of up to 3 kb, also allowing multiplex amplification of DNA. The arrayed DNA copies then act as templates for an *in situ* cell-free expression, yielding a protein microarray that presents the protein content of a particular tissue of an individual person. Expression control was conducted by a multiple spotting technique (MIST). C-terminus detection showed that translation was complete, yielding full-length proteins.

During the process of setting up the technique of producing individualized protein microarrays, the MIST technology was optimized concomitantly. The various steps involved were analyzed to determine optimal conditions for template preparation, protein expression and interaction detection. Protein microarrays of 3500 human proteins were produced with these procedures and their performance was tested in model studies of protein–protein interactions.

ZUSAMMENFASSUNG

Trotz erheblicher Fortschritte im biologischen Verständnis funktioneller, zellulärer Abläufe sowie der entsprechenden krankheitsrelevanten Störungen auf der molekularen Ebene der Nukleinsäuren, sind viele zugrunde liegende biochemische Prozesse häufig nur in Grundzügen oder teilweise bekannt, da ein Gutteil der Aktivität und Regulation durch Proteine, etwa durch Variation ihrer Expression oder ihrer Konformation, erfolgt oder beeinflusst wird. Speziell eine bessere Kenntnis der Effekte dieser Änderungen auf Proteinwechselwirkungen ist von elementarer Bedeutung für ein tieferes Verständnis biologischer Zusammenhänge und krankheitsrelevanter Störungen. Protein-Microarrays ermöglichen ebensolche Untersuchungen und sind eine wichtige Schlüsseltechnologie für das sich rasant entwickelnde Feld der Proteomik, mit einem großen Potenzial für analytische und funktionale Untersuchungen in Biologie und Medizin. Aufgrund vorhandener Sequenzinformationen von Einzelpersonen ist es nun möglich, krankheits-spezifische Proteinisoformen, die auf Mutationen, Polymorphismen, und Splice-varianten basieren, mittels personalisierter Protein-Microarrays zu charakterisieren. Im Rahmen der vorliegenden Doktorarbeit wurde ein solches Verfahren entwickelt und ausführlich erprobt.

Im ersten Schritt wird per Festphasen-PCR die RNA/cDNA eines Gewebes auf eine Microarray-Oberfläche kopiert. Dies geschieht für jedes Gen mit einem spezifischen Primer-Paar, das vorher an der Chip-Oberfläche befestigt wurde. Die entstehenden cDNA-Segmente sind auf der Array-Oberfläche lokalisiert und orientiert. Die Festphasen-PCR kann DNA bis zu einer Größe von 3 kb erfolgreich vervielfältigen; Multiplex-Amplifikation der DNA ist ebenfalls möglich. Die cDNA-Kopien auf dem Microarray dienen als Vorlage für eine zellfreie Expression *in situ*. Hierdurch entsteht ein Protein-Microarray, der das gewebespezifische Proteinmuster einer einzelnen Person widerspiegelt. Die Expressionskontrolle wurde mit der Multiple-Spotting Methode (MIST) gewährleistet. C-Terminus-Erkennung diente als Nachweis für die abgeschlossene Translation mit entsprechend vollständigen Proteinprodukten.

Im Zuge der Entwicklung der personalisierten Protein-Microarrays wurde gleichzeitig das MIST-Verfahren weiter optimiert. Die einzelnen Arbeitsschritte wurden verbessert in Bezug auf bestmögliche Ergebnisse der Herstellung der DNA-Vorlage, der

Protein-Expression sowie der Erkennung von Protein–Protein Wechselwirkungen. Protein-Microarrays mit 3500 Humanproteinen wurden mit den neu entwickelten Verfahrensschritten hergestellt und in Pilotstudien auf ihre Aussagekraft in Bezug auf Protein–Protein Wechselwirkungen hin geprüft.

1 INTRODUCTION

The deciphering of the human genome has provided the basic sequential information contained in the DNA, but there is still the need of elucidating the corresponding functional and mechanistic details of cellular processes leading up to an understanding of molecular interrelationships within the cells [1,2]. The response of individual cells to extracellular signals and the way cells associate in multicellular organisms can be understood by defining the activities of individual proteins. Each protein is involved in many interactions with other proteins as well as with nucleic acids, phospholipids, carbohydrates, and small ligand molecules [3]. As part of a complex system, a protein is not a solo actor. Protein complexes are established by stable or transient physical contacts forming large networks of interactomes controlling all cellular functions [4].

The development of high-throughput proteomics on the basis of protein microarrays is technically demanding due to the structural diversity and complexity of proteins (**Fig. 1**). Protein microarrays contain hundreds to thousands of different protein molecules that can be used for profiling analyses and functional screening.

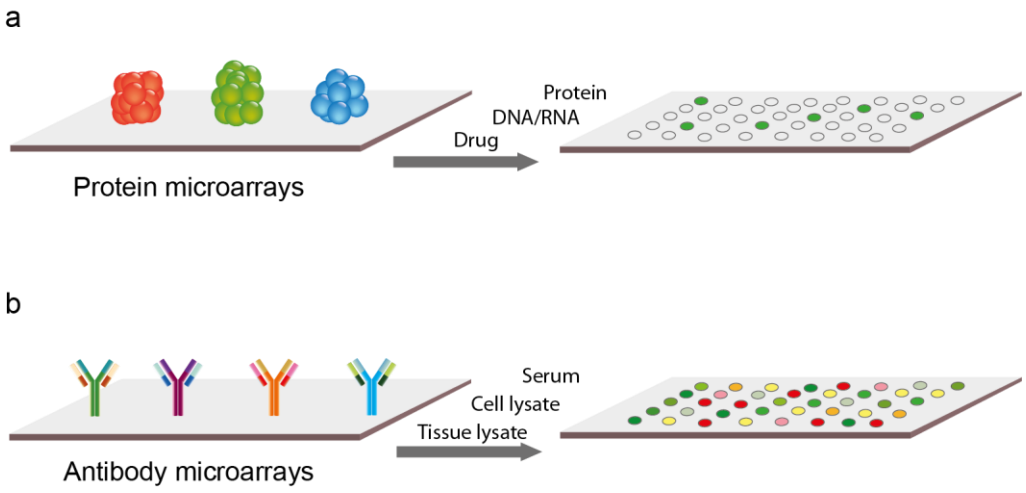


Figure 1. Protein microarrays. (a) Functional protein microarrays for studying protein interactions. A number of proteins or peptides are purified or synthesized, then immobilized onto a surface. The arrays are used to investigate protein interactions, for identifying the substrates of enzymes, and for studying post-translational modifications. (b) Analytical protein microarrays, represented by dual-color antibody arrays. Different types of capture molecules such as antibodies, aptamers, or antigens with high affinity and specificity are arrayed onto a surface. The procedure can be used to monitor expression levels and profiling of proteins from sample and reference.

Profiling protein microarrays are based on immobilized specific proteins such as antibodies that can recognize particular proteins from a sample mixture [5].

Functional protein microarrays contain entire proteins or protein domains that can be used to identify or screen for interactions with proteins, peptides, nucleic acids, and various types of smaller ligands [6]. To date, protein microarrays are mostly produced by spotting purified proteins onto solid surfaces [7]. Alternatively, protein arrays can be produced by using cell-free expression systems. Although most of the newly synthesized proteins do not retain their native three-dimensional structure, many binding sites will be of sufficient quality for individual applications. Cell-free expression systems use cell lysates from growing cells, either prokaryotic or eukaryotic, that contain the entire essential components for transcription and translation and can use mRNA or DNA as templates [8]. Cell-free-based protein microarrays can avoid the bottleneck of protein expression and purification [9,10].

Several cell-free protein microarray technologies have been developed. Nucleic acid programmable protein arrays (NAPPAs) simultaneously spot biotinylated cDNA plasmids encoding GST fusion proteins and anti-GST antibodies onto avidin-coated slides. *In situ* expression has been achieved by incubating the spotted molecules with cell-free lysate from rabbit reticulocytes. The expressed proteins were captured by anti-GST antibodies within each spot [10]. In a different strategy, two slides are used in a technology called DNA array to protein array (DAPA) – the first slide immobilizes the DNA templates, and the second captures the expressed proteins; cell-free lysate is placed between the two slides [9]. Another strategy uses the multiple spotting technique (MIST) where unpurified PCR products are spotted as DNA templates, followed by *Escherichia coli*-based lysate being placed directly on top of the spotted DNA [11]. Protein microarrays containing several hundred human proteins have been designed using MIST.

Alternatively, the attached DNA can be directly synthesized on the slide surface using on-chip PCR technology. The technology can be utilized to amplify RNA or cDNA encoding disease-specific protein isoforms (mutations, polymorphisms, and splice variants) on slide surfaces using DNA-specific primers attached to the slide surface. The DNA arrays act as templates for *in situ* cell-free expression to produce protein arrays of particular tissues of an individual person. On-chip PCR has been

introduced to synthesize DNA based on the extension of one or two primers covalently attached to the surface by DNA polymerase [12,13].

Aims and Scope. A new technique was developed to produce personalized protein microarrays from surface-bound DNA templates by on-chip PCR. The bound DNA is expressed using an *in situ* cell-free procedure to yield proteins on the array. Also, the MIST procedure was employed to produce and optimize protein microarrays of 3500 human proteins. Array quality was analyzed to determine and evaluate the technical usability and variability of the involved PCR, protein expression, and probing. These protein microarrays were used in various model studies of protein–protein interactions and screened for antibody specificity.

1.1 From Genomics to Proteomics

The final human genome sequence was published three years after the first announcement of a draft of the human genome in 2001. Smaller tasks still remain to be solved until sequencing technologies are able to read the highly repetitive sequences within heterochromatin [14]. Revealing the biological functions of 2.85 billion bases of the human genome from protein-coding genes, non-protein-coding genes, noncoding sequences, and regulatory elements is a key challenge for biomedical research [15,16].

The major human gene catalogs are Ensembl, RefSeq, UNSC, and GENCODE− each listing about 25,000 protein-encoding genes [17-19]. The ENCODE consortium has been able to determine the major biochemical functions of the human genome, while 20% of these functions remain unknown. Knowing the complete sequence has increased our understanding of the genome and in the near future this knowledge may lead to clinical applications. For example, gene expression profiling has provided several prognostic signatures in breast cancer (including NKI-70-gene, wound-response, and stem-cell signatures) [20]. These achievements are supported by high-throughput genomic analyses that can perform rapid and simultaneous interrogation of massive data.

The number of protein-coding genes in the human genome appears to be lower than the complexity of the biological functions of human cells [21]. The lack of correlation is explained by the transcriptional control mechanism of the regulating differential gene expression. Alternative transcripts and splicing hence generate

multiple transcript variants that may encode functionally distinct protein isoforms in addition to a diversity of structure modifications during protein synthesis [22-25].

1.1.1 Proteomic Diversity

Following high-throughput genomic analysis, correlations of mRNA expression levels and protein abundance can be surveyed [16]; such studies have been carried out in multicellular organisms [26]; larger proteomic datasets in progress may eventually provide the means for more detailed correlation studies. One study in the human Daoy medulloblastoma cell line found that mRNA expression only related to about 66% of the determined protein abundance and variation [27]. The diverse control mechanisms involved in production and maintenance of cellular proteins likely serve to provide a broad range of functional variability. Alternative splicing and protein modification contribute to this variability. Post-transcriptional, post-translational, and degradation mechanisms and their role in regulating protein abundance are still only barely understood [28-31]. Transcriptional control mechanisms regulate differential gene expression, alternative transcripts, and splicing thereby generating multiple transcript variants that encode functionally distinct protein isoforms [23,24,32].

Alternative splicing appears to serve to increase protein diversity thus enabling to manage the working machinery in multicellular eukaryotes without significantly increasing the genome size [29,33]. The mechanisms can be categorized based on five basic splicing patterns: casssete exons, alternative 5′ splice sites, alternative 3′ splice sites, mutually exclusive exons, and retained introns [34]. Furthermore, alternative splicing not only increases proteome size but contributes to the complexity of signaling and regulatory networks of higher organisms [28]. It has been estimated that between 35% and 60% of protein isoforms are derived from alternative splicing of mRNA [34]. The importance of alternative splicing as a regulatory process has become notably evident through its common occurrence regarding pre-mRNAs of regulatory and signaling proteins [35]. Alternative splicing can augment organism complexity, not only by effectively increasing proteome size and regulatory and signaling network complexity, but also by doing so in a time- and tissue-specific manner. Therefore, the mechanism can support cell differentiation, developmental pathways, and other

processes associated with multicellular organisms [35,36]. Alternative splicing has been proposed to occur in respect to intrinsic disorders [37].

There are examples of every kind of splicing in cancerous cells. Splice events that affect the protein-coding region of mRNAs lead to proteins that differ in their sequence and activities; splicing within the noncoding regions can result in changes in regulatory elements, such as translation enhancers or RNA stability domains, which may dramatically influence protein expression [34]. Cancer-specific splicing events have been reported in colon, bladder, pancreatic, and prostate tissue with diagnostic and prognostic implications [38,39]. Combined studies have utilized a spliceoform-specific microarray and PCR to evaluate all known splice variants in human pancreatic cancer cell lines representing a spectrum of differentiation, from near-normal HPDE6 to Capan-1 and poorly differentiated MiaPaCa2 cells. Data from both cell lines and human pancreatic tissue has revealed the largest diversity in alternative spliceoforms in normal cells and tissues, while pancreatic cancer and cancer cell lines showed a sharp reduction in the number of spliceoforms [40]. Therefore, proteome studies aiming to elucidate protein structures, functions, and interactions may very well lead to a better understanding of the biological functions of individual genes.

1.1.2 Protein–Protein Interactions (PPIs)

The identification of each protein and mapping of its interactions with stable or transiently interacting protein partner(s) will help us to better understand biological functions in the context of protein cellular networks [41]. The complete interactome approach has provided a valuable fundamental understanding of biological system network dynamics, and has helped in identifying relevant information on drug targets for applied biomedical research. This understanding of signaling protein interactions has allowed to elucidate certain cancer-related processes, like signaling, cell cycle control, and DNA repair mechanisms. The mapping of the human interactome as a complete network from both healthy and cancer patients can significantly help to identify potential drug targets [3, 42-44].

Around 130,000 binary interactions were estimated in the human interactome using a Monte Carlo simulation based on a combination of experimental data of high-throughput yeast two-hybrid system (Y2H), mammalian reverse two-hybrid technology

(MAPPIT), and Bayesian modeling [45]. Similarly, the combination of Y2H and small interference RNAs has been successful to map 2,269 protein interactions involved in the mitogen-activated protein kinase (MAPK) pathway, known as one of the main pillars of signal transduction in the mammalian cell [46]. Moreover, the statistical estimation could generate 4,000 types of complex proteins, but the protein structures available in the protein databank (PDB) at this time only covered 42% of the predicted protein complexes. This study also predicted that another 25 years of experimental work was needed to complete the repository [47]. Therefore, open-source human protein–protein interaction (PPI) databases have become a major resource for investigating cellular protein networks and signaling pathways. The Agile Protein Interaction DataAnalyzer (APID) has integrated the experimentally validated human interactome databases such as BIND, BioGRID, DIP, HPRD, IntAct, and MINT into a single interactive bioinformatic web tool [48].

PPI studies have revealed several effects that one interacting partner has over the other one – like, altering the kinetic properties of proteins, formation of new binding sites, change in the specificity of a protein for its substrate, and possible inactivation of a protein [49]. The interactions between proteins can occur in a transient or more permanent manner. Transient interactions are far less conserved and colocalized than stable interactions. They are the dominating players in cell processes and include several protein families like kinases, phosphatases, glycosyl transferases, and proteases, amongst others. They are involved in the recruitment and assembly of the transcription complex to specific promoters, the transport of proteins across membranes and breakdown and reformation of subcellular structures like the spindle apparatus and the nuclear pore complex during cell division [41]. Protein interactions can be obligate, where a protein constitutes an integral part of the structure (e.g., hemoglobin, ribosomes, and proteasomes) or nonobligate, where each individual component of a complex can exist free in solution [50]. However, there is no clear distinction between obligate and nonobligate interactions, rather there exists a continuum between them which depends on their physical localization and on various conditions like pH, protein concentration, etc.

Understanding the recognition events of PPIs requires the description of the binding thermodynamics that manage the associations under equilibrium conditions: the free energy and entropy of the binding and enthalpy of individual proteins involved in

the protein complex [51]. The strength of the PPI is defined by the dissociation constant (K_d) or as the affinity constant (K_a) which is the inverted value of K_d that is commonly employed for characterizing antibody–antigen interactions [50]. The K_d value needs to be specifically defined and be independent of the method used to measure it. Numerous proteomics methods have been used to provide thermodynamics data, kinetic analyses, and stoichiometric information [52-54]. We concisely describe the methods that have been commonly used to identify PPIs followed by description of several analyses regarding characteristic interactions studied in this project.

The methods were developed with different consideration related to the nature of the proteins and the type of PPIs. Instead of only finding single protein pairs, co-complex methods like co-immunoprecipitation (Co-IP) and affinity purification (AP) have the potential to catch sets of proteins and decipher the result as part of a protein network [55,56]. With the advantage of high-throughput performance and easy multiplexing, array-based technology is an attractive choice to detect PPIs attached to a surface. Furthermore, the protein microarray technology allows for high-throughput detection of PPIs and source protein bottlenecks can be avoided by the reduced labor input of cell-free protein expression system [5,7,57]. Ionization and separation of mass/charge ratios by mass spectrometry (MS) has advantageously been introduced to proteomics studies. As a golden method in proteomics, MS can be used both to identify protein modification (phosphoproteome) and for large-scale quantification of proteins (quantitative proteomics). MS-based quantitative proteomics can be applied for relative and absolute quantification [58].

Microscopic techniques have recently been employed to detect intact intracellular PPIs also providing useful information regarding their colocalization, especially through Förster or fluorescence resonance energy transfer (FRET), bimolecular fluorescence complementation (BiFC), and fluorescence correlation spectroscopy (FCS) [59-61]. These provide means for directly observing as to how, where, and when a protein interacts under natural conditions and natural expression levels without altering the physiological state of the cell. These techniques rely on the spectral emission from fluorescent and chemiluminiscent molecules that are either fused or conjugated [59,60]. The dye molecules can be combined with chemical and photo-crosslinking techniques, such as proximity-induced ligation or the recent protein complementation assay (PCA) [62].

1.2 Protein Microarrays

The emergence of protein microarrays followed in the footsteps of DNA microarrays. The basic concept of protein microarrays (**Fig. 1**) was first developed on the basis of the immunoassay concept scaling down the assay into micro-spot size [63]. This offers the advantage for a rapid, easily multiplexed, high-throughput procedure and only requires very small volumes of sample consumption over other proteomics methods. Depending on the mode of application, there are two different types of protein arrays: analytical and functional. The analytical protein microarray is still in an early stage of development and is based on binding molecules such as antibodies, antigens, or aptamers. These capture molecules are designed to specifically recognize particular proteins from complex mixtures of samples [64].

Antibody microarrays are most commonly employed in multiplexed high-throughput studies in cancer proteomics research. The development of detection methods for antibody microarrays has become a major challenge in particular regarding their sensitivity and reliable performance for future use as diagnostic tools [65]. Although ELISA-like traditional sandwich assays can be applied for antibody microarrays, the most common methods for detecting interacting proteins are by chemically labelled proteins from test and reference samples (healthy or cancerous) with two different fluorophores (e.g., Cy3 and Cy5). Antibody microarrays can provide profiles from protein extracts and from cell culture supernatants, protein extracts of tissue specimens, and from body fluids [66-68]. For comparing cancerous and normal cells, antibody microarrays can be used to identify cancer-specific biomarker candidates, to characterize the protein profiles, and to identify differentially expressed proteins associated with disease progression. The second type of protein microarrays will be discussed in more detail in the following section.

1.2.1 Functional Protein Microarrays

Functional protein microarrays are constructed by immobilizing a large number of individually purified proteins or protein domains onto a solid support, typically a glass surface, and are utilized for identifying the biochemical properties and the activity of the arrayed proteins. Functional protein microarrays (mostly just named protein microarrays) have been used to recognize protein–protein [12,69,70], protein–antibody

[64,71,72], protein–DNA, protein–RNA, protein–small molecules interactions [73], and for studying protein kinase substrates [2,5]. With this technology platform, one can easily examine the entire proteome by using cDNA from any biological source of interest. Having this potential, proteome arrays have developed rapidly within the last few years in the direction of obtaining biologically active protein arrays.

In the early stage, protein arrays were constructed as printed individually purified proteins that were expressed in *E. coli* or yeast by using contact or non-contact spotters [7,12]. One of this type of protein arrays has been used to print thousands of purified proteins for studying protein–protein interactions, kinase substrate identification, and to identify the small molecule-targeting proteins [7]. Another type of spotted protein array contained 82 purified coronavirus proteins expressed in yeast, purified by GST tags and spotted onto FAST slides, and then used to screen sera from patients with severe acute respiratory syndrome (SARS) infection and healthy controls [74]. However, using the traditional method for obtaining individual purified proteins by overexpression in host cells has some obvious limitations, such as toxicity, inhibitory effects, and low solubility of proteins. A further disadvantage of this process is the fact that it can take up to three weeks of labor-intensive work from expression to purification and finally spotting the proteins into an array [74].

In situ cell-free protein expression has helped to overcome this problem where the major bottleneck has been lessened with the labor and cost effective production of proteins of interest [5,57]. Moreover, cell-free expression systems from *E. coli*, wheat germ, insect and rabbit reticulocytes are rapidly gaining popularity with their clear benefit for protein expression over *in vivo* expression using living cells [75]. Thus, problems can be eliminated that are associated with *in vivo* conditions such as cell membranes, insolubility, cytotoxicity, and instability. Recent developments of prokaryotic and eukaryotic cell-free expression systems provide the means for the synthesis of full-length proteins with high yields and a high percentage of functional molecules [8]. The simultaneous development of cell-like compartments holds the promise of generating large numbers and amounts of biologically active proteins in a cell-free manner. This would also permit a simpler creation of molecules with internal modifications, such as the addition of non-natural amino acids or options for permutation experiments, which may support developments towards personalized medicine [76,77].

1.2.2 Cell-Free Expression Protein Microarrays

To explore this alternative strategy for protein production, several methods have attempted to fabricate cell-free protein micromicroarrays. Proteins were produced using the *in vitro* transcription and translation system directly from a DNA template, either as plasmid or linear DNA, encoding for full-length proteins or protein domains. Towards these means we had previously developed a cell-free expression strategy to generate protein microarrays called multiple spotting technique (MIST) where unpurified PCR products are spotted as DNA templates, followed by placing the cell-free expression system directly on top of the spotted DNA (**Fig. 2a**) [11]. As little as 35 fg DNA template can produce detectable expressed proteins. Since the templates are prepared by PCR, they can be modified by adding epitope tag sequences or by changing the base sequence of the templates. Protein microarrays containing several hundred human proteins have been produced by this method. Here we describe a novel strategy of an immobilized DNA template for *in situ* protein microarrays (**Fig. 2b**). The attached DNA can be directly synthesized on the slide surface using on-chip PCR technology.

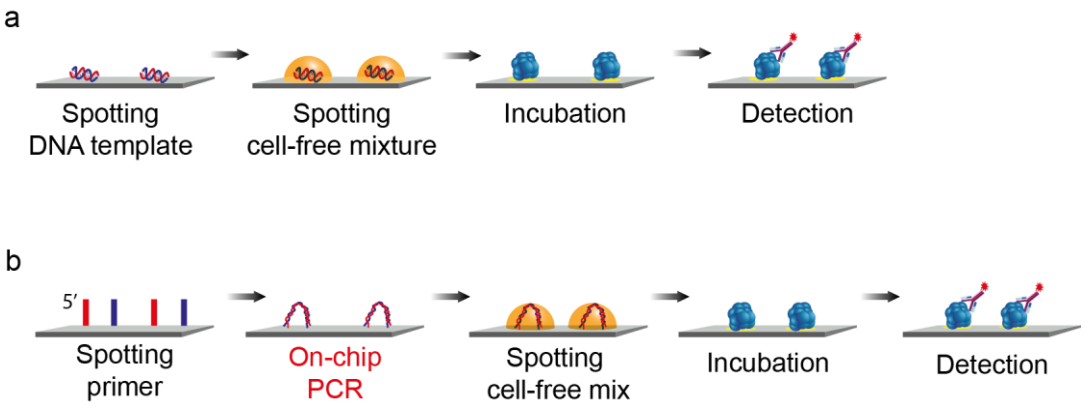


Figure 2. Schematic of MIST and personalized proteins microarrays. (a) MIST – spotting of DNA template and transfer of cell-free mixture on top of the every same spot in a second spotting run. (b) *In situ* cell-free expression of a personalized protein microarray.

A nucleic acid programmable protein array (NAPPA) has been used to generate protein microarrays by printing biotinylated plasmid DNA, avidin, and GST antibody simultaneously onto glass slides and then translating the target proteins by immersing the array with rabbit reticulocyte cell-free system. Epitope tags fused to the expressed proteins allowed them to be immobilized *in situ*. The technology has been applied to

map pairwise interactions among 29 human DNA replication initiation proteins, recapitulate the regulation of Cdt1 binding to select replication proteins, and to map its geminin-binding domain [10,78]. In a different strategy, two slides are used in a technology called DNA array to protein array (DAPA). Here, cell-free lysate is placed between the two slides, the first slide immobilizing the PCR-generated DNA as protein expression templates and the second capturing the expressed proteins [9]. Although, the immobilized DNA can be reused to produce several replicates of protein microarrays, only small numbers of proteins have been reported so far.

To achieve high-density *in situ* cell-free protein expression, the surface must be strong enough to immobilize the protein but still maintain its native function and also provide minimum disturbance of the background. Nonspecific or specific orientation are the two options that have been developed in this regard [11]. Chemically modified surfaces with active functions such as aldehyde or epoxy groups bind to $–NH_2$ on the protein molecules. Lysine residues provide an unprotonated primary amine that can interact with epoxy groups in moderate alkaline pH and form covalent bonds with protein molecules by nonspecific orientation [79]. The most general approach for specific orientation of immobilized protein is the use of recombinant tags, particularly polyhistidine (His tag). The addition of His tag either at the C- or N-termini have commonly been used in recombinant protein technology, because it generally does not interfere with the structure and function of proteins [80].

1.2.3 Detection Method and Data Analysis

Basically, there are two different detection methods used for specific detection of protein microarrays in high-throughput: label-dependent and label-free methods [57] Fluorescence and radio labeling of binding partner molecules can directly monitor protein interactions on the array [74]. The fluorescence labeling strategy has become widely used for detecting proteins on the array due to its effectiveness and its compatibility with the existing laser scanner technology [5,81]. For antibody microarrays, attomolar sensitivities have been reported [82], and even the detection of individual binding events is well in reach [83]. Functional protein microarrays use immobilized proteins mounted on a solid support for identifying protein–protein interactions. For example, fluorescent-labeled peptides have been used in protein

microarray printing of the purified Src homology 2 (SH2) domain and phosphotyrosine binding (PTB) domain encoded in the human genome to measure the according equilibrium dissociation constants; the peptides act as binding sites for tyrosine phosphorylation on the epidermal growth factor receptors [81].

Alternatively, label-free methods using biosensors, mass spectrometry, and atomic force microscopy technology have been applied on the array format [84]. For example, a label-free detection method has been developed to improve the sensitivity of protein microarrays in examining the interacting molecules by using high-density GMR (giant magnetoresistance) sensor arrays to quantify the kinetics of antibody–antigen binding at zeptomolar sensitivities [85]. Other techniques have been applied for protein microarray detection, one of them being surface plasmon resonance (SPR) [86]. The split-protein system is another alternative for detecting PPIs, and one study has successfully applied a cell-free split-protein system [87].

In case of cell-free protein microarrays, much technical variability during multistep production affected the array quality [88]. Learning from DNA microarrays, there are various methods that can be used to analyze protein microarrays either for one or dual channel data intensity. However, one must take into account that data analysis for protein microarrays are still not optimized in regard to than previous established microarrays. Protein microarray data analysis can be adapted to take advantage of several strategies from DNA microarrays, such as spot finding, Z-score calculation, and concentration-dependent analysis [89]. The normalization strategies that have been traditionally used in DNA microarrays, such as global normalization, quantile normalization, and robust linear model normalization, can reduce variation across the arrays [88]. A new approach for background correction called ProCAT is useful to correct background bias and spatial artifacts [90].

1.3 On-Chip PCR Technology

DNA microarrays are able to rapidly perform gene expression analyses; they are used for profiling in functional genomics, drug discovery, and diagnostics, for mutation analyses, and in various applications in microbiological and molecular medicine. DNA chips are produced by different means: usually immobilized oligonucleotides are spotted directly by a photolithographic process onto glass, silicone, or plastic as solid

support [91]. Primer extension of immobilized oligonucleotides at the 5′ end by DNA polymerase was introduced in the early 1990s [92]. DNA synthesis on the chip principally uses solid-phase amplification by means of DNA polymerase. On-chip PCR has been introduced to synthesize DNA based on the extension of one or two primers that are covalently attached to the surface by DNA polymerase. The immobilization of single- or double-stranded oligonucleotides at the 5′ end allows the free 3′ end to serve as a primer. The technology has also been referred to as solid-phase PCR [12] or bridge amplification [93]. Different strategies of primer attachment have been employed to obtain high amounts of amplified products immobilized on the solid support [12].

1.3.1 Solid-Phase PCR

One part of the basic concept of on-chip PCR is to amplify a given DNA template to the attached 5′-end primers through a polymerase reaction, that was introduced as solid-phase DNA amplification or solid-phase PCR [92]. Solid-phase PCR can make use of one immobilized primer with another one free in the solution [12] or two immobilized primers on a solid surface [93]. In this work, a pair of oligonucleotides was used as forward and reverse primers. The primers were covalently attached at the 5′ end to allow polymerase to copy the DNA template onto the immobilized oligonucleotides [12]. **Figure 3** illustrates the process of solid-phase PCR to generate covalently bound PCR products and the detection methods of DNA products.

The solid-phase PCR is similar in each step to standard PCR, except that primers are bound to the surface. The reactions are begun with denaturation of a freely diffusing DNA template into single strands. Then, a complementary region of the single-strand template anneals to the attached primers and DNA polymerase extends the oligonucleotides. The steps are repeated in an iterative way. Elongation products from solid-phase PCR (max. 6 kb long) have been used to detect single nucleotide polymorphisms (SNPs) in human mitochondrial DNA (mtDNA) [94] and it was shown that the entire human mtDNA genome can be synthesized on a chip [95]. This technology platform also has been used to detect and identify bacterial strains [96] and to perform on-chip reverse transcription PCR [97].

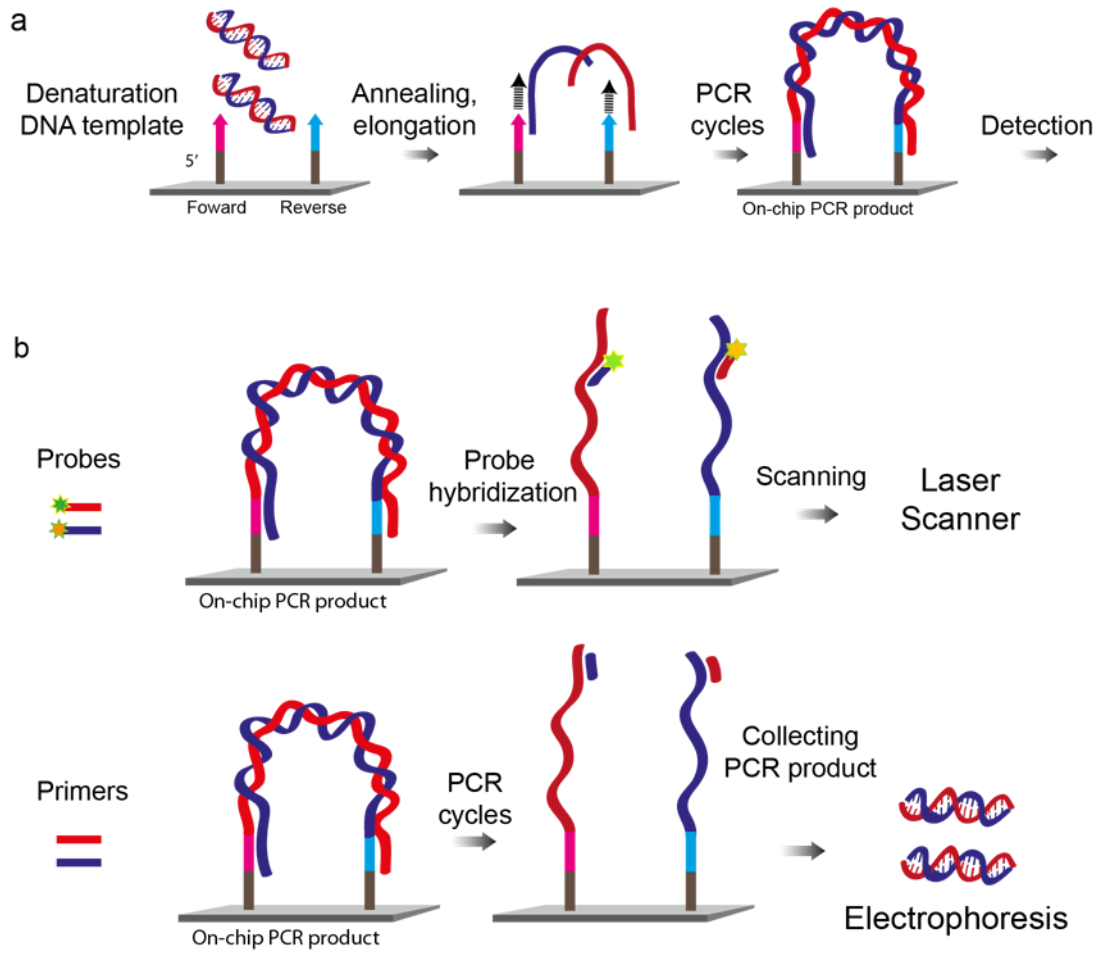


Figure 3. On-chip PCR. (a) Solid-phase PCR. The immobilized oligonucleotides anneals to a freely diffused DNA template; elongation takes place through DNA polymerase activity on a solid surface. (b) Probe detection and second run of solution PCR are the confirmation strategies of on-chip PCR products.

