

Phenotypic protein profiling of different B cell sub-populations using antibody CD-microarrays

P. Ellmark^{a,*}, C.M. Högerkorp^a, S. Ek^a, L. Belov^b, M. Berglund^c,
R. Rosenquist^d, R.I. Christopherson^b, C.A.K. Borrebaeck^a

^a Department of Immunotechnology, Lund University, BMC D13, SE-221 84 Lund, Sweden

^b School of Molecular and Microbial Biosciences, University of Sydney, Sydney, NSW 2006, Australia

^c Department of Oncology, Radiology and Clinical Immunology, Uppsala University, SE-751 85 Uppsala, Sweden

^d Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, SE-751 85 Uppsala, Sweden

Received 8 November 2007; received in revised form 28 January 2008; accepted 3 February 2008

Abstract

Antibody microarrays enable extensive protein expression profiling, and provide a valuable complement to DNA microarray-based gene expression profiling. In this study, we used DotScan™ antibody microarrays that contain antibodies against 82 different cell surface antigens, to determine phenotypic protein expression profiles for human B cell sub-populations. We then demonstrated that the B cell protein profile can be used to delineate the relationship between normal B cells and malignant counterparts. Principle component analysis showed that the lymphomas did not cluster with the normal memory B cells or germinal centre B cells, but they did cluster with germinal centre founder cells and naïve B cells. © 2008 Elsevier Ireland Ltd. All rights reserved.

Keywords: CD antigens; Antibody microarrays; DNA microarrays; B cell sub-populations; Lymphoma

1. Introduction

The integration of transcriptomics and proteomics is a key issue for the emerging field of systems biology. Transcriptomics has for the last 5–10 years been dominated by DNA microarray-based approaches, and today the entire human transcriptome can be analyzed on a single array. Transcriptomics has also proven useful for delineating the molecular mechanisms behind different diseases [1], development of new diagnostic approaches (www.rocche-diagnostics.com)

and identification of novel biomarkers [2]. However, mRNA analysis needs to be complemented by analysis of the proteome, as many gene products are post-translationally modified and regulated. Proteome analysis has traditionally been based on a sample separation technique, e.g. 2D gels or liquid chromatography, combined with mass-spectrometry [3]. Alternative approaches, such as antibody microarrays are rapidly emerging, providing valuable tools for multiplex detection of proteins [4–8].

However, antibody microarrays available today can only target a fraction of the proteome. For the analysis of leukocyte sub-populations and haematopoietic malignancies, the human cluster of differentiation (CD) antigens comprise a defined and

* Corresponding author. Tel.: +46 46 222 9818; fax: +46 46 222 4200.

E-mail address: peter.ellmark@immun.lth.se (P. Ellmark).

highly relevant part of the proteome [9]. The CD antigens are a well-characterised set of molecules united by their common expression on the surface of leukocytes and their essential roles in the immune system. The DotScan antibody microarray technique, developed by Belov et al. [10–12], is based on the capture of live cells on the microarray slide, in contrast to conventional antibody microarrays, which detect soluble antigens. DotScan microarrays contain antibodies against 82 different cell surface antigens (mainly CD-markers), and recently, these arrays have been used to analyze a large set of acute leukemias and B-lymphoproliferative disorders, demonstrating that expression profiles alone were sufficient to correctly classify these leukemias [13].

B cells are a sub-population of leukocytes that play a key role in the humoral immune system and their primary function is to produce antibodies against pathogens. On their developmental path to antibody-producing plasma and memory B cells, they pass through several different stages of differentiation. The phenotypic delineation of these different B cell sub-populations was outlined for the human lineage by Liu et al. [14] and recently revised by Högerkorp and Borrebaeck [15]. An example of this differentiation is when naïve B cells (IgD^+ , CD23^-) encounter an antigen in conjunction with the appropriate T cell stimulation, they become activated and differentiate into germinal centre (GC) founder B cells (IgD^+ , CD38^+). A few GC founder B cells will then expand clonally and form a GC, containing GC B cells (IgD^- , CD38^+) that will give rise to memory cells (IgD^- , CD38^-) and affinity matured antibody-producing plasma cells (IgD^- , CD38^{++}).

The different stages of B cell development involve clonal expansion, somatic hypermutation, and class switch recombination, events that make the B cell vulnerable to malignant transformation. Since several features of a non-malignant B cell are retained in the transformed malignant B cell counterpart [16], it is believed that B cell leukemias and lymphomas may resemble the B cell developmental stage from which they originated [17]. For this reason, it is important to delineate common features between malignant and non-malignant B cells to distinguish clinical markers for more accurate diagnosis and therapeutic strategies. In this study, we have used DotScanTM microarrays to immunophenotype several different mature B cell sub-populations and lymphoma samples. The resulting CD expression profiles distinguished the normal B cell sub-populations and were subsequently used for analysis of the

lymphoma samples. In addition, we have compared the CD antigen levels with transcriptional data obtained from high-density oligonucleotide arrays.

