Review Article

Open Access

Statistical Analysis of Protein Microarray Data: A Case Study in Type 1 Diabetes Research

Le TT An[1-3]#, Anna Pursiheimo[1,2]#, Robert Moulder[1] and Laura L Elo[1,2]*

[1]*Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, Turku, Finland*
[2]*Department of Mathematics and Statistics, University of Turku, Finland*
[3]*School of Applied Mathematics and Informatics, Hanoi University of Science and Technology, Vietnam*
#*Authors contribute equally*

Abstract

In this report we provide an overview of protein microarrays and devote particular consideration to the statistical methods used in data analysis with applications concerning the study of type 1 diabetes. The latter methodologies are illustrated with publically available data from a study that identified novel type 1 diabetes associated autoantibodies. Amongst the methods employed, Reproducibility-Optimized Test Statistic (ROTS) shows better detection over the widely used LIMMA. With the application of this analytical approach, we identify new protein biomarkers that were not previously reported in original investigation. This observation emphasises the benefit of using different methods to extract critical information in the analysis of microarray data.

Keywords: Protein microarray; Type 1 diabetes; Biomarker; Computational method; Reproducibility-Optimized Test Statistic (ROTS)

Abbreviations: T1D: Type 1 Diabetes; T2D: Type 2 Diabetes; NGT: Normal Glucose Tolerance; SAM: Significance Analysis of Microarrays; LIMMA: Linear Models for Microarray Data; RP: Rank Product; ROTS: Reproducibility-Optimized Test Statistic; FDR: False Discovery Rate; ROC: Receiver Operating Characteristics

Introduction

Protein microarrays

In recent years powerful high throughput microarray techniques for gene expression profiling have emerged and have been widely applied in comparative studies of cellular states and biological specimens [1,2]. These facilitate automated, paralleled analysis of thousands of genes and have created new possibilities in biomedical research. Despite such advantages, it should be noted that the correlation of RNA expression data and protein translation is variable [3] and transcriptomics measurement cannot take into account post-translational changes [4]. An important alternative is the use of protein microarrays, which can be used for protein detection, quantification and interaction measurements, thus providing a promising complementary approach to other systems biology approaches [5].

Akin to their oligonucleotide targeting analogues, protein microarrays are constructed on solid supports, such as a glass slide or nitrocellulose membrane, onto which small amounts of different probes (proteins) are bound at discrete locations [6]. These can range from high density chips containing thousands of proteins to specific arrays with tens or hundreds of antibodies. Currently there are three different types of protein microarrays: functional microarrays, reverse phase microarrays and analytical microarrays [6]. Functional protein microarrays generally incorporate a large panel of purified proteins or protein domains and are used to detect the biochemical activity and protein interactions [7]. With reverse phase protein microarrays a cell lysate is arrayed and then probed with antibodies against specific protein targets [8]. Analytical microarrays have included arrays of antibodies, aptamers, or affibodies and are typically used to measure binding affinities, specificities, and protein expression levels of proteins in complex mixtures [9,10]. Overall the main areas of application for protein microarrays include proteomics, protein functional analysis, antibody characterization, diagnostics, and treatment development. The listed range of specific targets and application reads like a who's who of protein orientated biological research, including protein-protein/peptide/RNA interactions [11-13], protein post-translational modifications [4] and biomarker identification [14]. The latter includes application towards detection of infectious disease [15], cancer [14,16] and autoimmune diseases, e.g. systemus erythematosus [17] and rheumatoid arthritis [18].

Amongst the methods of detection used with protein microarrays, fluorescent labelling is the most widely used. Other approaches include photochemical and radioisotope tags. The fluorescent label or tag is attached to the probe or secondary antibody and the interaction determined by, for example, a microarray scanner [6].

Type 1 diabetes

Type 1 diabetes (T1D) is an autoimmune disease that results in the destruction of the insulin producing beta cells of the islets of the Langerhans [19]. At this point the patient is dependent on a daily insulin substitution for the rest of his/her life and there is a high risk of developing acute and long-term complications. There is a strong genetic component for T1D risk, in addition to the role of the environment, diet and viral infections, which have been indicated as influential factors driving its onset. Whilst the genetic traits for T1D are common in some populations, the outcome is unpredictable. Early signs of its

onset are found with the detection of autoantibodies against islets cells (ICA). Currently the panel of autoantibodies that are regularly used to determine the development of the autoimmune reactions that underlie the development of T1D, consists of islet cell autoantibodies (ICA) and antibodies against insulin (IAA), glutamic acid decarboxylase (GADA), IA-2 protein (IA-2A) and Zinc transporter 8 (ZnT8) [20]. In prospective studies of T1D pathogenesis, serum samples are collected at regular intervals from genetically conferred risk groups and tested for these autoantibodies [21,22]. In this respect the use of protein microarrays is an attractive option for multiplexed detection of these known autoantibodies and for the detection of new markers.

Amongst the literature describing the use of protein microarrays in T1D research there have been a number of studies using commercial cytokine antibody arrays [23-27]. Broadly these include a group of studies using essentially the same cytokine array (RayBiotech) to study the effects of T1D autoantigen stimulation on cytokine production from peripheral blood mononuclear cells (PBMCs) from diabetics and children with diabetic parents. The PBMC samples used in these studies were from patients with cystic fibrosis related T1D [23], neonates with T1D parents [24], T1D patients and their relatives [26], T1D children [27] and children with mothers displaying maternal hyperglycaemia [25]. Miersch et al. [28] produced microarrays displaying 6000 proteins that were used to identify new T1D autoantibodies from the sera of T1D patients. Their analysis revealed 26 novel autoantibodies. In a similar fashion, Koo et al. [29] screened sera from type 1 and type 2 diabetes using arrays of 9,600 proteins. In their study, two novel autoantibodies were identified.