1.3.1 Surface Chemistry

On-chip PCR employs DNA polymerase to extend the free 3′ of a 5′ end immoblized primer, copying the given DNA template [92]. Full accessibility must be guaranteed for the DNA polymerase and for primer annealing. One or two oligonucleotides are attached covalently onto the solid support (microbeads, microtiter plates, or glass slides) [95,96,98]. The primer pair is covalently bound at the 5′ end to the slide surface and the 3′ end is left available for amplification reactions in the functionalized DNA microarray [12]. The covalently bound single or double strands can be detected by radioactive, enzymatic, or hybridization methods [95,99,100]. The covalent linkages are thermally stable as demonstrated by modifying the oligonucleotides at the 5′ thiol and

immobilizing them onto an amino-silanized surface using heterobifunctional crosslinker reagents [12]. Phosphorylated oligonucleotides covalently bound to polystyrene surfaces coated with carbodiimide groups and APTES silanized surfaces [95,100].

In order to perform on-chip PCR through solid-phase amplification technique, the immobilized oligonucleotides serving as PCR primers must be covalently linked via their 5′ end so as to not become affected by repeating thermal cycles [12]. A direct covalent attachment under mild conditions is preferable considering that downstream applications involve many enzymatic reactions. The 5′-amino-modified ends of oligonucleotides are immobilized onto the epoxy surface directly and used for detecting specific mutations via hybridization methods [101]. Additional linkers and spacers improve the efficiency of primer elongation by polymerase activity [98].

Incubation times of more than 4 h are required for immobilizing the probe molecules (like oligonucleotides) under the particular chemical and humidity conditions [102]. Betaine may be used in the spotting mixture of PCR products or oligonucleotides to reduce evaporation during transferring PCR products to the array surface [103]. A comparison of gel-coated and chemically modified non-gel-coated glass slides (with polylysine and aldehyde) showed that the latter provided certain advantages for protein immobilization regarding spot morphology and reproducibility [104]. Aminosilane surfaces are prepared by coating organosilane slides with primary amine. Primary amino-coated glass slides in aqueous solution < pH 9 carry positive charge and can form electrostatic interactions with anionic oligonucleotides or cDNA facilitated by UV exposure or high temperatures [105].

For the purpose of personalized protein microarrays, the slide surface plays a crucial role. Firstly, one of the three main considerations for choosing the type of solid support is that the surface must support specific covalent interactions of the oligonucleotides at the 5′ end. Nonspecific interactions with the PCR cocktail can be minimized using effective blocking methods. Finally, the surface must also be able to immobilize the newly synthesized protein from *in situ* cell-free expression. Specific orientations of the oligonucleotides serving as primer pairs are important, so the immobilization should not modify these structures nor interfere with base pairing or DNA polymerase activity. Proteins can be fused by epitope tagging, which can be applied for immobilization and for detection purposes.

1.3.2 Oligonucleotides and Conjugation Chemistry

The modification can be placed at the 5′ end, 3′ end, or internally using amino dC- or amino-dT-modified base. The purine and pyrimidine bases provide sites for nucleophilic and electrophilic displacements. A strong nucleophile, such as hydrazine and hydroxylamine, can be placed at the 5,6-double bond of pyrimidine in a spontaneous reaction generating a stable modification at the C-5 position. As the electrophilic sites are less advantageous for direct coupling, halogenations of the C-5 cytosine, for instance, with bromine, can be used to couple diamine spacers to the ring structure. Electrophilic alkylation reactions are more effective than nucleophilic reactions to modify purine [79].

The hydroxyl groups are effective for conjugation or modification reactions with alkylating agents or epoxide. In the DNA and RNA chain, the 5′ phosphate group can be modified using condensation agents such as carbodiimides or phosphoramidite derivatives to produce modified nucleotides containing fluorescent, biotin, chelating, or spacer groups. Nucleic acids provide many sites for crosslinking and modification although only one or two of these positions are suitable for specific covalent attachments of oligonucleotides to a solid surface. This internal modification can negatively interfere with oligonucleotide hybridization; modification at the 3′ and 5′ ends can help to prevent this [79].The efficacy of attaching oligonucleotides onto a chemically active surface for the most part depends on how these molecules are modified [106]. The functional groups usually linked to oligonucleotides are aldehydes, primary amines, phosphates, biotin, and thiols [79]. The modified oligonucleotides are attached to the glass surface coated with electrophilic groups (e.g., aldehyde, epoxy, isothiocyanate, and activated esters) through direct or indirect immobilization strategies. The most commonly employed oligonucleotides are modified by primary amines at the 5′ or 3′ ends and covalently immobilized onto the aldehyde- or epoxy-coated glass surface by direct coupling reactions [107]. The covalent attachment of aminated DNA on aldehyde-activated glass has shown better efficiency for hybridization compared to aminated, carboxylated, and phosphorylated DNA on carboxylic acid or amine-modified glass supports [107]. Similar conjugated chemistry of DNA on aldehyde-modified glass slides stabilized by dehydration allows covalent bond formation via Schiff base reaction and has been used to generate DNA microarrays [107].

A high efficiency of hybridization was also shown for 5′-end succinylated oligonucleotides covalently attached to amino-derivatized glass slides by carbodiimide-mediated coupling [108]. Alternatively, 5′-disulfide-conjugated oligonucleotides bound covalently to mercaptosilane-modified glass slides have provided a high-density substrate for nucleic acid hybridization. External thiol-modified DNA oligomers can covalently attach to silane-coated silica slides by using heterobifunctional crosslinkers [109]. Thiol modification also has been used to introduce disulfides to oligonucleotides and represents an efficient and specific covalent attachment chemistry for immobilization of DNA probes onto 3-mercaptopropyl silane-derivatized glass slides [67].

2 MATERIALS AND METHODS

2.1 Materials

Equipment

Name	Manufacturer
Centrifuge 5415D	Eppendorf, Germany
Centrifuge 5810R	Eppendorf, Germany
Heareus Oven	Thermo Scientific, USA
Hoefer TE 70 (Semi-dry Western-Blot)	Amersham Bioscience, USA
Mini-Protean®3 electrophoresis chamber and casting unit	BioRad, USA
NanoDrop ND 1000 Spectrophotometer	Thermo Scientific, USA
PCR Thermocycler PTC200	MJ Research BioRad, USA
pH-Meter MP 230 Mettler Toledo	Mettler Toledo, Germany
Photometer (UV/Vis) Ultrospec 2000	Pharmacia Biotech, Germany
Power Scanner	Tecan, Switzerland
ScanArray 5000	Perkin Elmer, USA

Labware

Name	Manufacturer
96-well culture plate	Greiner Bio-One, Germany
Cell culture microplate,96 and 384 well, F-bottom	Greiner Bio-One, Germany
Cell culture Petri dishes 96 × 20 mm	Corning Life Science, USA
Clean, covered 4-chamber incubation tray	Greiner Bio-One, Germany
Disposable cuvettes UV	Eppendorf, Germany
Eppendorf Safe-Lock tubes 1.5 mL and 2 mL	Eppendorf, Germany
Falcon 15 mL and 50 mL	Greiner Bio-One, Germany
Gene Frame®	Thermo Scientific-ABgene®, Germany
Latex gloves	Latex, Blossom Mexpo, USA
Lazy-L-Spreaders	Sigma-Aldrich, USA
Microtiter cover film	Nunc, Germany

NEXTERION® Slide E	Schott Technical Glass Solution GmbH, Jena, Germany
Nickel Chelate Slides	Xenopore, USA
Nitril gloves	Nitril, Microflex, Austria
Parafilm PM 996	Fisher Scientific, USA
Polypropylene Microplates 96 and 384 well for PCR	Greiner Bio-One, Germany
ProPlate 16-well slide module – 7 mm × 7 mm	Grace Bio-Labs, USA
ProPlate 1-well module, 16 mm × 70 mm	Grace Bio-Labs, USA
Sterile filter 500 mL	Nalgene, USA
Whatman® UNIPLATE microplates 384 well, V-bottom	GE Healthcare Life Sciences, Germany

Chemicals

Name	Manufacturer
1,2-Bis(dimethylamino)ethane (TEMED)	Roth, Germany
2-Ethanesulfonic acid (HEPES)	Roth, Germany
Acrylamide/Bisacrylamide	Roth, Germany
Agar	Roth, Germany
Agarose	Invitrogen, USA
Ammoniumpersulfate (APS)	Roth, Germany
Ampicillin (sodium salt)	Genaxxon, Germany
BSA (bovine serum albumin)	Sigma-Aldrich, Germany
Chloramphenicol	Genaxxon, Germany
Disodium phosphate (Na_2HPO_4)	Roth, Germany
DL-Dithiothreitol (DTT)	Sigma-Aldrich, Germany
Ethanol, absolute	Sigma-Aldrich, Germany
Ethidium bromide	AppliChem, Germany
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich, Germany
GeneRuler 1 kb DNA marker	MBI Fermentas, St. Leon-Rot
GeneRulerTM 100 bp DNA marker	MBI Fermentas, St. Leon-Rot
Glycerol	Sigma-Aldrich, Germany
GoTaq® DNA Polymerase	Promega, USA
Green fluorescent protein (GFP)	Millipore, Billirica, MA, USA

Imidazole	Roth, Germany
Isopropyl-β-D-thiogalactopyranoside (IPTG)	Roth, Germany
Kanamycin	Genaxxon, Germany
Methanol	VWR, Germany
Monosodium phosphate (NaH_2PO_4)	Roth, Germany
Oligonucleotides	Biomers, Germany
Oligonucleotides	Sigma-Aldrich, Germany
Protein-Marker: Broad Range	NEB, Germany
QIAGEN's 10× PCR buffer	Qiagen, Germany
Sodium chloride (NaCl)	Sigma-Aldrich, Germany
SuperScript® III Reverse Transcriptase	Invitrogen, Germany
SYBR® 555 Nucleic Acid Stain	Invitrogen, Germany
Taq DNA Polymerase	Qiagen, Germany
Tetracyclin	Genaxxon, Germany
Tris-base	Sigma-Aldrich, Germany
Tris-HCl	Sigma-Aldrich, Germany
Triton X 100	Gerbu, Germany
TRIzol® Reagent	Invitrogen, Germany
Tryptone/Peptone	Roth, Karlsruhe, Germany
Tween 20	Sigma-Aldrich, Germany
XL1-Blue MRF′ Supercompetent Cells	Stratagene, Agilent, USA
Yeast extract	Gerbu Biotechnik, Germany

Kits

Name	**Manufacturer**
Cloned AMV First-Strand cDNA Synthesis Kit	Invitrogen, Germany
Pierce HisPur cobalt spin columns	Thermo Scientific, USA
ProFoldin DNA Topoisomerase I Assay Kit	ProFoldin, USA
ProFoldin DNA Topoisomerase II (Gyrase) Kit	ProFoldin, USA
QIAGEN LongRange PCR Kit	Sigma-Aldrich, Germany
Qiagen QIAprep Spin Miniprep Kit	Qiagen, Germany
Qiagen QIAquick PCR Purification Kit	Qiagen, Germany
S30 T7 High-Yield Protein Expression System	Promega, USA

Buffers and media

Name	Composition
0.5 M EDTA pH 8.0	Dissolve 186.1 g $Na_2EDTA{\cdot}2H_2O$ in 800 mL dH_2O. adjust pH to 8.0 with NaOH (~20 g of NaOH pellets). EDTA will dissolve at pH 8.0. Adjust volume to 1 liter with dH_2O. Sterilize by autoclaving and store at room temperature
1 M Ethanolamine	60.5 ml ethanolamine in 1000 mL
1 M NaH_2PO_4 (monobasic)	138 g $NaH_2PO_4{\cdot}H_2O$ in sufficient H_2O to make a final volume of 1 L
1 M Tris-HCl pH 6–8	12.1 g Tris base in 100 mL H_2O, adjust pH with concentrated HCl
10 M NaOH	40 g NaOH in 100 ml H_2O
1M HEPES-NaOH pH 7.5	Dissolve 238.3 g HEPES in 1 L H_2O. Use NaOH pellets to adjust pH to 7.5. Start with about 5.5 g NaOH pellets
1M Na_2HPO_4 (dibasic)	142 g of Na_2HPO_4 in sufficient H_2O to make final volume 1 L
6× Gel-loading buffer	25 mg bromophenol blue and 4 g sucrose, make up volume to 10 mL with dH_2O, store at 4°C
6× Gel-loading buffer with Ficoll	25 mg bromophenol blue, 25 mg xylene cyanol FF and 1.5 g Ficoll (Type 400;Pharmacia) in 10 mL dH_2O, store at 4°C
Binding buffer 5×	0.1 M HEPES pH 8.0, 0.25M KCl, 25 mM DTT, 0.25 mM EDTA, 5 mM $MgCl_2$, 25% v/v glycerol
Ethidium bromide solution	0.5 µg/mL final concentration
Glutathione-free acid-reduced FORM*CRYST	Sigma-Aldrich Chemie GmbH, Germany
HEN buffer	250 mM HEPES pH 7.7, 1 mM EDTA, 0.1 mM neocuproine (2,9-dimethyl-1,10-phenanthroline). Store up to 1 year at room temperature wrapped in foil to protect from light
HMFM 10× (1 Liter)	36 mM K_2HPO_4, 13.2 mM KH_2PO_4, 0.4 mM $MgSO_4$, 1.7 mM $Na_3C_6H_5O_7$, 6.8 mM $(NH_4)_2SO_4$, 4.4% (v/v) glycerol: 0.76 g $MgSO_4 \times 7{\cdot}H_2O$, 4.5 g sodium citrate $2{\cdot}H_2O$, 9 g $(NH_4)_2SO_4$, 440 g glycerol, add water to 800 mL and autoclave. 18 g KH_2PO_4, 47 g K_2HPO_4, add water to 200 mL and autoclave. Mix both solution to make up final solution.

Laemmli buffer	30.1 g Tris base, 144.2 g glycine, 50 mL SDS (20%), add 1 L dH_2O
LB Agar	LB-medium + 1.5% (w/v) agar
LB Medium (1 Liter)	10 g Tryptone/Pepton, 5 g yeast extract, 10 g NaCl, pH 7.2
PBS 10× (1 Liter)	80 g NaCl, 2 g KCl, 26.8 g Na_2HPO_4, 2.4 g KH_2PO_4, pH 7.4
PBST 1× (1 Liter)	1× PBS/0.05% Tween-20
Spotting buffer (buffer E)	20 mM HEPES-KOH, 0.4mM EDTA, 1 mM DTT, 100 mM NaCl, 25% v/v glycerol, pH 8
SSC 20× (1 Liter)	3 M NaCl, 0.3 M sodium citrate buffer pH 7.0 (dissolve 175.3 g NaCl and 88.2 g sodium citrate in 800 mL dH_2O, adjust pH to 7.0 with 10 N NaOH) add dH_2O to 1000 mL
TBE 10× (1 Liter)	108 g Tris, 55 g boric acid, 40 mL 0.5M Na_2EDTA, pH 8
TBS 10× (1 Liter)	50 mM Tris, 150 mM NaCl with HCl, pH 7.5
TBST	1× TBS/0.05% Tween-20
Transfer buffer	150 mM glycine, 25 mM Tris base, 20% ethanol

Software

Name	Source
Bioconductor	Fred Hutchinson Cancer Research Center
ExPAsy	http://www.expasy.org/proteomics
GenePix Pro 6.0	Axon Instruments, Inc., Union City, USA
Primer-BLAST	http://www.ncbi.nlm.nih.gov/tools/primer-blast
PROTEIN CALCULATOR v3.3	http://www.scripps.edu/~cdputnam/protcalc.html
Protein molecular weight	http://www.bioinformatics.org/sms/prot_mw.html
Random DNA sequence generator	http://mkwak.org/oligorand
Recombinant protein solubility prediction	http://www.biotech.ou.edu/

Antibodies

Name	Manufacturer
AffiniPure Goat Anti-Human IgG, Cy3-conjugate antibody	Jackson ImmunoResearch, USA
Anti 6×His™ Alexa Fluor® 647 conjugates antibody	Qiagen, Germany
Anti-CLDN18, antibody	Sigma-Aldrich, Saint Louis, MO, USA
Anti-FLIP, N-Terminal antibody	Sigma-Aldrich, Saint Louis, MO, USA
Anti-FLIPα, C-Terminal antibody	Sigma-Aldrich, Saint Louis, MO, USA
c-Jun (1-169)-GST, soluble	Millipore, Temecula, CA
FOS (Human) recombinant protein (P01)	Abnova, Taipei, Taiwan
Goat Anti-Mouse IgG, Cy5 conjugate antibody	Jackson ImmunoResearch, USA
Goat Anti-Rabbit IgG, Cy3 conjugate antibody	Jackson ImmunoResearch, USA
Monoclonal Anti CDK2	Sigma-Aldrich, Saint Louis, MO, USA
Monoclonal Anti-BCL2L1 antibody	Sigma-Aldrich, Saint Louis, MO, USA
Monoclonal Anti-CFLAR antibody	Sigma-Aldrich, Saint Louis, MO, USA
Monoclonal Anti-TP53 antibody	Sigma-Aldrich, Saint Louis, MO, USA
Monoclonal Anti-V5, CY3 conjugate antibody	Sigma-Aldrich, Saint Louis, MO, USA

2.2 Methods

2.2.1 Constructed DNA Template for On-Chip PCR

DNA templates for on-chip PCR were constructed by standard PCR reaction. Full-length cDNA was constructed for transcription and translation by amplifying cDNA clones or first-strand cDNA products. The cDNA products were synthesized from isolated RNA from cell lines using TRIzol® Reagent. Basic forward and basic reverse primers were used for DNA template construction. Basic forward primers were designed with a ribosomal binding site (RBS), Kozak sequence, transcription promoter (T7 promoter), 6×His epitope tag sequence, and DNA-specific complementary sequences. Basic reverse primers contained V5 epitope sequence, stop codon, as well as complementary sequence. The constructed DNA templates were purified using QIAquick PCR Purification Kit in the standard protocol provided by the manufacturer. Afterwards, DNA template concentration was quantified using NanoDrop ND 1000 Spectrophotometer. The unique gene-specific primer sequences are listed in the Appendix (List of primers).

Basic primers	5′ → 3′
Forward_basic	TCCCGCGAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTAGCCACCATGGTAA GAAGGAGATATACCATGCATCATCATCATCATATGXXXXXXXXXXXXXX
Reverse_basic	CTGGAATTCGCCCTTTTATTACGTAGAATCGAGACCGAGGAGAGGGTTAGGGATAGGCyyyyyyyyyyyyyyyyy

Highlighted colors: *yellow* T7 promoter; *green* Kozak sequence; *blue* RBS.
Underlined sequences encode epitope tags 6×His (*red*) and V5 (*green*).

Solid-phase PCR primers were designed consisting of four main parts, from 5′ to 3′: 1) C6 amino linker, 2) 10T sequence as a spacer, 3) common sequences (blue), and 4) complementary sequences, respectively. Complementary sequences were 17–20 bp of downstream basic forward and reverse primers. A short primer pair was also synthesized for solution amplification. All the oligonucleotides were desalted and purified by HPLC.

Solid primers	5′ → 3′
Fow_comOCP	(NH_2-C6T)TTTTTTTTTTTTTCAGCAGTGTACAGATGGTGATGATGTCCCGCGAAATTAATACGACTC
Rev_comOCP	(NH_2-C6T)TTTTTTTTTTTTGCCACTTACCCCACAACCAGACGTTATTACGTAGAATCGAGACCGAGG

Short primers	5′ → 3′
Fow_Sol_OCP	TCCCGCGAAATTAATACGACTC
Rev_Sol_OCP	TTATTACGTAGAATCGAGACCGAGG

2.2.2 Primer Spotting, Slide Blocking, and Washing

Epoxy slides (NEXTERION® Slide E, standard dimensions of 75.6 mm × 25.0 mm × 1.0 mm) were chosen for this experiment as a solid support for covalent attachment of 5′-end amino-modified primers. Primers in spotting buffer were spotted using a NanoPlotter 2.0 (Gesim) noncontact spotting system. Primers were spotted in a concentration range from 0.1 to 20 µM and volumes of 0.3 (V1), 0.9 (V2), and 1.5 nL (V3). Primers were diluted in spotting buffer A: 150 mM phosphate buffer + 0.001% Tween-20, buffer B: 3× SSC, and buffer C: 3× SSC + 1.5 betaine. Printed slides were immobilized at room temperature at 70% humidity for 3 h, 24 h, and 72 h. Afterwards, slides can be stored within 8 weeks under inert condition in a desicator and protected from light at room temperature before applying to on-chip PCR. Prior to on-chip PCR, unreactive epoxy groups were deactivated using 150 mM ethanolamine, 100 mM Tris-HCl pH 9 for 1 h at 55°C. Alternatively, slides were blocked using 2% BSA in 1× PBST for 30 min at 30°C and then slides were rinsed in water and dried by pressurized air.

2.2.3 Solid-Phase PCR

Surface bound DNA for protein expression templates were generated by amplifying the gene of interest from the constructed DNA template. The solid-phase PCR cocktail mix contains all PCR components, but minus the primers. The adhesive gene frame (Thermo Scientific 1×1 cm) was fastened around the spotted primer, 25 µL of the PCR cocktail were placed inside the frame, and sealed with coverslip. The PCR cocktail contained, 1× LongRange Qiagen PCR buffer, 2.5 mM $MgCl_2$, 0.5 mM each dNTP, 5–30 ng/µL DNA template, and 1 unit of LongRange PCR enzyme mix per 25 µL reaction. The slides were then placed into a 16×16 dual-block thermocycler (PTC 200, MJ Research). Reactions were performed under the following conditions: 93°C for 3–5 min and 35 cycles (93°C for 45 sec, 50°C for 45 sec, and 68°C for 3–5 min) and 68°C for 10 min.

2.2.4 Probe Hybridization

After PCR, slides were rinsed with water to remove the PCR cocktail and pressured-air dried. A probe pair was always used to detect either immobilized primers or DNA from solid-phase PCR products. Fluorescently labeled (Cy3 and Cy5) probes were prepared with each concentration 1 µM in 3× SSC, 0.1% SDS, and hybridized onto glass slides

using lifter slip with following protocols: 95°C for 10 min and transfer immediately into water bath at 60°C for overnight. Slides were washed in 3× SSC, 0.1% SDS for 10 min continued with 3× SSC and 0.3× SSC for 10 min each. Slides were rinsed with water and pressured-air dried. The signal intensity of the labeled probe was detected using laser scanner.

2.2.5 Solution-Phase PCR

Slides from solid-phase amplification were rinsed in water for 5 min and dried. New adhesive gene frames were placed and filled in with PCR cocktail containing all components except the DNA template. PCR cocktail contained, 1× Qiagen PCR buffer, 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.2 µM of each short primer, and 1 unit of Taq polymerase. Slides were sealed with a flexible and removable coverslip and placed in the thermocycler. The PCR protocol followed that applied in solid-phase amplification. The PCR products were collected and resolved on 1–1.5% agarose gel by electrophoresis.

2.2.6 *In Situ* Cell-Free Expression of On-Chip PCR Products

For protein expression, slides from solid-phase PCR were washed using 1× SSC twice for 5 min each and dried. The multiple spotting technique (MIST) was applied for *in situ* cell-free expression [11]. Slides were arranged at the same position where primers were spotted. Firstly, 0.6 nL of 0.5 M betaine were spotted on top of bound PCR products and followed by 2.1 nL of S30 T7 High-Yield Protein Expression System (Promega). Slides were incubated for 1 h at 37°C and at 28°C overnight in humidified chamber. After incubation, slides were stored at –20°C for a minimum of 24 h before protein detection.

2.2.7 High-Throughput Culture of cDNA Library

3500 DNA samples of full-length human open reading frames (ORFs) human cDNA clones were obtained from The Gateway Full ORF Clones (German Cancer Research Center, Genomics & Proteomics Core Facilities (Appendix, List of genes). *Escherichia coli* with individual inserts of full-length human open reading frames (ORF) of individual cDNA libraries were subcultured as working libraries. Cells were inoculated

in LB medium containing 30 µg/mL kanamycin as selective antibiotic and 1× HMHM as cryoprotectant. Individual wells of 120 µL volumes in 96-well cell culture polystyrene microplates were inoculated with 1 µL of bacterial culture and grown at 37°C overnight. Bacterial cultures were stored at –80°C.

2.2.8 High-Throughput Colony PCR for High-Density Protein Microarray

An amount of 15 µL of overnight bacterial colony were diluted into 50 µL using sterile water, and heated at 85°C for 15 min. Five µL of the heat-killed bacteria was used as template for gene-specific amplification and the rest stored at −20°C for future use. The master mix for PCR reactions was prepared into the following final concentrations: 1× Qiagen PCR buffer, 100 µM of each dNTPs, 0.1 µM of each primer, 5 µL of pretreated bacterial colonies, and 2 units of DNA polymerase for every 25 µL reaction. Basic forward primers and basic reverse primers were used to construct complete full-length cDNA for *in vitro* transcription and translation. Thermal cycles were performed at 95°C for 5 min, continued with 95°C for 45 sec, 50°C for 45 sec, and 72°C for 3 min and repeated 35 times. Finally, the cycles were completed at 72°C for 10 min. The PCR products were validated by 1−1.5% agarose gel electrophoresis.

2.2.9 Small-Scale Standard PCR

The ORF of interest were isolated from the overnight culture in 2 mL LB medium containing 30 µg/µL kanamycin. These complete coding sequences lack stop codons. Plasmids carrying inserts of ORF were isolated from overnight cultures using QIAprep plasmid DNA purification (Qiagen, Germany) following the manufacturer protocol. The concentration of the isolated plasmid was determined using NanoDrop 1000 Spectrophotometer. Each isolated plasmid was aliquoted into working concentrations of 50 ng/µL and stored at –20°C. The master mix for PCR reactions was prepared into the following final concentrations: 1× Qiagen PCR buffer, 100 µM of each dNTP, 0.1 µM of each primer, 5 ng/µL DNA, and 2 units of Taq polymerase for every 25 µL reaction sample. Basic forward primers and basic reverse primers were used to construct complete full-length cDNA for *in vitro* transcription and translation. Thermal cycles as previously described above. The PCR products were validated by 1−1.5% agarose gel electrophoresis.

2.2.10 High-Density Protein Microarrays by MIST

The PCR products for protein microarray production were diluted 1:2 in nuclease-free water containing 0.5 M betaine, then transferred into 384-well spotting plates, followed by quick spindown in a centrifuge. The cell-free transcription and translation mixture was prepared as recommended by the manufacturer (S30 T7 High-Yield Protein Expression System, Promega). Ni-NTA- and epoxysilane-coated slides were used. The spotting area is defined for an area of 72 x 22 mm. PCR products were first spotted onto the slide surface; then the cell-free lysate was spotted on top of the PCR products in a second spotting run [11]. Slides were incubated for 1 h at 37°C and then at 28°C overnight in humidified chamber. After incubation, slides were stored at –20°C for a minimum of 24 h before protein detection.

2.2.11 Antibody Detection of Expressed Protein

The newly expressed proteins can be detected using antibodies that recognize N- and C-terminal tags as well as with protein-specific antibodies. Prior to antibody interactions, the slides were blocked using 2% BSA in 1× PBST for 30 min on orbital shaking at 30 rpm at room temperature. The fluorescent conjugated epitope tag antibody was diluted to 0.1 µg/mL in the blocking buffer, and for protein-specific antibodies to 0.5 µg/mL. This was incubated for 1 h in a rotary shaker at 30 rpm at room temperature. In the case of the epitope tag antibody, the antibody solution was removed and slides were washed; after 1 h of incubation, the antibody solution was replaced with a 1:10,000 dilution of fluorescent conjugated secondary antibody. Incubation was continued for another hour. After incubation, slides were washed three times with 1× PBST for 10 min each, rinsed with PBS, and dried. The expressed proteins were detected by laser scanner.

2.2.12 Data Acquisition and Analysis

Slides were scanned on a ScanArray 5000 or NimbleGen MS 200 Microarray Scanner. The instrument laser power (LP) and the photomultiplier tube (PMT) were adjusted to obtain visible signals for most spots, with only a small number of saturated signals. These scanner settings were maintained for all arrays in one experimental series. The recorded scanned image was converted into signal intensities by the spot

recognition software GenePix Pro. The false-color images were also generated with GenePix Pro. For on-chip PCR data analysis, the mean local background intensity was subtracted from the mean signal intensity for each spot to obtain background-corrected signal intensities. The background-corrected intensities for the red and green channels are designated *R* and *G*. The log converted intensities of the spots were calculated by $A = (\log_2 R + \log_2 G)/2$, a measure of the overall brightness of the spot [110]. The *A* value of each spot represents the amount of detected primer pairs and on-chip PCR products on the microarray.

The high-density protein microarrays were analyzed primarily using the software package LIMMA (Linear Models for Microarray Data) within an R-Bioconductor package [111,112]. Data loading via LIMMA for each spot consisted of the arithmetic mean value intensities of the pixels, R for N-terminal (red) and G for C-terminal detection (green), and the mean of the pixel values of the local background (Rb, Gb). In addition, quality information was added where available. The quality of one array from each batch of protein microarrays was verified using:

- Plot of C-terminus signal intensities of expressed proteins derived from purified and unpurified PCR products.
- Plot $\log_2$ N- and C-terminus signal intensities of expressed proteins on Ni-NTA and epoxy slides.
- Box plots of $\log_2$ signal intensities were made to examine the uniformity of signal intensities, signal intensity ratios, and coefficients of variation (CV) within (intra-) and across (inter-) arrays of each batch production.
- Density plots of $\log_2$ signaly intensities were examined to check whether the intensities from two-color detections evenly distribute.
- Scatter plots $\log_2$ C-terminus signal intensities and protein sizes were examined to obtain the distribution of expressed proteins based on their sizes.

3 RESULTS

3.1 On-Chip PCR and Personalized Protein Microarrays

To exploit the basic principle of solid-phase PCR for the production of DNA templates for cell-free on-chip expression, we optimized the solid-phase PCR method and generated arrays containing a DNA template encoding a full-length human protein fused with an N-terminal 6×His and C-terminal V5 tag, together with the required sequences.

3.1.1 Evaluation of Spotting Buffer and Primer Concentration

We optimized several parameters that are susceptible to affect covalent attachment of 5′-amino-modified oligonucleotides onto epoxy slides including primer concentration, spotting buffer, blocking, and washing conditions. The epoxy slide was selected because it has previously been reported to give stable covalent attachments of 5′-amino-modified oligonucleotides [101,102]. The attachment of covalently bound oligo-nucleotides was also relatively thermostable [12]. The spotting buffers described in the literature were selected because they readily allow covalent interactions between amino groups and opening epoxy groups [112]. The forward and reverse primers were diluted in three different spotting buffers and spotted using noncontact spotter in a concentration range from 0.1–20 µM and volumes of 0.3 (V1), 0.9 (V2), and 1.5 nL (V3) – spotting buffer A: 150 mM phosphate buffer +0.001% Tween-20, buffer B: 3× SSC, and buffer C: 3× SSC + 1.5 betaine. The 60mer primer contained 5′-C6 amino linker and 10Ts spacer to improve probe hybridization efficiency as well as primer extension during solid-phase amplification [98,107].