2. Materials and methods

2.1. B cell sub-populations

Human tonsils were obtained from 6 paediatric patients (aged between 2 and 16 years) undergoing routine tonsillectomy at Lund University Hospital (Lund, Sweden). Briefly, tonsils were minced and T cells were removed by rosetting with neuraminidase-treated sheep red blood cells. Mononuclear, T-cell-depleted, cells were isolated by density centrifugation, using Ficoll-Isopaque (Amersham Pharmacia Biotech AB, Uppsala, Sweden). The interphase fraction, containing predominantly mononuclear B cells, was washed in PBS containing FBS (10%). The total T cell depleted fraction, approximately $1\text{--}1.5 \times 10^9$ cells, was labeled with fluorescent antibodies in a stepwise process: (1) anti-IgM (BD Biosciences, Franklin Lakes, NJ), (2) goat anti-mouse-PE-Texas Red (Caltag, Carlsbad, CA), (3) anti-CD27-FITC (BD Biosciences), anti-IgD-PE (Dako, Glostrup Denmark), anti-CD38-PECy7 (BD Biosciences), anti-CD19-APCCy7 (BD Biosciences), anti-CD3-PB (BD Biosciences) and anti-CD14-PB (BD Biosciences). Cell sorting was performed on a FACS Aria (BD Biosciences). The different B cell populations were isolated as naïve (CD19^+ , CD3/14^- , IgD^+ , IgM^+ , CD38^- , CD27^-), GC founders (CD19^+ , CD3/14^- , IgD^+ , IgM^+ , CD38^+ , $\text{CD27}^{+/-}$), GC B cells (CD19^+ , CD3/14^- , IgD^- , IgM^- , CD38^+ , $\text{CD27}^{+/-}$), Plasma cells (CD19^+ , CD3/14^- , IgD^- , IgM^- , CD38^{++} , CD27^+) and Memory cells (CD19^+ , CD3/14^- , IgD^- , IgM^- , CD38^- , CD27^+). Purity was confirmed by reanalysis and was typically $>97\%$. Each population was sorted in three independent experiments, where the starting material for each round of selection was tonsillar lymphocytes pooled from two donors.

2.2. Isolation of lymphoma cells

Tumor samples were collected as diagnostic material from patients at Uppsala University Hospital. Seven lymphomas were analyzed, four Diffuse Large B Cell Lymphoma with non-GC phenotype [18] (DLBCL non-GC), one Chronic Lymphocytic Leukemia (CLL) and two Follicular Lymphomas (FL). The tumor samples were kept frozen as single-cell suspensions in media supplemented with 10% DMSO. After diagnosis, the cells were thawed at 37 °C and suspended in growth medium (1% L-glutamine/10% FCS/1% non-essential aminoacids/RPMI 1640). Cells were stained with anti-CD19-PB, anti-CD3-FITC (BD Biosciences) anti-immunoglobulin lambda light chain-PE (Dako) and anti-immunoglobulin kappa light chain-APC (BD Bioscience). The cells were sorted as CD19^+ , CD3^- , lambda^+ or kappa^+ (or lambda^- and kappa^-). An analytic

fraction of the cells was stained with PI to set a FSC and SSC-gate for purification of the viable cells. Purity was confirmed by reanalysis and was typically >97%.

2.3. Preparation of Affymetrix U133 plus 2.0 GeneChip microarrays

The samples and arrays were prepared according to the manufacturers protocols (Affymetrix, Inc., Santa Clara, CA, USA) [19]. Briefly, tRNA was isolated (Poly-A RNA control Kit), and cDNA synthesized and transcribed *in vitro* overnight (Two-cycle cDNA Synthesis Kit). The cRNA was cleaned (IVT cRNA Cleanup Kit) and subjected to a second round of amplification (Two-Cycle cDNA Synthesis Kit). After cleanup (Sample Cleanup Module) the cRNA was labeled with biotin (IVT Labeling Kit), and the biotinylated cRNA was further cleaned (Sample Cleanup Module) and fragmented for 1 h. After hybridizing the fragments to the array for 16 h, streptavidin–phycoerythrin and biotinylated anti-streptavidin antibodies were used for detection. The arrays were scanned using a GeneChip Scanner 3000 (Affymetrix).

2.4. Antibody microarray analysis

DotScan™ antibody microarray Chips (Leukaemia & Lymphoma Diagnostic Chip) were obtained from Medsaic Pty. Ltd. (Eveleigh, Australia). The construction of the antibody microarray Chips has been described previously [10,11] and a full list of the antibodies can be found on the website www.medsaic.com. Sorted B cell populations or lymphoma samples (2×10^6 cells) were washed with PBS, resuspended in 150 μ L PBS containing 1 mM EDTA and incubated for 1 h on the microarray Chip. Unbound cells were removed by gentle washing with PBS, and the

captured cells were fixed with 4% formaldehyde in PBS and imaged. The dot intensities for each duplicate antigen were quantified using the ScanArray Express software V2.0 (Perkin Elmer Life and Analytical Sciences, Waltham, MA). The background signals were subtracted, and the data were scaled on a 256 bits greyness scale, as described previously [11], and $\log_2$ values were calculated.

2.5. Statistical analysis

The logarithm (base 2) of the DotScan™ data was calculated and used for the ANOVA test, and thereafter Z-score normalized. Hierarchical clustering and Principal Component Analysis (PCA) were performed using the un-weighted pair-group method on the arithmetic mean, using the Spotfire 8.2 Decision Site software. The data from the GC sub-populations in the GeneChip microarray analysis [20] were merged to generate a representative comparison with the proteome data analysis. The GeneChip was treated identically to the DotScan™ data. [Supplementary Table 1](#) provides a detailed list of the detection antibodies and the Affymetrix probes used for this study. For calculation of the correlation coefficients for the lymphoma samples, the $\log_2$ data were normalized by dividing the data from each sample by the mean value of all samples. The relative levels of all CD-markers present and detected on both platforms were correlated and the *R*-values were calculated.

