

UPTEC X 03 003
JAN 2003

ISSN 1401-2138

HANNA GÖRANSSON

Identification of diagnostic markers for classification of thyroid tumours using expression array analysis

Master's degree project



Molecular Biotechnology Programme
Uppsala University School of Engineering

UPTEC X 03 003	Date of issue 2003-01-20	
Author	Hanna Göransson	
Title (English)	Identification of diagnostic markers for classification of thyroid tumours using expression array analysis	
Title (Swedish)		
Abstract	<p>The identification of the different subclasses of follicular thyroid tumours is important in the choice of thyroid cancer patients. Expression analysis may provide an alternative to morphology-based tumour classification. Using microarrays as a screening tool allows the gene expression profiles of thousands of genes to be monitored simultaneously. The aim of the project was to find markers in order to distinguish follicular carcinomas and adenomas on a molecular level. Tumour samples from 10 follicular adenomas and 7 follicular carcinomas were used. A handful of candidate genes were discovered and verified with real-time PCR. Supervised learning was performed providing tumour classification, with an error rate of only 5.8% using three genes. Further research will investigate whether these potential candidates may be used in a clinical application, in order to improve the diagnostic accuracy in thyroid cancer.</p>	
Keywords	Microarray, thyroid tumours, classification, diagnostic markers	
Supervisors	Anders Isaksson Department of Genetics and Pathology, Uppsala University	
Examiner	Ulf Pettersson Department of Genetics and Pathology, Uppsala University	
Project name	Thyroid project	
Sponsors		
Language	English	
Security		
ISSN 1401-2138	Classification	
Supplementary bibliographical information	Pages 34	
Biology Education Centre Box 592 S-75124 Uppsala	Biomedical Center Tel +46 (0)18 4710000	Husargatan 3 Uppsala Fax +46 (0)18 555217

Identification of diagnostic markers for classification of thyroid tumours using expression array analysis

Hanna Göransson

Sammanfattning

För att kunna ge cancerpatienter rätt behandling är det viktigt att upptäcka tumörerna i tid och att kunna klassificera dem. Tidigare har tumörers klassificering enbart baserats på vävnadens utseende. Med hjälp av en teknik som mäter genaktivitet i tumören kan man få ett nytt alternativ till nuvarande tumörklassificering. Med detta verktyg kan man för tusentals gener samtidigt mäta vilka gener som används i tumörcellerna, och i vilken utsträckning. Det finns två typer av follikulär sköldkörtelcancer, adenom (godartad) och carcinom (elakartad), som normalt är svåra att skilja från varandra. Målet med projektet var att bland tusentals gener hitta ett fåtal gener som används ovanligt mycket eller ovanligt lite i de två typerna. Dessa gener skulle kunna användas för att skapa en genprofil och därmed användas för att skilja mellan de två typerna av sköldkörteltumörer. Förhoppningen är att de generna man funnit ska kunna användas kliniskt i ett diagnostiskt test, som ett komplement när diagnos är svår att ställa.

Tumörprover från 10 follikulära adenom och 7 follikulära carcinom användes i studien. En handfull kandidatgener hittades och resultaten för dessa säkerställdes. En datamodell byggdes med hjälp av matematiska metoder och när den testades kunde de två typerna av sköldkörtelcancer skiljas åt med bara ett misstag. Vidare undersökning av resultaten krävs innan ett diagnostiskt test kan utvecklas i praktiken, men de visar ändå att information från genprofiler har stor potential att kunna användas för att klassificering av tumörer. De skulle därmed kunna bidra till möjligheten att skraddarsy behandling till patienter i framtiden.