The spotted primers were detected after 24-hour immobilization in 60–70% humidity by hybridization using fluorescent complementary probe. Prior to probe hybridization, slides were blocked and washed to deactivate unreacted epoxy groups and remove unbound primers. Spotting buffer A and B showed perfect spot morphology, but spotting buffer C produced a doughnut-like spot shapes at concentrations lower than 5 µM (**Fig. 4a**).

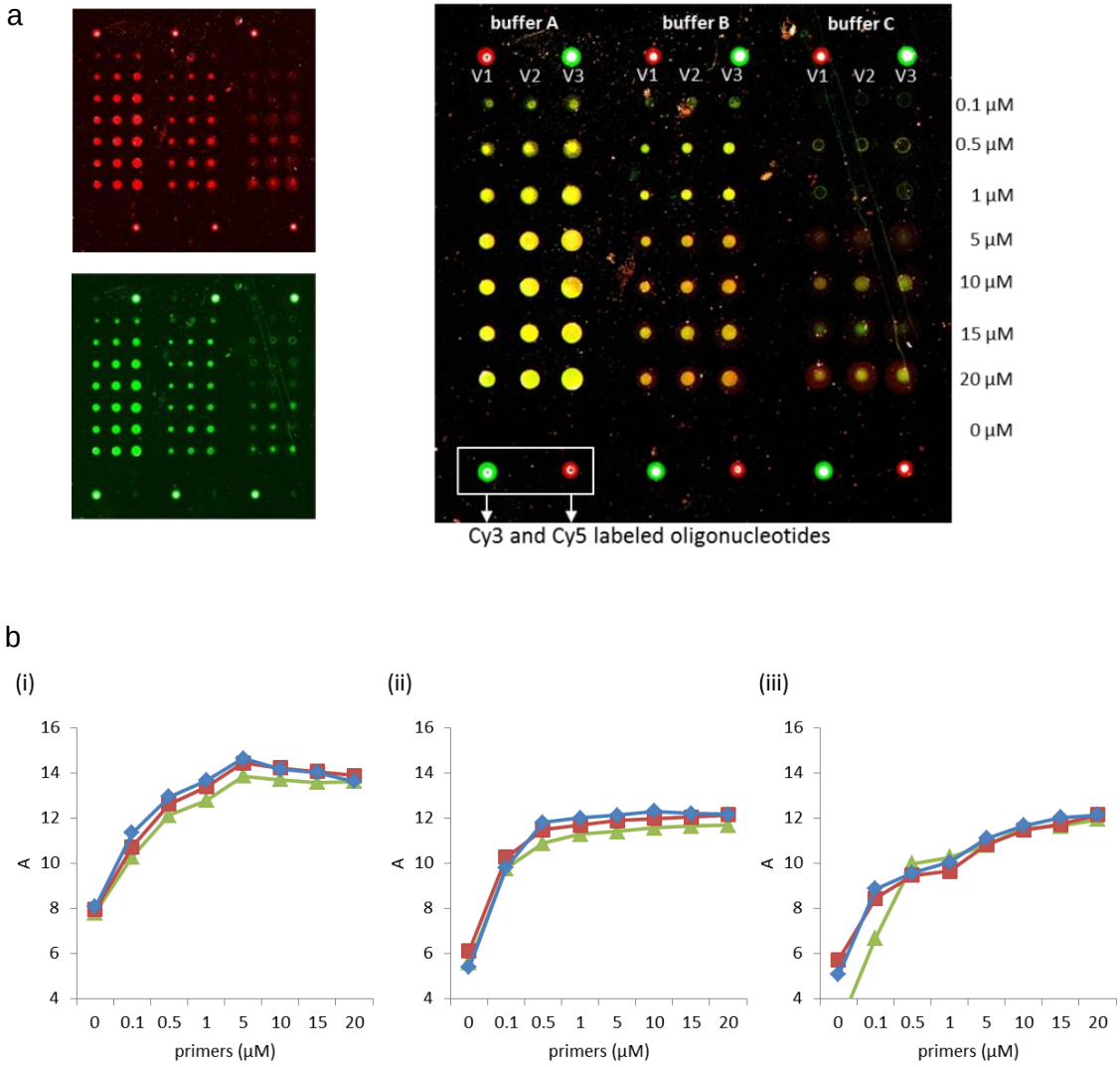


Figure 4. Dependence of detected solid-phase PCR products on spotting buffer and primer concentration. (**a**) False-color images from probe detection of elongation of forward primers (*red spots*), reverse primers (*green spots*), and overlay image from both strands (*yellow spots*). (**b**) Overall signal intensity of spotted primer in three different spotting buffers. Spotting buffer A: 150 mM phosphate buffer pH 8.5 +0.001% Tween-20, buffer B: 3× SSC, and buffer C: 3× SSC+1.5 betaine. Plots of average log red and green signal intensities (*A* value) for buffers A (i), B (ii), and C (iii) in three different spotting volumes are represented by *blue* (V1), *red* (V2), and *green* (V3) lines, respectively.

The regular shape of spots indicated that the chemical composition in these spotting buffers facilitates an homogeneous distribution of primer molecules in the spot. The spot size depends on spotting volume and concentration, particularly for spotting buffer A. Betaine solution in buffer SSC (buffer C) did not significantly affected the spot morphology of spotted primers on epoxy slides. However, it has been reported that combination of SSC and betaine can improve DNA immobilization efficieny on poly-L-lysine slides by preventing evaporation [103].

Spotting buffer A gave bigger spot diameters than the other two buffers. The presence of nonionic detergent in spotting buffer A increased the distribution of primer molecules in the spot. The signal intensities of detected primers spotted by means of buffer A were higher than for the other two buffers reaching saturated spotted primer concentrations at 5 µM, but not with the other buffers (**Fig. 4b**). Optimal concentrations must be determined for each type of molecule, because the spotting concentration is inversely dependent on the molar mass of molecules [102]. The chosen blocking method was suitable to reduce the background of the array. Any unreacted epoxy group has to be chemically deactivated to prevent nonspecific binding of the probe to the slide surface. A good spotting buffer should give spots of equal size, and with a nice round morphology, no change over time, and be resistant to evaporation during long print runs. The overall result from this work suggests the most suitable spotting buffer for 5′-amino-modified primers onto epoxy slides as being phosphate buffer pH 8.5, supplemented with 0.001% Tween-20 and at an optimum concentration of 5 µM.

3.1.1 Dependence of On-Chip PCR Products on Primer Immobilization Time

We have optimized the immobilization strategy to obtain thermostable covalent attachment of the 5′-end amino-modified oligonucleotides onto epoxy slides. Immobilization of molecules through specific chemical reactions requires sufficient time to complete the coupling chemistry [102]. The time to immobilize the primers was optimized. We also tested the optimum concentration of immobilized primer for optimum yield of solid-phase PCR product. Three different immobilization periods were chosen and primers were spotted onto the slide surface. The primer pairs were diluted in a concentration range from 0 to 20 µM using 150 mM phosphate buffer pH 8.5 + 0.001% Tween-20. Then, primers were spotted and incubated in humidity chamber (60–70%) for 3 h, 24 h, and 72 h. Afterwards the slides were blocked and washed. Three independent probe detections were carried out: first, by complementary fluorescence probes before on-chip PCR; then, the immobilized primers were checked after the PCR cycles; and finally, checked with DNA-specific probes. For every solid-phase PCR reaction we always used DNA target at concentration 1 ng/µL. **Figure 5** shows the scanned image and the plot of concentrations of the spotted primers *versus* log intensities from three different immobilization times.

The amount of attached primers and DNA products were represented as the log of fluorescence signal intensity (**A**) of both red- and green-labeled probes. Comparing line graphs for 3-h, 24-h, and 72-h immobilization, the amount of detected primers before being subjected to a thermocycling protocol increased with longer incubation times (*blue line*). This showed that inadequate time for coupling reactions reduced the amount of bound primers on the surface. The signal intensities from detected primers after thermal cycles (*red line*) displayed that unbound primers were washed away during on-chip PCR, particularly for very short immobilization times (**Fig. 5a**). Post-hybridization washing processes also contributed to detaching unstable molecules from the surface. This was a major reason for the low yield of detected PCR products (**Fig. 5a**, *green line*). Although, there was a moderate decrease of detected bound DNA (**Fig. 5b**), one full-day immobilization under humid conditions gave relatively stable attachment of 5′-end amino-modified oligonucleotides onto epoxy slides.

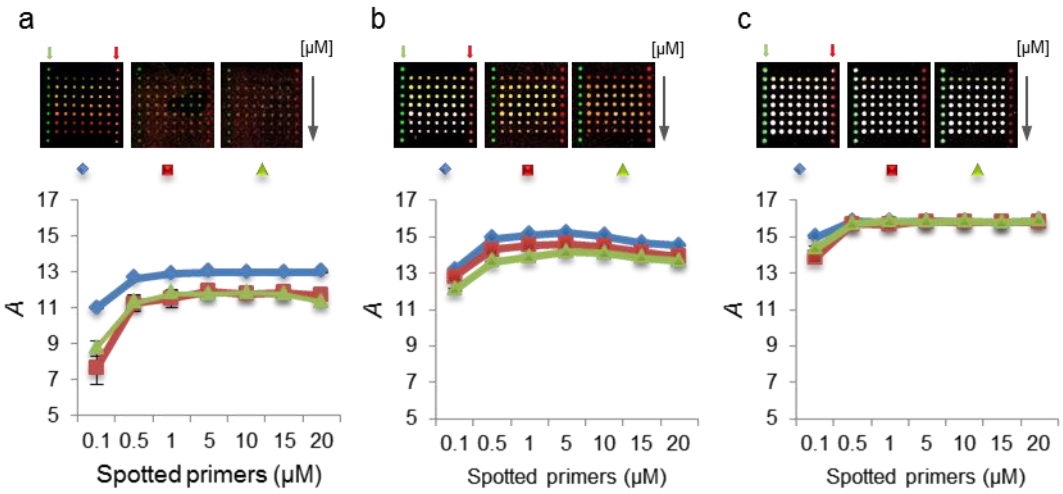


Figure 5. Dependence of on-chip PCR products on primer immobilization time. False-color image of fluorescent probe hybridization and graph of the spotted primers immobilized for 3 h (**a**), 24 h (**b**), and 72 h (**c**). Graph lines are the concentration of spotted primers plotted against the log of overall intensities, *A*. Three independent probe hybridizations of the immobilized primers were carried out by complementary fluorescence probes before on-chip PCR (blue line/♦) and after the PCR cycles (red line/ ■); then detections of the PCR products with DNA-specific probes (green line/▲). Positional controls are indicated by *red* and *green* arrows (Cy5- and Cy3 oligonucleotides).

The log signal intensity for detected DNA after PCR cycles in **Fig. 5c** showed that the amount of detected PCR products was higher for primers that were immobilized for 72 h than only for 3 h or 24 h. There were plateau lines for all three probe

hybridizations. With the exception of the lowest primer concentration, all spotted primers yielded saturated signals for all primer concentrations. However, the result from 24-h immobilization indicated that 5 µM was the optimum primer concentration that yielded optimum detected PCR product. This result also suggested that higher densities of immobilized primers will not increase the yield of DNA amplification on the solid surface.

For on-chip PCR purposes, a minimum of 24 h of incubation in humidity chamber was necessary for primer immobilization followed by dry storage for 24 h in order to achieve stable covalent attachments and reduce the loss of molecules during washing, blocking, and thermal cycles. Carefully adjusted immobilization conditions prevented the detachment of primers from the surface during thermal cycling.

3.1.2 Full-Length DNA by On-Chip PCR

We have generated arrays containing a DNA template encoding a full-length human protein fused with an N-terminal 6×His and C-terminal V5 tag, together with the required sequences. The immobilized primers can reach extension lengths above 5 kb and with an according polymerase system the entire genome of human mitochondrial DNA can be amplified, comparing well with the product length produced with standard PCR [13,113]. In this work, we generated different sizes of full-length DNA by on-chip PCR with a primer pair bound to the surface. The DNA-specific solid-phase primers were spotted based on optimum conditions that had been previously described. The primers were arrayed onto 30 spots for each PCR grid system. For DNA templates, human full-length ORFs were used. Four different lengths of DNA targets were selected as proof of principle for the efficiency of DNA polymerase activity in the solid-phase PCR.

The newly synthesized DNA on the surface was detected with DNA-specific probes using the hybridization protocol described in the Methods sections (**Fig. 6a**). Subsequently, this DNA was amplified using solution-phase primers and the PCR products were detected by 1–1.5% agarose gel electrophoresis (data not shown). Epoxy slides were used to obtain thermostable covalent attachments of 5′-end amino-modified primers and to ensure the free 3′ end is vacant for DNA polymerase activity. The 10Ts linker is added to primer sequences to improve hybridization efficiency [107,114].

Also, for the second run of PCR cycles, we added 20mer nucleotides referred to as common sequences. These oligonucleotides were noncomplementary to any given DNA template. As little as 7.5 fmole of primer pairs was sufficient to obtain an adequate density of immobilized oligonucleotide molecules for effective DNA polymerase activity.

We were able to extend the attached primers into full-length DNA up to 3 kb. The amplified DNA products bound to the slide surface were examined by means of Cy-5 and Cy3-labeled 25–30mer probe complements to the plus and minus strands of the corresponding DNA, then visualized by laser scanner. Alternatively, another PCR step was implemented to amplify on-chip DNA product using a pair of 18–20mer primers. Then, the newly synthesized DNA was generated as PCR products in solution phase (data not shown). Mean deviations from the mean signal intensity of detected solid-phase PCR products for upper and lower strands ranged between 10–20% (**Fig. 6b**). The amount of detected PCR products was independent of the length of amplified DNA.

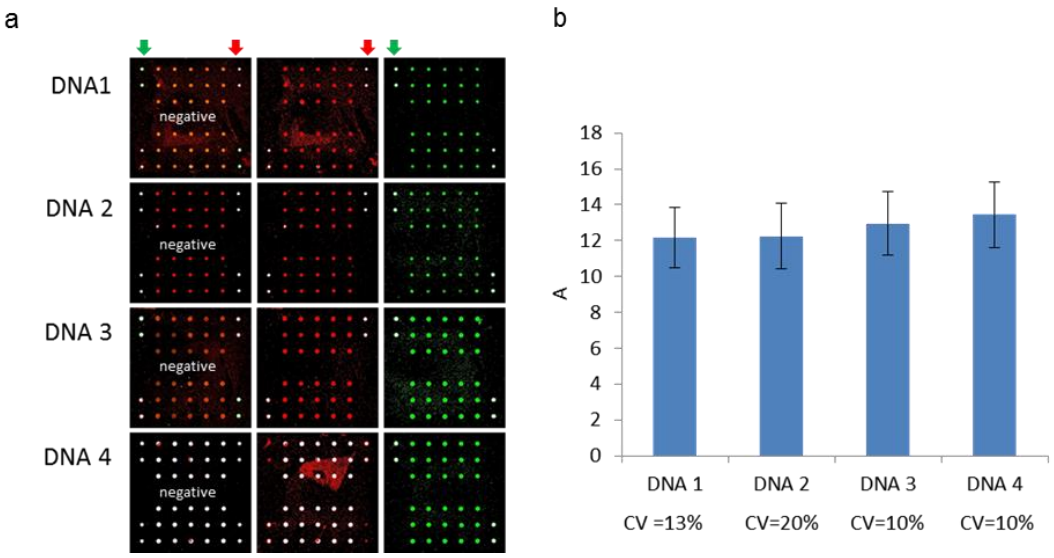


Figure 6. Covalently bound DNA detected by a fluorescent-labeled complementary probe. (**a**) Overlay red and green channels from both DNA strands are shown in the *left vertical panel*, and individual probe detection of single-strand DNA in the *center* and *right panel*. Four human full-length ORFs were used as DNA templates, DNA 1-4: *DIABLO* (717 bp), *IDH3B* (1155 bp), *CHMP2A* (666 bp), and *CREB3L1* (1557 bp). Positional controls are indicated by *red* (Cy5-oligonucleotides) and *green* (Cy3-oligonucleotides) *arrows*. Negatives were spots of spotting buffer only. (**b**) Bar chart showing the mean of the log *red* and *green* intensities; CV was calculated for 30 spots in one array.

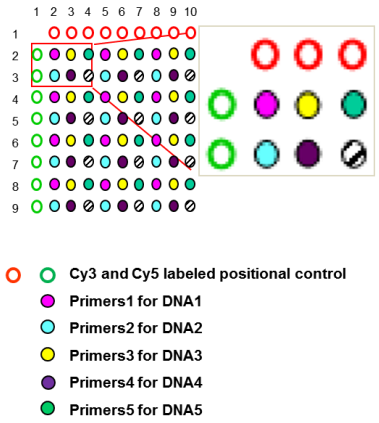
Newly synthesized DNA was detected by DNA-specific probes that recognize both strands of DNA. Each probe pair hybridized at a position of approx. two-thirds along the DNA. An alternative detection method on second-run PCR used short primers. These primers basically have the same DNA complementary sequences that were previously used for solid-phase PCR. Since, the second polymerase reaction was completed without using new DNA targets, this ensured that the PCR products came from the amplification of bound DNA. By using two detection strategies, the full-length DNA can be confirmed. This result indicates solid-phase PCR can copy full-length DNA into bound primers. The solid-phase PCR method generates covalently bound DNA arrays in specific orientation.

3.1.3 Multiplex Reaction Using On-Chip PCR

As proof of principle, on-chip PCR was used to amplify double-stranded DNA in a multiplex reaction. Highly specific primer pairs were designed with similar melting temperatures. These primers were tested for their specificity using the Primer-BLAST program [115]. Multiplex solid-phase PCR has been used for a rapid detection of bacteria from clinical samples based on the recognition of the 23S ribosomal DNA sequences [116]. For this work, five pairs of solid-phase primers were immobilized together in one on-chip PCR system. We immobilized five primer pairs in several replicates. **Figure 7a** shows our spotting layout with five primer pairs on the array outlining results of primer-pair verification using standard PCR. In parallel, these primers were also tested by normal PCR (**Fig. 7b**, right panel). Then, solid-phase PCR was performed; only one DNA target was added for each PCR mixture. The DNA targets (DNA1–DNA5) were 300–500 bp fragments of *ILR1RN*, *AK1*, *CHMP2A*, *IDH3B*, and *CREB3L1*, respectively. After polymerase reaction, slides were detected with DNA-specific probes for the validation of primer specificity in multiplex solid-phase PCR (**Fig. 7b**, left panel).

As can be seen in the scanned image for DNA 2 and DNA 3, the solid-phase PCR reaction only amplified its own specific DNA target. Primers 2 and 3 specifically elongated the given PCR template. This result was consistent with the DNA products generated using standard PCR (**Fig. 7a, b**; right panel).

a



	primer1	primer2	primer3	primer4	primer5
DNA1	+	-	-	+	-
DNA2	-	+	-	-	-
DNA3	-	-	+	-	-
DNA4	-	-	-	+	+
DNA5	+	-	-	+	+

b

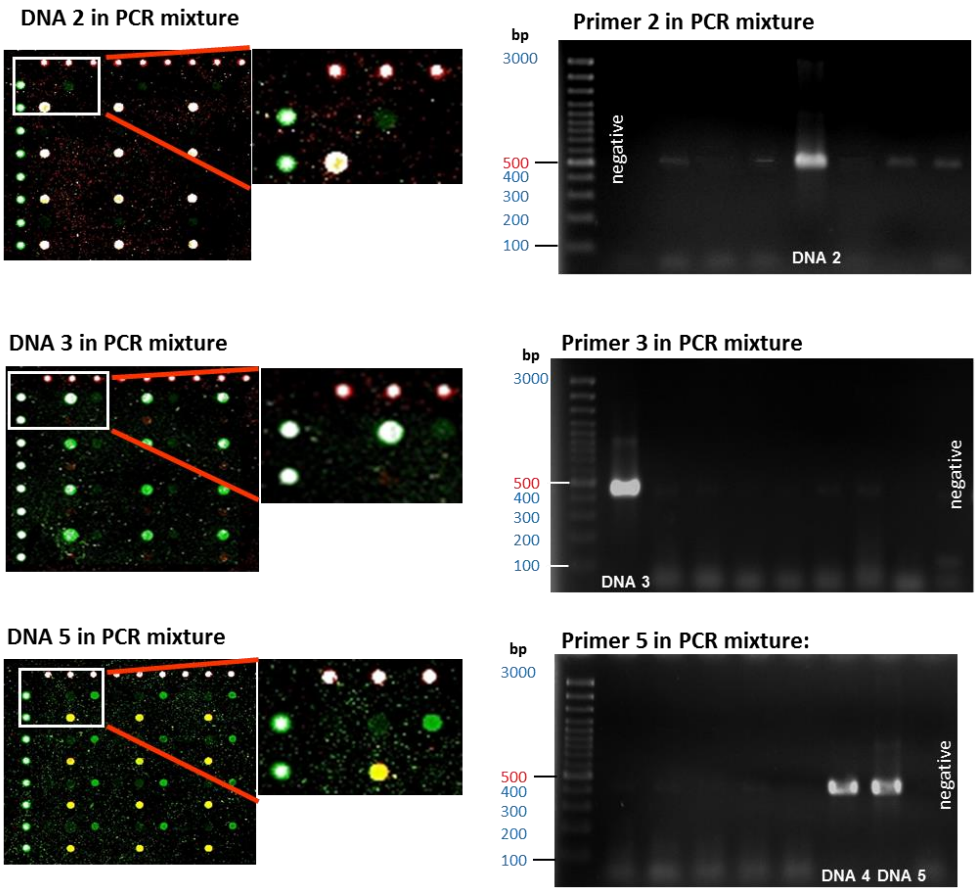


Figure 7. Multiplex solid-phase PCR. (**a**) Spotting layout of solid-phase primers (*left panel*) and results of primer verification using standard PCR (*right panel*). (**b**) False-color images of DNA-specific probes (*left panel*) and gel images of tested primers (*right panel*).

For all given DNA in the PCR mixture, the electrophoresis gel for primers 2 and 3 only showed one band for its own DNA target (**Fig. 7b**, right panel). Though the specificity of the tested primers have been validated *in silico*, still it showed a cross-

reaction to nontarget DNA during PCR reaction. Three of the five primer pairs showed unspecific results either in solid-phase PCR or standard PCR. The solid-phase PCR system can amplify multitargets simultaneously which requires the use of highly specific primers to their DNA target.

In this work, we showed that highly specific primers amplify only their own DNA target without significant cross-reactions to other DNA templates. However, less stringent primers may exert cross-amplification to nontarget DNA.

3.1.4 *In Situ* Protein Synthesis from Covalently Bound of cDNA

We have utilized a novel strategy of immobilizing DNA templates for *in vitro* transcription and translation directly onto a chip (slide surface). The solid phase of specific primers has been tested to produce specific cDNAs from isolated mRNAs by using reverse transcription PCR (RT-PCR) [117]. The solid-phase primer pairs were used to generate attached full-length DNA on the chip surface, and these served as templates for *in situ* transcription and translation to produce protein microarrays. The DNA-encoded proteins were constructed with the necessary sequence for transcription and translation as described in the Methods section. Additionally, sequences for 6×His and V5 epitope were inserted to encode fusion tags at N- and C-termini. Then, the construct was used as template in solid-phase PCR to produce covalently bound DNA on the surface. The amplified DNA was validated by hybridization of a complementary probe. The full length of PCR products was also validated by another PCR step detecting the dissolved products by gel electrophoresis. Afterwards, protein expression was carried out by spotting these newly synthesized cDNAs with cell-free expression lysate and incubating at 37°C and at RT (28°C) in the humidity chamber for transcription and translation.

The expressed proteins were detected by fluorescent conjugate antibodies that recognize fusion epitope tags, as described in the Methods section. Then slides were scanned to visualize the expressed proteins (**Fig. 8a**). In this array, DNA encoding CREB3L1, TP53, MAPK1, and DIABLO was synthesized on the slide surface and subsequently used for *in situ* cell-free expression. The expression of bound DNA was observed by the fluorescent signal intensity from Cy3- and Cy5-labeled antibodies

(*red* and *green* spots). The negative controls of solid-phase primers did not emit fluorescent signals from any of the antibodies.

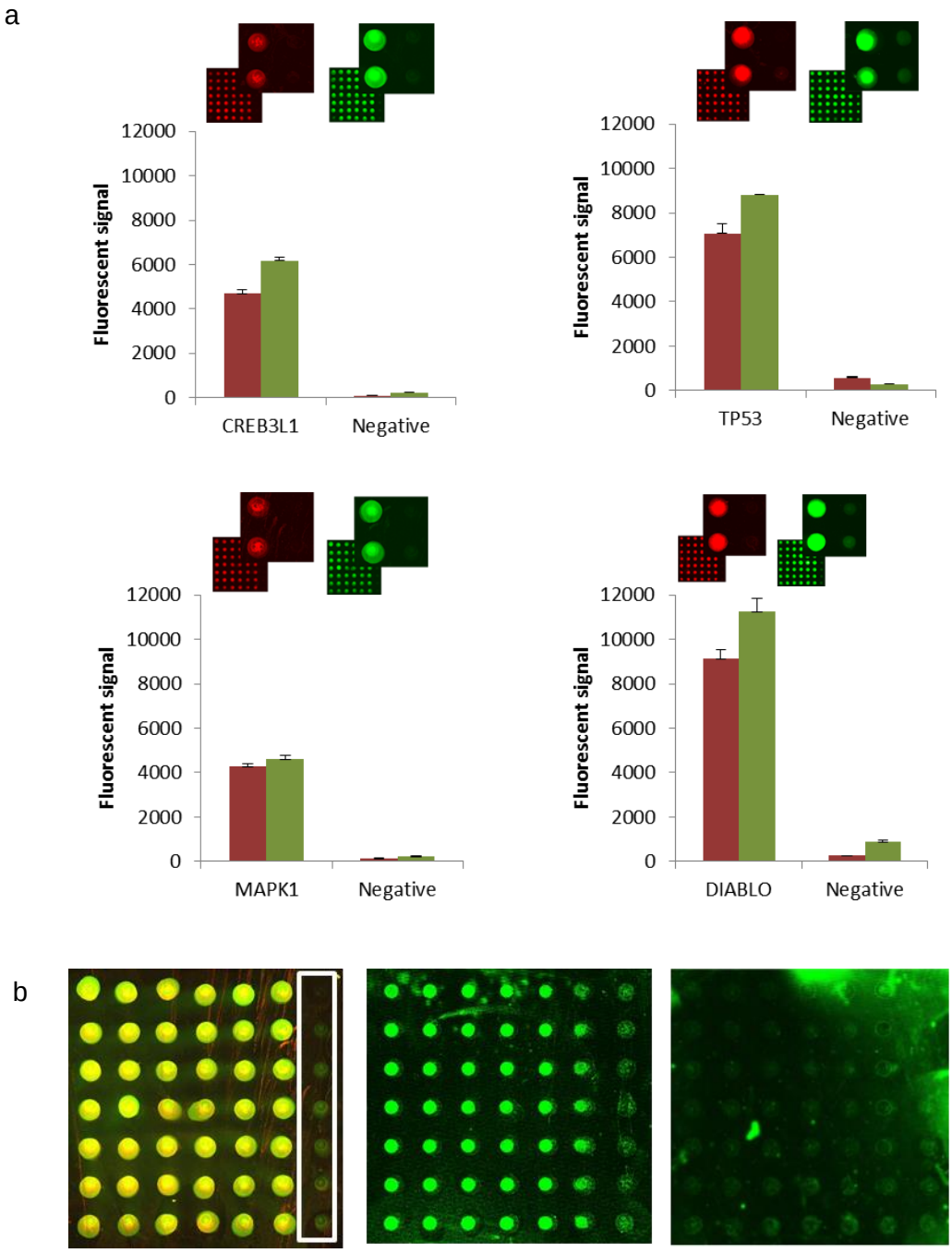


Figure 8. Protein expression from bound DNA templates. (**a**) False-color images and plot signal intensities from epitope detection using fluorescent conjugate antibodies against fusion epitope in N- and C-termini of expressed CREB3L1, TP53, MAPK1, and DIABLO. Individual scanned images from N- and C-termini are shown by *red* and *green* arrays, respectively. The *inset* in each panel shows spots of expressed protein and negative control. Background-corrected signal intensities of each expressed protein were plotted together with the respective negative spots. (**b**) False-color image of terminal epitope detection of MAPK1 (*left panel*), anti-MAPK1 antibody and fluorophore-conjugated secondary antibody (*middle panel*), and secondary antibody only (*right panel*). *White box* indicates spots of negative controls.

The arrays were also detected using protein-specific antibodies. MAPK1 was expressed *in situ* from bound cDNA as template. One array was used to verify newly synthesized protein with epitope tag antibody. The interaction of anti-MAPK1 antibody and protein microarray was shown using fluorescent conjugate secondary antibodies (**Fig. 8b**). The expressed proteins were also incubated with BSA and secondary antibody only as negative control.

The expressed protein on the chip was accessible across both terminus sites. C-terminus detection showed that the translation was complete, yielding a full-length expressed protein. The protein–antibody interactions showed that the *in situ* expressed proteins obtained directly from covalently bound double-strand cDNA at the 5′ end were able to interact with the specific protein antibodies.

3.1.5 Personalized *In Situ* Protein Microarrays

The solid-phase PCR system can be used to perform multiplex reactions (**Fig. 7**). In order to check whether the solid-phase PCR system with multiple primer pairs would specifically produce multiple DNA targets, we generated various sequences of DNA bound onto the slide surface. To use the optimized construction of the immobilized primers that had worked successfully under our solid-phase PCR conditions, we did not change the key elements of the immobilized primers, i.e., spacer, linker, and common sequences. With an additional unique 5′ end for each DNA template, the immobilized primers would specifically recognize their own targets in the PCR mixture. These unique oligonucleotide pairs were designed with a very high stringency, but since they were added as part of primers, it was easier in setting a uniform melting temperature (T_m).

Then, with this strategy, multiple primer arrays could be transformed into multiple bound DNA templates at a specific orientation. The DNA arrays can be stored dry for a reasonable time prior to expressing the proteins. Then, in this particular instance we expressed seven coding PCR products as proof of principle to generate personalized protein arrays. Using anti-6×His and anti-V5 antibodies which recognize the fused N- and C-terminal tags, respectively, in each coding sequence, we were able to confirm full-length expressed proteins (**Fig. 9a**). The detected protein signals for all tested coding sequences were above the signals from negative control spots (**Fig. 9b**). For testing this concept of personalized protein microarrays we used seven bound DNA

templates encoding full-length proteins: DIABLO, CFLAR, TP53, AKT1, MAPK1, BCL2L1, CDK2. The success rate of expression varied between these proteins. For negative spots yielded signals in the background range. Two of the expressed proteins were also confirmed by means of protein-specific antibodies (**Fig. 9c, d**). The antibodies were confirmed for selectivity using high-density protein microarrays (**Fig. 16**).

Each DNA requires a highly specific melting temperature (T_m) at which a primer targets the original sequence of DNA. Additional flanking primer sequences are routinely used in next generation sequencing, usually known as adapters [118]. With uniform T_m, primer annealing can be easily controlled and nonspecific PCR products can be avoided.