3. Results and discussion

3.1. CD-marker profiling of B cell sub-populations

We used antibody CD-microarray Chips to analyze sub-populations of B cells, to identify phenotypic expression profiles that discriminate between these populations.

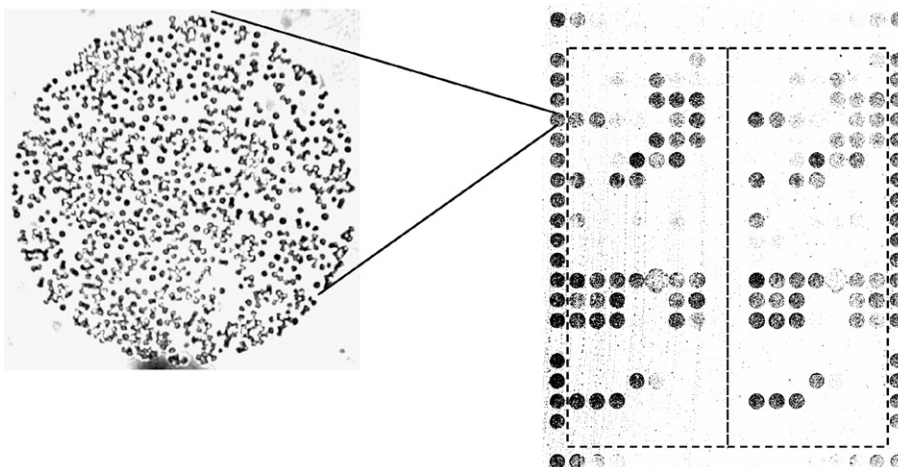


Fig. 1. Image of a B cell sample analyzed on a DotScan™ antibody microarray, where each antibody is spotted in two identical sub-arrays (marked as dashed boxes). Anti-CD44 spots applied down the left and right sides of the duplicate arrays provide array alignment define antibody addresses. Anti-CD44 titrations are shown at the top and bottom of the array. The magnified image shows individual cells captured on a dot.

Four different B cell sub-populations were isolated from tonsils, using flow cytometry based cell sorting. Naïve B cells, GC founder B cells (GCF), GC B cells, and memory B cells were purified. The isolated B cell populations were

>97% pure, and 2×10^6 cells from each population were used for the Dotscan™ microarray analysis. To generate a general immunophenotype for the different sub-populations, each population was sorted and analyzed on the