**Examensarbete 20p i Molekylär bioteknikprogrammet
Uppsala universitet Januari 2003**

Contents

1 INTRODUCTION	2
2 DIAGNOSIS AND TREATMENT OF THYROID CANCER	3
3 MICROARRAY TECHNOLOGY	5
3.1 APPLICATIONS	5
3.2 MICROARRAYS	6
3.3 EXPERIMENTAL PROCEDURE	6
4 MATERIALS AND METHODS	8
4.1 ARRAY FABRICATION	8
4.2 RNA EXTRACTION	8
4.3 PROBE PREPARATION, HYBRIDISATION AND IMAGE PROCESSING	8
4.4 DATA ANALYSIS	9
4.5 REAL-TIME PCR	10
5 RESULTS	11
6 DISCUSSION	16
REFERENCES	21
THEORETICAL APPENDIX	23
NORMAL FUNCTION OF THE THYROID	23
DATA ANALYSIS	24
LOW-LEVEL ANALYSIS	24
HIGH-LEVEL ANALYSIS	25
REAL-TIME PCR	27
APPENDIX I	28
APPENDIX II	29

1 Introduction

The field of cDNA microarray technology allows the efficient measurement of expression levels of several thousand genes simultaneously in a high-throughput fashion. Global expression analysis offers new opportunities to obtain molecular profiles in patient samples, thereby providing invaluable information about disease progression leading to improved diagnosis and treatment.

Today, pathological diagnosis and thyroid cancer classification is primarily based on the microscopic view of the solid tumour, even though this has certain limitations. Tumours with similar appearance can follow significantly different clinical courses, making the prediction of responses to treatments and clinical outcomes difficult.

There is an urge to change the emphasis from morphology-based tumour classification to molecular. Measuring the expression profiles provides a snapshot of mRNA levels at the time of sample preparation and could describe variation in cells and tissues, thereby providing an alternative way of tumour classification. This genomic-scale approach has already identified genes useful for predicting clinical behaviour, which is particularly relevant for cases where malignant and benign lesions can not be clearly distinguished morphologically. Further investigation during the post genomic era is likely to provide genes useful as diagnostic, prognostic and treatment response biomarkers that may lead to personalised management of patients.

In thyroid cancer, the differentiation between follicular carcinomas and adenomas is difficult because of overlapping morphology. In unclear cases, radical, surgical treatment is the alternative chosen, even though many tumours are unnecessarily removed since the true frequency of malignancy is low. The identification of molecular markers that would improve the preoperative diagnosis is likely to have a significant clinical impact, but no follicular thyroid cancer-specific gene transcript has been found so far.

The aim of this project is to find diagnostic markers that could correctly identify follicular thyroid cancer preoperatively, through the use of cDNA microarrays as a screening tool. The ideal thyroid malignancy marker would be both sensitive and specific, allowing for accurate categorisation of follicular adenomas and carcinomas providing a useful tool for presurgical diagnosis that can guide the choice of treatment.

2 Diagnosis and treatment of thyroid cancer

The thyroid is a small hormone-producing butterfly-shaped gland located in the neck. It has a weight of around 20 grams and consists of two connected lobes. It is formed by a large number of follicles of varying size, which regulate the metabolism by production of the thyroid hormones triiodothyronine (T_3) and thyroxine (T_4). C-cells are situated between the follicles. They produce the hormone calcitonin, which is a calcium-regulating hormone. For more information on normal function of the thyroid see Theoretical appendix.

Thyroid cancer is fairly uncommon, accounting for 1.2% of all new cancers in the United States annually. Women face the greatest risk of developing thyroid tumours, seven times more often than men. Although thyroid cancer is still a cancer that requires treatment and life-long monitoring, it is considered a “good cancer” since survival rates are high, with 95% of all thyroid cancer patients surviving without reoccurrence. There are four types of thyroid cancer; papillary, follicular, medullary, and anaplastic.

Papillary cancer is the most common type of thyroid cancer, maybe because papillaries are quite common in the thyroid gland. It usually involves one side of the thyroid and the cure rate is very high.

The second most common type, follicular cancer, is considered to be more malignant. It is more common in older people, but the long-term survival rate is still high. Mortality is related to the degree of vascular invasion, and distant metastases are rather common. Age is a very important factor in terms of prognosis, for example patients over 40 have a more aggressive disease. Unlike the papillary histological type, follicular carcinoma usually exists as a solitary thyroid tumour that is more or less encapsulated. Depending on the degree of invasiveness, the tumour is classified as either the more common minimally invasive or widely invasive [1].

Fine needle aspiration (FNA) has become the principal method for preoperative evaluation of thyroid nodules, and it is well suited for identifying papillary carcinomas. On the other hand, about 80% of patients with follicular lesions undergo thyroidectomy for benign follicular adenomas. FNA interpretation demands an experienced pathologist, and in that aspect molecular analysis offers a potentially more objective way, useful as a complement in difficult cases [2].