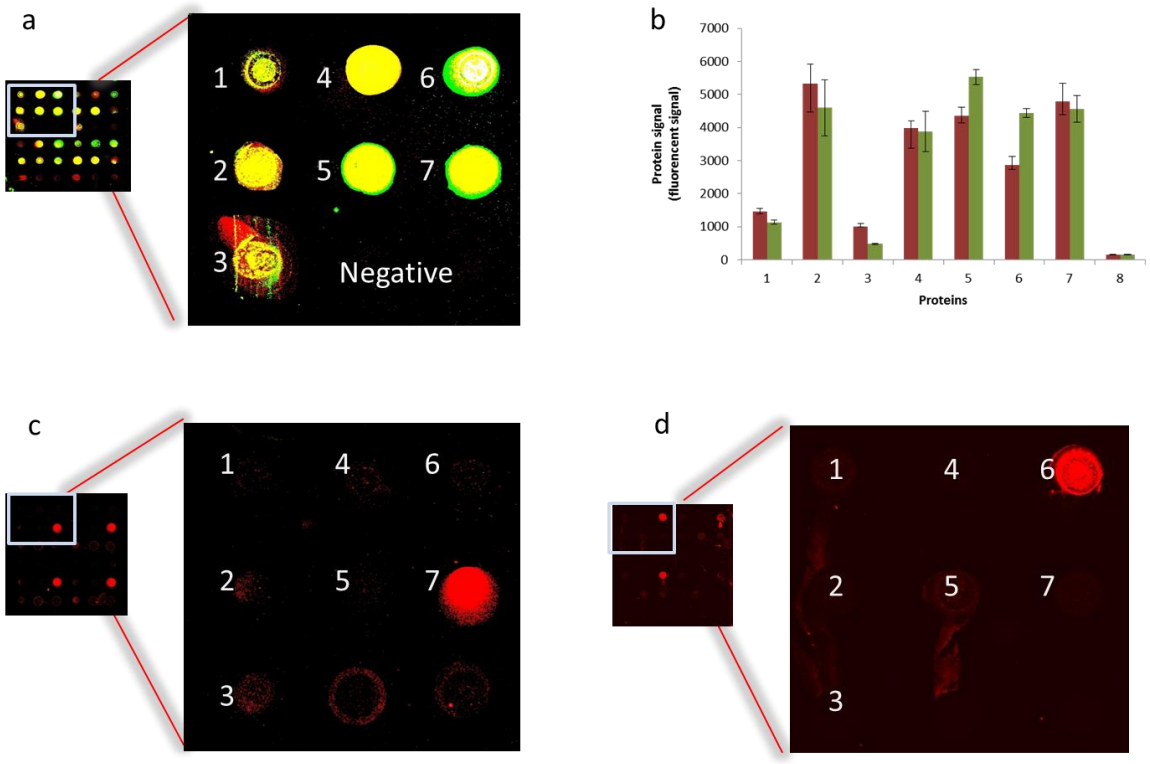


Figure 9. False-color images of *in situ* protein expression. (**a**) The superimposed images are obtained from epitope detection using fluorescent conjugate antibodies against fusion epitopes in the N- and C-termini of the expressed proteins. *Negative* indicates position of nontranscribed DNA as negative control. (**b**) Individual signals from N- and C-termini depicted in a bar chart. (**c**, **d**) False-color images of CDK2-specific antibodies and BCL2L1 detection.

3.2 High-Density Protein Microarrays

We have optimized the conditions of MIST for larger-scale production of protein microarrays [11]. The array quality was examined to determine and evaluate the technical usability and variations induced by PCR, protein expression, and probing. Furthermore, protein microarrays produced by the MIST method were tested for protein–protein and protein–RNA interactions, as well as for antibody and binder selection.

3.2.1 Purified or Unpurified PCR Products

Unpurified PCR products were used for our microarray production. The decision was supported by the level of expressed proteins from unpurified and purified PCR products from the same proteins produced under similar conditions. One volume of PCR product (25 µL) was purified using ethanol-based DNA purification methods and eluted with 25 µL of water. The PCR purification process removed primers, nucleotides, enzymes, salts, and other impurities after polymerase reaction [119]. Purified and unpurified PCR products were diluted in a similar fashion with spotting buffer. Afterwards, these PCR products were spotted to generate protein microarrays. The newly synthesized proteins were compared based on signal intensities from Cy3-conjugated V5 antibody recognizing the V5 epitope at the C-termini of the proteins (**Fig. 10**).

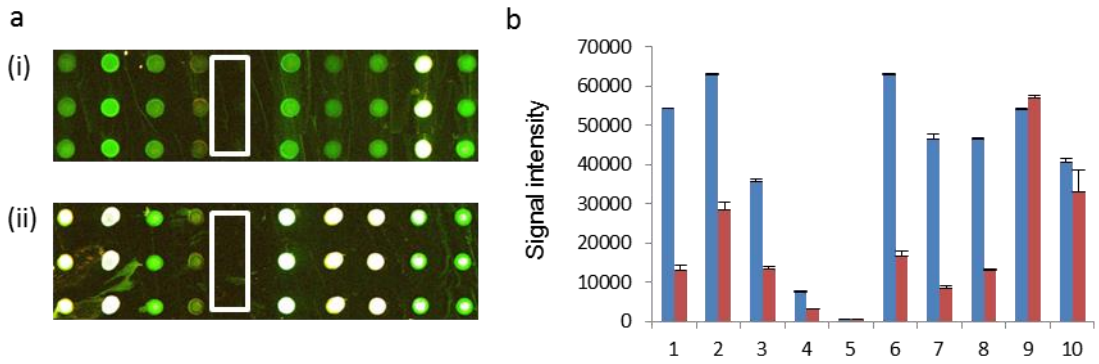


Figure 10. Purified and unpurified PCR products used for protein expression on microarrays. (**a**) Protein microarrays of nine proteins and one negative controls, produced from purified (i) and unpurified (ii) PCR products. False-color images detect expressed proteins using Cy3-conjugated anti-V5 antibody, recognizing V5-fused tags at the C-termini. (**b**) Signal intensities of detected proteins on the array from unpurified (*blue*) and purified (*red*) expression templates.

Table 1. Proteins and purified PCR product concentrations

Position	1	2	3	4	5	6	7	8	9	10
Name	BCL2A1	VSNL1	BDNF	GEM	negative	ACSBG2	NGEF	OGT	PHKA2	ICAM3
Lenght (bp)	525	573	741	888	N/A	1998	2130	3108	3705	3561
ng/µL	64.0	50.3	51.2	70.6	1.05	70.7	72.6	50.3	95.3	96.7

For all tested genes, unpurified PCR products gave higher detection signal intensities, indicating that the amount of expressed protein was higher than if purified. Only two out of nine tested proteins gave similar expression levels of unpurified and purified samples, which might have been caused by higher spot concentrations (**Table 1**). Interestingly, the negative controls of both purified and unpurified DNA showed similar signal intensities. The primers themselves were 120mer and 60mer, respectively, that may produce primer dimers of 120–170mer length. The chosen buffer in the purification system only removed DNA fragments smaller than 100 bp. This indicates that PCR purification will not be able to remove the excess primers and its dimers. However, some components usually find in the PCR mixture can inhibit cell-free expression, such as glycerol at >1% and $MgCl_2$ at >3 mM. As a rule of thumb for cell-free expression, higher DNA concentrations will produce higher yields of expressed proteins, while not taking protein properties into account.

3.2.2 Specific or Nonspecific Orientation of Immobilized Proteins

Two different slide surfaces were used to compare the spot morphology of expressed proteins on the array, the amount of detected proteins, and stability of protein immobilization. We have expressed 94 human proteins and 12 GFPs on nickel-nitrilotriacetic acid (Ni-NTA) and epoxy slides. The full-length human proteins and GFPs contained N-terminal 6×His and C-terminal V5 tag. These two different types of coated glass slides immobilized proteins by specific and nonspecific orientation, respectively. The addition of His-tag either to the C- or N-termini have commonly been used in recombinant protein technology, because they generally do not interfere with the structure and function of proteins [80]. The yield of expressed human proteins and GFPs were compared for both slides (**Fig. 11a**). Ni-NTA slides produced smaller spot diameters than epoxy slides. Low levels of expressed protein gave doughnut-like shaped spots on the epoxy slides.

The raw signal detection of expressed GFP was used to compare the background signal intensities of the slides (**Fig. 11b**). After background subtraction, the mean signal intensity of detected GFPs on epoxy slides was higher than detected GFPs on Ni-NTA slides, as the latter emitted an almost three times higher background signal. All detected signal intensities were included in the analysis, both expressed proteins and negative controls (**Fig. 11c**). all experiments were performed in triplicates.

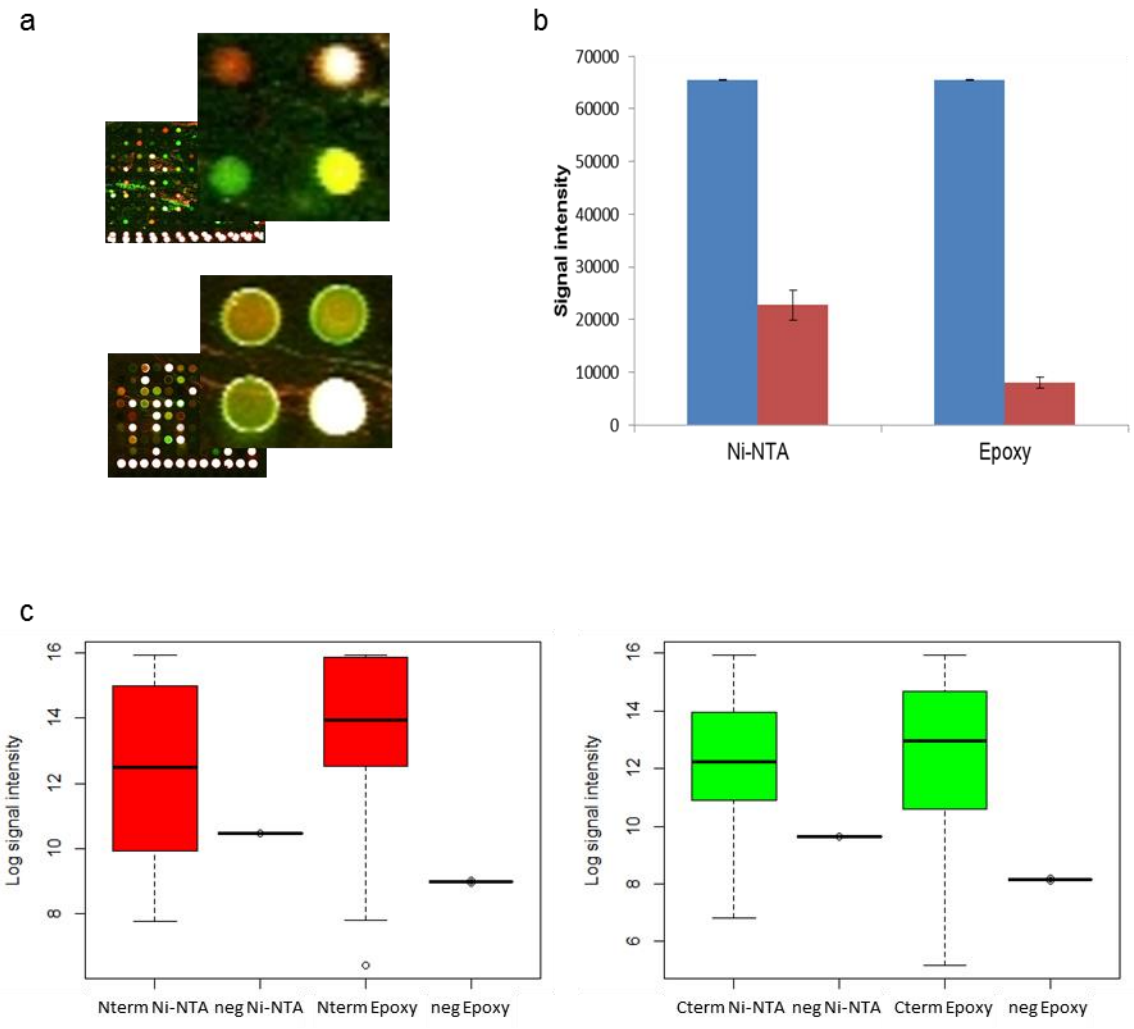


Figure 11. Comparison of protein microarrays on Ni-NTA and epoxy slides. (**a**) Spot morphology of expressed proteins of Ni-TNA slide (*upper panel*) and epoxy slide (*lower panel*). (**b**) Comparison of raw signal intensities of detected GFP *(blue)* and respective back-grounds *(red).* (**c**) Comparison of background-corrected signal intensities of 90 proteins expressed on Ni-NTA and epoxy slides. Detection of N-termini (*red*) and C-termini (*green*) from two slides were compared, as well as negative controls.

A comparison of the expression of 94 human proteins on Ni-NTA and epoxy-coated surfaces based on signal intensities from N-termini shows that there is no significant difference in the the mean expression on both surfaces (*p*-value = 0.063). Similarly, C-termini detection of V5 epitope tags gave similar results on both slides, although there was a slight increase in the number of detected proteins on the epoxy surface but no significant difference (*p*-value = 0.66). This result indicates that both specific and nonspecific orientation of protein binding to the surface allows for similar interactions with other molecules, in this case antibody. The same result have also been reported for direct detection of expressed GFP on APTES and Ni-NTA slide [11].

3.2.3 Consistency of Cell-Free Mixture Spotting

One of the key strategies of MIST is to transfer cell-free expression mix directly on top of each DNA spot, and to keep the amount of applied cell-free mix as small as possible. To test the variability of spotting process and the consistency of protein expression on the array, we expressed 384 GFPs. Intra-array variability was predicted by computing the coefficient variation (CV) from the detected signal intensities of expressed GFP. Each GFP had three intra- and three interarray replicates. An ANOVA model was used to determine the cell-free mixture spotting variation associated with protein expression that are resulted from repeated spotting processes. The variance estimate of the random errors was used to determine the consistency of cell-free mixture spotting across the arrays.

The expressed GFP was detected using anti-6×His antibody recognizing C-terminal fusion His-tags. The local background was subtracted from the raw mean signal intensity to obtain the background-corrected signal intensity. One single uptake by the spotter pin (volume capacity, 10 µL) was sufficient to transfer cell-free mixture to every spot in all three blocks of the arrays (**Fig. 12**). We found the expressed proteins CV to be 12%, 11%, and 13%, for sub-array 1, 2, and 3 respectively. Variation in the number and levels of the expressed protein on high-density GFP microarrays using MIST was statistically insignificant at the 1% confidence interval (*p*-value = 0.014).

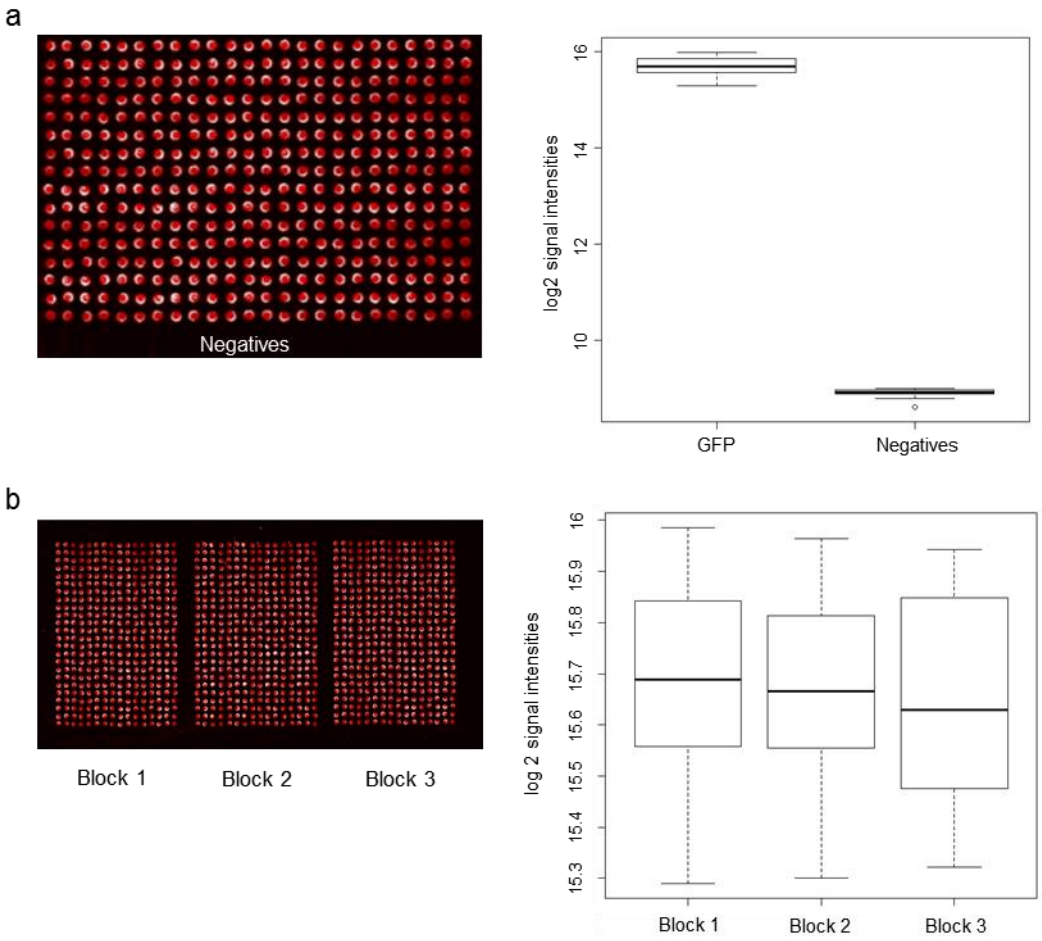


Figure 12. Expression of GFP microarrays. (**a**) False-color image of one block of GFP protein microarrays (*left*) and plot of mean signal intensity with negative controls. (**b**) Three intrareplicates of the expressed GFP and plot of their signal intensities.

3.2.4 High-Density Protein Microarrays

About 3500 full-length human cDNAs were arrayed and expressed *in situ* using a cell-free expression system directly on the slide surface (**Fig. 13**). These proteins were arrayed onto two different slides, one slide containing 1656 spots and the other 2070 spots. Currently, one slide can hold up to 2500 spots. To fabricate protein microarrays, a non contact-printing robot was used to deliver nanoliter volumes of PCR products and cell-free expression mixture to the slides, yielding spots about 480 to 500 µm in diameter (225 spots per square centimeter). The protein microarrays were produced from unpurified colony PCR products and expressed using multiple spotting techniques. Regulatory sequences for transcription and translation were constructed as overhangs of PCR primers, as well as the sequence to encode 6×His and V5 epitope tag. PCR

reactions were 384-well PCR format, and 25% of randomly chosen PCR products from each plate were confirmed in 1.5% gel agarose electrophoresis.

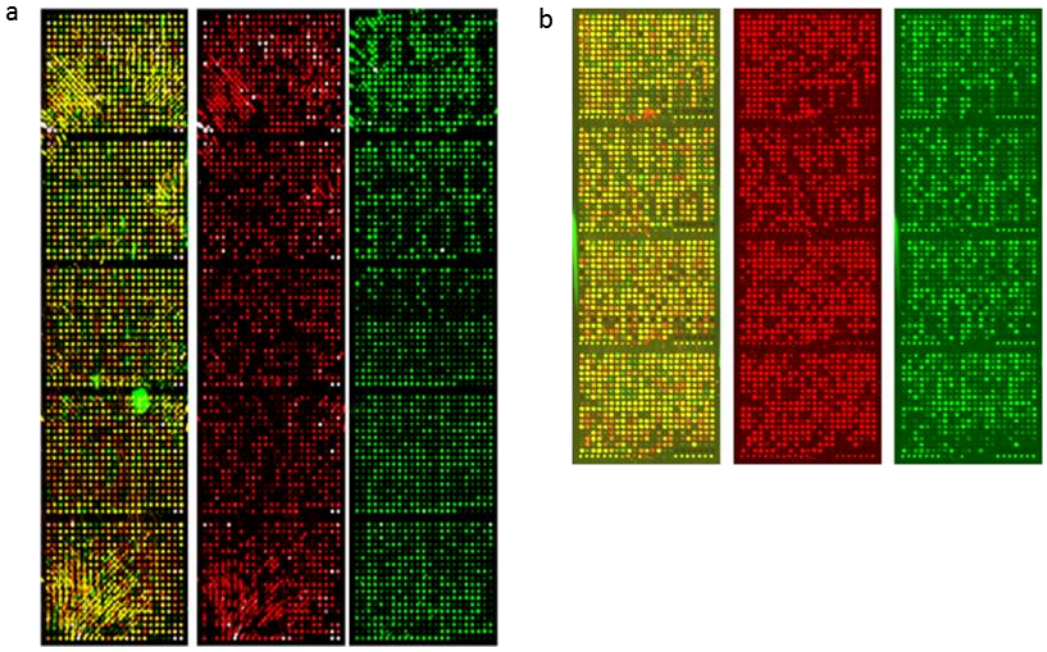


Figure 13. Protein microarray of human proteome. Three and a half thousand human proteins were expressed and arrayed directly onto two seperate glass slides. Newly expressed proteins were confirmed using antibody detection against N- and C- terminal fusion epitopes.

The PCR products supplemented with 0.5 M betaine were spotted onto epoxy slides in volumes of 0.6 nL followed by spotting of 3.6 nL of cell-free expression lysate on top of every DNA spot. Immediately after finishing the spotting of the cell lysate, the slides were placed in the humidity chamber and incubated for one hour at 37°C and removed to room temperature (28°C) for overnight incubation. On the next day, slides were stored in a slide box at –20°C minimum for three days before being used for the functional assay. Each expressed protein was detected by antibody recognizing fusion epitopes in the N- and C- termini. The antibodies were Alexa Fluor 647 conjugated-anti 6×His antibody and Cy3-conjugated anti V5 antibody. The quality of production was checked by statistical analysis by randomly choosing slides from one batch of the production (**Fig. 14** and **Fig 15**).

The success of protein expression was determined from the ratio of the foreground signal intensity from which the background was subtracted. The proteins

were considered to be expressed if the signal intensity was 1.5 times higher than the negative controls which contained primer dimer, spotting buffer, and cell-free expression lysate (**Fig. 14a**). Using this threshold, the success rate of full-length protein expression on the microarray was shown to be more than 90%. Some of the proteins could not be detected at their C-terminal tag since they did not pass the threshold limit which was defined as three times higher than the mean signal intensity of negative controls.

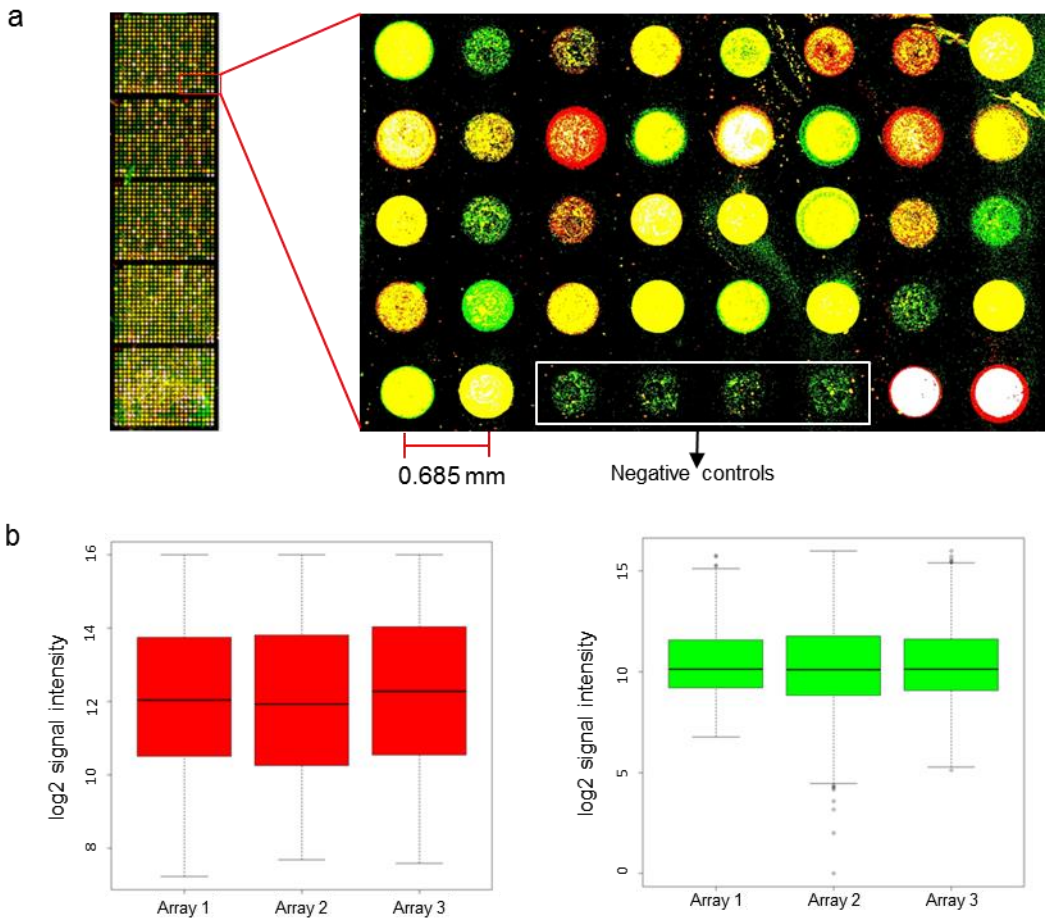


Figure 14. Analysis of high-density protein microarrays. (**a**) High-density array representing 2030 different human proteins. (**b**) Log2 signal intensities of triplicate arrays, *red box plot* for N-terminal detection and *green box plot* for C-terminal detection.

We compared the signal intensities of the N- and C-termini detections from interarray triplicates (**Fig. 14b**) using one-way ANOVA. There was no significant variantions between arrays of the same batch of production (*p*-values of 0.0026 and 0.023, respectively). More than 99% of expressed proteins can be replicated in all

array that were created at single time batch production. ANOVA analysis of N-termini detection indicates that most of expressed proteins were successful to be expressed among intearray replicates, but not with C-termini detection results. This result might be caused by incomplete translation or degradation of the nascent peptide

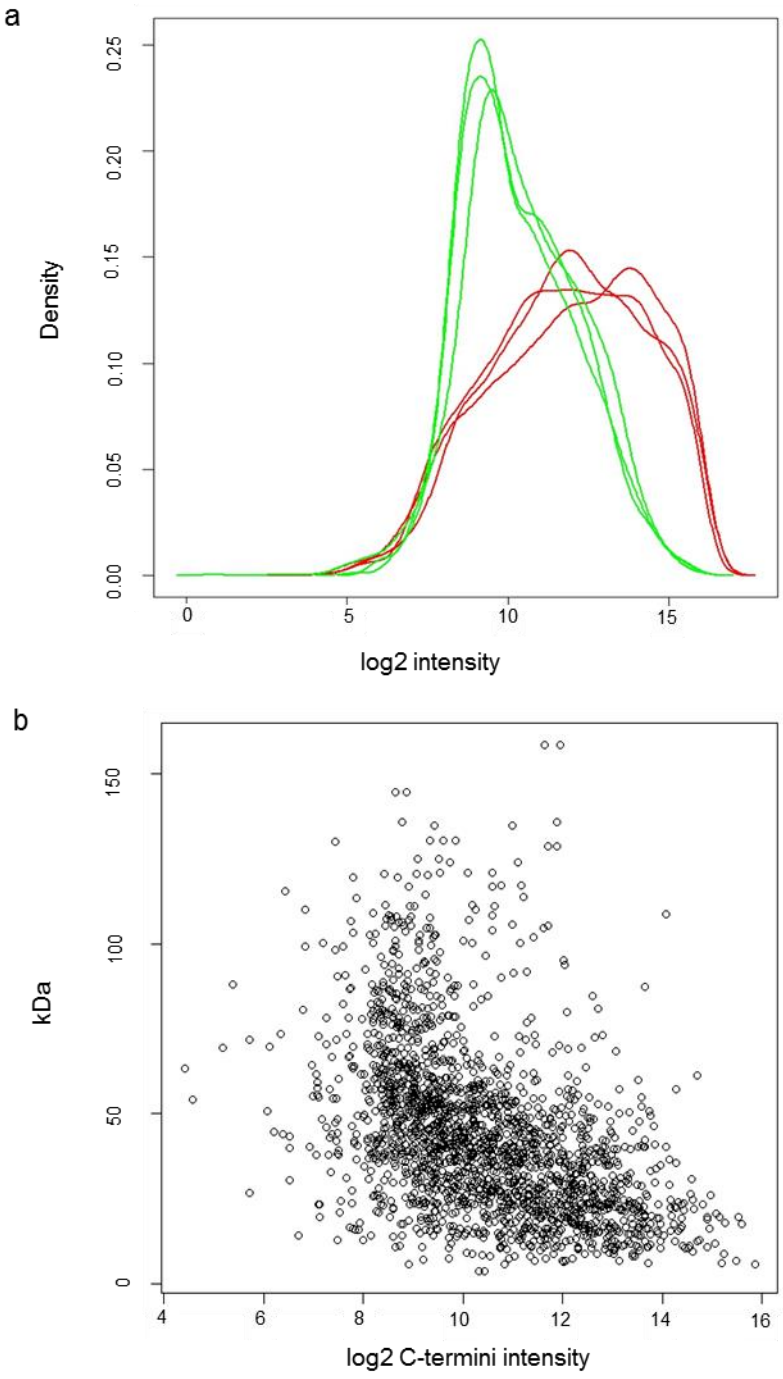


Figure 15. Analysis of high-density protein microarrays. (**a**) Density plot of of N- and C-terminus detection. (**b**) Plot of full-length detected proteins against the protein size.

These results show that the difference in the levels of detected proteins between different arrays that were produced in one single batch is statistically significant. More than 99% of expressed proteins can be replicated in all array that were created at single time batch production. ANOVA analysis of N-termini detection indicates that most of expressed proteins were successful to be expressed among intearray replicates, but not with C-termini detection results. This result might be caused by incomplete translation or degraded of newly synthesize protein.

The signal detection of N- and C-terminus shows dissimilar distributions (**Fig. 15a**). The C-teminal tag were often detected at lower value of intensities than N-terminal tag. Pearson's correlation between the protein size and intensities of full-length expressed proteins showed a moderate negative correlation ($r = -0.46$). A trend was shown in the plot of C-terminus intensity vs. protein size, bigger proteins had lower levels of expression (**Fig. 15b**).

3.2.5 The autofluorescence of Green Fluorescent Protein (GFP)

GFP was expressed and the emission of its autofluorescence was measured by direct scanning at 488 nm. We also compared this with purified recombinant GFP expressed *in vivo*. Purified recombinant GFP was also used to generate standard curves that were then used to semi-quantitatively determine the amount of expressed GFP on the microarray (**Fig. 16).** The estimated amount of expressed GFP on each spot was equal to 248 µg/mL of purified GFP. Up to the concentration of 500 ng/nL of spotted GFP did not show a saturation of mean signal intensity as well as with the expressed GFP. The fluorescence of GFP can be detected when it reaches the native stage condition [120]. We have shown that expressed protein on the microarray can retain its native function, in this case autofluorescence of GFP.

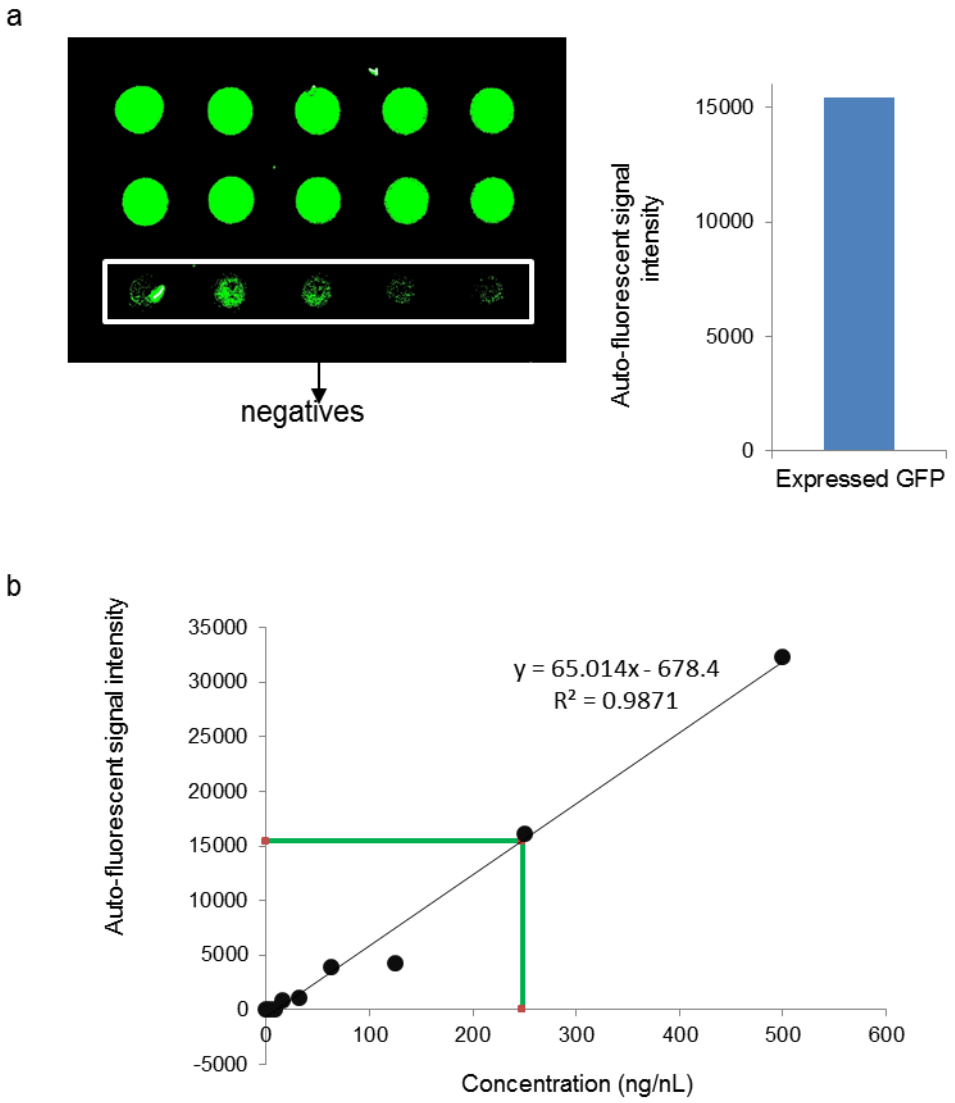


Figure 16. Autofluorescence of expressed GFP and standard GFP. (**a**) Array and signal intensity of expressed GFP, generated by detection of its autofluorescent emission. (**b**) Plot of autofluorescent emission of serial dilution standard GFP.