Protein microarrays: statistical and computational approaches

There are a number of reviews on differential expression analysis and feature selection using microarray data [30-35]. Most of them have focused on gene expression microarrays. Here we provide a brief review of differential expression analysis in the context of protein microarrays. In particular, we describe the essential steps in the analysis of protein microarray data and a number of computational tools for determining statistically significant differences between distinct sample groups. Finally, we provide a case study with a recently published protein microarray data from a study of T1D, which demonstrates the favourable performance of our reproducibility-optimized test statistic ROTS in comparison to six other methods, including Rank Product [36], T-test [37], SAM [38], LIMMA [39], Wilcoxon rank sum test [40] and M-score [41].

Pre-processing

After data generation pre-processing is typically needed. This step includes removing unwanted outliers, damaged microarrays and normalizing data distribution [42]. With the determination of biological differences between different sample groups (e.g. cases and controls), normalization is needed to avoid systematic errors and other artificial differences. In general, normalization is used to adjust the expression values so that the measurements across the samples can be compared. The most common methods for protein microarrays are quantile normalization, variance stabilizing normalization, cyclic loess and robust linear model normalization [43-47]. The normalization methods can affect the results of the analysis, but currently there is no global consensus on the best solution for this [42,48,49].

Analysis of differential expression

A common goal in the analysis of experimental results is to identify the features that distinguish different conditions. This often begins by using statistical tools to compare the expression levels of the different conditions to find differentially expressed proteins. Several different approaches have been employed for protein microarrays, including Rank Product (RP) [50], Wilcoxon rank sum test [26,50], T-test [51,52], Significance Analysis of Microarrays SAM [53], Linear Models for Microarray Data LIMMA [54], M statistic [29] and many more. The available tools often have some critical points, for instance, they can be time consuming to apply, they do not adapt well to the intrinsic properties of new data sets or the results show poor reproducibility across data sets. For example, the T-test does not work well for data sets with only few replicate samples and it relies on the assumption that the data are normally distributed, which is not usually the case [55,56]. When the distribution of the data is not known, non-parametric methods are preferred. However, they can also be dependent on the characteristics of the data. To deal with small sample sizes, it has been proposed to use relevant background knowledge [57], improved statistical tests [38,39,58,59] or more than one method to obtain better detections [45,60]. Overall, there have been strong reasons to develop different tools to improve statistical power and identify reliable features from differential expression analysis.

To address the problem of deciding which statistical test is suitable for a particular data and best adaptive to the data characteristics, we have introduced a Reproducibility-Optimized Test Statistic ROTS [58,59], which learns the optimal test statistic directly from the given data. More specifically, ROTS gives more freedom for the standard deviation term in T-test, which enables the optimization process toward maximal overlap among top-ranked proteins across bootstrap resamples. Table 1 summarizes ROTS together with several other widely used tools for differential expression analysis: Rank Product, Wilcoxon rank sum test, the ordinary T-test, SAM and LIMMA. In brief, SAM and LIMMA are modifications of the ordinary T-test whereas Wilcoxson rank sum test and Rank Product are common non-parametric methods based on ranks.

False discovery rate

Statistical testing is based on setting a null hypothesis (e.g. there is no difference in protein expression between two groups) and testing if it is true or not. Statistical significance is determined by p-value, which is the probability that we reject the hypothesis while it is actually true. If the p-value is small (e.g. 0.05 or less) then there is strong evidence against the null hypothesis, whilst a large p-value means that the evidence is weak or the test is not significant.

In protein microarray studies, a large number of statistical tests are made simultaneously, one for each protein on the array. With such multiple testing it is necessary to apply corrections when assessing protein differential expression. For example, if there are 1000 proteins on the array and we use the p-value of 0.05 to determine differential expression, then by random chance alone we would expect 50 false positive discoveries. To reduce the number of false positives, the p-value needs to be corrected. Traditionally, Bonferroni correction has been used, but it is often too conservative and may also discard many true discoveries [61]. Therefore, the False Discovery Rate (FDR) approach has been developed that is less conservative. Common methods to control FDR include the Benjamini–Hochberg procedure and permutation-based procedures [62,63].