The way the tumour is examined under the microscope may be unreliable to make definitive diagnosis of follicular cancer after the time of surgery, which is a problem not seen with other types of thyroid cancer. Cytologically, minimally invasive carcinomas can not be distinguished from adenomas and even fine needle biopsy (FNB) is incapable of differentiating them. The inability depends on the fact that FNB only provides a visualisation of cells and not the organisation of cells or the whole tissue, which would be required since by definition a follicular carcinoma is characterised by a growth through the capsule, not detectable when examining individual cells.

As stated above, the major limitation of FNA or FNB, without aspiration, is when results are not sufficient to permit diagnosis. This is the case for 5-10% of follicular tumours, which raises the question of what treatment to prescribe. Generally, surgery is the most common treatment. It is not very risky, and the main concern is potential indirect damage to the vocal

cords. During surgery, either part or all of the thyroid gland is removed. The side effects of thyroidectomy include damage to parathyroid glands with hypoparathyroidism, which causes low calcium levels in the bloodstream and thereby elevated excitability of nerves [2]. Some recommend surgical treatment for all, while others limit surgery to nodules considered suspect by other clinical or laboratory findings. Even when only these cases are chosen for surgery, the malignancy is as low as between 10% and 20% [1].

After surgery, most patients receive radioactive iodine (RAI) treatment to kill any remaining thyroid tumour tissue. Side effects of RAI used today include nausea, inflammation of the salivary glands, vomiting, exhausting and bone marrow suppression. Many patients are also classified as hypothyroid, requiring thyroid hormone for the rest of their lives.

The follicular tumours are even more problematic from another view point. It is the fact that a follicular adenoma may progress to a follicular carcinoma [3]. In rat an adenoma with time became mixed, containing both benign and malignant components, and at the end only malignant components were observed [4]. Only indirect evidence is available in humans, for example areas resembling follicular adenoma are found in follicular carcinomas. There are markers considered to be indicators of tumour progression, signalling a tendency for conversion of adenomas to carcinomas [3]. It is therefore suggested that some morphologically suspicious follicular adenomas could be considered as potential early cancers [5].

Considering the unpleasant side effects, and the fact that a large percentage of the tumours turn out to be benign, unnecessarily extensive thyroid surgery should and could be avoided if more reliable markers for the presence of thyroid cancer were available for preoperative evaluation. Moreover, the potential progression of follicular tumours will be further analysed through development of reliable molecular diagnostic tests, promoting the gain of knowledge in the difficult area of follicular thyroid tumours. That is why our main concern during this study was to find markers for identification and separation of the different subclasses of follicular thyroid tumours.

3 Microarray technology

3.1 Applications

Microarrays offer a new approach to study the cellular mechanisms of pathways in anticancer drug resistance and to predict drug sensitivity and unexpected side effects. However, the range of applications for microarray technology is enormous and not limited to cancer research only. For example, the temporal impact on gene expression by drugs, environmental toxins, or oncogenes may be mapped and regulatory networks and coexpression patterns can then be investigated. Expression patterns of many previously uncharacterised genes may provide clues to their possible function by comparison, and combined with metabolic schemas they will contribute to the understanding of how pathways are changed under varying conditions.

Initially, microarray technology was developed to study differential gene expression [6], and nowadays these methods also allow the analysis of copy number imbalances [7] and SNP analysis concerning mutation or polymorphism detection [8]. Many of the principles of global analysis using microarrays are in principle applicable at the RNA, DNA or even protein level [9].

In the past couple of years, microarrays have been used to investigate tumour classification. In order to find new and individually adjusted approaches, tumour classification of today still has to be improved a lot and microarray studies are likely to have a profound effect on this field. Microarrays have a potential of providing a way to combine large-scale molecular analysis of expression profiles with classic morphologic and clinical methods for better cancer diagnosis and outcome prediction.

Today there is an exponential growth in microarray-based studies giving new insight into tumour biology. In the future, microarray technology is likely to be used for both diagnostic and prognostic purposes. A study by Kim *et al.* [10], where the identification of biomarker genes in cancer is subsequently tested by conventional biochemical assays, demonstrates the potential value of cDNA microarray studies.