3.2.6 Protein Microarrays for Binder and Antibody Selection

Monoclonal antibodies were tested using high-density human protein microarrays. In addition, we also confirmed a pair of protein isoforms by antibodies and interaction of DARPin binders with their target proteins. To analyze antibody specificity we screened six tested antibodies from our collaboration partners within the Affinomics and six commercially available antibodies against human protein microarrays. These antibodies were validated using Western blot against purified recombinant antigen. The synthesized antigens for producing antibodies were included on the array as a control

for antibody detection. The expressed proteins were confirmed using both terminus detections. Antibody incubation was carried out on the array after being blocked by 2% BSA in 1× PBST buffer for 1 hour. The antibody interactions were confirmed using Cy3-conjugated goat anti-human antibody. Primary antibodies were diluted to 1:1,000 and secondary antibodies to 1:5,000 dilution. After washing 4 times with 1× PBST, every 5 min, slides were dried and scanned.

Two of six tested antibodies on the protein microarrays demonstrated their binding with positive controls and expressed target proteins. These antibodies were created targeting IL6 (Tol40.8-H8) and MAPK1 (SH511-IIB1). However, only SH511-IIB1 showed mono-specificity to its target proteins.

The antibody against MAPK1 (SH511-IIB1) showed strong interactions to its target peptide and target proteins on the array. Although, protein of MAP kinase 11 showed very weak interaction with this antibody, but with another proteins it did not give any interaction. These findings confirmed the previous result that did not show cross-interactions between antibody SH511-IIB1 and other mitogen-activating kinase proteins.

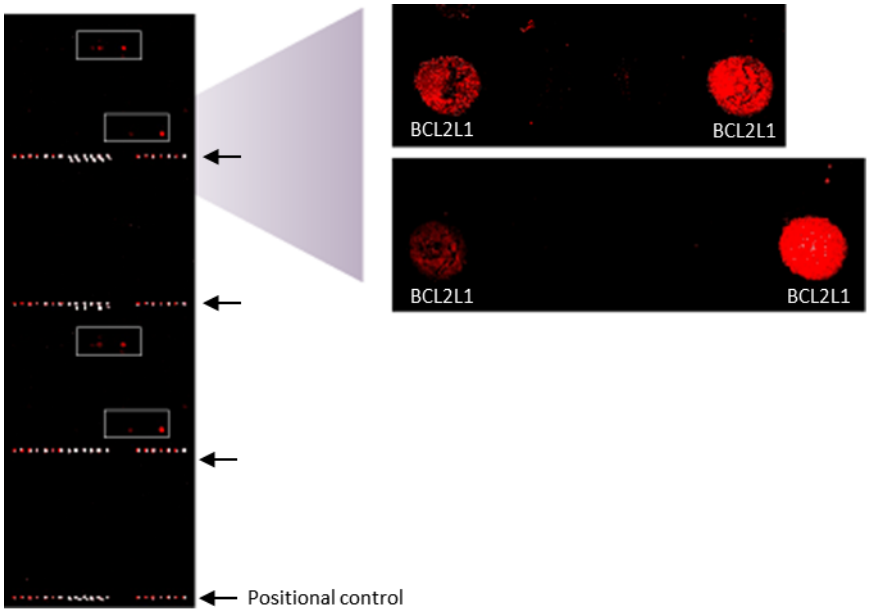


Figure 17. Antibody specificity by human protein microarray. The protein microarray was incubated with monoclonal anti-BCL2L1 and detected with Cy5-conjugated anti-mouse antibody. Positional controls are Cy5-conjugated peptides.

Antibody against IL6 (TOL40.8-H8) gave many nonspecific interactions to most spotted control recombinant proteins as well as to proteins on the array. The antibody showed strong binding to its corresponding purified recombinant proteins that were used as positive controls of the interaction. TOL40.8-H8 showed strong binding to the expressed IL6 but also with many other proteins, proving poor selection of the antibody for its interaction partner.

Three out of six commercially available antibodies gave either insensitive or non-specific for detection of target proteins on the arrays. For insensitive antibody increasing the antibody concentration and incubation time did not increase the detection signal. The monoclonal monoclonal anti-TP53, anti-CDK2 and anti-BCL2L1 antibodies exhibited strong signal intensities only in expressed of correspond proteins and displayed no cross reactivity with non-target proteins. The overall results of antibody screening was summarized in **Table 2**.

Table 2. Selectivity and cross-reactivity of the selected antibodies

Antibody name	Protein target	Positive control	Target expressed proteins	Nontarget expressed proteins
ToL40.8-H8	IL6	+++	+++	yes
SH511-IIB1	MAPK1	+++	+++	no
SH473-IIIF7	PRKAR1A	– – –	– – –	yes
SH474-IIC12	PRKAR1B	– – –	– – –	yes
SH475-IID11	PRKAR2A	– – –	– – –	yes
SH476-IIIE4	PRKAR2B	– – –	– – –	yes
Monoclonal anti-TP53 antibody	TP53 (a.a 94−202)	N/A	+++	no
Monoclonal anti-CDK2 antibody	CDK2 (a.a. 211-298)	N/A	+++	no
Monoclonal anti-CFLAR antibody	CFLAR (a.a 178−227)	N/A	---	yes
Monoclonal anti-BCL2L1 antibody	BCL2L1(full length)	N/A	+++	no
Anti-FLIP, N-terminal antibody	CFLAR (a.a 2-18)	N/A	– – –	yes
Antibody 9A	CYP1A1 and isoform2	+++	+++	no
Antibody 9B	CYP1A1 isoform2	+++	+++	no

– – – = no/weak signal intensity; +++ = strong signal intensity

Protein microarrays were used to validate the specific recognition of antibodies against protein isoforms. The antibodies have been tested before by Western blotting. One pair of the tested new antibodies can distinguish between two protein isoforms resulting from splicing variants of the coding sequence (**Fig. 18a**). The binding activity was confirmed on protein microarrays.

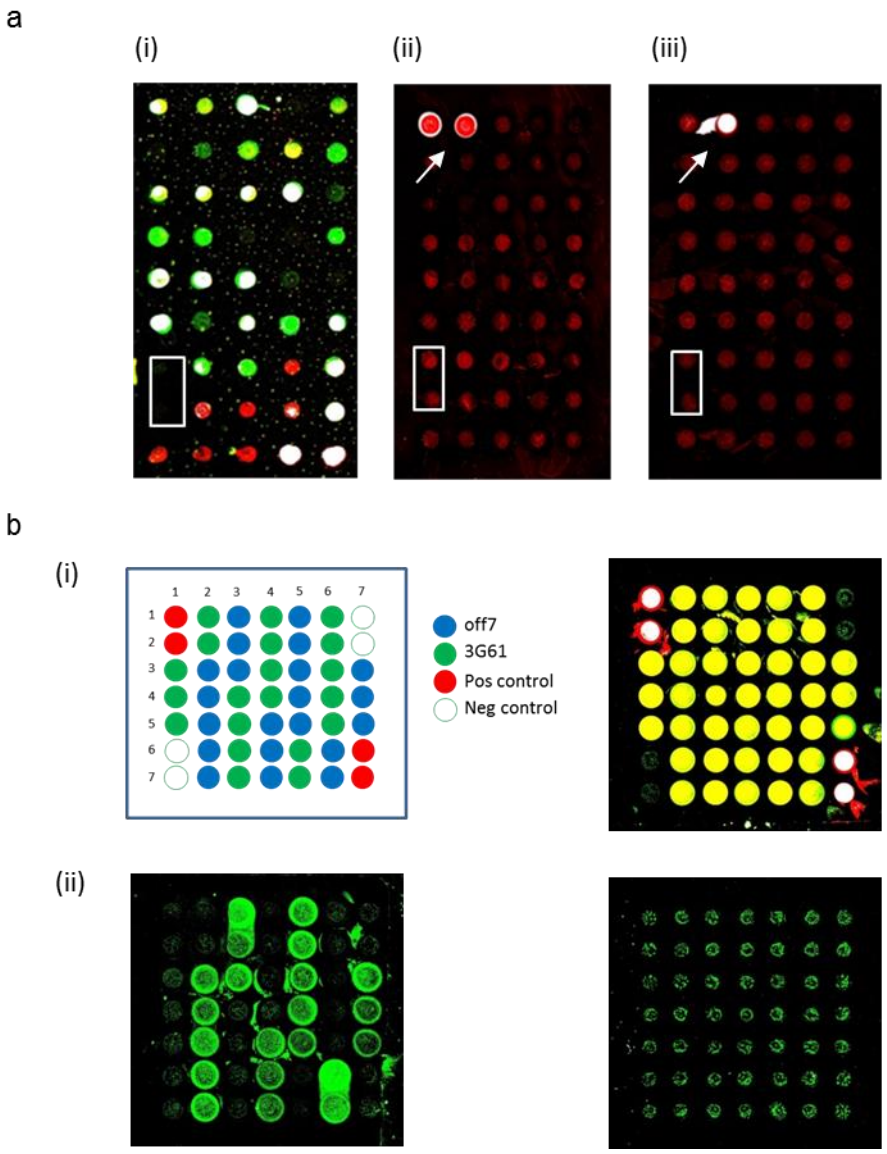


Figure 18. Protein isoform-specific antibody and binder selection. (**a**) Twenty pairs of protein isoforms were expressed on microararys. The arrays were confirmed with terminal tag detection (i) and used for antibody selection. One antibody recognizes both isoforms (ii) and another is specific only to one isoform, the spots are indicated by white arrow (iii). White box indicates the negative controls (**b**) DARPin arrays for binder selection. (i) Two DARPin proteins were expressed and arrayed using MIST technology. (ii) Specific interaction between biotinylated MBP with DARPin_off7 (*left*) and false interaction of sfGFP with all proteins on the array.

The ankyrin repeat protein DARPin, as a high-affinity binder, was also tested to screen the specificity of the DARPin–protein construct. Two constructs of DARPin proteins were amplified and arrayed into protein microarrays with many replicates in one block using this technology. The *in situ* expressed proteins were confirmed using antibody against theit terminal epitopes. The arrays were incubated with two different binder proteins separately. Each protein was specific only to one DARPin protein. Maltose binding protein (MBP) showed a strong binding signal to DARP_off7. On the other hand, sfGFP did not show any specific interactions to DARPin_3G61 (**Fig. 18b**).

3.2.7 Detection of PPI Using Protein Microarrays

JUN-FOS interactions served as another model to be tested on protein microarrays. JUN and FOS proteins were expressed and arranged randomly on the array together with other proteins. The newly expressed proteins were confirmed by N- and C-terminal fusion epitope detection. All proteins on the arrays were fused with 6×His and V5 epitope. The slides were blocked using phosphate saline buffer containing 2% BSA and 20% glycerol. For the binding partner, JUN protein was fused with 6×His and c-Myc epitope and expressed *in vitro* using the same expression system and purified with Ni-NTA agarose beads. The purified protein was checked using SDS-PAGE. The binding proteins were diluted in binding buffer and incubated to block the slide. After one hour of incubation, the slides were washed three times with binding buffer to remove unbound protein. **Figure 18** shows the interactions of proteins were confirmed with Cy3-conjugated anti-c-Myc antibody. Controls were carried out with BSA and anti-cMyc antibody.

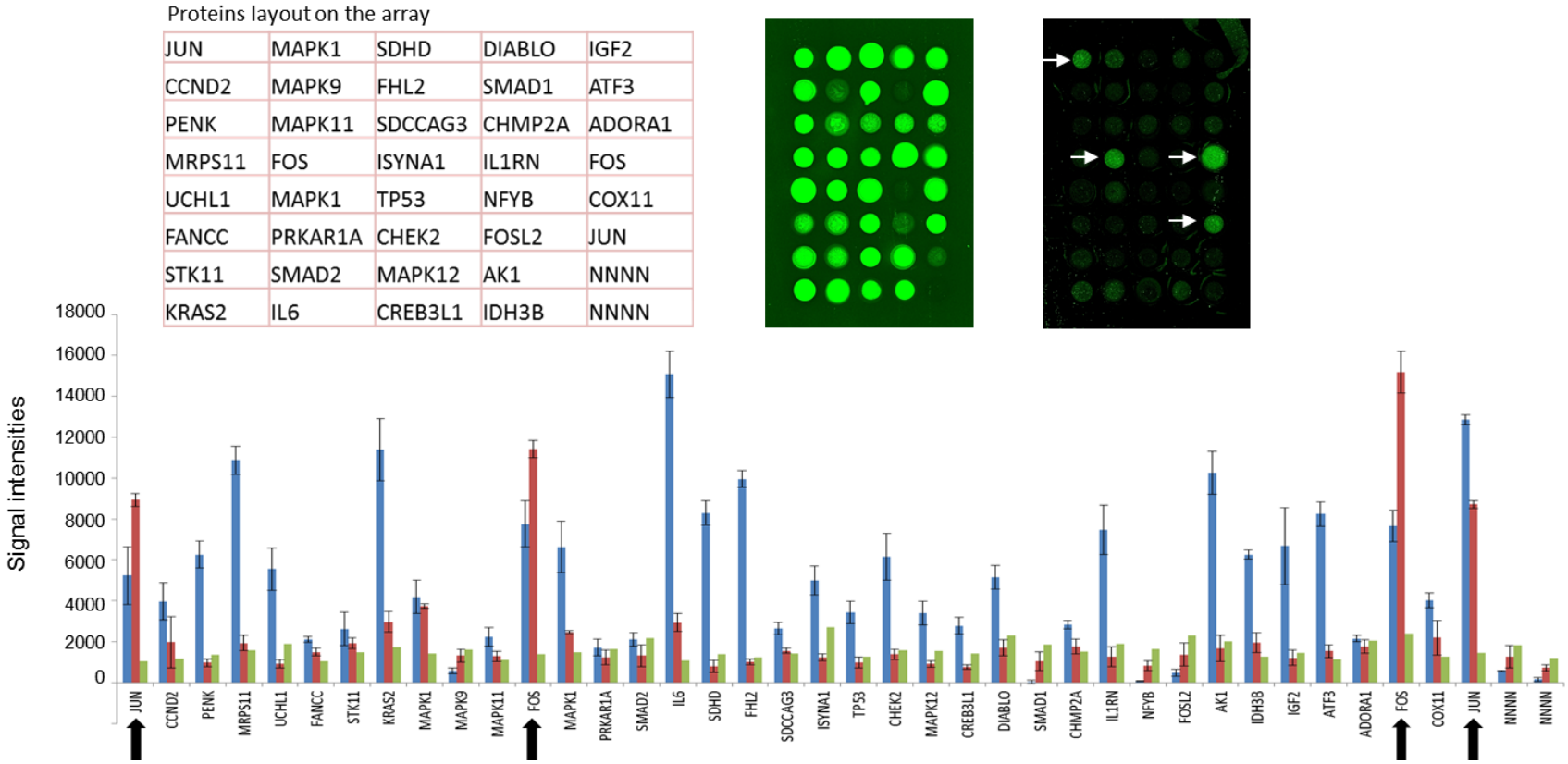


Figure 19. PPI detection by protein microarrays. C-terminus confirmation of expressed target protein on the array using Cy3-conjugated anti-V5 antibody (*upper-left array*). The interactions of JUN-proteins were detected using anti-c-Myc antibody (*upper-right array)*. The white and black arrows indicate the detection and the signal intensity of PPI, respectively. Plot of signal intensities of expressed proteins *(blue)*, detection of JUN-proteins interactions *(red)* and control of detection *(green)*.

4 DISCUSSION

4.1 Generation of Surface-Bound DNA for Cell-Free Expression Templates

The production of covalently bound DNA via polymerase reaction requires thermostable 5′ covalent attachment of primer molecules. As in DNA microarray technology, one of the most commonly used modifications for oligonucleotides is by primary amines at the 5′ or 3′ ends. The 5′-amino-modified oligonucleotides can directly be coupled onto the aldehyde and epoxy-coated glass surface to form covalent attachments. Amino-modified oligonucleotides covalently interacting with aldehyde-modified glass slides were stabilized by dehydration to allow bond formation via Schiff base reaction. The reaction of epoxide with amine groups requires mild alkaline buffer conditions between pH 8–9 [101]. Therefore, the type of spotting buffer determines the success of direct covalent coupling of 5′ amino-modified oligonucleotides with epoxy groups of the surface.

The success of primer immobilization on the slide surface is indicated by spot size and morphology [102]. The most suitable conditions in these studies were sodium phosphate buffer at pH 8.5 supplemented with 0.001% of nonionic detergent Tween-20 (**Fig. 4**). The presence of nonionic and zwitterionic detergents has been reported to yield perfect quality of spots in terms of morphology, size, homogeneity, and signal reproducibility, as well as overall intensity [121]. This buffer gave uniform spot size and also showed dependency with the spotted volume and concentration of primers. The perfectly round spot morphology showed that the primers had been immobilized with the right combination of surface properties, spotting buffer, and working concentration.

Then, for more quantitative comparison of spotting conditions, the experiment can be evaluated by direct or indirect detection of immobilized molecules. We used fluorescence-labeled probe hybridization for this purpose. With this method, the optimal number of molecules immobilized on the slide surface can be experimentally determined. Saturation can cause smearing near the spots or even across the entire slide surface. In most cases, however, extensive washing with buffer and finally with water can effectively reduce this effect.

The efficiency of primer immobilization is crucial for the subsequent steps of solid-phase PCR. The capacity of the surface area limits the number of molecules attaching to the active groups. For overall purposes of producing personalized protein microarrays, three types of molecules must be accommodated for: the primer, the solid-phase PCR product, and the expressed protein. The unreacted epoxy groups can be chemically deactivated or sealed by blocking agents. This blocking solution should not dissolve the spotted molecules.

By applying previous experience with immobilization chemistry of oligonucleotides with active groups on surfaces we were able to obtain perfect primer spot morphology. Good quality of spot morphology is an indication that the primer molecules are evenly distributed in the spot. Thermostable covalent attachments were also optimized by simple adjustment of timing of immobilization. It has been reported that covalent attachment of 5′-thiol-modified primers to amino-modified glass slides using a heterobifunctional cross-linker is thermally more stable than other conjugate chemistry combinations [12]. With optimum primer concentration, stable covalent attachments of 5′-amino-modified primer will also produce optimum amount of bound DNA by solid-phase PCR (**Fig. 5**). The detachment of primers is essentially not due to the stability of covalent interactions between oligonucleotides and the active groups on the surface, but is also determined by the reaction system, i.e., buffer, additives, and humidity. Since covalent bonds are not entirely formed in solution, the coupling reactions took longer to complete. The optimum primer spotting conditions for solid-phase PCR are 5 μM primer and a coupling time of 24 h at 60–70% humidity, followed by dry storage for at least an additional 24 h.

Choosing epoxy groups for surface chemistry has the dual advantage of providing direct covalent coupling of primers as well as the solid-phase PCR-generated DNA which are used for the generation of personalized protein microarrays. Epoxy groups can be deactivated by ethanolamine allowing for efficient solid-phase PCR as well as for fluorescent probe detection. One disadvantage of this treatment is poor immobilization of newly expressed proteins, however. An alternative of blocking method is to cover epoxy groups with inert molecules like BSA. BSA is known to compete with DNA polymerase for active adsorption on surfaces. Unfortunately, these blocking molecules can burry small-molecular probes that are used for DNA detection. In this case, both blocking methods were used in parallel.

Novel strategies for immobilizing DNA templates for *in situ* cell-free protein expression employ polymerases. PCR has been the main method for enriching copy numbers of DNA of interest [122]. DNA polymerase starts its activity whenever a short sequence of nucleotides has primed the replication process – these oligonucleotides are called primers. DNA polymerase only moves in the 5′-3′ direction, hence the 3′ end of the oligonucleotide primer must be free for DNA polymerase activity. Solid-phase PCR is also based on activity of DNA polymerase, to copy single-strand DNA into tethering primer which then generates the complementary strand. The most critical part for solid-phase PCR is to have the 3′ end of the attached primer free in solution so that the DNA polymerase can elongate it. Using both upper and lower primers attached, this type of solid-phase PCR has been introduced as a bridge amplification [93]. Another solid-phase PCR method allows one primer to diffuse freely in the PCR mix while the other is attached to the surface [12].

In this study, we attached both primers to the slide surface after solid-phase PCR has produced an inverse U-shape type of DNA product (**Fig. 3a**). The generated PCR products can be detected by various means, such as by cutting with restriction enzymes [98], fluorescent labeling of primers (only for single attached primer) [12], and probe hybridization [123]. In this study we used two different detection methods. First, with fluorescent-labeled probe detection of bound DNA products the amount of PCR products were interpreted based on the signal intensity of the fluorophore. We detected both strands of PCR products using a pair of probes (**Fig. 6**). With this detection method we ensured that the PCR products are double-strand DNA. The alternative detection method used second primer pairs, running a second PCR. The PCR mix contains all PCR components but minus DNA template, generating the PCR products in solution. By electrophoresis one can tell whether the DNA product is complete or not. By these means we have been able to produce bound DNA on epoxy slide surfaces by solid-phase PCR with sizes up to 3 kb.

4.2 Cell-Free Expression of Personalized Protein Microarrays

The procedure of immobilizing DNA produced by solid-phase PCR supports efficient transcription and translation processes. Linear and circular DNA templates have been routinely used for producing cell-free-based protein microarrays [11,78]. To our

knowledge, solid-phase PCR has not yet been used for producing protein microarrays. A recent application reports the production of mRNA immobilized on a slide surface [117]. Using solid-phase PCR, the immobilized primer can generate bound cDNAs on the chip surface that can be subsequently expressed into personalized protein microarrays. The particular conformation of the DNA will determine the eventual success of protein expression. The critical step is the removal of excess PCR mixture after solid-phase PCR, which can be achieved by washing; this should not denaturize the DNA. Salt buffer has been used for post-hybridization washing in DNA microarrays, both HEPES and SSC are suitable for this purpose [124]. After proper washing, the DNA array can be stored dry for some time or subsequently expressed in MIST by placing cell-free mixture on top of each spot of bound DNA [11].

We have generated covalently bound DNA arrays encoding human proteins on epoxy slide surface by solid-phase PCR and expressing them by the MIST method (**Fig. 8**). DNA encoding CREB3L1, TP53, MAPK1, and DIABLO was synthesized on the slide surface and subsequently used for *in situ* cell-free expression. All protein were fused at their N-terminal to 6×His, and their C-terminal to V5 epitope tag. Using anti-6×His and anti-V5 antibodies which recognize the fused N- and C-terminal tags, respectively, in each coding sequence, we were able to confirm full-length expressed proteins. Terminal epitope tags can be easily replaced with other epitope tags in designing primers for DNA template construction. The system can also be adjusted for eukaryotic expression, as, e.g., rabbit reticulocytes, insect cells, and wheat germ cells by changing the regulatory elements in the primer sequences.

We were able to express a multiple number of proteins on microarrays at the same time, leading up to personalized protein microarrays. To use the optimized construction of the immobilized primers that had worked successfully under our solid-phase PCR conditions, we did not change the key elements of the immobilized primers, i.e., spacer, linker, and common sequences. Production of multiple bound cDNAs on the surface required highly specific primer pairs, since we were to perform multiplex solid-phase PCR (**Fig. 7**). We added unique 5′ ends for each DNA target, through adapter-mediated primers by standard PCR. Thus, it was easier in setting a uniform T_m during solid-phase PCR. With uniform T_m, primer annealing can be easily controlled and nonspecific PCR products can be avoided.

The expressed protein on the chip was accessible across both terminus sites (**Fig. 8**). C-terminus detection indicated that translation was complete, yielding the full-length expressed protein. The protein–antibody interactions revealed that the *in situ* expressed proteins obtained directly from covalently bound double-strand cDNA at the 5′ end were able to interact with the protein-specific antibodies (**Fig. 9**). We could prove the success of the personalized protein microarrays by expressing DIABLO, CFLAR, TP53, AKT1, MAPK1, BCL2L1, and CDK2 from their cDNA products. The success rate of expression varied between these proteins. Negative spots yielded signals in the background range. Two of the expressed proteins were also confirmed by means of protein-specific antibodies (**Fig. 9c, d**).

This novel protein microarray offers advanced strategies in personalized proteomics, reflecting the protein content of particular tissues of an individual. The array can be used for basic analysis of protein interactions with other proteins, DNA, RNA, or small molecules. The overall objectives of our efforts have been in detecting disease-specific protein isoforms (caused by mutation, polymorphisms, and splice variants), developing personalized diagnosis methods, and identifying compounds relevant for therapy.

4.3 High-Density Protein Microarrays and Their Application

We produced high-density protein microarrays by MIST with unpurified PCR products as DNA templates. Unpurified PCR products had no apparent disadvantage in terms of yield of expressed proteins and their following detection (**Fig. 10**). All PCR components were added at concentrations which do not inhibit cell-free expression [125]. The PCR purification buffer or kit was designed to clean the DNA from PCR residues by removing side products of sizes below 100 bp, with the exception of the particular system for purifying small fragment PCR products [119]. With long primers for DNA template construction (129- and 87mer) the available PCR purification system is not labor efficient.

We compared Ni-NTA and epoxy slides in the production of protein microarrays. The expressed proteins were fused with N-terminal 6×His tags and C-terminal V5 tags. When His-tagged proteins are immobilized on Ni-NTA, each bivalent Ni ion forms a complex with two His residues. Protein-fused N-terminal 6×His tags provide a specific

site for oriented immobilization. Enzymatic studies showed that proteins fused with 6×His and thus becoming attached in control orientation to the solid surface had comparable activities as free proteins in solution [126]. However, the drawback of this protein immobilization approach is the low affinity of the His-tag to the Ni-NTA complex ($K = 10^7\ M^{-1}$) causing unwanted dissociation of immobilized proteins [127]. We found the same to be the case for our production of protein microarrays. Protein spots tend to present a comet-tail effect that affected the spot morphology of microarrays. Ni-NTA surfaces gave higher background than epoxy surfaces (**Fig. 11b**). Until now, protein microarrays based on site-specific immobilization of His-tag proteins on Ni-NTA slides have been reported only by Zhu and coworkers [12]. Salts and certain common chemicals in protein-binding buffers can affect the stability of protein binding to Ni-NTA surfaces [38,80].

Epoxy slides immobilized proteins via randomly covalent interactions with primary amino groups in the protein structure. The interaction of proteins with the epoxy surface can be mediated by primary amino groups at N-termini or lysine residues that mostly are located at the surface of proteins [127]. The first protein microarrays were successfully produced by immobilizing proteins onto aldehyde slides in random orientation of protein binding [7]. Little evidence is available to claim the presence of covalent attachments of proteins to epoxy slides. The formation of covalent coupling requires pH 8−9, which was not possible for the protein expression system. However, the spot morphology of expressed proteins on microarrays was satisfactory and the immobilized proteins could be used for the testing of protein interactions. Based on the result of terminal detection for expressed proteins, we found no difference between Ni-NTA and epoxy slides in terms of the amount of detected proteins on the slides (**Fig. 11c**). Random or control orientation for the binding of proteins to the surface gave similar detection yields of expressed proteins.

With careful adjustment of spotter pin capacity, high-density protein microarrays can be multiplexed without losing the consistency of protein expression across the arrays (**Fig. 12**). The CV of expressed GFP on the array was between 11 to 13%. This value indicated a slight variability in GFP yield per spot on the array. However, it should be noted that the amount of DNA template was not quantitatively equal, since we used unpurified PCR products. The previous report showed the average CV of spotted proteins and antibodies on epoxy slides was 34% [118]. As a basic rule of

thumb for cell-free protein expression: the more DNA template, the more yield of expressed protein. Furthermore, we found that the spotting replication of cell-free mixture did not interfere on the amount of GFP expression. Our fabrication of high-density protein microarrays yields consistent and satisfactory spot morphology which has been reproducible in each case.

The high-density human protein microarrays showed similar array quality as GFP microarrays. In MIST protein microarrays, each spot serves as a microscopic compartment providing individual environments for protein expression. Any cross contamination to the neighboring spots due to dissolving or diffusing can be prevented [9]. The newly expressed proteins can be retained within their own spots either by using specific or nonspecific interactions to the slide surface. The characteristics of a protein determine whether it is detectable or not. For this work, proteins were immobilized through nonspecific orientation covalent binding onto the slide surface. This nonorientated position may also limit access of antibodies to the fused epitope tags. By means of high-density protein microarrays we were able to express over 3500 proteins (**Fig. 13**) with successful expression reaching 90%. We show that multiplexing of production does not affect the amount of expressed proteins on the microarrays (**Fig. 14b**). With the existing noncontact spotter available, in single batch productions we are able to produce 25 slides of high-density protein microarrays containing more than 2000 proteins on each slide.

N-terminus signal intensities more frequently occurred in the lower range than C-termini, as detected in the red (Alexa Fluor 647 conjugated 6×His antibody) and green (Cy3-conjugated V5 antibody) laser channels, respectively (**Fig. 15a**). The difference of signal detection might indicate incomplete protein translation. The results can be interpreted in several ways, however: firstly, the C-terminal region is much less accessible than the N-terminal region. It is also possible, and potentially likely, that the C-terminal region is not fully available for V5 antibody interaction. The type of conjugated fluorophore molecules can also be a relevant factor. Unfortunately the antibodies were conjugated with different types of dyes. The Alexa Fluor dyes were more resistant to photobleaching and more brightly fluorescent than the Cy dyes [128].

The range of the signal intensities of the expressed proteins differs depending on protein size (**Fig. 15b**). The protein size has moderate negative association with the

level of expressed protein on the array. The signal intensities of expressed protein was found to decrease moderately with the increase of protein size. The success of protein expression in the cell-free-based system may vary depending on the nature of the expressed proteins and the type and conditions of the applied expression system. Certain limitations arise from the prokaryotic origin of the S30 T7 protein expression system regarding the translation and folding of eukaryotic proteins, whether *in vivo* or *in vitro* [75]. While the *E. coli* system produces high yields of synthesized proteins as compared to rabbit, wheat, and PURE systems, the *E. coli* system has not been able to effectively express very large (>100 kDa) or very small (<5 kDa) proteins [76]. A possible explanation is that small polypeptides with minor secondary structures tend to be rapidly degraded and that the generation of incomplete lengths of proteins is due to pausing of the *E. coli* ribosome [77].

Our *in situ* protein microarray method can be used in combination with any cell-free expression system of interest. DNA template construction can be adjusted for prokaryotic or eukaryotic systems through the PCR reaction. There is no need for complicated template manipulation; any known or unknown cDNA sequence can be used to easily construct protein expression templates by PCR. All necessary elements for each lysate can be added to upper primers as well as lower primers. Eukaryotic proteins expressed in the *E. coli* system lack most post-translational modifications, of course. However, because our technology is an open system, it is possible to add modifying enzymes, to modify the DNA template construction, and even to change the cell-free expression system not limited to the *E. coli* system. Although not effective for producing *in situ* protein microarrays, the noncoupled cell-free expression system can also be used in the MIST method.

The density of the proteins on the surface determines the limit of detection and the sensitivity [93]. The epoxysilane coating of the NEXTERION® Slide E did not produce a saturated autofluorescence signal from spotted GFP at concentrations of up to 500 ng/nL per spot, which is equivalent to ~ 6 pmol (**Fig. 16**), as well as for the expressed GFP. However, expressed GFP exhibited saturation signal intensities when detected by Alexa Fluor 647-conjugated anti 6×His antibody.

A systematic comparison of protein microarrays on epoxy surface and other tested slides such as those coated with PEG-epoxy, amine, and dendrimer showed similar

detection limits. The authors of that report also demonstrated that all tested surfaces exert saturation of mean signal intensities of spotted proteins and antibodies in the range of 2–2.5 fmol/spot [129].