Data visualization

For each processing step, good visualization is helpful in the

Methods	Formula with relevant details	FDR computation	Note
T-test	$t(i) = \dfrac{\bar{x}_1(i) - \bar{x}_2(i)}{s(i)}$ where $\bar{x}_j(i)$ is the average level of protein i in sample group j; $s(i)$ is the pooled standard error: $s(i) = \sqrt{\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right) \dfrac{(n_1 - 1)s_1^2(i) + (n_2 - 1)s_2^2(i)}{n_1 + n_2 - 2}}$ $s_j^2(i)$ is the variance of protein i in sample group j and n_j denotes the number of samples in group j.	Benjamini-Hochberg method	Assumes normal distribution
SAM	$d(i) = \dfrac{\bar{x}_1(i) - \bar{x}_2(i)}{s(i) + s_0}$ where $\bar{x}_j(i)$ and $s(i)$ are the same as in T-test. $s_0 > 0$ is a constant added to avoid the dependency of $d(i)$ on the protein abundance level.	permutation-based approach	
LIMMA	$t^*(i) = \dfrac{\bar{x}_1(i) - \bar{x}_2(i)}{u(i)\tilde{s}(i)}$ where $\bar{x}_j(i)$ is the same as in T-test, $\tilde{s}^2(i)$ is the posterior residual variance, a weighted average of prior and residual variance; $u(i)$ is the unscaled standard deviation.	Benjamini-Hochberg method	A linear model to determine differential expression
Wilcoxon rank sum test	$z = \dfrac{U - E(U)}{\sigma}$ where $U = n_1 n_2 + \dfrac{n_1(n_1 + 1)}{2} - T_1, E(U) = \dfrac{n_1 n_2}{2}$, $\sigma^2 = \dfrac{n_1 n_2(n_1 + n_2 + 1)}{12}$ and n_j is sample size for group j, T_1 is sum of ranks for group 1.	Benjamini-Hochberg method	Non parametric, normal distribution approximation
Rank Product	$RP^{up}(i) = \prod_{l=1}^{n} \left(\dfrac{r_l^{up}(i)}{k_l}\right)$, where $r_l^{up}(i)$ (or $r_l^{down}(i)$, respectively) are the ranks of protein i in the ordered protein lists of length k_l sorted in decreasing or increasing order for replicate l; n is the total number of replicates.	permutation-based approach	Non parametric
ROTS (maximize the reproducibility of detections in a family of modified t-statistics)	$d_\alpha(i) = \dfrac{\bar{x}_1(i) - \bar{x}_2(i)}{\alpha_1 s(i) + \alpha_0}, \alpha_0 \geq 0, \alpha_1 \in \{0,1\}$ where $\bar{x}_j(i)$ and $s(i)$ are the same as in T-test. The parameters α_0 and α_1 are determined so that they maximize the reproducibility Z-score $Z_{k,\alpha} = (R_{k,\alpha} - R^0_{k,\alpha})/s_{k,\alpha}$, where k is the top list size; $\alpha = (\alpha_0, \alpha_1)$; $R_{k,\alpha}$ and $R^0_{k,\alpha}$ are the observed and random reproducibility; $s_{k,\alpha}$ is the standard deviation of the bootstrap distribution. Reproducibility is the overlap of the top-ranked proteins across bootstrap re-samples.	permutation-based approach	Special cases: the T-test ($\alpha_0 = 0$, $\alpha_1 = 1$), the signal log-ratio ($\alpha_0 = 1$, $\alpha_1 = 0$), the SAM statistic ($\alpha_1 = 1$, α_0 is a percentile of the standard deviation).

Table 1: Summary of six statistical tools used to determine differential expression: T-test, SAM, LIMMA, Wilcoxon rank sum test, Rank Product and ROTS.

interpretation of the results, especially with high-dimensional data. Before and after normalization, histograms or boxplots can show the overall change in the shape of the data. This can also indicate possible outliers, which should be removed from the data before further analysis. Next, unsupervised clustering such as hierarchical clustering and heat maps can be used to explore known patterns or suggest new ones to be considered in order to obtain the full picture of the data. After detection of differential expression, volcano plots or the receiver operating characteristics (ROC) curves are often drawn to aid the interpretation the results. For example, volcano plots can show the fold change and the significance, while ROC curves visualize the relationships between the sensitivity and specificity of classification.

Case study: Identification of autoantibodies for T1D

To illustrate the performance of the different statistical methods in discovering new autoantibody biomarkers of T1D, we re-analysed the recently published protein microarray data by Koo et al. [29]. The data

were downloaded from Gene Expression Omnibus (GEO) database (accession number GSE50866), including measurements from serum samples of 16 T1D patients, 16 T2D patients, and 27 healthy controls with normal glucose tolerance (NGT). The data were from the ProtoArray protein microarrays, containing 9480 human proteins. The data were log transformed (base 2) and quantile normalized before the statistical analysis. The readily normalized data were downloaded from GEO.

Using six different statistical methods, we identified proteins that showed significant differences between two groups of samples at false discovery rate FDR<0.05. Following the approach of the original study, three comparisons were considered: T1D vs. NGT, T1D vs. NGT and T2D (NGT+T2D), and T1D vs.T2D. Additionally, we compared the obtained results to those of the original study using the M-statistic with $P < 0.05$ and an additional Z-score criterion [29]. Table 2 illustrates the numbers of detections with the different methods.

Overall, the widely used LIMMA and the Wilcoxon test detected only one protein in the three comparisons, whereas Rank Product resulted in very long lists of detections, suggesting that these methods may not suit the present data. In general, many more findings were made in the original study than in our comparisons. This is in line with the fact that the original study did not control the FDR levels, and was therefore more liberal than our FDR controlling strategy and thus more prone to false positive detections. SAM and ROTS detected similar numbers of proteins as significant. T-test detected more proteins in the comparisons T1D vs. NGT, and T1D vs. NGT+T2D, but none in the comparison T1D vs. T2D.

Investigation of the common detections between SAM, T-test and ROTS suggested that the overlap of the detections between these methods was often relatively small (Figure 1). In the comparison T1D vs.NGT, only ~15% (6/39) of the detections made by the T-test were found with at least one other method. With SAM the overlap was ~50% (10/19), and with ROTS 60% (9/15). In the comparison T1D vs. NGT+T2D, the overlap was ~5% (1/23) with T-test, 20% (2/10) with SAM, and 40% (2/5) with ROTS. Finally, in the comparison T1D vs. T2D, the overlap was ~50% (5/11) and ~80% (5/6) for SAM and ROTS, respectively, whereas T-test did not find any proteins. Taken together, ROTS gave the highest proportion of detections that were also found by at least one of the other methods. This supports the potential relevance of the proteins detected using ROTS, as it has been found in various contexts that detections made simultaneously by multiple different statistics are more likely to be true than those made by a single statistic [64,65]. Furthermore, out of all the 18 detections made with at least two methods across the three comparisons (11 in T1D vs. NGT, 2

in T1D vs. NGT+T2D, and 5 in T1D vs. T2D; Table 3) only two were not detected by ROTS (~10%), and these two were quite close to the borderline (FDR=0.056 and FDR=0.125).