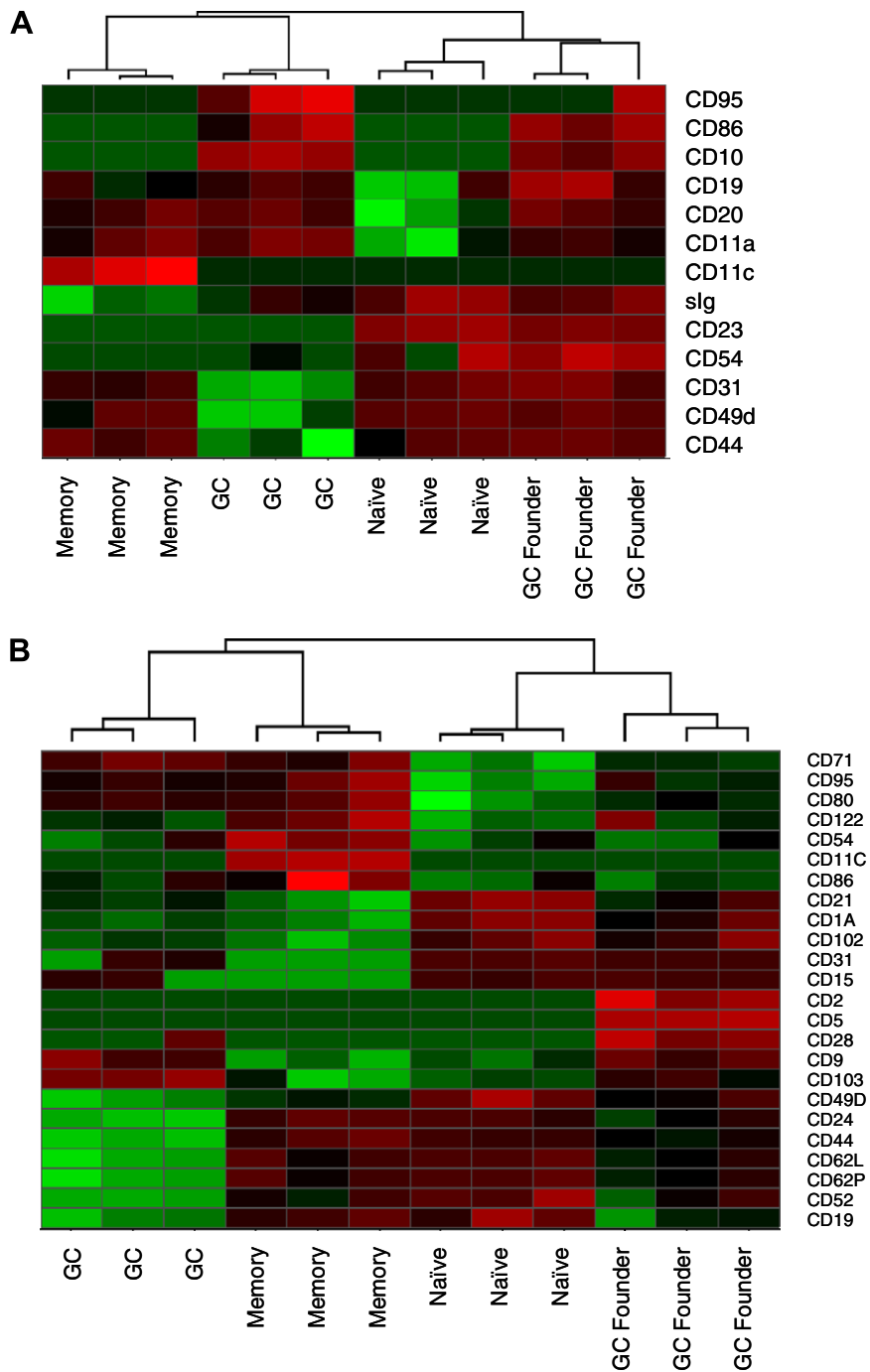


Fig. 2. Hierarchical clustering of CD antigen profiles showing differential expression between B cell sub-populations. (A) Clusters obtained from the antibody microarray data, and (B) clusters from the DNA microarray data. Red indicates up-regulation, green down-regulation and black no change.

Dotscan microarrays in three independent experiments, where each round of sorting was based on pooled cells from two donors. The four B cell sub-populations were characterized on Dotscan™ microarray Chips and the resulting dot-patterns were quantified and analyzed. Image analysis revealed highly specific dot-patterns for each cell type, with low background binding (Fig. 1).

3.2. Hierarchical cluster analysis of the B cell sub-populations

An ANOVA test was used to identify CD-markers that were significantly expressed differentially between the different B cell sub-populations. This selective CD-marker profile was then used for the hierarchical cluster analysis (Fig. 2A), which revealed four well-separated clusters that differentiated between the sub-populations. The naïve B cells display lower levels of CD19, CD20 and CD11a, and like the memory B cells, they do not express the typical activation markers of B cells such as CD95 or CD86. In contrast, these markers are highly expressed by the activated GC founder cells, and GC B cells. The GC B cells have down-regulated adhesion markers CD44, CD49d and CD31 compared to the other populations, whereas the memory B cell profile is distinguished by up-regulation of CD11c. These phenotypic profiles, summarized in Table 1, include some CD-markers that have not been fully described for the different B cells sub-populations, and others that were previously validated by flow cytometry [15], and additional studies will be needed to fully validate the profiles. For example, CD31 (PECAM-1) is down-regulated in follicular B cells [21]; however, its expression levels have not been compared to those of naïve GC founder and memory cells. Further-

more, the expression levels of CD49d (VLA-4 alpha chain) and CD54 (ICAM-1) have not been associated with the B cell sub-populations described in this study, although they may have a prognostic value for B cell malignancies [22,23].

Having demonstrated that the antibody microarrays can distinguish between B cell sub-populations, we analyzed the corresponding mRNA levels for the same CD-markers. To facilitate a comparison of the antibody microarray data with GeneChip microarray data for the same cell populations, the CD-markers shared by Dot-Scan™ and the GeneChip microarrays were identified and analyzed (summarized in Supplementary Table 1). Again we performed an ANOVA to identify the most significant differentially expressed genes between the B cell sub-populations, followed by hierarchical cluster analysis (Fig. 2B). In concordance with the antibody microarray analysis, the B cell populations were clearly separated into four clusters. Interestingly, only eight CD-markers from the gene profile overlapped with the protein profile, as shown in the Venn diagram (Fig. 3), demonstrating the importance of membrane protein profiling for identifying individual B cell sub-populations and their functional associations.

3.3. Analysis of lymphomas using the CD-marker protein expression profile

B cell lymphomas preserve in many cases the phenotypic markers characteristic of the differentiation stage at which the malignant transformation occurred [16,17]. The CD-marker expression profiles (Fig. 2A and Table 1) could be a valuable tool for investigating how different lymphomas relate to normal B cell sub-populations. The protein expression profiles would be of particular relevance, as they could be compared to the diagnostic profiles obtained previously with Dotscan™ for other haematological malignancies [13].

Table 1
Phenotypic protein profile of normal B cell populations

Cell surface antigen	Naïve	GC founder	GC	Memory
CD95	–	–	+	–
CD86	–	+	+	–
CD10	–	+	+	–
CD19	–	+	+	–
CD20	–	+	+	+
CD11a	–	+	+	+
CD11c	–	–	–	+
sIg	+	+	–	–
CD23	+	+	–	–
CD54	+	+	–	–
CD31	+	+	–	+
CD49d	+	+	–	+
CD44	+	+	–	+

The average intensity of the three replicates for each sub-population was calculated from the Z-score normalized data. The CD antigens in the protein expression profile were classified as positive (+) if the average was >0, or negative (–) if the average was <0.