The expressed GFP on the microarray can retain its autofluorescence function. The fluorophore is generated by a sequential mechanism in an autocatalytic process. No cofactors or enzymatic components are required. The residues Ser65, dehydroTyr66, and Gly67 of the GFP form the imidazolidone ring. The reaction is initiated by a rapid cyclization between Ser65 and Gly67 to form an imidazolin-5-one intermediate which is followed by a much slower rate-limiting oxygenation of the Tyr66 side chain by oxygen on a time scale of hours. Gly67 is required for formation of the fluorophore and no other amino acid can replace Gly in this role. Once produced, GFP is quite thermostable, though [120].

We have tested the expressed protein microarrays in protein–antibody and protein–protein interaction studies. High-density protein microarrays were shown to be useful for rapid screening of antibody specificity (**Fig. 17**). We have detected highly specific recognition of antibodies to their target proteins and of antibodies cross binding to nontarget proteins (**Table 2**). An antibody was made which recognizes denatured, linear proteins in Western blot, while flow cytometry, sandwich ELISA, and functional assays require antibodies that recognize the natural, folded shape of the proteins. Several commercially available monoclonal antibodies have been validated by Western blotting and immunohistochemistry. The antibodies can be designed against partial regions of the target proteins while antibodies also can recognize full-length proteins.

Three of these antibodies gave consistent results with those indicated in the product information of the manufacturers. One of these is anti-BCL2L1 antibody, made to recognize full-length BCL2L1; using a similar concentration as that recommended for ELISA, this antibody was able to detect newly expressed BCL2L1 on our arrays. In the case for anti-TP53 antibodies, it was designed to recognize internal regions of amino acid sequences. The anti-CDK2 and and anti-CFLAR antibodies were designed to regonize C-terminal regions of the respective proteins. Anti-CFLAR antibody failed to recognize the expressed CFLAR, but anti-CDK2 antibody effectively bound the expressed CDK2 on the arrays.

One out of six tested antibodies from our collaboration partner bound exclusively to its target peptide antigen as well as to the expressed protein (MAPK1). The anti-MAPK1 antibody showed no cross-reactivity with nontarget proteins. For the rest of tested antibodies, they are either insensitive or nonspecific for detection of target proteins on the arrays.

We expressed two DARPins using MIST and tested for their interaction with target proteins (**Fig. 18**). As a binder, DARPin typically exhibits highly specific binding and high affinity to target proteins; in contrast with traditional antibodies this binder does not require disulfide bonds for its active conformation [130]. One out two of the tested DARPins showed strong binding to maltose-binding proteins (MBP), confirming a previous report on DARPin interactions [130]. Protein microarrays can be used as an alternative selection method of DARPin libraries.

Lastly, protein microarrays have been used for protein dimer interactions such as JUN-JUN and JUN-FOS (**Fig. 19**). JUN in its native form is a homodimer complex in solution, and forms a more stable complex with FOS. The dimer formation is mediated by coiled-coil-like leucine zippers. JUN-FOS interacts with the activator protein-1 (AP1) binding site and regulates gene transcription in response to extracellular signals [131]. With a suitable choice of protein-binding buffer, protein microarrays can be used for a rapid and high-throughput detection of protein interactions. Salts can disguise the electrostatic attractions [132]. Additional macromolecules as the crowding agents can simulate the intracellular conditions – for *in vitro* PPI detection this can be achieved by adding 60−70% glycerol, 10% PEG 12,000, or 7% Ficoll 70 [133].

5 CONCLUSION

We have developed a new kind **personalized protein microarrays** for detecting protein interactions starting from surface-bound DNA templates obtained by solid-phase PCR. Our optimized conditions of spotting buffer, primer concentration, and immobilization time in solid-phase PCR have allowed for surface-bound DNA with sizes up to 3 kb. The PCR products were detected using a pair of probes to hybridize both strands of DNA and run another PCR using the second primer pairs as alternative detection method for confirming the full-length DNA.

Multiplex amplification could be optimized by solid-phase PCR. By these means protein-encoding DNA was copied from cDNA products and bound to the slide surface. Then, the arrayed DNA was expressed using a cell-free system to yield proteins on the array. **Personalized protein microarray technology thus yields cell/tissue-specific, full-length proteins from cDNA/RNA**. The expressed proteins on the chip can be confirmed by both terminal sites using antibodies corresponding to the fused tag. C-terminus detection was able to show that the *in situ* transcription and translation processes were complete. *In situ* expressed proteins obtained directly from covalently bound double-strand cDNA at the 5′ ends were able to interact with the specific protein antibodies, as shown by protein–antibody interactions. We conclude that personalized protein microarrays yield full-length expressed proteins with conformations adequate for specific protein–antibody interactions.

The MIST technology has been optimized to produce **high-density protein microarrays** of a total of 3500 human proteins. The array quality was proven to be excellent in terms of technical usability, variability of the involved template preparation, protein expression conditions, and detection methods. Our protein microarrays can thus be of great value *i*) in screening for antibody specificity, *ii*) in DARPin arrays for selective target binding, and *iii*) for model studies of protein–protein interactions.

REFERENCES

1. Steinmetz LM, Davis RW. Maximizing the potential of functional genomics. *Nat. Rev. Genet.* 5(3), 190-201 (2004).
2. Uzoma I, Zhu H. Interactome mapping: using protein microarray technology to reconstruct diverse protein networks. *Genomics, Proteomics & Bioinformatics* 11(1), 18-28 (2013).
3. Rual JF, Venkatesan K, Hao T *et al.* Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437(7062), 1173-1178 (2005).
4. Bonetta L. Protein-protein interactions: Interactome under construction. *Nature,* 468(7325), 851-854 (2010).
5. Schweitzer B, Predki P, Snyder M. Microarrays to characterize protein interactions on a whole-proteome scale. *Proteomics* 3(11), 2190-2199 (2003).
6. Coleman MA, Beernink PT, Camarero JA, Albala JS. Applications of functional protein microarrays: identifying protein-protein interactions in an array format. In: Floriano PN (Ed) *Microchip-Based Assay Systems. Methods in Molecular Biology* 385, pp 121-130. Humana Press, Totowa, NJ (2007).
7. MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. *Science* 289(5485), 1760-1763 (2000).
8. Jackson AM, Boutell J, Cooley N, He M. Cell-free protein synthesis for proteomics. *Brief Funct. Genomic Proteomic* 2(4), 308-319 (2004).
9. He M, Stoevesandt O, Palmer EA, Khan F, Ericsson O, Taussig MJ. Printing protein arrays from DNA arrays. *Nat. Methods* 5(2), 175-177 (2008).
10. Ramachandran N, Raphael JV, Hainsworth E *et al.* Next-generation high-density self-assembling functional protein arrays. *Nat. Methods* 5(6), 535-538 (2008).
11. Angenendt P, Kreutzberger J, Glokler J, Hoheisel JD. Generation of high density protein microarrays by cell-free *in situ* expression of unpurified PCR products. *Mol. Cell. Proteomics* 5(9), 1658-1666 (2006).
12. Adessi C, Matton G, Ayala G *et al.* Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic Acids Res.* 28(20), E87 (2000).
13. Nickisch-Rosenegk Mv, Marschan X, Andresen D, Abraham A, Heise C, Bier FF. On-chip PCR amplification of very long templates using immobilized primers on glassy surfaces. *Biosensors Bioelectron.* 20(8), 1491-1498 (2005).
14. Stein LD. Human genome: end of the beginning. *Nature* 431(7011), 915-916 (2004).
15. Finishing the euchromatic sequence of the human genome. *Nature* 431(7011), 931-945 (2004).
16. Collins FS, Green ED, Guttmacher AE, Guyer MS. A vision for the future of genomics research. *Nature* 422(6934), 835-847 (2003).
17. Consortium EP, Dunham I, Kundaje A *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* 489(7414), 57-74 (2012).
18. Clamp M, Fry B, Kamal M *et al.* Distinguishing protein-coding and noncoding genes in the human genome. *Proc. Natl. Acad. Sci. USA* 104(49), 19428-19433 (2007).

19. Harrow J, Frankish A, Gonzalez JM *et al.* GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* 22(9), 1760-1774 (2012).
20. Normanno N, De Luca A, Carotenuto P, Lamura L, Oliva I, D'Alessio A. Prognostic applications of gene expression signatures in breast cancer. *Oncology* 77 Suppl 1, 2-8 (2009).
21. Vogel C, Chothia C. Protein Family Expansions and Biological Complexity. *PLoS Comput. Biol.* 2(5), e48 (2006).
22. Pal S, Gupta R, Davuluri RV. Alternative transcription and alternative splicing in cancer. *Pharmacol. Ther.* 136(3), 283-294 (2012).
23. Shabalina SA, Spiridonov AN, Spiridonov NA, Koonin EV. Connections between alternative transcription and alternative splicing in mammals. *Genome Biol. Evol.* 2, 791-799 (2010).
24. Sultan M, Schulz MH, Richard H *et al.* A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* 321(5891), 956-960 (2008).
25. Taft RJ, Pheasant M, Mattick JS. The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays* 29(3), 288-299 (2007).
26. de Sousa Abreu R, Penalva LO, Marcotte EM, Vogel C. Global signatures of protein and mRNA expression levels. *Mol. Biosyst.* 5(12), 1512-1526 (2009).
27. Vogel C, Abreu Rde S, Ko D *et al.* Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. *Mol. Syst. Biol.* 6, 400 (2010).
28. Andreadis A, Gallego ME, Nadal-Ginard B. Generation of protein isoform diversity by alternative splicing: mechanistic and biological implications. *Annu. Rev. Cell Biol.* 3(1), 207-242 (1987).
29. Breitbart RE, Andreadis A, Nadal-Ginard B. Alternative splicing: a ubiquitous mechanism for the generation of multiple protein isoforms from single genes. *Annu. Rev. Biochem.* 56(1), 467-495 (1987).
30. Seo J, Lee KJ. Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *J. Biochem. Mol. Biol.* 37(1), 35-44 (2004).
31. Vogel C, Marcotte EM. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* 13(4), 227-232 (2012).
32. Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* 12(1), 32-42 (2011).
33. Graveley BR. Alternative splicing: increasing diversity in the proteomic world. *Trends Genet.* 17(2), 100-107 (2001).
34. Stamm S, Ben-Ari S, Rafalska I *et al.* Function of alternative splicing. *Gene* 344, 1-20 (2005).
35. Lareau LF, Green RE, Bhatnagar RS, Brenner SE. The evolving roles of alternative splicing. *Curr. Opin. Struct. Biol.* 14(3), 273-282 (2004).
36. Ast G. How did alternative splicing evolve? *Nat. Rev. Genet.* 5(10), 773-782 (2004).
37. Romero PR, Zaidi S, Fang YY *et al.* Alternative splicing in concert with protein intrinsic disorder enables increased functional diversity in multicellular organisms. *Proc. Natl. Acad. Sci. USA* 103(22), 8390-8395 (2006).
38. Brinkman B. Splice variants as cancer biomarkers. *Clin. Biochem.* 37(7), 584-594 (2004).

39. Thorsen K, Sørensen KD, Brems-Eskildsen AS *et al.* Alternative splicing in colon, bladder, and prostate cancer identified by exon array analysis. *Mol. Cell. Proteomics* 7(7), 1214-1224 (2008).
40. Carrigan PE, Bingham JL, Srinvasan S, Brentnall TA, Miller LJ. Characterization of alternative spliceoforms and the RNA splicing machinery in pancreatic cancer. *Pancreas* 40(2), 281-288 (2011).
41. De Las Rivas J, Prieto C. Protein interactions: mapping interactome networks to support drug target discovery and selection. In: Larson RS (Ed) *Bioinformatics and Drug Discovery. Methods in Molecular Biology* 910, pp 279-296, Humana Press/Springer, Heidelberg New York (2012).
42. Chu LH, Chen BS. Construction of a cancer-perturbed protein-protein interaction network for discovery of apoptosis drug targets. *BMC Syst. Biol.* 2, 56 (2008).
43. Ramani AK, Bunescu RC, Mooney RJ, Marcotte EM. Consolidating the set of known human protein-protein interactions in preparation for large-scale mapping of the human interactome. *Genome Biol.* 6(5), R40 (2005).
44. Wu G, Feng X, Stein L. A human functional protein interaction network and its application to cancer data analysis. *Genome Biol.* 11(5), R53 (2010).
45. Venkatesan K, Rual JF, Vazquez A *et al.* An empirical framework for binary interactome mapping. *Nature Methods* 6(1), 83-90 (2009).
46. Bandyopadhyay S, Chiang CY, Srivastava J *et al.* A human MAP kinase interactome. *Nat. Methods* 7(10), 801-805 (2010).
47. Garma L, Mukherjee S, Mitra P, Zhang Y. How many protein-protein interactions types exist in nature? *PLoS One* 7(6), 13 (2012).
48. Prieto C, De Las Rivas J. APID: Agile Protein Interaction DataAnalyzer. *Nucleic Acids Res.* 34 (Web Server issue), W298-302 (2006).
49. Reichmann D, Rahat O, Cohen M, Neuvirth H, Schreiber G. The molecular architecture of protein-protein binding sites. *Curr. Opin. Struct. Biol.* 17(1), 67-76 (2007).
50. Nooren IM, Thornton JM. Diversity of protein-protein interactions. *EMBO J.* 22(14), 3486-3492 (2003).
51. Perozzo R, Folkers G, Scapozza L. Thermodynamics of protein-ligand interactions: history, presence, and future aspects. *J. Recept. Signal Transduct. Res.* 24(1-2), 1-52 (2004).
52. Chavez JD, Liu NL, Bruce JE. Quantification of protein-protein interactions with chemical cross-linking and mass spectrometry. *J. Proteome Res.* 10(4), 1528-1537 (2011).
53. Hwang S, Schmitt AA, Luteran AE, Toone EJ, McCafferty DG. Thermodynamic characterization of the binding interaction between the histone demethylase LSD1/KDM1 and CoREST. *Biochemistry* 50(4), 546-557 (2011).
54. Bornhop DJ, Latham JC, Kussrow A, Markov DA, Jones RD, Sorensen HS. Free-solution, label-free molecular interactions studied by back-scattering interferometry. *Science* 317(5845), 1732-1736 (2007).
55. Bildl W, Haupt A, Muller CS *et al.* Extending the dynamic range of label-free mass spectrometric quantification of affinity purifications. *Mol. Cell Proteomics* 11(2), M111 007955 (2012).
56. Geetha T, Langlais P, Luo M *et al.* Label-free proteomic identification of endogenous, insulin-stimulated interaction partners of insulin receptor substrate-1. *J. Am. Soc. Mass Spectr.* 22(3), 457-466 (2011).

57. Berrade L, Garcia AE, Camarero JA. Protein microarrays: novel developments and applications. *Pharm. Res.* 28(7), 1480-1499 (2011).
58. Yates JR, Ruse CI, Nakorchevsky A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev.Biomed. Eng.* 11, 49-79 (2009).
59. Kerppola TK. Bimolecular fluorescence complementation: visualization of molecular interactions in living cells. *Method. Cell Biol.* 85, 431-470 (2008).
60. Padilla-Parra S, Tramier M. FRET microscopy in the living cell: different approaches, strengths and weaknesses. *BioEssays* 34(5), 369-376 (2012).
61. Lakowicz JR. *Principles of fluorescence spectroscopy*. Springer, New York, (2010).
62. Lowder MA, Appelbaum JS, Hobert EM, Schepartz A. Visualizing protein partnerships in living cells and organisms. *Curr. Opin. Struct. Biol.* 15(6), 781-788 (2011).
63. Ekins RP. Multi-analyte immunoassay. *J. Pharm. Biomed. Anal.* 7(2), 155-168 (1989).
64. Gaster RS, Hall DA, Wang SX. Autoassembly protein arrays for analyzing antibody cross-reactivity. *Nano Lett.* 11(7), 2579-2583 (2011).
65. Kopf E, Zharhary D. Antibody arrays – An emerging tool in cancer proteomics. *Int. J. Biochem. Cell Biol.* 39(7–8), 1305-1317 (2007).
66. Schaaij-Visser TB, de Wit M, Lam SW, Jimenez CR. The cancer secretome, current status and opportunities in the lung, breast and colorectal cancer context. *Biochim. Biophys. Acta* 1834(11), 2242-2258 (2013).
67. Rogers YH, Jiang-Baucom P, Huang ZJ, Bogdanov V, Anderson S, Boyce-Jacino MT. Immobilization of oligonucleotides onto a glass support via disulfide bonds: A method for preparation of DNA microarrays. *Anal. Biochem.* 266(1), 23-30 (1999).
68. Shi W, Meng Z, Chen Z, Luo J, Liu L. Proteome analysis of human pancreatic cancer cell lines with highly liver metastatic potential by antibody microarray. *Mol. Cell. Biochem.* 347(1-2), 117-125 (2011).
69. Hurst R, Hook B, Slater MR, Hartnett J, Storts DR, Nath N. Protein-protein interaction studies on protein arrays: effect of detection strategies on signal-to-background ratios. *Anal. Biochem.* 392(1), 45-53 (2009).
70. Tom I, Lewin-Koh N, Ramani SR, Gonzalez LC. Protein microarrays for identification of novel extracellular protein-protein interactions. *Curr. Protoc. Protein Sci.* Chap. 27, Unit 27 23 (2013).
71. Kijanka G, Ipcho S, Baars S *et al.* Rapid characterization of binding specificity and cross-reactivity of antibodies using recombinant human protein arrays. *J. Immunol. Methods* 340(2), 132-137 (2009).
72. Michaud GA, Salcius M, Zhou F *et al.* Analyzing antibody specificity with whole proteome microarrays. *Nat. Biotechnol.* 21(12), 1509-1512 (2003).
73. Kraus PR, Meng L, Freeman-Cook L. Small molecule selectivity and specificity profiling using functional protein microarrays. In: Chittur SV (Ed) *Microarray Methods for Drug Discovery. Methods in Molecular Biology* 632, pp 251-267. Humana Press/Springer, New York (2010).
74. Espina V, Woodhouse EC, Wulfkuhle J, Asmussen HD, Petricoin EF, 3rd, Liotta LA. Protein microarray detection strategies: focus on direct detection technologies. *J. Immunol. Methods* 290(1-2), 121-133 (2004).

75. Bernhard F, Tozawa Y. Cell-free expression – making a mark. *Curr. Opin. Struct. Biol.* 23(3), 374-380 (2013).
76. Shimizu Y, Kuruma Y, Ying BW, Umekage S, Ueda T. Cell-free translation systems for protein engineering. *FEBS J.* 273(18), 4133-4140 (2006).
77. Whittaker JW. Cell-free protein synthesis: the state of the art. *Biotechnol. Lett.* 35(2), 143-152 (2013).
78. Ramachandran N, Hainsworth E, Bhullar B *et al.* Self-assembling protein microarrays. *Science* 305(5680), 86-90 (2004).
79. Hermanson GT. *Bioconjugate techniques.* Academic Press, Amsterdam, Boston (2008).
80. Jarvik JW, Telmer CA. Epitope tagging. *Annu. Rev. Genet.* 32(1), 601-618 (1998).
81. Kaushansky A, Allen JE, Gordus A *et al.* Quantifying protein-protein interactions in high throughput using protein domain microarrays. *Nat. Protoc.* 5(4), 773-790 (2010).
82. Hoheisel JD, Alhamdani MS, Schroder C. Affinity-based microarrays for proteomic analysis of cancer tissues. *Proteomics Clin. Appl.* 7(1-2), 8-15 (2013).
83. Schmidt R, Jacak J, Schirwitz C *et al.* Single-molecule detection on a protein-array assay platform for the exposure of a tuberculosis antigen. *J. Proteome Res.* 10(3), 1316-1322 (2011).
84. Ramachandran N, Larson DN, Stark PR, Hainsworth E, LaBaer J. Emerging tools for real-time label-free detection of interactions on functional protein microarrays. *FEBS J.* 272(21), 5412-5425 (2005).
85. Gaster RS, Xu L, Han SJ *et al.* Quantification of protein interactions and solution transport using high-density GMR sensor arrays. *Nat. Nanotechnol.* 6(5), 314-320 (2011).
86. Ray S, Mehta G, Srivastava S. Label-free detection techniques for protein microarrays: prospects, merits and challenges. *Proteomics* 10(4), 731-748 (2010).
87. Porter JR, Stains CI, Jester BW, Ghosh I. A general and rapid cell-free approach for the interrogation of protein-protein, protein-DNA, and protein-RNA interactions and their antagonists utilizing split-protein reporters. *J. Am. Chem. Soc.* 130(20), 6488-6497 (2008).
88. Sboner A, Karpikov A, Chen G *et al.* Robust-linear-model normalization to reduce technical variability in functional protein microarrays. *J. Proteome Res.* 9(1), 636-636 (2009).
89. Díez P, Dasilva N, González-González M *et al.* Data analysis strategies for protein microarrays. *Microarrays* 1(2), 64-83 (2012).
90. Zhu X, Gerstein M, Snyder M. ProCAT: a data analysis approach for protein microarrays. *Genome Biol.* 7(11), R110 (2006).
91. Heller MJ. DNA microarray technology: devices, systems, and applications. *Annu. Rev. Biomed. Eng.* 4, 129-153 (2002).
92. Kohsaka H, Carson DA. Solid-phase polymerase chain reaction. *J. Clin. Lab. Anal.* 8(6), 452-455 (1994).
93. Bing DH, Boles C, Rehman FN *et al.* Bridge amplification: a solid phase PCR system for the amplification and detection of allelic differences in single copy genes. *Genetic Identity. Conf. Proc., 7th Intl. Symp. Human Identification* (1996).

94. Erdogan F, Kirchner R, Mann W, Ropers HH, Nuber UA. Detection of mitochondrial single nucleotide polymorphisms using a primer elongation reaction on oligonucleotide microarrays. *Nucleic Acids Res.* 29(7), E36 (2001).
95. Nickisch-Rosenegk Mv, Marschan X, Andresen D, Abraham A, Heise C, Bier FF. On-chip PCR amplification of very long templates using immobilized primers on glassy surfaces. *Biosens. Bioelectron.* 20(8), 1491-1498 (2005).
96. Mitterer G, Schmidt WM. Microarray-based detection of bacteria by on-chip PCR. In: O'Connor L (Ed) *Diagnostic Bacteriology Protocols. Methods in Molecular Biology* 345, pp 37-51. Humana Press, Totowa, NJ (2006).
97. Nickisch-Rosenegk Mv, Marschan X, Andresen D, Bier FF. Reverse transcription-polymerase chain reaction on a microarray: the integrating concept of "active arrays". *Anal. Bioanal. Chem.* 391(5), 1671-1678 (2008).
98. Carmon A, Vision TJ, Mitchell SE, Thannhauser TW, Muller U, Kresovich S. Solid-phase PCR in microwells: effects of linker length and composition on tethering, hybridization, and extension. *BioTechniques* 32(2), 410, 412, 414-418, 420 (2002).
99. Andreadis JD, Chrisey LA. Use of immobilized PCR primers to generate covalently immobilized DNAs for in vitro transcription/translation reactions. *Nucleic Acids Res.* 28(2), e5 (2000).
100. Oroskar AA, Rasmussen SE, Rasmussen HN, Rasmussen SR, Sullivan BM, Johansson A. Detection of immobilized amplicons by ELISA-like techniques. *Clin. Chem.* 42(9), 1547-1555 (1996).
101. Beattie WG, Meng L, Turner SL, Varma RS, Dao DD, Beattie KL. Hybridization of DNA targets to glass-tethered oligonucleotide probes. *Mol. Biotechnol.* 4(3), 213-225 (1995).
102. Sobek J, Aquino C, Schlapbach R. Optimization workflow for the processing of high quality glass-based microarrays: applications in DNA, peptide, antibody, and carbohydrate microarraying. In: Rampal JB (Ed) *Microarrays*, Vol. 2, *Applications and Data Analysis*, 2nd edn. *Methods in Molecular Biology* 382, pp 33-51. Humana Press, Totowa, NJ (2007).
103. Diehl F, Grahlmann S, Beier M, Hoheisel JD. Manufacturing DNA microarrays of high spot homogeneity and reduced background signal. *Nucleic Acids Res.* 29(7), E38 (2001).
104. Angenendt P, Glokler J, Murphy D, Lehrach H, Cahill DJ. Toward optimized antibody microarrays: a comparison of current microarray support materials. *Anal. Biochem.* 309(2), 253-260 (2002).
105. Grainger DW, Greef CH, Gong P, Lochhead MJ. Current microarray surface chemistries. In: Rampal JB (Ed) *Microarrays*, 2nd edn. *Methods in Molecular Biology* 381, pp 37-57. Humana Press, Totowa, NJ (2007).
106. Lindroos K, Liljedahl U, Raitio M, Syvanen AC. Minisequencing on oligonucleotide microarrays: comparison of immobilisation chemistries. *Nucleic Acids Res.* 29(13), E69-69 (2001).
107. Zammatteo N, Jeanmart L, Hamels S *et al.* Comparison between different strategies of covalent attachment of DNA to glass surfaces to build DNA microarrays. *Anal. Biochem.* 280(1), 143-150 (2000).
108. Joos B, Kuster H, Cone R. Covalent attachment of hybridizable oligonucleotides to glass supports. *Anal. Biochem.* 247(1), 96-101 (1997).
109. Chrisey LA, Lee GU, O'Ferrall CE. Covalent attachment of synthetic DNA to self-assembled monolayer films. *Nucleic Acids Res.* 24(15), 3031-3039 (1996).

110. Steel A, Levicky R, Herne T, Tarlov MJ. Immobilization of nucleic acids at solid surfaces: effect of oligonucleotide length on layer assembly. *Biophys. J.* 79(2), 975-981 (2000).
111. Shchepinov M, Case-Green S, Southern E. Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays. *Nucleic Acids Res.* 25(6), 1155-1161 (1997).
112. Rozenberg BA. Kinetics, thermodynamics and mechanism of reactions of epoxy oligomers with amines. In: Dušek K (Ed) *Epoxy Resins and Composites II.* Springer, Berlin Heidelberg (1986) pp 113-165.
113. Erdogan F, Kirchner R, Mann W, Ropers HH, Nuber UA. Detection of mitochondrial single nucleotide polymorphisms using a primer elongation reaction on oligonucleotide microarrays. *Nucleic Acids Res.* 29(7), e36-e36 (2001).
114. Guo Z, Guilfoyle RA, Thiel AJ, Wang R, Smith LM. Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res.* 22(24), 5456-5465 (1994).
115. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13, 134 (2012).
116. Mitterer G, Huber M, Leidinger E *et al.* Microarray-based identification of bacteria in clinical samples by solid-phase PCR amplification of 23S ribosomal DNA sequences. *J. Clin. Microbiol.* 42(3), 1048-1057 (2004).
117. Nickisch-Rosenegk Mv, Marschan X, Andresen D, Bier FF. Reverse transcription-polymerase chain reaction on a microarray: the integrating concept of "active arrays". *Anal. Bioanal. Chem.* 391(5), 1671-1678 (2008).
118. Quail MA, Smith M, Coupland P *et al.* A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13(1), 341 (2012).
119. Sambrook J, Russell DW. Purification of PCR products in preparation for cloning. *CSH Protoc.* 2006(1) (2006).
120. Tsien RY. The green fluorescent protein. *Annu. Rev. Biochem.* 67, 509-544 (1998).
121. Rickman DS, Herbert CJ, Aggerbeck LP. Optimizing spotting solutions for increased reproducibility of cDNA microarrays. *Nucleic Acids Res.* 31(18), e109-e109 (2003).
122. Saiki RK, Gelfand DH, Stoffel S *et al.* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239(4839), 487-491 (1988).
123. Sjöroos M, Ilonen J, Lövgren T. Solid-phase PCR with hybridization and time-resolved fluorometry for detection of HLA-B27. *Clin. Chem.* 47(3), 498-504 (2001).
124. Poulsen L, Søe MJ, Snakenborg D, Møller LB, Dufva M. Multi-stringency wash of partially hybridized 60-mer probes reveals that the stringency along the probe decreases with distance from the microarray surface. *Nucleic Acids Res.* 36(20), e132 (2008).
125. Iskakova MB, Szaflarski W, Dreyfus M, Remme J, Nierhaus KH. Troubleshooting coupled in vitro transcription-translation system derived from *Escherichia coli* cells: synthesis of high-yield fully active proteins. *Nucleic Acids Res.* 34(19), e135 (2006).

126. Zhu X, Guo A. The critical role of surface chemistry in protein microarrays. In: Predki PF (Ed.) *Functional Protein Microarrays in Drug Discovery* (2010) pp. 53-71
127. Jonkheijm P, Weinrich D, Schröder H, Niemeyer CM, Waldmann H. Chemical strategies for generating protein biochips. *Angew. Chem. Intl. Edn.* 47(50), 9618-9647 (2008).
128. Berlier JE, Rothe A, Buller G *et al.* Quantitative comparison of long-wavelength Alexa Fluor dyes to Cy dyes: fluorescence of the dyes and their bioconjugates. *J. Histochem. Cytochem.* 51(12), 1699-1712 (2003).
129. Angenendt P, Glökler J, Sobek J, Lehrach H, Cahill DJ. Next generation of protein microarray support materials:: Evaluation for protein and antibody microarray applications. *J. Chromatogr.* 1009(1), 97-104 (2003).
130. Binz HK, Amstutz P, Kohl A *et al.* High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat. Biotechnol.* 22(5), 575-582 (2004).
131. Patel L, Abate C, Curran T. Altered protein conformation on DNA binding by Fos and Jun. *Nature* 347, 572-575 (1990).
132. Sheinerman FB, Norel R, Honig B. Electrostatic aspects of protein–protein interactions. *Curr. Opin. Struct. Biol.* 10(2), 153-159 (2000).
133. Schlarb-Ridley BG, Mi H, Teale WD, Meyer VS, Howe CJ, Bendall DS. Implications of the effects of viscosity, macromolecular crowding, and temperature for the transient interaction between cytochrome f and plastocyanin from the cyanobacterium *Phormidium laminosum*. *Biochemistry* 44(16), 6232-6238 (2005).