Figure 2 further illustrates the relationship between the significant proteins detected using ROTS in the different comparisons. A total of four proteins were detected both in the comparison T1D vs. NGT and in the comparison T1D vs. NGT+T2D. These included EEF1A1 (eukaryotic translation elongation factor 1 alpha 1), EDIL3 (EGF-like repeats and discoidin I-like domains 3), ZADH1 (PTGR2, prostaglandin reductase 2), and MGC72080 (MGC72080 pseudoprotein). The latter two were not detected in the original study or with the other statistical tests considered in this study (Table 4). Prostaglandin reductase 2 (ZADH1) is involved in the metabolism of prostaglandins and has been implicated in relation to insulin sensitivity [66]. MGC72080 appears to be the product of a pseudo gene. It should be noted, however, that both of these proteins were detected with low intensity signals, and thus the interpretation of the results should be treated with caution. One alternative in such circumstances, is to filter out the low abundant proteins using, for instance, the overall average intensity or variance across the samples [67] or the combination of Z score, Chebyshev inequality precision value and coefficient of variation [68,69]. However, such implementations can be subjective and result in the loss of data describing potentially important proteins, such as lower abundance signalling molecules or receptors.

Four different proteins were detected by all the three methods (T-test, SAM and ROTS) across the comparisons. These were EEF1A1, EDIL3, SFRS3 (serine/arginine-rich splicing factor 3), and CPEB1 (cytoplasmic polyadenylation element binding protein 1). EEF1A1 was the key finding in the original study, where as SFRS3 was not detected in the original study despite the overall lower stringency used there. The T1DBase database http://www.t1dbase.org/ shows that these proteins are highly expressed in T1D related cells, such as in pancreatic islets. Our consistent findings with the different methods suggest that these proteins could be useful candidates for further experimental studies to validate their role in T1D, such as using validation methods shown in [28,29]. In addition it was notable that UBE2L3, the other validated protein in the original study, was not consistently detected by multiple statistical tests in our comparisons. This is likely due to the fact that it was detected with a wide range of signal intensities in the individual

Method	T1D vs. NGT	T1D vs. NGT+T2D	T1D vs. T2D
M-test (original study)	103	79	27
T-test	39	23	0
SAM	19	10	11
LIMMA	0	1	0
Wilcoxon rank sum test	0	1	0
Rank Product	736	885	548
ROTS	**15**	**5**	**6**

Table 2: Number of differentially expressed proteins detected with the different tools (FDR<0.05). The data contains measurements from 16 T1D patients, 16 T2D patients, and 27 healthy controls with normal glucose tolerance (NGT).
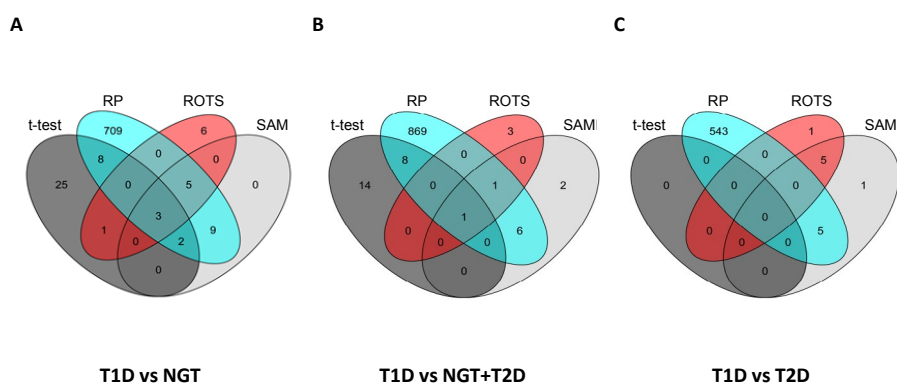