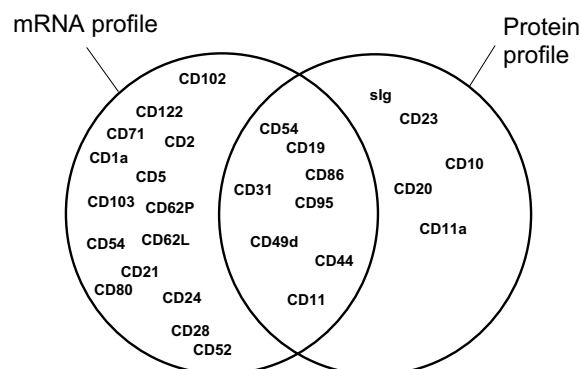


Fig. 3. Venn diagram showing the relationship between the differentially expressed CD antigen profiles of B cell sub-populations identified by antibody and DNA microarrays.

To demonstrate the utility of the B cell protein profile for correlating B cell lymphomas with normal B cell populations, seven lymphoma samples were purified by flow cytometry, and analyzed on the Dotscan™ antibody microarray Chips and Affymetrix DNA microarrays: four non-GC diffuse large B cell lymphomas (non-GC DLBCL), two follicular lymphomas (FL) and one chronic lymphocytic leukemia (CLL). The expression profiles of the lymphomas were compared with the normal B cell sub-populations, using PCA (Fig. 4). PCA is a means of reducing the dimensionality of a data set by assessing the covariance within the data and representing that as new parameters (reviewed by [24]). The outcome is a number of new parameters, in which the greatest variance is represented in the first few principal components. In Fig. 4, the resulting values from the first two principal components are displayed (PCA 1 and PCA 2). Thereby, the PCA reduced 13 dimensions, i.e. the signal intensities from the 13 CD-markers differentially expressed between the B cell sub-populations, to two dimensions. The proximity of the different samples on the PCA-plot indicates their relationship (i.e. how they cluster with each other), demonstrating that the normal B cell sub-populations forms four separated clusters (Fig. 4).

The PCA (Fig 4) showed that none of the malignant B cells in this study clustered with the normal memory B cells or GC B cells. Although none of the lymphomas were expected to be of memory B cell origin, the FLs are considered to have a GC origin. Interestingly, however, they clustered with GC founders and naïve B cells (Fig 4). FLs have been previously reported to display an aberrant phenotype, diverging from the normal GC lineage [25], and gene expression analysis has also revealed a lack of resemblance to GC B cells [26]. It is possible that this phenotypic difference may be attributed to the effect of disease related changes in the follicular stroma, which has been shown to differ from normal germinal centre follicles in many cases [27]. The analysis also showed that three of the four non-GC-DLBCLs clustered with GC founder cells, while one sample had a distinctly different profile. This correlates with a previous investigation, demonstrating that the gene expression pattern of non-GC-DLBCL resembles activated B cells more than GC B cells [26].

Due to the sample number it is not yet possible for a general conclusion of the relative levels of different CD-markers on the lymphomas in relation to the normal B cell populations, and the classification approach will be further validated by analyzing a larger set of lymphoma

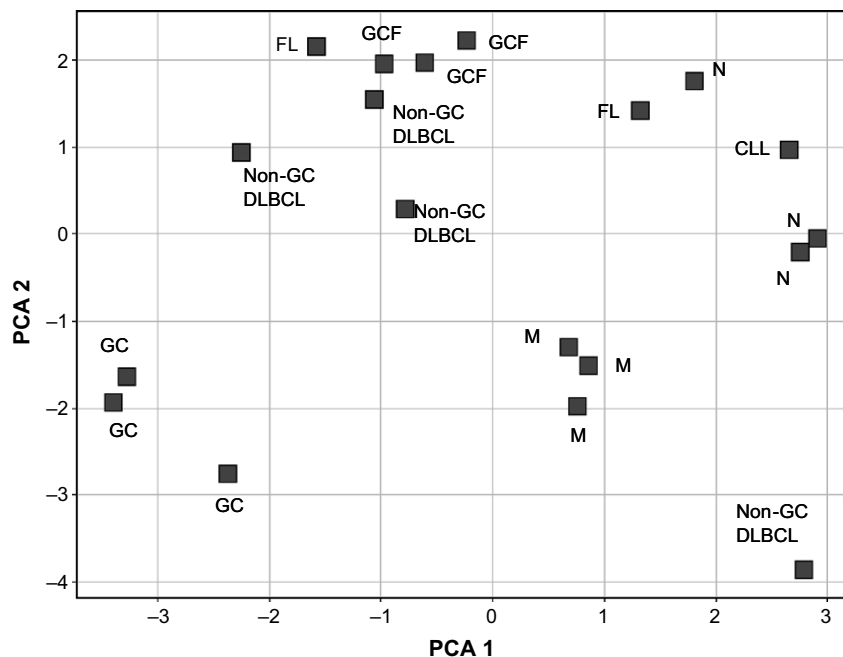


Fig. 4. Principal component analysis (PCA) of CD antigen protein profiles for lymphoma samples and B cell sub-populations, using the phenotypic B cell protein profile. The axes represent scores in the principal component space for the first (PCA 1) and second (PCA 2) principal component. Four B cell sub-populations: memory B cells (M), germinal centre founder B cells (GCF), germinal centre B cells (GC) and naïve B cells (N) forms separate clusters. The proximity of the of the B cell sub-populations to the different lymphoma samples, diffuse large B cell lymphoma with non-GC phenotype (non-GC DLBCL), follicular lymphomas (FL) and chronic lymphocytic leukemia (CLL) on the PCA-plot indicates their similarities, i.e. how they cluster with each other.