APPENDIX

List of Primers

Unmodified

Gene	Primer name	Sequence 5′→3′
AKT1	FOPU4_AKT1	tccgaacattggactggcacgggtaatacgactcactatagggagaccacaacggtttccctctagaaataattaggagatatacctaatgcatcatcat catcatcatgctcccctcaacaacttctc
AKT1	ROPU4_AKT1	ctgttggtccgatgccacttgcgttattacgtagaatcgagaccgaggagagggttagggataggcttaccgtcatcttggtcaggtggtgt
ATF3	Fow_ATF3	cttcaacacccaggccag
ATF3	Rev_ATF3	ctgcaatgttccttcttttatct
BCL2L1	Fow_BCL2L1	tctcagagcaaccgggag
BCL2L1	Rev_BCL2L1	ccgactgaagagtgagcc
BCL2L1	FOPU1_BCL2L1	tcatggaaccgtactcgccgtgtaatacgactcactatagggagaccacaacggtttccctctagaaataattaggagatatacctaatgcatcatcatca tcatcattctcagagcaaccgggagct
BCL2L1	ROPU1_BCL2L1	ctggcgtagagtcgcatccagtcttattacgtagaatcgagaccgaggagagggttagggataggcttaccccgactgaagagtgagcccag
CASP7	Fow_CASP7	gcagatgagcagggctg
CASP7	Rev_CASP7	ttgactgaagtagagttccttgg
CASP7	Fow_CASP7	gcagatgaacagggctg
CASP7	Rev_CASP7	ttgactgaagtagagttccttgg
CDK2	FOPU7_CDK2	tcccagtcaacactggaggctcgtaatacgactcactatagggagaccacaacggtttccctctagaaataattaggagatatacctaatgcatcatcatc atcatcatagaaacaagttgacgggagag
CDK2	ROPU7_CDK2	ctgagtgagacgctcgagtagcattattacgtagaatcgagaccgaggagagggttagggataggcttaccatggggtactggcttggtc

CFLAR	Fow_CFLAR	tctgctgaagtcatccatca
CFLAR	Rev_CFLAR	tgtgtaggagaggataagtttctt
CFLAR	FOPU2_CFLAR	tcctaccctcgtggttgcgctaataatacgactcactatagggagaccacaacggtttccctctagaaataattaggagatatacctaatgcatcatcatc atcatcattctgctgaagtcatccatcag
CFLAR	ROPU2_CFLAR	ctgtagaccctgtcggctacggattattacgtagaatcgagaccgaggagagggttagggataggcttaccgtgctgcagccagacataata
CLDN18	Fow_CLDN18	atgtccaccaccacatgc
CLDN18	Rev_CLDN18	cacatagtcgtgcttggaag
CLDN18	FOPU3_CLDN18tv2	tccgcgtcttatacggaaccgcgtaatacgactcactatagggagaccacaacggtttccctctagaaataattaggagatatacctaatgcatcatcatc atcatcatgtgactgcctgtcagggcttg
CLDN18	ROPU3_CLDN18tv2	ctgtatcgtagggccactcaggcttattacgtagaatcgagaccgaggagagggttagggataggcttaccgtcgtgcttggaaggataagattg
DARPin	Fow_6HisDARPin	tcccgcgaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattagccaccatggtaagaaggagatataccatgactatga gaggatcgcatcac
DARPin	Rev_V5DARPin	ctggaattcgccctttttattacgtagaatcgagaccgaggagagggttagggataggcttaccattaagcttttgcaggatttc
DARpin	Fow_DARpin	actatgagaggatcgcatcac
DIABLO	FOPU6_DIABLO	tcccgatctgggaggtatggctttaatacgactcactatagggagaccacaacggtttccctctagaaataattaggagatatacctaatgcatcatcatc atcatcatttgtgtgttcctgttgtggc
DIABLO	ROPU6_DIABLO	ctgcccatcggaatcgagtgtggttattacgtagaatcgagaccgaggagagggttagggataggcttacctttctgacggagctcttctatc
MAPK1	FOPU5_MAPK1	tccgaccttggcagtcgctaacgtaatacgactcactatagggagaccacaacggtttccctctagaaataattaggagatatacctaatgcatcatcatc atcatcatgagatggtccgcgggcagg
MAPK1	ROPU5_MAPK1	ctggggatgaacggctatcgcagttattacgtagaatcgagaccgaggagagggttagggataggcttaccgggctcgtcactcgggtcgtaata
MCL1	Fow_MCL1	aacgcggtaatcggactc
MCL1	Rev_MCL1	tcttattagatatgccaaaccag
MDM2	Fow_MDM2	gtgaggagcaggcaaatg
MDM2	Rev_MDM2	ggggaaataagttagcacaat
MDM2	Fow_6His_MDM2	tcccgcgaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattagccaccatggtaagaaggagatataccatgcatcatc atcatcatcatgtgaggagcaggcaaatg

MDM2	Rev_V5_MDM2	ctggaattcgccctttttattacgtagaatcgagaccgaggagagggttagggataggcttaccggggaaataagttagcacaat
MDM2	Rev_Myc_MDM2	ctggaattcgccctttttattacagatcctcttcagagatgagtttctgctcggggaaataagttagcacaat
MDM2	Fow_MDM2_ovh	aaagcaggctccaccatggtgaggagcaggcaaatg
MDM2	Rev_MDM2_ovh	ggggaaataagttagcacaataactttgtacaagaaagctgggtc
pDON221, human cDNA	Rev_fullV5	ctggaattcgccctttttattacgtagaatcgagaccgaggagagggttagggataggcttaccaactttgtacaagaaagctgggtc
pDON221, human cDNA	Fow_full6His	tcccgcgaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattagccaccatggtaagaaggagatataccatgcatcatc atcatcatcataaagcaggctccaccatg
pDON221, human cDNA	Fow_full/Fow_full6His	attaatacgactcactatagggagaccacaacggtttccctctagaaataattagccaccatggtaagaaggagatataccatgcatcatcatcatcatc ataaagcaggctccaccatg
pDON221, human cDNA	Fow_full/Fow_fullNoTag	attaatacgactcactatagggagaccacaacggtttccctctagaaataattagccaccatggtaagaaggagatataccatgaaagcaggctccaccatg
pDON221, human cDNA	Fow_fullNoHis	tcccgcgaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattttgtttaactttaagaaggagatataccatgaaagcag gctccaccatg
pDON221, human cDNA	Rev_full6xHis	ctggaattcgccctttttattaatgatgatgatgatgatgaactttgtacaagaaagctgggtc
pDON221, human cDNA	Rev_fullV5	ctggaattcgccctttttattacgtagaatcgagaccgaggagagggttagggataggcttaccaactttgtacaagaaagctgggtc
pDON221, human cDNA	Rev_full_Myc	ctggaattcgccctttttattacagatcctcttcagagatgagtttctgctcaactttgtacaagaaagctgggtc
piVEX, wtGFP	Fow_full6His_GFP	tcccgcgaaattaatacgactcactatagggagaccacaacggtttccctctagaaataattagccaccatggtaagaaggagatataccatgcatcatc atcatcatcatttaagaaggagatataccatg
piVEX, wtGFP	Rev_fullV5_GFP	ctggaattcgccctttttattacgtagaatcgagaccgaggagagggttagggataggcttaccatgatgatgatgatgagaac
PPM1A	Fow_PPM1A	ggagcatttttagacaagcc
PPM1A	Rev_PPM1A	ccacatatcatctgttgatgtaga
Rev_DARpin	Rev_DARpin	attaagcttttgcaggatttc
SMAD2	Seq_revSMAD2	cagacttttcactgctttc
SMAD2	Seq_fowSMAD2	cagcagtgtaattaatacgac
SMAD	Seq_revSMAD2 order	tgggatacctggagacgac

Amino-modified

Target	Primer name	Sequence NH_3-C_6 5′→3′
ConstSMAD2	Sld_Fow6His_SMAD2sp	ttttttttttcagcagtgta*att*aatacgactcactatagggagaccacaacggtttccctctagaaataattagccaccatggtaagaaggagatataccatgcatcatcatcatcatcattccatcttgccattcacg
ConstSMAD2	Sld_FownoTag_SMAD2sp	ttttttttttcagcagtgtacagatggtgatgatgattaatacgactcactatagggagaccacaacggtttccctctagaaataattagccaccatggtaagaaggagatataccatgtccatcttgccattcacg
CosntSMAD2	Sld_RevV5_SMAD2sp	ttttttttttgccacttaccccacaaccagacgctggaattcgccctttttattacgtagaatcgagaccgaggagagggttagggataggcttaccgcttgagcaacgcactg
ConstpDON,human cDNA	Solid FowpDON	ttttttttttaggtgcgtgtgggttggatctcccgcgaaattaatacg
ConstpDON,human cDNA	Solid RevpDON	ttttttttttgattagaaagtaacctcagtctggaattcgccctttttatta
ConstU1	Sfow_U1	ttttttttttagatgcatgtgagttagatcatggaaccgtactcgccgtg
ConstU1	Srev_U1	ttttttttttgattagaaagtaacctcagtgcgtagagtcgcatccagtc
ConstU2	Sfow_U2	ttttttttttagatgcatgtgagttagatctaccctcgtggttgcgctaa
ConstU2	Srev_U2	ttttttttttgattagaaagtaacctcagttagaccctgtcggctacgga
ConstU3	Sfow_U3	ttttttttttaggtgcgtgtgggttggatcgcgtcttatacggaaccgcg
ConstU3	Srev_U3	ttttttttttgattagaaagtaacctcagttatcgtagggccactcaggc
ConstU4	Sfow_U4	ttttttttttagatgcatgtgagttagatcgaacattggactggcacggg
ConstU4	Srev_U4	ttttttttttgattagaaagtaacctcagtttggtccgatgccacttgcg
ConstU5	Sfow_U5	ttttttttttagatgcatgtgagttagatcgaccttggcagtcgctaacg
ConstU5	Srev_U5	ttttttttttgattagaaagtaacctcagtgggatgaacggctatcgcag
ConstU6	Sfow_U6	ttttttttttagatgcatgtgagttagatccgatctgggaggtatggctt
ConstU6	Srev_U6	ttttttttttgattagaaagtaacctcagtcccatcggaatcgagtgtgg
ConstU7	Sfow_U7	ttttttttttagatgcatgtgagttagatccagtcaacactggaggctcg
ConstU7	Srev_U7	ttttttttttgattagaaagtaacctcagtagtgagacgctcgagtagca

Fluorophore-modified

Gene	Probe name	Sequence Cy3/Cy5 5′→3′
ILR1RN	prbn_revILR1RN	cgagaacagaaagcaggacaagcgct
ILR1RN	prbn_fowILR1RN	gcacatcttccctccatggattcccaag
CYCS	nprb_revCYCS	caataagaacaaaggcatcatctggggaga
CYCS	nprb_fowCYCS	aaagagaccatggagatttggcccagtc
DIABLO	nprb_revDIABLO	tatcaaactggcgcagatcaggcctctata
DIABLO	nprb_fowDIABLO	tcttcctcctctgaattcattttcccaagt
CHMP2A	nprb_revCHMP2A	ctcaagtccaacaactcgatggcacaagc
CHMP2A	nprb_fowCHMP2A	accaagtcttttgccatgatgcgaacagc
GFP	nprb_revGFP	actacaagacacgtgctgaagtcaagtttg
GFP	nprb_fowGFP	ataagagaaagtagtgacaagtgttggcca
SORT1	nprb_fowSORT1	tgttgtctgatccccatttggccaaac
SORT1	nprb_revSORT1	agattatgaagatggctgcattttgggctaca
SORT1	prb_fowSORT1	gtttggccaaatggggatcagacaaca
SORT1	prb_revSORT1	tgtagcccaaaatgcagccatcttcataatct
ICAM3	nprb_fowICAM3	ggagagcgaggtgacagaggtgcagaaag
ICAM3	nprb_revICAM3	ccacgagcacatcatcaagtacaagggc
ICAM3	prb_fowICAM3	ctttctgcacctctgtcacctcgctctcc
ICAM3	prb_revICAM3	gcccttgtacttgatgatgtgctcgtgg

List of Genes

76P	*ACTRT2*	*AKR1A1*	*ANXA5*	*ARHGAP1*	*ATP6AP1*
A4GALT	*ACVR1C*	*AKR1B1*	*ANXA8L1*	*ARHGAP9*	*ATP6V0C*
AAAS	*ACY1*	*AKR1B10*	*ANXA9*	*ARHGDIA*	*ATP6V0D1*
AARS	*ADA*	*AKR1C2*	*AOAH*	*ARHGDIB*	*ATP6V1B2*
AASDHPPT	*ADAM2*	*AKR1C3*	*AP1S2*	*ARHGEF16*	*ATXN10*
AATF	*ADAM30*	*AKR1C4*	*AP2A1*	*ARHGEF2*	*ATXN7L1*
ABAT	*ADAM32*	*AKR7A2*	*AP2A2*	*ARHGEF5*	*AUH*
ABCB6	*ADAMTS1*	*AKT1*	*AP2B1*	*ARIH2*	*AURKA*
ABCD1	*ADAT1*	*ALB*	*AP2M1*	*ARL1*	*AVPI1*
ABCD4	*ADC*	*ALDH1A1*	*AP3B1*	*ARL3*	*AXL*
ABCF2	*ADCK4*	*ALDH1A2*	*AP3M1*	*ARL5A*	*AZIN1*
ABCG2	*ADFP*	*ALDH3A2*	*AP3S1*	*ARL5B*	*B2M*
ABHD11	*ADH5*	*ALDH3B1*	*AP3S2*	*ARMCX1*	*B3GNT1*
ABHD5	*ADMR*	*ALDH3B2*	*AP4S1*	*ARMCX2*	*B3GNT2*
ABI2	*ADORA1*	*ALDH4A1*	*APBB1*	*ARMCX3*	*B3GNT4*
ABI3	*ADORA2B*	*ALDH5A1*	*APEH*	*ARNT2*	*B3GNT5*
ACAA1	*ADORA3*	*ALDOA*	*APEX1*	*ARPC1B*	*B4GALT3*
ACAD8	*ADRBK1*	*ALDOC*	*APEX2*	*ARPM1*	*B4GALT7*
ACAD9	*ADRM1*	*ALG1*	*API5*	*ARR3*	*BAALC*
ACADL	*ADSL*	*ALG10*	*APLN*	*ARRB2*	*BAAT*
ACADM	*AFP*	*ALG12*	*APLP1*	*ARSA*	*BACE1*
ACADS	*AGA*	*ALG2*	*APOA1*	*ART3*	*BACE2*
ACO2	*AGER*	*ALG3*	*APOA2*	*ASAH1*	*BAD*
ACOT2	*AGMAT*	*ALG9*	*APOC3*	*ASB2*	*BAG3*
ACOT8	*AGPAT2*	*ALKBH1*	*APOD*	*ASB3*	*BAIAP2*
ACP2	*AGPAT5*	*ALS2CR2*	*APOH*	*ASCL1*	*BAK1*
ACPP	*AGT*	*ALS2CR8*	*APOL2*	*ASGR1*	*BANK1*
ACSBG1	*AGTR1*	*AMD1*	*APPBP1*	*ASGR2*	*BANP*
ACSBG2	*AGTR2*	*AMY1A*	*AQP1*	*ASNS*	*BAP1*
ACSL4	*AGTRAP*	*ANAPC2*	*AQP2*	*ASPA*	*BASP1*
ACSM3	*AHCY*	*ANGPTL7*	*AQP3*	*ASPSCR1*	*BAT1*
ACTA1	*AHCYL1*	*ANKH*	*AQP5*	*ATF3*	*BAT3*
ACTA2	*AICDA*	*ANKRA2*	*AQP9*	*ATF4*	*BAT4*
ACTB	*AIFM2*	*ANKRD1*	*ARAF*	*ATG10*	*BAT5*
ACTC1	*AIG1*	*ANKRD5*	*ARC*	*ATG3*	*BBOX1*
ACTG2	*AIPL1*	*ANKRD9*	*AREG*	*ATG4C*	*BBS2*
ACTN1	*AK1*	*ANP32A*	*ARF3*	*ATIC*	*BBS4*
ACTN4	*AK3*	*ANXA1*	*ARF4*	*ATOH7*	*BBS7*
ACTR10	*AK3L1*	*ANXA11*	*ARF6*	*ATP4B*	*BCAT2*
ACTR1A	*AK5*	*ANXA13*	*ARFGAP3*	*ATP5A1*	*BCHE*
ACTR1B	*AKAP10*	*ANXA2*	*ARFIP2*	*ATP5B*	*BCKDK*
ACTR3B	*AKAP8L*	*ANXA4*	*ARG1*	*ATP5G2*	*BCL2*

BCL2A1	*C11orf73*	*C6orf55*	*CASP4*	*CCNG2*	*CD81*
BCL2A1	*C12orf24*	*C7orf36*	*CASP7*	*CCNH*	*CD82*
BCL2L1	*C12orf45*	*C7orf36*	*CASQ1*	*CCNI*	*CD83*
BCL2L13	*C14orf58*	*C8orf38*	*CASQ2*	*CCR6*	*CD84*
BCL7B	*C14orf94*	*C9orf116*	*CAST*	*CCR7*	*CD86*
BCS1L	*C15orf43*	*C9orf156*	*CAV1*	*CCRL2*	*CD8A*
BDKRB2	*C15orf48*	*C9orf71*	*CBFB*	*CCS*	*CD9*
BDNF	*C16orf33*	*C9orf80*	*CBLB*	*CCT3*	*CD93*
BECN1	*C16orf61*	*C9orf89*	*CBLC*	*CCT5*	*CD99L2*
BET1	*C18orf21*	*CA1*	*CBR1*	*CCT7*	*CDC20*
BFAR	*C18orf55*	*CA12*	*CBR3*	*CD160*	*CDC23*
BGLAP	*C19orf28*	*CA14*	*CBS*	*CD164*	*CDC25A*
BGN	*C19orf43*	*CA2*	*CBX6*	*CD177*	*CDC27*
BHMT2	*C19orf53*	*CA7*	*CCDC102A*	*CD19*	*CDC34*
BIK	*C19orf6*	*CA9*	*CCDC104*	*CD1A*	*CDC34*
BIRC2	*C1orf128*	*CACNG4*	*CCDC107*	*CD2*	*CDC37*
BIRC4	*C1orf162*	*CACNG6*	*CCDC110*	*CD200*	*CDC42*
BIRC7	*C1orf182*	*CAGE1*	*CCDC120*	*CD207*	*CDC42EP1*
BLZF1	*C1orf62*	*CALM2*	*CCDC120*	*CD24*	*CDC42EP2*
BMP7	*C1orf76*	*CALML3*	*CCDC16*	*CD247*	*CDC42SE1*
BMPR1A	*C1QB*	*CALN1*	*CCDC22*	*CD300C*	*CDC45L*
BMX	*C1QBP*	*CALR3*	*CCDC47*	*CD300LB*	*CDC5L*
BNIP2	*C1QBP*	*CALU*	*CCDC54*	*CD33*	*CDC6*
BOK	*C1QTNF1*	*CAMK1D*	*CCDC72*	*CD34*	*CDCA4*
BOP1	*C1QTNF2*	*CAMK1G*	*CCDC86*	*CD36*	*CDCA5*
BPIL1	*C1QTNF6*	*CAMK2D*	*CCIN*	*CD37*	*CDH13*
BRE	*C22orf28*	*CAMK2G*	*CCKBR*	*CD4*	*CDH15*
BTBD1	*C2orf15*	*CAMK2G*	*CCL11*	*CD40*	*CDH16*
BTBD14A	*C2orf18*	*CAMK4*	*CCL13*	*CD46*	*CDH18*
BTF3	*C2orf30*	*CAMKK1*	*CCL19*	*CD47*	*CDIPT*
BTG1	*C2orf4*	*CAMKK2*	*CCL2*	*CD48*	*CDK2*
BUB3	*C2orf40*	*CAMKV*	*CCL21*	*CD5*	*CDK2AP1*
BUD13	*C2orf49*	*CAMP*	*CCL22*	*CD52*	*CDK2AP2*
BXDC2	*C2orf52*	*CANX*	*CCL5*	*CD55*	*CDK4*
BYSL	*C3AR1*	*CAP1*	*CCNA1*	*CD59*	*CDK9*
BZW1	*C3orf37*	*CAPG*	*CCND1*	*CD5L*	*CDK9*
BZW2	*C4BPA*	*CAPN1*	*CCND2*	*CD63*	*CDKN2C*
C10orf10	*C4orf14*	*CAPN10*	*CCND3*	*CD68*	*CDKN2D*
C10orf33	*C4orf22*	*CAPN6*	*CCNDBP1*	*CD69*	*CDR2*
C10orf63	*C5AR1*	*CAPZA1*	*CCNE1*	*CD7*	*CDS2*
C11orf45	*C5orf15*	*CASP2*	*CCNE2*	*CD7*	*CDSN*
C11orf64	*C6*	*CASP3*	*CCNG1*	*CD74*	*CEACAM5*

CEACAM8	*CITED2*	*COG8*	*CPT1C*	*CTGF*	*CYP3A5*
CEBPG	*CKLF*	*COIL*	*CPXCR1*	*CTNNA1*	*CYP3A5*
CENPH	*CKM*	*COL8A1*	*CRABP1*	*CTNS*	*CYP4B1*
CENPK	*CKMT1A*	*COL9A3*	*CRADD*	*CTSA*	*CYP4F11*
CES1	*CKS1B*	*COMMD9*	*CREB1*	*CTSD*	*CYP4F12*
CES2	*CKS1B*	*COMT*	*CREB3L1*	*CTSF*	*CYP4X1*
CESK1	*CLCC1*	*COPB1*	*CREG1*	*CTSG*	*CYP51A1*
CETN1	*CLCF1*	*COPE*	*CRIP1*	*CTSH*	*CYR61*
CFB	*CLCN7*	*COPS5*	*CRIP2*	*CTSK*	*CYTL1*
CFHR1	*CLDN1*	*COPS7A*	*CRK*	*CTSL*	*D21S2056E*
CFHR2	*CLDN10*	*COPZ2*	*CRMP1*	*CTSL2*	*DAD1*
CFLAR	*CLDN11*	*COQ2*	*CRSP9*	*CUGBP1*	*DAG1*
CFP	*CLDN14*	*COQ6*	*CRTAP*	*CUL4B*	*DAO*
CGB	*CLDN2*	*CORO1B*	*CRTC1*	*CX3CL1*	*DAPP1*
CGREF1	*CLDN8*	*CORO2B*	*CRX*	*CX3CR1*	*DARC*
CH25H	*CLDND2*	*COX10*	*CRYAB*	*CXADR*	*DARS*
CHEK1	*CLEC1A*	*COX11*	*CRYM*	*CXCL1*	*DBP*
CHEK2	*CLEC2B*	*COX15*	*CRYZL1*	*CXCL10*	*DBR1*
CHEK2	*CLEC3B*	*COX4I1*	*CSDA*	*CXCL11*	*DC2*
CHES1	*CLEC4D*	*COX5A*	*CSF2RA*	*CXCL5*	*DCAMKL2*
CHFR	*CLEC4M*	*COX5B*	*CSF3*	*CXCL6*	*DCBLD2*
CHGA	*CLGN*	*COX6A1*	*CSH1*	*CXCR3*	*DCC1*
CHI3L1	*CLIC1*	*COX6A2*	*CSK*	*CXCR6*	*DCK*
CHI3L2	*CLIC2*	*COX6B1*	*CSNK1A1L*	*CXCR7*	*DCN*
CHIC2	*CLIC3*	*COX6C*	*CSNK1E*	*CXorf20*	*DCP1A*
CHMP2A	*CLIC5*	*COX7A1*	*CSNK1E*	*CXorf38*	*DCT*
CHMP5	*CLPP*	*COX7A2L*	*CSNK1G2*	*CXXC1*	*DCTN3*
CHMP5	*CLPTM1*	*COX7C*	*CSNK2A1*	*CYB5A*	*DCTN4*
CHN1	*CLPTM1L*	*CP110*	*CSNK2A2*	*CYB5D2*	*DDIT4*
CHODL	*CLSTN1*	*CPA1*	*CSNK2B*	*CYB5R4*	*DDOST*
CHP	*CLU*	*CPA2*	*CSRP2*	*CYBA*	*DDT*
CHRAC1	*CLUL1*	*CPA3*	*CSRP2BP*	*CYBB*	*DDX20*
CHRM1	*CMPK*	*CPA6*	*CSRP3*	*CYC1*	*DDX21*
CHRNA5	*CNIH4*	*CPB1*	*CST3*	*CYCS*	*DDX23*
CHRNB1	*CNKSR1*	*CPE*	*CST7*	*CYFIP1*	*DDX24*
CHST10	*CNN3*	*CPEB3*	*CST9L*	*CYP11A1*	*DDX39*
CHST12	*CNOT10*	*CPM*	*CSTB*	*CYP1A1*	*DDX5*
CHST4	*CNP*	*CPN1*	*CSTF1*	*CYP20A1*	*DDX50*
CHST8	*CNTFR*	*CPNE1*	*CSTF2*	*CYP20A1*	*DDX56*
CIB1	*CNTROB*	*CPNE4*	*CTBP1*	*CYP2C8*	*DEFA1*
CILP	*COG1*	*CPNE6*	*CTCF*	*CYP2J2*	*DENND1B*
CITED1	*COG2*	*CPSF1*	*CTDSPL2*	*CYP2S1*	*DES*

DXI	*DNMT3L*	*E2F4*	*EIF3S6IP*	*ERAF*	*FAM13C1*
DGAT1	*DNPEP*	*EAF2*	*EIF3S7*	*ERAL1*	*FAM46D*
DGCR14	*DNTT*	*EBAG9*	*EIF3S8*	*ERCC1*	*FAM48A*
DGCR2	*DOK2*	*EBI2*	*EIF4A1*	*ERCC3*	*FAM55C*
DGKA	*DOK4*	*EBP*	*EIF4A2*	*ERCC5*	*FAM55D*
DHCR24	*DOM3Z*	*EBPL*	*EIF4E*	*ERCC8*	*FAM76B*
DHCR7	*DPF2*	*ECD*	*EIF4EBP1*	*EREG*	*FAM83G*
DHDH	*DPH1*	*ECGF1*	*EIF4EBP2*	*ERRFI1*	*FAM84A*
DHFR	*DPP10*	*ECH1*	*EIF4ENIF1*	*ESAM*	*FAM84A*
DHRS3	*DPP3*	*ECM1*	*EIF4H*	*ET*	*FAM89B*
DHRSX	*DPP7*	*EDEM1*	*EIF5*	*ETF1*	*FANCC*
DHX16	*DPP8*	*EDF1*	*EIF5A*	*ETS2*	*FANCG*
DHX30	*DPP9*	*EDG1*	*EIF5A2*	*ETV3*	*FAP*
DHX32	*DPT*	*EDG8*	*ELA3A*	*ETV4*	*FARSLA*
DHX38	*DRD1*	*EDN1*	*ELAC1*	*ETV5*	*FAS*
DHX40	*DRD2*	*EDNRA*	*ELAVL2*	*EVI1*	*FASLG*
DIABLO	*DRD5*	*EEF1A1*	*ELF1*	*EVI2B*	*FASTK*
DIDO1	*DRG1*	*EEF1A2*	*ELF2*	*EXOD1*	*FATE1*
DIMT1L	*DRG2*	*EEF1G*	*ELF3*	*EXOSC3*	*FAU*
DIMT1L	*DSCR1*	*EEF2K*	*ELF4*	*EXOSC8*	*FBL*
DIP13B	*DSCR1L1*	*EEFSEC*	*ELF5*	*EXT1*	*FBLIM1*
DIRAS1	*DSCR1L2*	*EFEMP1*	*ELK3*	*EXTL3*	*FBLN1*
DIRAS3	*DSCR2*	*EFHC1*	*ELL2*	*EZH2*	*FBXL13*
DKC1	*DSTN*	*EFNA1*	*ELMO3*	*F11R*	*FBXL2*
DKK1	*DTX2*	*EFNA3*	*ELP4*	*F13A1*	*FBXL5*
DKK3	*DTYMK*	*EFNB1*	*EML1*	*F2R*	*FBXO17*
DLAT	*DULLARD*	*EFTUD2*	*EMP3*	*F2RL1*	*FBXO2*
DLL3	*DUSP10*	*EHD2*	*ENC1*	*F3*	*FBXO5*
DLX3	*DUSP11*	*EHD4*	*ENG*	*F8A1*	*FBXW11*
DMAP1	*DUSP12*	*EI24*	*ENO1*	*FABP1*	*FBXW2*
DMRTC1	*DUSP13*	*EIF1*	*ENO2*	*FABP3*	*FCAR*
DNAJB1	*DUSP14*	*EIF1AY*	*ENO3*	*FABP5*	*FCER1A*
DNAJB11	*DUSP14*	*EIF2A*	*ENPP2*	*FABP7*	*FCGR2B*
DNAJB4	*DUSP19*	*EIF2B2*	*ENPP4*	*FADS1*	*FCGRT*
DNAJB5	*DUSP22*	*EIF2B5*	*ENTPD2*	*FADS3*	*FCN1*
DNAJB9	*DUSP3*	*EIF2S1*	*ENTPD4*	*FAF1*	*FCN3*
DNAJC7	*DUSP6*	*EIF2S2*	*EPB41L1*	*FAIM*	*FCRL3*
DNASE1	*DUT*	*EIF2S3*	*EPB49*	*FAM103A1*	*FCRLA*
DNASE1L1	*DVL3*	*EIF3S1*	*EPHA2*	*FAM107A*	*FDFT1*
DNASE1L3	*DYNLT1*	*EIF3S2*	*EPHX1*	*FAM122A*	*FDPS*
DND1	*DYRK1B*	*EIF3S4*	*EPHX2*	*FAM126A*	*FEM1C*
DNM1L	*DYRK4*	*EIF3S6*	*EPSTI1*	*FAM126B*	*FES*

FEV	*FNDC4*	*GAD1*	*GGTLA1*	*GNGT1*	*GPS2*
FEZ1	*FNTB*	*GADD45A*	*GHRL*	*GNL1*	*GPSM2*
FFAR3	*FOLH1*	*GAGE1*	*GIF*	*GNLY*	*GPSN2*
FGD2	*FOLR1*	*GAGE2*	*GIPC1*	*GNMT*	*GPT*
FGF10	*FOS*	*GAL3ST1*	*GJA1*	*GNPTG*	*GPX7*
FGF12	*FOSB*	*GALC*	*GJA5*	*GOLGA7*	*GRAP*
FGF13	*FOSL1*	*GALE*	*GJB1*	*GORASP2*	*GRAP2*
FGF16	*FOSL2*	*GALK2*	*GJB2*	*GOSR2*	*GRB10*
FGF18	*FOXA2*	*GALNACT-2*	*GJB4*	*GOT2*	*GRB2*
FGF21	*FPRL1*	*GALNT10*	*GK*	*GP1BA*	*GRB7*
FGF7	*FRAT2*	*GALNT6*	*GK2*	*GP9*	*GRM3*
FGFR1	*FRK*	*GALNTL2*	*GK5*	*GPAA1*	*GRP*
FGFR1OP	*FRS3*	*GALNTL5*	*GKAP1*	*GPBAR1*	*GRPEL1*
FGFR4	*FRZB*	*GAP43*	*GLA*	*GPBP1L1*	*GRWD1*
FGFRL1	*FSCN1*	*GAPDH*	*GLB1*	*GPC2*	*GSG1*
FHIT	*FSCN3*	*GAPDHS*	*GLE1L*	*GPC3*	*GSG1L*
FHL2	*FSD1*	*GARS*	*GLIPR1L1*	*GPC4*	*GSK3A*
FHL3	*FST*	*GAS7*	*GLIPR1L2*	*GPD1*	*GSK3B*
FHL3	*FTH1*	*GATA2*	*GLRA3*	*GPHN*	*GSS*
FHL5	*FTL*	*GBAS*	*GLRB*	*GPI*	*GSTA2*
FKBP10	*FTSJ3*	*GBL*	*GLRX*	*GPM6A*	*GSTA3*
FKBP14	*FUBP1*	*GBP1*	*GLTSCR2*	*GPM6B*	*GSTA4*
FKBP2	*FUK*	*GBP5*	*GLUL*	*GPNMB*	*GSTM4*
FKBP3	*FUS*	*GCG*	*GLYAT*	*GPR114*	*GSTP1*
FKBP4	*FUT11*	*GCH1*	*GM2A*	*GPR135*	*GSTT1*
FKBP8	*FUT4*	*GCK*	*GMCL1L*	*GPR146*	*GTF2A2*
FKBPL	*FUT9*	*GCLM*	*GMDS*	*GPR155*	*GTF2F1*
FKHL18	*FXYD5*	*Gcom1*	*GMFB*	*GPR157*	*GTF2H1*
FKRP	*FXYD6*	*GDF3*	*GMFG*	*GPR161*	*GTF2H4*
FLAD1	*FXYD7*	*GDI1*	*GMPPA*	*GPR17*	*GTF2IRD1*
FLI1	*FZD9*	*GDI2*	*GMPS*	*GPR171*	*GTF3C5*
FLII	*G0S2*	*GDNF*	*GNA13*	*GPR173*	*GTPBP3*
FLJ10324	*G3BP1*	*GEM*	*GNA14*	*GPR3*	*GTPBP8*
FLJ10815	*G6PD*	*GENX-3414*	*GNAI1*	*GPR30*	*GTSE1*
FLJ35773	*GABARAPL1*	*GFI1*	*GNAI2*	*GPR34*	*GUCA1A*
FLOT1	*GABARAPL2*	*GFPT2*	*GNAI3*	*GPR55*	*GUK1*
FLOT2	*GABPB2*	*GFRA1*	*GNAT2*	*GPR65*	*GUSB*
FLRT3	*GABRA3*	*GFRA3*	*GNB2L1*	*GPR84*	*GYG2*
FMO2	*GABRA4*	*GGA2*	*GNB3*	*GPR89A*	*GYLTL1B*
FMO3	*GABRA5*	*GGCX*	*GNG10*	*GPR92*	*GZMA*
FMO4	*GABRB1*	*GGN*	*GNG11*	*GPRC5B*	*GZMH*
FMOD	*GABRG1*	*GGT1*	*GNG3*	*GPRC5C*	*GZMK*