Figure 1: Overlap of the detections between the methods. The Venn diagrams illustrate the numbers and overlaps of the detections (FDR<0.05) made by the ordinary T-test, the Rank Product (RP), the SAM method, or the reproducibility-optimized test statistic (ROTS) in the different comparisons: **(A)** T1D vs NGT, **(B)** T1D vs NGT+T2D and **(C)** T1D vs T2D. LIMMA and Wilcoxon rank sum test detected almost nothing and they were excluded from here for the clarity of illustration. The numbers of detections with all the methods are shown in Table 2.
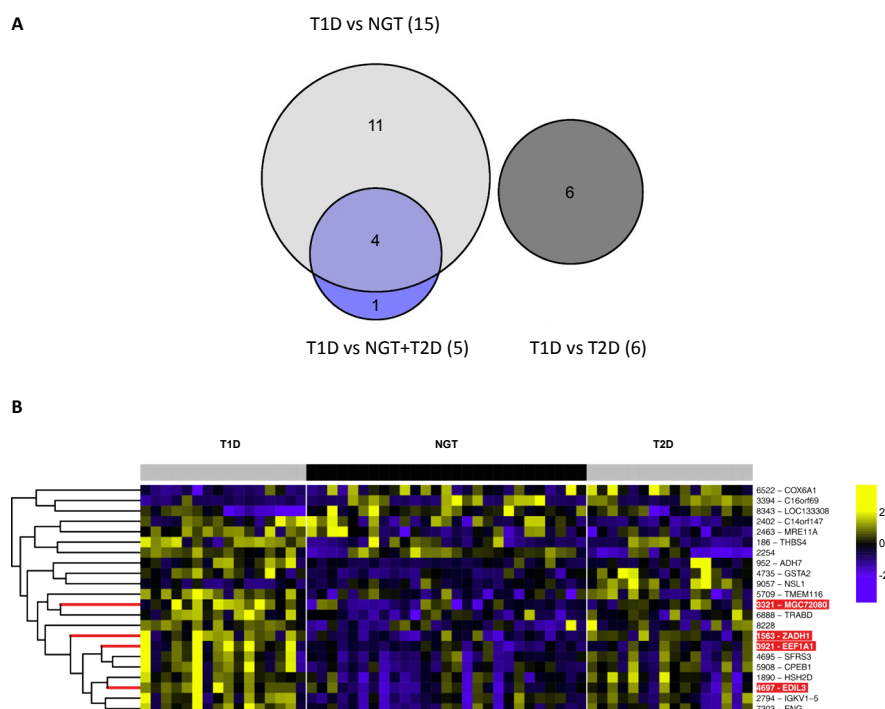
Figure 2: ROTS detections. (A) Numbers and overlaps of the ROTS detections (FDR < 0.05) in the three comparisons: T1D vs NGT, T1D vs NGT+T2D, and T1D vs T2D. **(B)** A heat map representation of all these 22 proteins detected by ROTS. The columns show the total of 59 samples from three groups (T1D, NGT, T2D) and the rows correspond to the 22 detected proteins. The 4 proteins overlapped between T1D vs NGT and T1D vs NGT+T2D are highlighted in red: EEF1A1 (eukaryotic translation elongation factor 1 alpha 1), EDIL3 (EGF-like repeats and discoidin I-like domains 3), ZADH1 (PTGR2, prostaglandin reductase 2), and MGC72080 (MGC72080 pseudoprotein). The ZADH1 and MGC72080 were not found in original study [29]. The signal intensities were $\log_2$-transformed and autoscaled similarly as in the original study [29]. Yellow represents increase and blue means decrease in abundance.

samples measured. This further highlights the importance of careful validation in independent sample cohorts.

Conclusions

This report provides an overview of the statistical and computational tools available for protein microarray data analysis and demonstrates how they can be used to help to study T1D. In addition to clarifying the expression changes and activity of known proteins relevant to T1D, protein microarrays can enable the discovery of new biomarkers to predict the onset of T1D. From the computational viewpoint, the existing literature reveals limitations in the current practices. In particular, our reanalysis of the recently published T1D data [29] demonstrated how the choice of the statistical test can have a large impact on the results obtained. For instance, the widely-used methods for differential expression analysis LIMMA, Rank Product and Wilcoxon rank sum test did not perform well in these data. To overcome such limitations, we propose to adjust the test statistic to the properties of the data by optimizing the reproducibility of detection by using bootstrap resampling. Our ROTS package performed well for the given data set, yielding the highest proportion of detections that were also found by at least one of the other methods, supporting their potential relevance.

Another important issue in the analysis to be highlighted is the use of FDR to reduce the number of false positive discoveries. For instance, in the original study of the T1D data [29], the authors used nominal p-values to determine significance, which does not control FDR and

is likely to produce several false positive findings. Accordingly, they found a large set of detections that was eventually reduced to two candidates validated in independent experiments [29]. Controlling FDR helps to eliminate many of the false positive detections.

In prospective studies of T1D risk cohorts, diabetes has been diagnosed in subjects who have not displayed any of the known autoantibodies [70,71]. Noticeably in one diabetes study, 19% of the children were negative for all autoantibodies and this significantly increased with the age of diagnosis [70]. Therefore, there is a growing demand to discover and validate new autoantibodies which can better predict the disease onset. The capabilities of protein microarray technology present many possibilities for T1D research, including the search for new autoantibodies. Moreover, proteomics markers, derived from discovery experiments in T1D research, could be profiled using targeted antibody assays to assist in risk classification, as has been investigated in the context of Systemic lupus erythematosus [72]. In such studies, flexibility in the statistical approaches employed can help to fully utilize the data. With more biological information about the relevant proteins, more complex dimensions could be integrated for further study, for example, connecting the detected proteins with their interactive pathways or networks to enhance the markers and practical applications in clinical T1D.

Finally, increasing the public availability of protein microarray data sets, in formats suitable for reanalysis, would greatly benefit the research community. If the collected data from most of the studies were