Problems encountered have been how to determine which genes are relevant to the diagnostic system. Therefore, combining expression microarray analysis with other approaches can complement the situation at hand, and offer the possibility to focus on significantly smaller subsets of relevant genes. One example is using microarray technology as a screening tool with the aim to identify genes useful in a diagnostic test, which will serve to support the pathologist in situations of diagnosis difficulties. Finally, validation of the relative expression obtained from microarrays is critical, e.g. by RT-PCR [11], to make sure the correct genes are investigated, since there is a risk of cross-hybridisation on the array.

3.2 Microarrays

Several methods have been used to generate DNA microarrays for genome-wide gene expression studies.

Spotted cDNA microarrays are typically made by the mechanical deposition of solutions containing individual PCR-amplified cDNA fragments (~500-2000 bps) in small spots on glass slides by a robot. Spot sizes range between 80 and 150 μm in diameter, and arrays containing up to 80000 spots can be obtained. Gene sequences are chosen from public databases, which contain resources for obtaining clones to be amplified and purified before spotting on the array. Another possibility is to use gene specific oligonucleotides when spotting cDNA microarrays.

The Affymetrix Genechip [12] is produced by using a modification of semiconductor photolithography for *in situ* synthesis of tens of thousands of oligonucleotides onto silicone chips. Generally they contain 16-20 oligonucleotides representing each gene, which is also matched with another almost identical oligonucleotide differing only by a central single base mismatch. This is used to determine non-specific binding and results in the ability of Affymetrix Genechips to measure absolute expression. Since spotted microarrays do not contain an absolute quantity of cDNA, there is a need to use two samples simultaneously in order to obtain a ratio, which can then be used to quantify relative expression. One of the main advantages of spotted arrays is the lower price and flexibility in design, which is not possible for purchased Affymetrix Genechips, even though this is compensated by a number of different types being available on the market.

3.3 Experimental procedure

The sample of interest is typically compared with a common reference, which should have an adequate representation of the majority of genes on the array. It is stated by Yang *et al.* [13] that pooled reference samples should not be designed only based on the representation of individual genes in each cell line. It is also necessary to consider expression levels for genes within the cell lines. The reason for doing this is that a larger number of cell lines do not necessarily improve overall gene representation since rare transcripts become undetectable when diluted in a pool of many cell lines. Using a common reference is also beneficial to be able to compare microarrays from different projects and to research groups interested in common aspects.

The importance of replicates is also worth mentioning, because the variability can be very high in microarray experiments. In order to reduce variation due to the efficiency of dye incorporation when different dyes are used, many groups perform “dye-swap” experiments.

A standard microarray experiment involves reverse transcription of the isolated RNA into target cDNA. This is carried out through a process of direct labelling, with fluorescent or radio-labelled deoxynucleotides, but can also be done in an indirect way where the fluorophores are associated with the samples after cDNA synthesis or even after the hybridisation. Washing under stringent conditions takes place after hybridisation, to remove unspecific target binding, followed by scanning. Fluorescence intensities are measured using a

laser scanner and the two fluorescent images (Cy3 and Cy5) constitute the raw data from which differential gene expression ratio values are calculated.