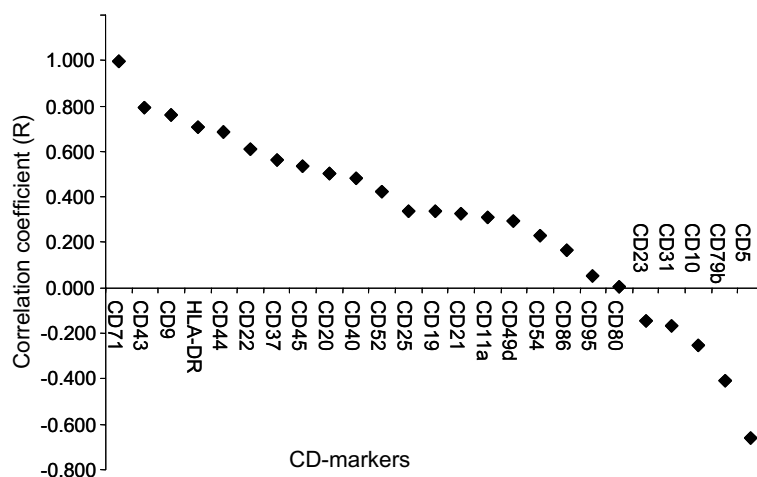


Fig. 5. Correlation coefficients (R) for protein and mRNA expression levels in the lymphoma samples, showing values for each CD antigen detected on both platforms averaged for seven different lymphoma samples.

samples. Supervised learning methods, such as support vector machine (SVM) or similar methods, may be used for automated classification of larger sample sets.

3.4. Correlation between mRNA and protein expression level

To assess the correlation between the mRNA and protein expression levels, we calculated correlation coefficients for the expression levels of the different CD-markers for the lymphoma samples (Fig. 5). The average correlation coefficient (R) of 0.29 was similar to that reported by Strömberg et al [28]. Four of the markers (16%; CD71, CD43, CD9, HLA-DR) correlated significantly ($R \geq 0.707$), similar to the correlation reported by Chen et al. [29]. We have previously demonstrated a good correlation between phenotypes obtained by FACS analysis and mRNA data, using a limited set of CD antigens [15].

However, since transcription, translation and expression of a surface antigen are dependent on sequence of events subject to cellular regulation, the transcriptional analysis would detect events before they are detectable at the protein level, which has also been seen in other studies [29–32]. For example, the activation-induced expression of CD95 on memory B cells was detected at the transcriptional level before the cell surface (Fig. 2), which resulted in a poor correlation between gene expression and protein expression for CD95 (Fig. 5). Furthermore, the expression of a cell surface antigen such as a receptor is also dependent on external factors such as the level of cell activation. CD3 for instance, is known to be down-regulated upon activation and is best visualised by intracellular staining [33]. In addition, there may be technical issues such as different limits of detection or degrees of validation for the different array platforms (reviewed by Hegde et al [31]).

In this study, we have determined a CD antigen expression profile for human B cell sub-populations, using CD-antibody microarray Chips. The results have been used for investigating the relationship between normal B cells and their malignant counterparts, and provide phenotypic information complementary to gene expression analysis. We have demonstrated a correlation between the non-GC-DLBCL and GC founder B cell profiles, and a lack of correlation between the lymphomas and the memory B cell or GC B cell profiles. These findings will be further validated using a larger set of lymphoma samples before more general conclusions are made.

Acknowledgments

This investigation was supported by grants from VR-NT and Magn. Bergvalls foundation. We also thank Pauline Huang for making the CD antibody microarray Chips, provided by Medsaic Pty. Ltd. (Eveleigh, NSW 1430, Australia).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.canlet.2008.02.006](https://doi.org/10.1016/j.canlet.2008.02.006).

References

- [1] D.R. Rhodes, A.M. Chinnaiyan, Integrative analysis of the cancer transcriptome, *Nat. Genet.* 37 (Suppl.) (2005) S31–S37.
- [2] S. Ek, U. Andreasson, S. Hober, C. Kampf, F. Ponten, M. Uhlen, H. Merz, C.A. Borrebaeck, From gene expression analysis to tissue microarrays: a rational approach to