T1D vs. NGT							
SwissProt	Symbol	T-test	Wilcoxon	SAM	LIMMA	ROTS	Koo et al. [29]
P84103	**SFRS3**	*0,042*	0,262	*0,000*	0,141	*0,000*	0
O43854	**EDIL3**	*0,038*	0,146	*0,000*	0,140	*0,000*	*1*
Q9BZB8	**CPEB1**	*0,031*	0,213	*0,000*	0,140	*0,000*	*1*
P40394	ADH7	0,586	0,278	*0,000*	0,141	*0,000*	*1*
Q96JZ2	HSH2D	0,096	0,199	*0,000*	0,141	*0,000*	*1*
Q6PIH6	IGKV1-5	0,128	0,395	*0,000*	0,249	*0,050*	*1*
P68104	EEF1A1	0,060	0,146	*0,000*	0,140	*0,000*	*1*
P17813	ENG	0,173	0,262	*0,000*	0,141	*0,000*	*1*
P09210	GSTA2	*0,043*	0,146	1,000	0,141	*0,000*	0
Q8NB37	PDDC1	*0,022*	0,393	*0,000*	0,231	0,056	*1*
NA	NA	*0,041*	0,395	*0,051*	0,389	0,125	0
T1D vs. NGT+T2D							
SwissProt	Symbol	T-test	Wilcoxon	SAM	LIMMA	ROTS	Koo et al. [29]
P68104	**EEF1A1**	*0,038*	0,251	*0,000*	*0,025*	*0,000*	*1*
O43854	EDIL3	0,058	0,359	*0,000*	0,237	*0,000*	*1*
T1D vs. T2D							
SwissProt	Symbol	T-test	Wilcoxon	SAM	LIMMA	ROTS	Koo et al. [29]
NA	NA	0,281	0,186	*0,000*	0,069	*0,000*	0
Q969W0	C14orf147	0,654	0,241	*0,088*	0,354	*0,000*	0
Q5R7J7	C16orf69	0,839	0,146	*0,000*	0,257	*0,000*	0
P12074	COX6A1	0,435	0,168	*0,000*	0,354	*0,000*	0
Q96IY1	NSL1	0,907	0,112	*0,000*	0,354	*0,000*	0

Table 3: Differential expression detections made with at least two methods excluding Rank Product (FDR<0.05). Significant detections with FDR<0.05 are highlighted in bold text. The first two columns are the detected proteins and the next five columns show their corresponding FDR values from the different tests. The last column indicates whether our detections were found in the original study of Koo et al. [29]: "1" means "detected" and "0" means "not detected".

T1D vs. NGT							
SwissProt	Symbol	T-test	Wilcoxon	SAM	LIMMA	ROTS	Koo et al. [29]
Q8N8N7	**ZADH1**	0,13	0,15	1,00	0,14	*0,00*	0
Q4G0Q6	**MGC72080**	0,34	0,15	1,00	0,14	*0,00*	0
Q8NCL8	TMEM116	0,13	0,26	1,00	0,19	*0,00*	0
Q19CC5	TRABD	0,21	0,31	1,00	0,21	*0,00*	0
NA	NA	0,18	0,15	1,00	0,20	*0,00*	0
Q86UD5	LOC133308	0,99	0,52	1,00	0,17	*0,00*	0
T1D vs. NGT+T2D							
SwissProt	Symbol	T-test	Wilcoxon	SAM	LIMMA	ROTS	Koo et al. [29]
P35443	THBS4	0,15	0,36	1,00	0,24	*0,00*	0
Q8N8N7	**ZADH1**	0,24	0,36	1,00	0,26	*0,00*	0
Q4G0Q6	**MGC72080**	0,42	0,25	1,00	0,24	*0,00*	0
T1D vs. T2D							
SwissProt	Symbol	T-test	Wilcoxon	SAM	LIMMA	ROTS	Koo et al. [29]
P49959	MRE11A	0,66	0,26	1,00	0,35	*0,00*	0

Table 4: ROTS detections that were not found by the other methods (FDR<0.05). The first two columns are the detected proteins and the next five columns show their corresponding FDR values from the different tests. The last column indicates whether our detections were found in the original study of Koo et al. [29]: "1" means "detected" and "0" means "not detected".

made available, one could utilize several computational and statistical methods to identify and suggest a smaller set of relevant candidate biomarkers for further validation experiments, which would essentially save laboratorial effort and cost.

Acknowledgement

References

1. Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270: 467-470.

2. Pandey AK, Williams RW (2014) Genetics of gene expression in CNS. Int Rev Neurobiol 116: 195-231.

3. Gygi SP, Rochon Y, Franza BR, Aebersold R (1999) Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 19: 1720-1730.

4. Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, et al. (2000) Analysis of yeast protein kinases using protein chips. Nat Genet 26: 283-289.

5. Yang L, Guo S, Li Y, Zhou S, Tao S (2011) Protein microarrays for systems biology. Acta Biochim Biophys Sin (Shanghai) 43: 161-171.

6. Hall DA, Ptacek J, Snyder M (2007) Protein microarray technology. Mech Ageing Dev 128: 161-167.

7. Hu S, Xie Z, Qian J, Blackshaw S, Zhu H (2011) Functional protein microarray technology. Wiley Interdiscip Rev Syst Biol Med 3: 255-268.

8. Paweletz CP, Charboneau L, Bichsel VE, Simone NL, Chen T, et al. (2001) Reverse phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. Oncogene 20: 1981-1989.

9. Seidel M, Niessner R (2008) Automated analytical microarrays: a critical review. Anal Bioanal Chem 391: 1521-1544.

10. Wingren C, Borrebaeck CA (2009) Antibody-based microarrays. Methods Mol Biol 509: 57-84.

11. Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, et al. (2001) Global analysis of protein activities using proteome chips. Science 293: 2101-2105.

12. Zhu J, Gopinath K, Murali A, Yi G, Hayward SD, et al. (2007) RNA-binding proteins that inhibit RNA virus infection. Proc Natl Acad Sci U S A 104: 3129-3134.

13. Jones RB, Gordus A, Krall JA, MacBeath G (2006) A quantitative protein interaction network for the ErbB receptors using protein microarrays. Nature 439: 168-174.