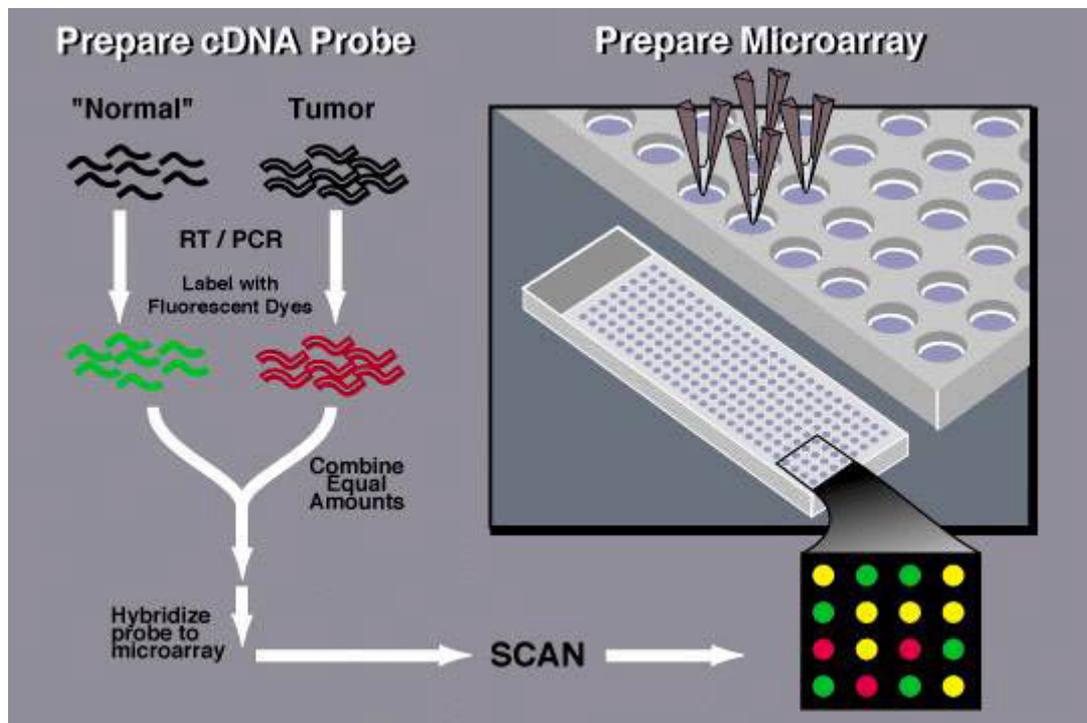


Fig 1. The procedure of standard microarray experiments involves reverse transcription of the isolated RNA into target cDNA through a process of direct labelling, in the presence and incorporation of fluorescent or radio-labelled deoxynucleotides, but can also be done in an indirect way where the fluorophores are associated with the samples after cDNA synthesis or even after the hybridisation. The arrays are then washed under stringent conditions, to remove unspecific target binding, followed by scanning. The illustration was adapted from <http://microimm.queensu.ca/micr930/> (Jan 2003).

Efforts are being made to propose standards for documentation and availability of information about microarray experiments, and it is likely that they will be seen in the near future [14].

4 Materials and methods

4.1 Array fabrication

cDNA microarrays containing about 8000 genes spotted at the WCN Expression platform were used. cDNA clones were obtained from Research Genetics sequence-verified human cDNA clones. They were amplified with PCR, and all clones were printed in duplicate using printing solutions with 3xSSC and 0.005%SDS for Batch 4 and 5, and using 30% DMSO for Batch 6 onto GAPII-slides (Corning), and crosslinked at 950 μ J with a Stratalinker.

4.2 RNA extraction

RNA from small pieces of thyroid tissue samples was used. Normal thyroid tissue from a pool of patients was used as a common reference for all samples. The RNA extraction was performed with Trizol (Life Technologies) and carried out according to the manufacturers protocol slightly modified and the samples were stored at -80°C. For protocol details see Appendix I.

A microcapillary-based device, the Agilent Bioanalyzer (Agilent Technologies), was used for RNA quality control. Spectrophotometry was used for concentration determination.

4.3 Probe preparation, hybridisation and image processing

cDNA synthesis, hybridisation and detection were carried out using the TSA Labelling and Detection Kit (NEN Life Science Products), which includes all the necessary reagents. The TSA probe labelling and array hybridization were performed as described in the instruction manual (*MICROMAX Human cDNA Microarray System*, NEN Life Science Products) with minor modifications. For protocol details see Appendix II [15].

After total RNA extraction the RNA, without further amplification, is converted into Fluorescein (FL) or Biotin labelled cDNA respectively, using reverse transcriptase and nucleotide analogs, for use as individually traceable gene targets. The cDNA probes are then combined and allowed to hybridise to the array in an overnight incubation. The sequential Cy3 and Cy5 TSA detection process follows hybridisation and stringency washes. The microarray is first incubated with anti-FL-HRP (horseradish peroxidase). This antibody-enzyme conjugate specifically binds to the hybridised FL labelled cDNA probe. HRP catalyses the deposition of Cy3 labelled tyramide amplification reagent, resulting in numerous Cy3 labels adjacent to the immobilised HRP. The amount of tyramide relative to cDNA FL or Biotin is greatly amplified because of the enzymatic process. HRP inactivation takes place before the next TSA step, in which streptavidin-HRP binds to the hybridised Biotin labelled cDNA probe. This time deposition of Cy5 labelled tyramide is amplified by the HRP portion of the enzyme. Fluorescence detection is next and the arrays are scanned using the GenePix 4000B scanner from Axon Instruments.