H1F0	*HIST3H3*	*HOXA5*	*ICAM2*	*IKZF5*	*ISL1*
H2AFV	*HK1*	*HOXA9*	*ICAM3*	*IL10*	*ISOC1*
H2AFZ	*HK2*	*HOXB13*	*ICAM4*	*IL10RA*	*ISYNA1*
H3F3A	*HK3*	*HOXD1*	*ICMT*	*IL10RB*	*ITFG1*
HAAO	*HLA-B*	*HOXD3*	*ICT1*	*IL11RA*	*ITFG2*
HABP2	*HLA-C*	*HPCAL4*	*ID1*	*IL13RA1*	*ITGAX*
HACL1	*HLA-DMB*	*HPD*	*ID2*	*IL13RA2*	*ITGB1*
HADH	*HLA-DOA*	*HPN*	*ID3*	*IL15*	*ITGB1BP1*
HADHA	*HLA-DOB*	*HPRT1*	*IDH1*	*IL15RA*	*ITGB2*
HADHB	*HLA-DPA1*	*HPS3*	*IDH2*	*IL17RA*	*ITGB4BP*
HAGH	*HLA-DPB1*	*HRASLS3*	*IDH3A*	*IL17RB*	*ITGB5*
HAO2	*HLA-DQA1*	*HRSP12*	*IDH3B*	*IL18*	*ITGB7*
HAS1	*HLA-DQB1*	*HSD17B2*	*IDH3B*	*IL1A*	*ITM2B*
HAVCR1	*HLA-DRA*	*HSD17B3*	*IER2*	*IL1B*	*ITPA*
HBA1	*HLA-E*	*HSD17B4*	*IER3IP1*	*IL1F5*	*ITPK1*
HBB	*HLA-G*	*HSD3B1*	*IFI16*	*IL1RN*	*IVD*
HBG1	*HLX1*	*HSD3B1*	*IFI27*	*IL21R*	*IZUMO1*
HBZ	*HM13*	*HSF1*	*IFI30*	*IL24*	*JAGN1*
HCLS1	*HMGB2*	*HSF4*	*IFI35*	*IL27RA*	*JPH3*
HDAC1	*HMGCL*	*HSH2D*	*IFIT1*	*IL2RB*	*JUN*
HDAC3	*HMGCR*	*HSP90AB1*	*IFITM1*	*IL2RG*	*JUNB*
HDGF2	*HMGN1*	*HSPA1A*	*IFITM2*	*IL6*	*JUP*
HEBP1	*HMOX1*	*HSPA2*	*IFITM3*	*IL8*	*K-ALPHA-1*
HEMK1	*HMOX1*	*HSPA5*	*IFNA2*	*IL8RA*	*KARS*
HES1	*HMOX2*	*HSPA8*	*IFNAR1*	*IL8RB*	*KATNAL1*
HES4	*HN1*	*HSPB1*	*IFNAR2*	*ILF2*	*KAZALD1*
HEXIM1	*HN1*	*HSPB7*	*IFT57*	*ILK*	*KBTBD10*
HGS	*HNF4G*	*HSPB8*	*IGBP1*	*ILKAP*	*KBTBD7*
HHLA2	*HNMT*	*HSPBAP1*	*IGF1R*	*IMP4*	*KCMF1*
HIBADH	*HNRPD*	*HSPBP1*	*IGF2*	*IMPDH2*	*KCNA3*
HIF1AN	*HNRPDL*	*HSPC152*	*IGF2*	*INA*	*KCNE1*
HIGD2A	*HNRPF*	*HSPC171*	*IGFALS*	*INHA*	*KCNE1L*
HINT3	*HNRPK*	*HSPC171*	*IGFBP3*	*INHBA*	*KCNE2*
HIST1H2AD	*HNRPR*	*HSPE1*	*IGFBP6*	*INPP1*	*KCNG4*
HIST1H2AG	*HNRPU*	*HSPH1*	*IGL@*	*INPP4A*	*KCNIP3*
HIST1H2AK	*HOMER1*	*HTR1D*	*IGL@*	*IPO11*	*KCNIP3*
HIST1H2AM	*HOMER2*	*HTR3A*	*IGLL1*	*IQWD1*	*KCNJ10*
HIST1H3B	*HOMER3*	*HTRA2*	*IGSF4B*	*IRAK4*	*KCNJ13*
HIST1H3E	*HOOK1*	*HTRA3*	*IGSF6*	*IRF1*	*KCNJ3*
HIST1H3F	*HOOK2*	*HYAL1*	*IGSF9*	*IRF3*	*KCNJ8*
HIST1H4H	*HOXA1*	*HYAL2*	*IHPK2*	*IRF4*	*KCNK1*
HIST2H2AA3	*HOXA10*	*ICAM1*	*IKBKG*	*ISCU*	*KCNK13*

KCNK17	*KLK7*	*LCP1*	*LPXN*	*MAL2*	*MCHR1*
KCNK6	*KLRC1*	*LCP2*	*LRCH3*	*MALL*	*MCL1*
KCNMB2	*KLRC4*	*LDB1*	*LRCH4*	*MALT1*	*MCM3*
KCNN4	*KLRD1*	*LDB2*	*LRFN3*	*MAN1B1*	*MCM7*
KCNS3	*KLRG1*	*LDHA*	*LRP12*	*MAN2B1*	*MCMDC1*
KCNV1	*KNS2*	*LDLR*	*LRRC29*	*MAOA*	*MCOLN1*
KCTD13	*KNTC2*	*LECT1*	*LRRC39*	*MAP1LC3A*	*MDH2*
KCTD14	*KPNA1*	*LENG1*	*LRRC59*	*MAP1LC3B*	*MDK*
KCTD4	*KPNA2*	*LENG1*	*LRRN5*	*MAP2K1IP1*	*ME3*
KDELC1	*KPNA3*	*LENG8*	*LSAMP*	*MAP2K2*	*MED28*
KDELR1	*KPNA4*	*LEPROTL1*	*LSM2*	*MAP2K3*	*MED4*
KDELR2	*KPNA6*	*LGALS1*	*LSM7*	*MAP2K5*	*MED6*
KDELR3	*KPNB1*	*LGALS2*	*LSMD1*	*MAP2K6*	*MEST*
KEAP1	*KRAS*	*LGALS3*	*LSP1*	*MAP2K7*	*METAP2*
KEL	*KREMEN2*	*LGALS3BP*	*LSS*	*MAP3K14*	*METTL6*
KERA	*KRR1*	*LGALS8*	*LTA*	*MAP3K7*	*MFAP1*
KHDRBS1	*KRT10*	*LGI1*	*LTB4R*	*MAP4K5*	*MFAP2*
KHDRBS3	*KRT17*	*LGMN*	*LTBR*	*MAP7*	*MFN2*
KIF22	*KRT18*	*LGR4*	*LTC4S*	*MAPBPIP*	*MFSD1*
KIF2A	*KRT20*	*LGTN*	*LTF*	*MAPK1*	*MGAT1*
KIF2C	*KRT23*	*LHFP*	*LUM*	*MAPK11*	*MGC29506*
KIF9	*KRT33B*	*LHFPL5*	*LY6H*	*MAPK12*	*MGC33302*
KIFAP3	*KRT5*	*LHX4*	*LYPD1*	*MAPK13*	*MGC34821*
KIFC3	*KRT5*	*LIG4*	*LYPLA1*	*MAPK15*	*MGC35154*
KIR3DL1	*KRT5*	*LILRA3*	*LYSMD2*	*MAPK3*	*MGC35361*
KISS1	*KRT7*	*LILRB1*	*LYZ*	*MAPK6*	*MGC39372*
KLC2	*KRT8*	*LILRB3*	*M6PRBP1*	*MAPK9*	*MGC40069*
KLC4	*KRTCAP2*	*LIN28*	*MAB21L1*	*MAPKAPK3*	*MGC40574*
KLF12	*KSP37*	*LMAN1*	*MAD2L2*	*MAPRE2*	*MGC42105*
KLF15	*L3MBTL2*	*LMBR1L*	*MAGEA1*	*MAPRE3*	*MGC45491*
KLF4	*LAD1*	*LMCD1*	*MAGEA10*	*MARCKSL1*	*MGC5139*
KLF5	*LAIR1*	*LMO2*	*MAGEA2*	*MARCO*	*MGC5139*
KLHDC2	*LALBA*	*LMO6*	*MAGEA3*	*MARK2*	*MGC7036*
KLHL1	*LAMP1*	*LNX1*	*MAGEA6*	*MARK3*	*MGLL*
KLHL2	*LAMP2*	*LOC200420*	*MAGEA9*	*MAT1A*	*MGMT*
KLHL6	*LANCL1*	*LOH12CR1*	*MAGEB1*	*MATR3*	*MGP*
KLK1	*LASP1*	*LONP1*	*MAGEB4*	*MB*	*MGST2*
KLK10	*LASS4*	*LOXL2*	*MAGEC2*	*MBP*	*MIF*
KLK2	*LBP*	*LOXL4*	*MAGED2*	*MC1R*	*MIF4GD*
KLK3	*LBR*	*LPIN1*	*MAGEH1*	*MCEE*	*MINK1*
KLK5	*LCK*	*LPL*	*MAGOH*	*MCF2L*	*MIPEP*
KLK6	*LCN2*	*LPPR2*	*MAL*	*MCFD2*	*MIZF*

MKI67IP	*MS4A4A*	*NANS*	*NFKBIB*	*NRIP1*	*OSGEPL1*
MKNK1	*MS4A6A*	*NAP1L1*	*NFKBIL2*	*NRSN1*	*OSM*
MKRN2	*MS4A7*	*NAP1L3*	*NFYB*	*NSDHL*	*OTX1*
MLANA	*MSH2*	*NAPSA*	*NGB*	*NSFL1C*	*OXCT1*
MLF2	*MSH4*	*NASP*	*NGEF*	*NSL1*	*OXCT2*
MLLT11	*MSI2*	*NAT13*	*NGLY1*	*NSUN5*	*OXR1*
MMD	*MSL3L1*	*NAT5*	*NIF3L1*	*NT5C2*	*OXSR1*
MMP1	*MSMB*	*NCBP1*	*NIP30*	*NTF5*	*P2RX4*
MMP10	*MSN*	*NCBP2*	*NIP7*	*NUCKS1*	*P2RX7*
MMP2	*MSX2*	*NCDN*	*NIPA1*	*NUDT2*	*P2RY12*
MMP9	*MT1A*	*NCF1*	*NIT1*	*NUDT21*	*P2RY2*
MND1	*MT1H*	*NCF2*	*NIT2*	*NUDT3*	*P2RY6*
MOAP1	*MT1M*	*NCKAP1*	*NKG7*	*NUDT4*	*P4HA1*
MOBKL2B	*MT1X*	*NCLN*	*NKIRAS2*	*NUDT6*	*PA2G4*
MOCOS	*MT3*	*NCOA4*	*NKX2-5*	*NUP133*	*PABPC1*
MOCS3	*MTAP*	*NDRG1*	*NKX2-5*	*NUP54*	*PABPC3*
MOG	*MTCH2*	*NDUFA10*	*NME1*	*NUP62*	*PACSIN1*
MOG	*MTERF*	*NDUFA3*	*NME6*	*NXF2*	*PACSIN3*
MORF4L1	*MTHFD1*	*NDUFA6*	*NMNAT3*	*NXF3*	*PAF1*
MOSPD1	*MTHFD2*	*NDUFB3*	*NMT2*	*OAS1*	*PAFAH2*
MOV10	*MTMR9*	*NDUFB4*	*NMU*	*OCLN*	*PAGE1*
MOXD1	*MUC7*	*NDUFB7*	*NMUR2*	*ODC1*	*PAGE4*
MPG	*MVD*	*NDUFB8*	*NNMT*	*OGG1*	*PAICS*
MPP1	*MVK*	*NDUFS2*	*NOC4L*	*OGN*	*PAK1IP1*
MPP2	*MVP*	*NDUFS3*	*NOL3*	*OGT*	*PAK4*
MPP6	*MX2*	*NDUFV1*	*NOLA1*	*OLFM2*	*PAK6*
MPST	*MXD1*	*NDUFV2*	*NOLC1*	*OLFML3*	*PAK7*
MPV17	*MXD3*	*NEDD1*	*NOSTRIN*	*OLIG1*	*PAMCI*
MPZ	*MYBL2*	*NEDD9*	*NOV*	*OLR1*	*PAOX*
MRAS	*MYC*	*NEFL*	*NPC2*	*OMG*	*PAPOLA*
MRAS	*MYCBP*	*NEIL1*	*NPEPL1*	*OPRL1*	*PAPSS1*
MRCL3	*MYF6*	*NEK3*	*NPEPL1*	*OPRS1*	*PAPSS2*
MRFAP1	*MYL2*	*NEU1*	*NPLOC4*	*OPTN*	*PAQR5*
MRP63	*MYL6B*	*NEUROD1*	*NPM1*	*OR2C3*	*PARD6A*
MRPL10	*MYNN*	*NEUROG1*	*NPPA*	*OR51E2*	*PARL*
MRPL13	*MYOC*	*NF2*	*NPPB*	*ORC2L*	*PARP1*
MRPL17	*MYOZ1*	*NFATC3*	*NPTX2*	*ORC5L*	*PAX6*
MRPL28	*MYST2*	*NFE2*	*NPY*	*ORMDL1*	*PAX8*
MRPL40	*NAGA*	*NFE2L1*	*NPY1R*	*OS9*	*PBK*
MRPS11	*NAGK*	*NFIB*	*NQO1*	*OSBPL2*	*PCDHB14*
MRPS5	*NALP12*	*NFIC*	*NR4A1*	*OSCAR*	*PCDHB5*
MS4A1	*NALP2*	*NFIL3*	*NRBP1*	*OSGEP*	*PCDHGC3*

PCK1	*PEX14*	*PIR*	*PNKP*	*PPP1R14A*	*PRPH*
PCNA	*PEX19*	*PISD*	*PNLIPRP2*	*PPP1R14C*	*PRPS1*
PCNP	*PEX3*	*PITX2*	*PNOC*	*PPP2R1A*	*PRR14*
PCOLCE	*PEX5*	*PITX3*	*PNOC*	*PPP2R1B*	*PRRG4*
PCSK2	*PEX7*	*PKD2L1*	*PNPLA4*	*PPP2R2B*	*PRSS1*
PCSK4	*PFKFB4*	*PKIA*	*PNPLA8*	*PPP2R4*	*PRSS8*
PCSK5	*PFKL*	*PKNOX1*	*PODN*	*PPP2R5A*	*PRUNE*
PCSK7	*PFKM*	*PKP3*	*POLA2*	*PPP4C*	*PSAP*
PCTK2	*PFKP*	*PLA2G12A*	*POLD1*	*PPT1*	*PSAT1*
PCTK3	*PFN1*	*PLA2G2A*	*POLD2*	*PPY*	*PSCD2*
PCYOX1	*PGAM1*	*PLA2G3*	*POLI*	*PPYR1*	*PSCD3*
PDCD10	*PGCP*	*PLA2G4D*	*POLR2C*	*PQBP1*	*PSEN2*
PDCD2	*PGDS*	*PLAC1*	*POLR2I*	*PRAME*	*PSENEN*
PDCD4	*PGF*	*PLAT*	*POLR2J2*	*PRC1*	*PSG1*
PDCD5	*PGLS*	*PLAU*	*POLR2K*	*PRCP*	*PSG9*
PDCD6	*PGM1*	*PLAUR*	*POLR2L*	*PRDM4*	*PSMA1*
PDCL	*PGRMC1*	*PLCD4*	*POLR2L*	*PRDX1*	*PSMA3*
PDE1A	*PGRMC2*	*PLCG2*	*POLR3D*	*PRDX2*	*PSMA4*
PDE1B	*PGS1*	*PLEK*	*POLR3F*	*PRDX4*	*PSMA6*
PDE4A	*PHB*	*PLEK*	*POLR3K*	*PRELP*	*PSMA7*
PDE6B	*PHF1*	*PLEK2*	*POMC*	*PREP*	*PSMB2*
PDE6D	*PHF10*	*PLEKHA1*	*POMP*	*PRG1*	*PSMB4*
PDGFRL	*PHF11*	*PLEKHA8*	*POP4*	*PRG2*	*PSMB6*
PDHA1	*PHF17*	*PLEKHB2*	*PORCN*	*PRKACG*	*PSMC1*
PDHB	*PHF7*	*PLEKHF2*	*POT1*	*PRKAR1A*	*PSMC3*
PDIA3	*PHKA2*	*PLEKHO1*	*POU2AF1*	*PRKCDBP*	*PSMC4*
PDIA4	*PHYH*	*PLK1*	*PPAP2B*	*PRKCH*	*PSMC5*
PDIA6	*PHYH*	*PLK2*	*PPARG*	*PRKCZ*	*PSMC6*
PDK1	*PI3*	*PLK4*	*PPBP*	*PRL*	*PSMD12*
PDK4	*PI4KII*	*PLN*	*PPCDC*	*PRM1*	*PSMD13*
PDLIM5	*PIAS2*	*PLS3*	*PPFIBP2*	*PRMT2*	*PSMD5*
PDLIM7	*PIAS3*	*PLSCR1*	*PPIA*	*PRMT5*	*PSME1*
PDPK1	*PIAS4*	*PLSCR3*	*PPIB*	*PRMT8*	*PSME2*
PEA15	*PICK1*	*PLTP*	*PPID*	*PRNP*	*PSME3*
PELO	*PIGF*	*PMCH*	*PPIH*	*PROC*	*PSTPIP1*
PENK	*PIGQ*	*PMP2*	*PPIL1*	*PROCR*	*PSTPIP2*
PEPD	*PIK3R3*	*PMP22*	*PPIL2*	*PROK1*	*PTDSS1*
PEPP-2	*PIM2*	*PMP22*	*PPIL5*	*PROM1*	*PTDSS2*
PEX10	*PIN1*	*PMPCB*	*PPM1A*	*PROS1*	*PTEN*
PEX11A	*PINK1*	*PMS2L3*	*PPM1G*	*ProSAPiP1*	*PTGES2*
PEX11B	*PIP5K2C*	*PMVK*	*PPME1*	*PRPF19*	*PTGES3*
PEX12	*PIPOX*	*PNCK*	*PPOX*	*PRPF3*	*PTGFR*

PTGS1	*RAB9*	*RCN1*	*RNF182*	*RPS14*	*S100A11*
PTHLH	*RABAC1*	*RCVRN*	*RNF183*	*RPS16*	*S100A16*
PTMA	*RABEPK*	*RDBP*	*RNF2*	*RPS17*	*S100A4*
PTMS	*RABGGTA*	*REG3A*	*RNF26*	*RPS2*	*S100A6*
PTP4A1	*RABL2B*	*REG4*	*RNF38*	*RPS20*	*S100A7*
PTPN1	*RAC2*	*RENBP*	*RNF40*	*RPS21*	*S100A8*
PTPN6	*RAC3*	*REPS1*	*RNF6*	*RPS25*	*S100P*
PTPRA	*RACGAP1*	*RERG*	*RNF7*	*RPS26*	*SAA1*
PTS	*RAD1*	*REXO4*	*RNF8*	*RPS28*	*SAA2*
PTTG1	*RAD17*	*RFC3*	*RNPEP*	*RPS29*	*SAE1*
PTTG1IP	*RAD18*	*RFC4*	*RNPS1*	*RPS3*	*SAMM50*
PUS1	*RAD23B*	*RFX2*	*RNUXA*	*RPS3A*	*SARS*
PUS3	*RAD51L1*	*RFX4*	*RORC*	*RPS4X*	*SATB1*
PVR	*RAD51L3*	*RFX5*	*RP11-444E17.2*	*RPS4Y1*	*SC4MOL*
PVRL4	*RAD54L2*	*RFXAP*	*RPA2*	*RPS5*	*SCAMP3*
PWP1	*RAG2*	*RGR*	*RPL10*	*RPS6*	*SCAMP4*
PX19	*RAMP1*	*RGS1*	*RPL10A*	*RPS6KA1*	*SCARB2*
PYCR2	*RAMP2*	*RGS13*	*RPL11*	*RPS6KA2*	*SCG5*
PYGB	*RAN*	*RGS18*	*RPL13*	*RPS7*	*SCGB1A1*
PYGO2	*RAPGEF3*	*RGS19*	*RPL14*	*RPS9*	*SCGB1A1*
QDPR	*RAPGEF4*	*RGS2*	*RPL17*	*RPSA*	*SCGB3A2*
RAB11A	*RARA*	*RGS5*	*RPL18*	*RPUSD1*	*SCHIP1*
RAB11B	*RARRES1*	*RGS5*	*RPL23*	*RRAD*	*SCIN*
RAB11FIP5	*RARRES2*	*RGS7*	*RPL24*	*RRAGC*	*SCLY*
RAB13	*RARRES3*	*RGS7*	*RPL26L1*	*RRAS*	*SCMH1*
RAB1A	*RASD2*	*RHAG*	*RPL28*	*RRM1*	*SCN2B*
RAB20	*RASSF5*	*RHOBTB1*	*RPL31*	*RRM2*	*SCNM1*
RAB24	*RB1*	*RHOC*	*RPL32*	*RRP9*	*SCO1*
RAB27A	*RBBP4*	*RHOD*	*RPL34*	*RSAD2*	*SCYE1*
RAB2B	*RBBP5*	*RHOH*	*RPL35*	*RSPO3*	*SCYL3*
RAB30	*RBBP7*	*RHPN2*	*RPL36*	*RSU1*	*SDC1*
RAB31	*RBBP8*	*RICTOR*	*RPL39*	*RTN4R*	*SDC3*
RAB32	*RBL1*	*RILP*	*RPL4*	*RTP1*	*SDC4*
RAB34	*RBM38*	*RIN1*	*RPL41*	*RTP4*	*SDCCAG3*
RAB35	*RBM8A*	*RLN1*	*RPL7A*	*RUSC1*	*SDF2*
RAB38	*RBMX*	*RNASEH1*	*RPL9*	*RUTBC3*	*SDHD*
RAB3IL1	*RBMY1F*	*RNASEH2C*	*RPLP0*	*RUVBL1*	*SDPR*
RAB4A	*RBP1*	*RND3*	*RPN2*	*RUVBL2*	*SDSL*
RAB5A	*RBP4*	*RNF10*	*RPP30*	*RWDD1*	*SEC22B*
RAB5B	*RBP5*	*RNF12*	*RPS10*	*RWDD2*	*SEC22C*
RAB7L1	*RCBTB2*	*RNF125*	*RPS11*	*RYBP*	*SEC23A*
RAB8A	*RCC1*	*RNF138*	*RPS13*	*S100A10*	*SEC23B*

SEC24D	*SGCA*	*SLC22A4*	*SLC6A13*	*SNUPN*	*SSB*
SELL	*SGCB*	*SLC22A5*	*SLC6A8*	*SNX1*	*SSBP1*
SELPLG	*SGCD*	*SLC22A6*	*SLC7A3*	*SNX10*	*SSBP2*
SEMA3C	*SGCE*	*SLC22A7*	*SLC7A4*	*SNX2*	*SSR3*
SEMA3D	*SH2D1A*	*SLC22A8*	*SLC7A5*	*SNX4*	*SSR4*
SEMA4F	*SH2D1B*	*SLC23A3*	*SLC7A7*	*SNX8*	*SSRP1*
SENP5	*SH2D3A*	*SLC25A1*	*SLC7A9*	*SNX9*	*SSTR1*
SENP8	*SH3BGRL3*	*SLC25A1*	*SLCO3A1*	*SOAT1*	*SSTR2*
SEPHS1	*SH3BP1*	*SLC25A10*	*SLCO4A1*	*SOCS3*	*SSU72*
SEPTIN4	*SH3GL1*	*SLC25A11*	*SLCO6A1*	*SOCS5*	*SSX2IP*
SEPTIN9	*SH3GL2*	*SLC25A13*	*SLFN5*	*SOD2*	*ST14*
SERF1A	*SH3GLB1*	*SLC25A15*	*SLPI*	*SOD3*	*ST3GAL1*
SERF2	*SH3KBP1*	*SLC25A16*	*SLU7*	*SORD*	*ST5*
SERINC3	*SH3PX3*	*SLC25A17*	*SMA3*	*SORT1*	*ST7*
SERPINA10	*SHC3*	*SLC25A18*	*SMA3*	*SOX10*	*STAC*
SERPINA4	*SHMT2*	*SLC25A19*	*SMAD1*	*SOX15*	*STAM2*
SERPINA5	*SIAH1*	*SLC25A20*	*SMAD2*	*SP1*	*STAMBP*
SERPINA7	*SIGIRR*	*SLC25A22*	*SMAD2*	*SPAG6*	*STAR*
SERPINB2	*SIGLEC9*	*SLC25A4*	*SMAF1*	*SPAM1*	*STAR*
SERPINB3	*SIRT6*	*SLC27A3*	*SMARCAL1*	*SPANXB1*	*STARD3*
SERPINB4	*SIRT7*	*SLC29A1*	*SMARCE1*	*SPARC*	*STARD3NL*
SERPINB6	*SKAP2*	*SLC2A3*	*SMC5*	*SPATA2*	*STAT3*
SERPINB9	*SKIP*	*SLC2A5*	*SMCP*	*SPATA22*	*STAT4*
SERPIND1	*SKP1A*	*SLC2A6*	*SMN1*	*SPATA6*	*STAT5A*
SERPINE1	*SLBP*	*SLC2A8*	*SMO*	*SPFH1*	*STAT5A*
SERPINE2	*SLC12A4*	*SLC2A9*	*SMOC1*	*SPIN*	*STATH*
SERPINF2	*SLC13A3*	*SLC30A3*	*SMPDL3A*	*SPINK2*	*STAU2*
SERPING1	*SLC13A4*	*SLC31A1*	*SMTN*	*SPINT1*	*STC1*
SERTAD1	*SLC16A1*	*SLC31A2*	*SMU1*	*SPN*	*STC2*
SERTAD4	*SLC16A5*	*SLC33A1*	*SMYD3*	*SPOCK1*	*STIM1*
SESN2	*SLC16A7*	*SLC35B2*	*SNAI1*	*SPOCK2*	*STK11*
SET	*SLC17A3*	*SLC38A2*	*SNAP23*	*SPP1*	*STK11IP*
SF3A1	*SLC18A1*	*SLC38A5*	*SNAPC1*	*SPR*	*STK17B*
SF3A2	*SLC19A1*	*SLC39A1*	*SNCB*	*SPTLC2*	*STK25*
SF3B4	*SLC19A3*	*SLC39A7*	*SNCG*	*SPZ1*	*STK32B*
SFMBT1	*SLC1A1*	*SLC39A8*	*SND1*	*SQLE*	*STK36*
SFN	*SLC1A4*	*SLC3A1*	*SNF1LK*	*SQRDL*	*STK38L*
SFRP4	*SLC1A5*	*SLC40A1*	*SNN*	*SRD5A1*	*STMN2*
SFRS2	*SLC1A7*	*SLC44A5*	*SNRPA1*	*SRPK2*	*STMN3*
SFRS4	*SLC20A1*	*SLC4A2*	*SNRPB2*	*SRPR*	*STMN4*
SFTPD	*SLC20A2*	*SLC5A12*	*SNRPD2*	*SRrp35*	*STOM*
SFXN2	*SLC22A15*	*SLC6A1*	*SNTA1*	*SS18*	*STOML2*

STON1	*TAGLN*	*TGM4*	*TMEM153*	*TPI1*	*TRIP6*
STRA13	*TAGLN2*	*THAP10*	*TMEM161B*	*TPI1*	*TROVE2*
STRAP	*TAP1*	*THAP11*	*TMEM175*	*TPM1*	*TRPV1*
STRBP	*TARBP2*	*THAP3*	*TMEM31*	*TPM2*	*TRPV2*
STUB1	*TBC1D2*	*THEG*	*TMEM5*	*TPM4*	*TRPV6*
STX5	*TBCB*	*THEX1*	*TMEM66*	*TPP1*	*TRUB1*
STX6	*TBL2*	*THRA*	*TMEM8*	*TPSAB1*	*TRUB2*
STX8	*TBRG4*	*THRA*	*TMOD1*	*TPSAB1*	*TSEN34*
STXBP1	*TBX3*	*THRAP4*	*TMPIT*	*TPST1*	*TSG101*
STXBP6	*TCAP*	*THRAP5*	*TNF*	*TPST1*	*TSGA10*
STYXL1	*TCEAL1*	*THRAP6*	*TNFAIP1*	*TPTE*	*TSPAN13*
SUFU	*TCEB1*	*THY1*	*TNFRSF10A*	*TRA2A*	*TSPAN31*
SUGT1	*TCEB3*	*THY1*	*TNFRSF11B*	*TRADD*	*TSPAN4*
SULT1A1	*TCF19*	*TIFA*	*TNFRSF12A*	*TRAF2*	*TSPAN5*
SULT2A1	*TCL1A*	*TIMM13*	*TNFRSF14*	*TRAF4*	*TSPAN6*
SULT4A1	*TCN1*	*TIMM44*	*TNFRSF19L*	*TRAF6*	*TSPAN7*
SUMO1	*TCN2*	*TIMP1*	*TNFRSF21*	*TRAFD1*	*TSPYL1*
SUMO2	*TCP1*	*TIMP3*	*TNFRSF9*	*TRAM1*	*TSSC1*
SUPT5H	*TCP11L2*	*TIMP4*	*TNFSF10*	*TRAM1L1*	*TSSK1*
SUPV3L1	*TDRD7*	*TINF2*	*TNFSF13*	*TRAM2*	*TSSK3*
SURF1	*TDRKH*	*TIRAP*	*TNFSF14*	*TRAP1*	*TSSK6*
SURF2	*TEAD2*	*TITF1*	*TNIP1*	*TRAPPC1*	*TTC1*
SURF4	*TEF*	*TLE1*	*TNK1*	*TRAPPC2*	*TTC33*
SUSD3	*TEKT1*	*TLK1*	*TNNC1*	*TRAPPC2*	*TTK*
SUV39H1	*TEKT2*	*TLR2*	*TNNC2*	*TRAT1*	*TTL*
SUV39H2	*TEKT3*	*TLR7*	*TNNI1*	*TREM1*	*TTLL1*
SYNGR2	*Tenr*	*TLX3*	*TNNI2*	*TREM2*	*TUBA1*
SYNGR3	*TERF1*	*TLX3*	*TNNI3*	*TRIAP1*	*TUBA2*
SYT12	*TESC*	*TM4SF5*	*TNNT3*	*TRIM13*	*TUBA3*
SYT3	*TESK1*	*TMBIM1*	*TOB1*	*TRIM13*	*TUBB*
SYTL1	*TFAP2A*	*TMC4*	*TOB2*	*TRIM14*	*TUBB2B*
TAC1	*TFAP2B*	*TMC6*	*TOLLIP*	*TRIM22*	*TUBB2B*
TAC3	*TFAP4*	*TMCO6*	*TOMM22*	*TRIM23*	*TUBB6*
TACSTD1	*TFB1M*	*TMED1*	*TOMM40*	*TRIM27*	*TUBE1*
TACSTD2	*TFDP1*	*TMED2*	*TOMM70A*	*TRIM31*	*TUBG1*
TAF11	*TFF3*	*TMEFF1*	*TOR1A*	*TRIM32*	*TUBGCP3*
TAF12	*TFG*	*TMEM115*	*TOR1B*	*TRIM38*	*TUFT1*
TAF1A	*TFIP11*	*TMEM123*	*TOX*	*TRIM44*	*TUSC3*
TAF1B	*TFPI2*	*TMEM126B*	*TP53*	*TRIM7*	*TWF1*
TAF6	*TFRC*	*TMEM138*	*TPD52*	*TRIM74*	*TWF2*
TAF7	*TGIF*	*TMEM14C*	*TPD52*	*TRIM8*	*TWIST1*
TAF9B	*TGIF2*	*TMEM15*	*TPD52L2*	*TRIP10*	*TWIST2*

TXN2
TXNDC1
TXNDC12
TXNDC5
TXNL1
TYROBP
U2AF1L4
UBADC1
UBAP1
UBD
UBE1C
UBE1L
UBE2A
UBE2B
UBE2C
UBE2D1
UBE2I
UBE2J1
UBE2L6
UBE2T
UBE2U
UBE2V2
UBOX5
UBXD2
UCHL1
UCHL5IP
UCN2
UCP2
UFC1
UGDH
UGT2B4
UGT2B7
UNC45A
UQCRC2
URG4
UROD
UROS
USP10
USP12
USP13
USP14
USP16
USP2
USP3
USP30
USP33
USP44
USP5
UTP14A
UXS1
VAMP5
VCL
VDAC1
VDP
VEGFB
VEGFC
VEGFC
VHL
VIL2
VILL
VIPR2
VPREB3
VPS18
VPS54
VPS72
VRK2
VSIG4
VSNL1
VTI1B
VTN
WBP1
WBP11
WBP2
WBSCR27
WDR12
WDR18
WDR20
WDR4
WDR45
WDR51B
WDR57
WDR68
WFDC1
WIF1
WNT2
WNT5B
WNT7A
WRNIP1
WWP2
WWTR1
XAB1
XAB2
XAGE2
XBP1
XLKD1
XPA
XRCC3
XRCC5
YARS
YARS2
YBX1
YKT6
YKT6
YTHDF2
YWHAG
YWHAZ
ZBTB44
ZBTB48
ZBTB48
ZCCHC17
ZCRB1
ZDHHC4
ZDHHC5
ZDHHC7
ZFAND3
ZFP36L1
ZFP36L2
ZG16
ZMYND11
ZMYND12
ZMYND19
ZNF307
ZNF350
ZNF426
ZNF461
ZNF496
ZNF524
ZNF559
ZNF567
ZNF581
ZNHIT1
ZPBP
ZSCAN5
ZYX