14. Anderson KS, Sibani S, Wallstrom G, Qiu J, Mendoza EA, et al. (2011) Protein microarray signature of autoantibody biomarkers for the early detection of breast cancer. J Proteome Res 10: 85-96.

15. Zhu H, Hu S, Jona G, Zhu X, Kreiswirth N, et al. (2006) Severe acute respiratory syndrome diagnostics using a coronavirus protein microarray. Proc Natl Acad Sci U S A 103: 4011-4016.

16. Parracino A, Neves-Petersen MT, di Gennaro AK, Pettersson K, Lövgren T, et al. (2010) Arraying prostate specific antigen PSA and Fab anti-PSA using light-assisted molecular immobilization technology. Protein Sci 19: 1751-1759.

17. Carlsson A, Wuttge DM, Ingvarsson J, Bengtsson AA, Sturfelt G, et al. (2011) Serum protein profiling of systemic lupus erythematosus and systemic sclerosis using recombinant antibody microarrays. Mol Cell Proteomics 10.

18. Urbanowska T, Mangialaio S, Hartmann C, Legay F (2003) Development of protein microarray technology to monitor biomarkers of rheumatoid arthritis disease. Cell Biol Toxicol 19: 189-202.

19. Polychronakos C, Li Q (2011) Understanding type 1 diabetes through genetics: advances and prospects. Nat Rev Genet 12: 781-792.

20. Ziegler AG, Rewers M, Simell O, Simell T, Lempainen J, et al. (2013) Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. JAMA 309: 2473-2479.

21. Salminen K, Sadeharju K, Lönnrot M, Vähäsalo P, Kupila A, et al. (2003) Enterovirus infections are associated with the induction of beta-cell autoimmunity in a prospective birth cohort study. J Med Virol 69: 91-98.

22. Johnson SB, Lee HS, Baxter J, Lernmark B, Roth R, et al. (2011) The Environmental Determinants of Diabetes in the Young (TEDDY) study: predictors of early study withdrawal among participants with no family history of type 1 diabetes. Pediatr Diabetes 12: 165-1671.

23. Stechova K, Kolouskova S, Sumnik Z, Cinek O, Kverka M, et al. (2005) Anti-GAD65 reactive peripheral blood mononuclear cells in the pathogenesis of cystic fibrosis related diabetes mellitus. Autoimmunity 38: 319-323.

24. Bohmova K, Hladikova Z, Cerny M, Flajsmanova K, Vrabelova Z, et al. (2007) Cord blood cytokine profile detection in neonates with T1D parents--monitoring of cellular auto-reactivity using protein microarray. Scand J Immunol 66: 563-571.

25. Stechova K, Spalova I, Durilova M, Bartaskova D, Cerny M, et al. (2009) Influence of maternal hyperglycaemia on cord blood mononuclear cells in response to diabetes-associated autoantigens. Scand J Immunol 70: 149-158.

26. Vrabelova Z, Kolouskova S, Böhmova K, Faresjö MK, Sumnik Z, et al. (2007) Protein microarray analysis as a tool for monitoring cellular autoreactivity in type 1 diabetes patients and their relatives. Pediatr Diabetes 8: 252-260.

27. Ryden A, Stechova K, Durilova M, Faresjö M (2009) Switch from a dominant Th1-associated immune profile during the pre-diabetic phase in favour of a temporary increase of a Th3-associated and inflammatory immune profile at the onset of type 1 diabetes. Diabetes Metab Res Rev 25: 335-343.

28. Miersch S, Bian X, Wallstrom G, Sibani S, Logvinenko T, et al. (2013) Serological autoantibody profiling of type 1 diabetes by protein arrays. J Proteomics 94: 486-496.

29. Koo BK, Chae S2, Kim KM3, Kang MJ4, Kim EG3, et al. (2014) Identification of novel autoantibodies in type 1 diabetic patients using a high-density protein microarray. Diabetes 63: 3022-3032.

30. Liu H, Li J, Wong L (2002) A comparative study on feature selection and classification methods using gene expression profiles and proteomic patterns. Genome Inform 13: 51-60.

31. Jeffery IB, Higgins DG, Culhane AC (2006) Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data. BMC Bioinformatics 7: 359.

32. Saeys Y, Inza I, Larrañaga P (2007) A review of feature selection techniques in bioinformatics. Bioinformatics 23: 2507-2517.

33. Cui X, Churchill GA (2003) Statistical tests for differential expression in cDNA microarray experiments. Genome Biol 4: 210.

34. Schwämmle V, León IR, Jensen ON (2013) Assessment and improvement of statistical tools for comparative proteomics analysis of sparse data sets with few experimental replicates. J Proteome Res 12: 3874-3883.

35. Liang P, Pardee AB (2003) Analysing differential gene expression in cancer. Nat Rev Cancer 3: 869-876.

36. Breitling R, Armengaud P, Amtmann A, Herzyk P (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett 573: 83-92.

37. Box JB (1987) Guinness, Gosset, Fisher, and Small Samples. Statistical Science 2: 45-52.

38. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 98: 5116-5121.

39. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3: Article3.

40. Mann HB, Whitney DR (1947) On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Ann Math Stat 18: 50-60.

41. Love B (2007) The Analysis of Protein Arrays in Functional Protein Microarrays in Drug Discovery. USA: {CRC} Press.

42. Quackenbush J (2002) Microarray data normalization and transformation. Nat Genet 32: 496-501.

43. Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19: 185-193.

44. Rudy J, Valafar F (2011) Empirical comparison of cross-platform normalization methods for gene expression data. BMC Bioinformatics 12: 467.