4.4 Data analysis

GenePix 4.0 Microarray Analysis Software (Axon Instruments, California, USA) was used to identify spots and extract information from the TIFF files obtained from slide scanning. Spots were flagged as good, bad, absent or not found in order to be able to discard values from further analysis.

Normalisation was performed using the robust scatter-plot smoother lowess function in the program R, a freeware statistics language, which together with the SMA [16, 17] package can perform various normalisations and analysis. Print-tip normalisation was the alternative chosen, i.e. the normalisation algorithm was applied locally, to help correct for spatial variation in the array between subsets spotted with a single pin. For more details on normalisation see Theoretical Appendix.

Before further analysis of the data, certain filters were applied to reduce the noise and experimental variation. The following filtering criteria were used:

- spots flagged bad, absent or not found were discarded
- a mean/median filter proposed by Tran *et al.* [18] was used, where spots with ratios between mean and median less than a threshold value were discarded, a value of 0.25 was appropriate for our dataset
- genes whose dye-swap experiments had a deviation of more than two-fold were discarded, with the exception of experiments with intensities ± 3 -fold
- genes with less than 3 out of 10 values for adenomas and 3 out of 7 values for carcinomas were discarded, considered not reliable due to too few values
- genes considered containing unchanged expression patterns, with a max ratio/min ratio < 2 , were discarded.

This was a rather stringent filtering and reduced the data dimensions from ~ 7400 genes down to approximately 2400. Further analysis of the pre-processed data was performed with both SAM (statistical analysis of microarrays) [19] and supervised machine learning approaches.

The statistical technique SAM was used to identify candidate genes in the data set. It is an Excel Add-In that correlates gene expression data to various clinical parameters such as treatment and diagnosis categories. This is accomplished by computing a statistic d_i for each gene i , measuring the strength of relationship between gene expression and the response variable. It uses repeated permutations of the data to determine if the expression of any genes is significantly related to the response. A tuning parameter *delta*, adjusted by the user, determines the cut-off for significance. SAM can also be used as a method for ranking genes in a non-wrapper approach to find significantly expressed genes in a set of microarray experiments, and thereby discover potential candidate genes.

For the supervised learning we used Fisher's linear discriminant function, which has been shown to provide a practical and accurate method for tumour classification using gene expression profiles [20]. The Fisher approach is based on the idea to transform the multivariate observations present in a dataset into univariate observations such that the projected observations are separated as much as possible.

The feature selection procedure "Greedy pairs" was chosen. It is a method based on evaluating genes in pairs and how well a pair in combination distinguishes two experiment classes [21]. According to the results presented in the article, pairs of genes reveal

information about class differences not discovered otherwise, i.e. when using the genes individually, and taking advantage of this extra information will improve class prediction. The feature selection was performed with a wrapper approach, i.e. the classification accuracy was used as a criterion function when ranking genes. This makes the feature selection dependent on the learning algorithm, which was the same algorithm as the one used to build the final classifier.

Prediction performance was estimated by leaving one sample out at each round, a so called leave-one-out cross-validation. A permutation test was also applied to the dataset, i.e. randomly distributing the class labels, in order to investigate if the results could have arisen by chance.

4.5 Real-time PCR

Quantitative RT-PCR assays were performed on the ABI Prism 7700 Sequence Detection System with protocols supplied by manufacturers to confirm microarray data results. Primers used were ordered from the Assay-On-Demand service available from Applied Biosystems for the cases of available primers, and in the other cases primers were designed using the Primer Express software. Data were normalised using human β -actin. For more general information and information on the mechanism of real-time PCR, see Theoretical appendix.