- identify therapeutic and diagnostic targets in lymphoid malignancies, *Mol. Cell. Proteomics* 5 (2006) 1072–1081.
- [3] C. Delahunty, J.R. Yates 3rd, Protein identification using 2D-LC-MS/MS, *Methods* 35 (2005) 248–255.
- [4] A. Aguilar-Mahecha, S. Hassan, C. Ferrario, M. Basik, Microarrays as validation strategies in clinical samples: tissue and protein microarrays, *OMICS* 10 (2006) 311–326.
- [5] C.A. Borrebaeck, Antibody microarray-based oncoproteomics, *Expert Opin. Biol. Ther.* 6 (2006) 833–838.
- [6] B.B. Haab, Antibody arrays in cancer research, *Mol. Cell. Proteomics* 4 (2005) 377–383.
- [7] C. Wingren, C.A. Borrebaeck, Antibody microarrays: current status and key technological advances, *OMICS* 10 (2006) 411–427.
- [8] C. Wingren, C.A. Borrebaeck, High-throughput proteomics using antibody microarrays – an update, *Expert Rev. Mol. Diagn.* 7 (2007) 673–686.
- [9] A. Woolfson, P. Ellmark, J.S. Chrisp, M.A. Scott, R.I. Christopherson, The application of CD antigen proteomics to pharmacogenomics, *Pharmacogenomics* 7 (2006) 759–771.
- [10] L. Belov, O. de la Vega, C.G. dos Remedios, S.P. Mulligan, R.I. Christopherson, Immunophenotyping of leukemias using a cluster of differentiation antibody microarray, *Cancer Res.* 61 (2001) 4483–4489.
- [11] L. Belov, P. Huang, N. Barber, S.P. Mulligan, R.I. Christopherson, Identification of repertoires of surface antigens on leukemias using an antibody microarray, *Proteomics* 3 (2003) 2147–2154.
- [12] P. Ellmark, L. Belov, P. Huang, C.S. Lee, M.J. Solomon, D.K. Morgan, R.I. Christopherson, Multiplex detection of surface molecules on colorectal cancers, *Proteomics* 6 (2006) 1791–1802.
- [13] L. Belov, S.P. Mulligan, N. Barber, A. Woolfson, M. Scott, K. Stoner, J.S. Chrisp, W.A. Sewell, K.F. Bradstock, L. Bendall, D.S. Pascovici, M. Thomas, W. Erber, P. Huang, M. Sartor, G.A. Young, J.S. Wiley, S. Juneja, W.G. Wierda, A.R. Green, M.J. Keating, R.I. Christopherson, Analysis of human leukaemias and lymphomas using extensive immunophenotypes from an antibody microarray, *Br. J. Haematol.* 135 (2006) 184–197.
- [14] V. Pascual, Y.J. Liu, A. Magalski, O. de Bouteiller, J. Banchereau, J.D. Capra, Analysis of somatic mutation in five B cell subsets of human tonsil, *J. Exp. Med.* 180 (1994) 329–339.
- [15] C.M. Högerkorp, C.A. Borrebaeck, The human CD77⁺ B cell population represents a heterogeneous subset of cells comprising centroblasts, centrocytes, and plasmablasts, prompting phenotypical revision, *J. Immunol.* 177 (2006) 4341–4349.
- [16] R. Kuppers, Mechanisms of B-cell lymphoma pathogenesis, *Nat. Rev. Cancer* 5 (2005) 251–262.
- [17] V. Pascual, Y.J. Liu, J. Banchereau, Normal human B cell sub-populations and their malignant counterparts, *Baillieres Clin. Haematol.* 10 (1997) 525–538.
- [18] M. Berglund, U. Thunberg, R.M. Amini, M. Book, G. Roos, M. Erlanson, J. Linderöth, M. Dictor, M. Jerkeman, E. Cavallin-Stahl, C. Sundström, S. Rehn-Eriksson, C. Backlin, H. Hagberg, R. Rosenquist, G. Enblad, Evaluation of immunophenotype in diffuse large B-cell lymphoma and its impact on prognosis, *Mod. Pathol.* 18 (2005) 1113–1120.
- [19] S. Ek, C.M. Högerkorp, M. Dictor, M. Ehinger, C.A. Borrebaeck, Mantle cell lymphomas express a distinct genetic signature affecting lymphocyte trafficking and growth regulation as compared with subpopulations of normal human B cells, *Cancer Res.* 62 (2002) 4398–4405.
- [20] C.M. Högerkorp, C.A. Borrebaeck, The genetic programs of subsets of GC founder B cells and sub-epithelial B cells support their role as extrafollicular proliferating plasma blasts (submitted for publication).
- [21] A. Stacchini, R. Chiarle, V. Antinoro, A. Demurtas, D. Novero, G. Palestro, Expression of the CD31 antigen in normal B-cells and non Hodgkin's lymphomas, *J. Biol. Regul. Homeost. Agents* 17 (2003) 308–315.
- [22] B.T. Pittner, T.D. Shanafelt, N.E. Kay, D.F. Jelinek, CD38 expression levels in chronic lymphocytic leukemia B cells are associated with activation marker expression and differential responses to interferon stimulation, *Leukemia* 19 (2005) 2264–2272.
- [23] A. Zucchetto, P. Sonigo, M. Degan, R. Bomben, M. Dal Bo, S. Russo, V. Attadia, M. Rupolo, F. Buccisano, A. Steffan, G. Del Poeta, C. Pucillo, A. Colombatti, R. Campanini, V. Gattei, Surface-antigen expression profiling (SEP) in B-cell chronic lymphocytic leukemia (B-CLL): identification of markers with prognostic relevance, *J. Immunol. Methods* 305 (2005) 20–32.
- [24] B. Norden, P. Broberg, C. Lindberg, A. Plymoth, Analysis and understanding of high-dimensionality data by means of multivariate data analysis, *Chem. Biodivers* 2 (2005) 1487–1494.
- [25] S. Ray, F.E. Craig, S.H. Swerdlow, Abnormal patterns of antigenic expression in follicular lymphoma: a flow cytometric study, *Am. J. Clin. Pathol.* 124 (2005) 576–583.
- [26] A.A. Alizadeh, M.B. Eisen, R.E. Davis, C. Ma, I.S. Lossos, A. Rosenwald, J.C. Boldrick, H. Sabet, T. Tran, X. Yu, J.I. Powell, L. Yang, G.E. Marti, T. Moore, J. Hudson Jr., L. Lu, D.B. Lewis, R. Tibshirani, G. Sherlock, W.C. Chan, T.C. Greiner, D.D. Weisenburger, J.O. Armitage, R. Warnke, R. Levy, W. Wilson, M.R. Grever, J.C. Byrd, D. Botstein, P.O. Brown, L.M. Staudt, Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling, *Nature* 403 (2000) 503–511.
- [27] K.C. Chang, X. Huang, L.J. Medeiros, D. Jones, Germinal centre-like versus undifferentiated stromal immunophenotypes in follicular lymphoma, *J. Pathol.* 201 (2003) 404–412.
- [28] S. Strömberg, M.G. Björklund, C. Asplund, A. Sköllerö, A. Persson, K. Wester, C. Kampf, P. Nilsson, A.C. Andersson, M. Uhlen, J. Kononen, F. Ponten, A. Asplund, A high-throughput strategy for protein profiling in cell microarrays using automated image analysis, *Proteomics* 7 (2007) 2142–2150.
- [29] G. Chen, T.G. Gharib, C.C. Huang, J.M. Taylor, D.E. Misek, S.L. Kardia, T.J. Giordano, M.D. Iannettoni, M.B. Orringer, S.M. Hanash, D.G. Beer, Discordant protein and mRNA expression in lung adenocarcinomas, *Mol. Cell. Proteomics* 1 (2002) 304–313.
- [30] B. Bartling, H.S. Hofmann, T. Boettger, G. Hansen, S. Burdach, R.E. Silber, A. Simm, Comparative application of antibody and gene array for expression profiling in human squamous cell lung carcinoma, *Lung Cancer* 49 (2005) 145–154.