45. Turewicz M, May C, Ahrens M, Woitalla D, Gold R, et al. (2013) Improving the default data analysis workflow for large autoimmune biomarker discovery studies with ProtoArrays. Proteomics 13: 2083-2087.

46. Sboner A, Karpikov A, Chen G, Smith M, Mattoon D, et al. (2009) Robust-linear-model normalization to reduce technical variability in functional protein microarrays. J Proteome Res 8: 5451-5464.

47. Huber W, von Heydebreck A, Sültmann H, Poustka A, Vingron M (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics 18 Suppl 1: S96-104.

48. Wu Z (2009) A review of statistical methods for preprocessing oligonucleotide microarrays. Stat Methods Med Res 18: 533-541.

49. Zhu X, Gerstein M, Snyder M (2006) ProCAT: a data analysis approach for protein microarrays. Genome Biol 7: R110.

50. Pollard HB, Srivastava M, Eidelman O, Jozwik C, Rothwell SW, et al. (2007) Protein microarray platforms for clinical proteomics. Proteomics Clin Appl 1: 934-952.

51. Sreekumar A, Nyati MK, Varambally S, Barrette TR, Ghosh D, et al. (2001) Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins. Cancer Res 61: 7585-7593.

52. Sundaresh S, Doolan DL, Hirst S, Mu Y, Unal B, et al. (2006) Identification of humoral immune responses in protein microarrays using DNA microarray data analysis techniques. Bioinformatics 22: 1760-1766.

53. Hueber W, Kidd BA, Tomooka BH, Lee BJ, Bruce B, et al. (2005) Antigen microarray profiling of autoantibodies in rheumatoid arthritis. Arthritis Rheum 52: 2645-2655.

54. Gong W, He K, Covington M, Dinesh-Kumar SP, Snyder M, et al. (2008) The development of protein microarrays and their applications in DNA-protein and protein-protein interaction analyses of Arabidopsis transcription factors. Mol Plant 1: 27-41.

55. Shapiro SS, Wilk MB (1965) An Analysis of Variance Test for Normality (Complete Samples). Biometrika 52: 591-596.

56. D'Agostino DB, Belanger A, D'Agostino RB (1990) A Suggestion for Using Powerful and Informative Tests of Normality. Am Stat 44: 316-332.

57. Nikulin V (2014) On a solution for the high-dimensionality-small-sample-size regression problem with several different microarrays. Int J Data Min Bioinform 9: 221-234.

58. Elo LL, Hiissa J, Tuimala J, Kallio A, Korpelainen E, et al. (2009) Optimized detection of differential expression in global profiling experiments: case studies in clinical transcriptomic and quantitative proteomic datasets. Brief Bioinform 10: 547-555.

59. Elo LL, Filén S, Lahesmaa R, Aittokallio T (2008) Reproducibility-optimized test statistic for ranking genes in microarray studies. IEEE/ACM Trans Comput Biol Bioinform 5: 423-431.

60. Kallionpää H, Elo LL, Laajala E, Mykkänen J, Ricaño-Ponce I, et al. (2014) Innate immune activity is detected prior to seroconversion in children with HLA-conferred type 1 diabetes susceptibility. Diabetes 63: 2402-2414.

61. Hochberg Y (1988) A Sharper Bonferroni Procedure for Multiple Tests of Significance. Biometrika 75: 800-802.

62. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57: 289-300.

63. Xie Y, Pan W, Khodursky AB (2005) A note on using permutation-based false discovery rate estimates to compare different analysis methods for microarray data. Bioinformatics 21: 4280-4288.

64. Yang YH, Xiao Y, Segal MR (2005) Identifying differentially expressed genes from microarray experiments via statistic synthesis. Bioinformatics 21: 1084-1093.

65. Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, et al. (2012) Wisdom of crowds for robust gene network inference. Nat Methods 9: 796-804.

66. Zhu X, Walton RG, Tian L, Luo N, Ho SR, et al. (2013) Prostaglandin A2 enhances cellular insulin sensitivity via a mechanism that involves the orphan nuclear receptor NR4A3. Horm Metab Res 45: 213-220.

67. Bourgon R, Gentleman R, Huber W (2010) Independent filtering increases detection power for high-throughput experiments. Proc Natl Acad Sci USA 107: 9546-9551.

68. Auger I, Balandraud N, Rak J, Lambert N, Martin M, et al. (2009) New autoantigens in rheumatoid arthritis (RA): screening 8268 protein arrays with sera from patients with RA. Ann Rheum Dis 68: 591-594.

69. Le Roux S, Devys A, Girard C, Harb J, Hourmant M (2010) Biomarkers for the diagnosis of the stable kidney transplant and chronic transplant injury using the ProtoArray® technology. Transplant Proc 42: 3475-3481.

70. Wang J, Miao D, Babu S, Yu J, Barker J, et al. (2007) Prevalence of autoantibody-negative diabetes is not rare at all ages and increases with older age and obesity. J Clin Endocrinol Metab 92: 88-92.

71. Tiberti C, Buzzetti R, Anastasi E, Dotta F, Vasta M, et al. (2000) Autoantibody negative new onset type 1 diabetic patients lacking high risk HLA alleles in a caucasian population: are these type 1b diabetes cases? Diabetes Metab Res Rev 16: 8-14.

72. Borrebaeck CA, Sturfelt G, Wingren C (2014) Recombinant antibody microarray for profiling the serum proteome of SLE. Methods Mol Biol 1134: 67-78.