- [31] P.S. Hegde, I.R. White, C. Debouck, Interplay of transcriptomics and proteomics, *Curr. Opin. Biotechnol.* 14 (2003) 647–651.
- [32] L. Nie, G. Wu, W. Zhang, Correlation of mRNA expression and protein abundance affected by multiple sequence features related to translational efficiency in *Desulfovibrio vulgaris*: a quantitative analysis, *Genetics* (2006).
- [33] Y.D. Mahnke, M. Roederer, Optimizing a multicolor immunophenotyping assay, *Clin. Lab. Med.* 27 (2007) 469–485, v.

Supplementary Table 1. Cell surface antigens and Affymetrix probes

Cell surface antigen	Cell specificity	Affymetrix probe
CD2	T-Cell	205831_at
CD3	T-Cell	205456_at
CD4	T-Cell	216424_at
CD5	T-Cell	230489_at
CD7	T-Cell	214049_x_at
CD8	T-Cell	205758_at
CD9	B-Cell	201005_at
CD10	B-Cell	203434_s_at
CD13	Myeloid	234576_at
CD14	Myeloid	201743_at
CD15	Myeloid	209892_at
CD16	NK	204007_at
CD19	B-Cell	206398_s_at
CD20	B-Cell	228592_at
CD21	B-Cell	205544_s_at
CD22	B-Cell	38521_at
CD23	B-Cell	206759_at
CD24	Other	266_s_at
CD25	T-Cell	206341_at
CD28	T-Cell	206545_at
CD29	Other	215879_at
CD30	Other	NA
CD31	Other	208981_at
CD32	Other	1565674_at
CD33	Myeloid	215856_at
CD34	SC	209543_s_at
CD36	Myeloid	228766_at
CD37	B-Cell	204192_at
CD38	B-Cell	234187_at
CD40	Other	215346_at
CD41	Myeloid	217472_at
CD43	Other	211567_at
CD44	Other	212063_at
CD45	Additional	212588_at
CD52	Other	34210_at
CD54	Other	202638_s_at
CD56	NK	227394_at
CD57	NK	219521_at
CD60	Other	NA
CD61	Myeloid	211579_at
CD64	Myeloid	216951_at
CD65	Myeloid	NA
CD71	Other	208691_at
CD77	B-Cell	NA
CD80	Other	1554519_at
CD86	Other	205685_at
CD88	Myeloid	220088_at
CD95	Other	204780_s_at

CD102	Other	204683_at
CD103	Additional	205055_at
CD117	SC	205051_s_at
CD122	Other	205291_at
CD126	Other	226333_at
CD128	Other	207094_at
CD130	Other	234967_at
CD134	Other	208023_at
CD135	SC	206674_at
CD138	B-Cell	201286_at
CD154	Other	207892_at
CD11a	Other	1554240_a_at
CD11b	Myeloid	205785_at
CD11c	Other	210184_at
CD120a	Other	207643_s_at
CD1a	T-Cell	210325_at
CD235a	Myeloid	1559520_at
CD42a	Myeloid	206883_x_at
CD49d	Other	213416_at
CD49e	Other	201389_at
CD62E	Other	206211_at
CD62L	Other	204563_at
CD62P	Myeloid	204563_at
CD66c	Myeloid	203757_s_at
CD79a	B-Cell	1555779_a_at
CD79b	B-Cell	1555746_at
HLA-DR	B-Cell	208894_at
kappa	B-Cell	NA
lambda	B-Cell	NA
slg	B-Cell	NA
TCRa/b	T-Cell	NA
TCRg/d	T-Cell	NA
CD45RA	Additional	NA
CD45RO	Additional	NA