5 Results

The situation of follicular thyroid tumours is problematic and there is a need for finding molecular markers discriminating them. Many research groups have focused on finding differences in expression between healthy and tumour tissues to identify patterns and biomarkers for tumours [22]. These patterns can be used to find differences between classes of tumours, and thereby constitute the basis upon which classification can be performed. Since the results so far have been promising, the arguments for using expression array analysis for finding markers to facilitate tumour classification in our study are well-founded.

In this project expression studies of thyroid tumours were performed. Tissue samples from ten adenomas and seven carcinomas were used in the experiments, and dye-swap experiments were also carried out. All the RNA samples passed the quality control, performed after extraction and purification, to detect contamination or degradation. All in all, 34 arrays were taken through the experimental and analytical procedure.

To define and distinguish the spots from non-specific background signals, the scanned images were overlaid with a grid specifying each target location. All clone information was attached to the spots, as well as quantification of signal intensities. Next, the lowess normalisation process was performed to adjust the individual hybridisation intensities in order to balance them appropriately so that meaningful comparisons could be made.

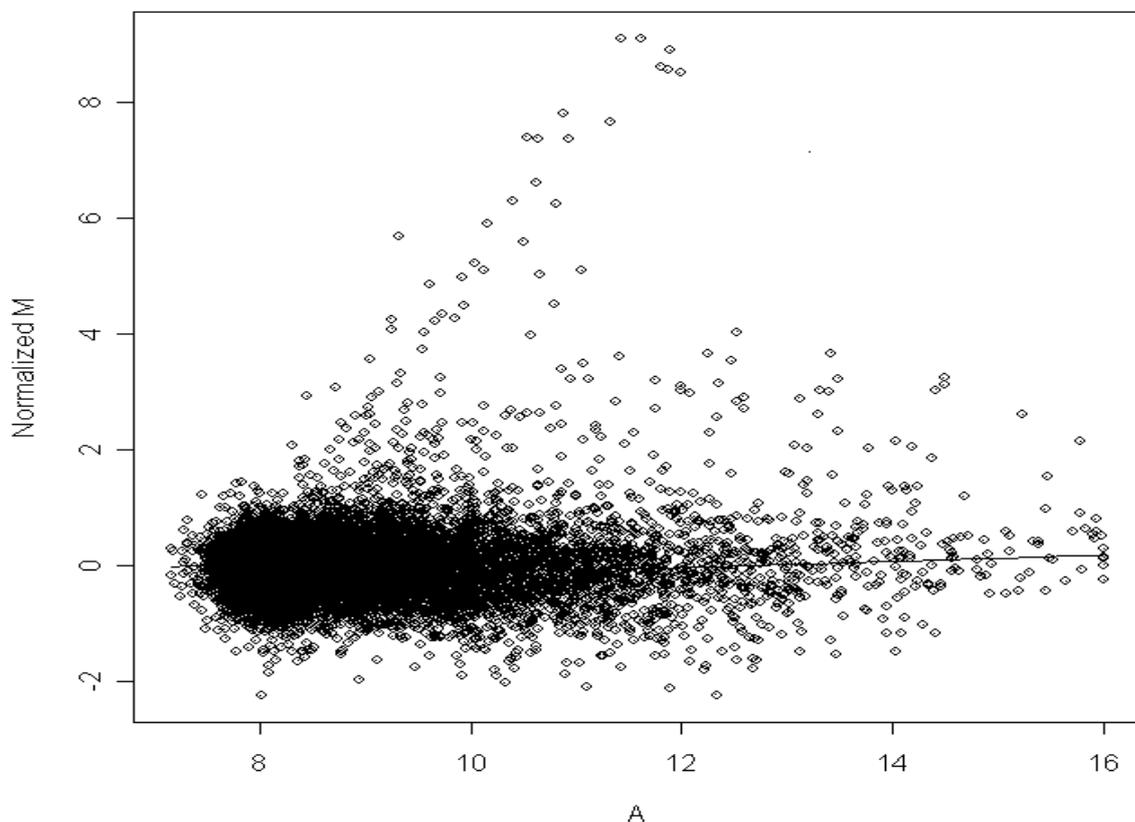


Fig. 1 shows that application of print-tip lowess normalisation has corrected for both systematic variation as a function of intensity and spatial variation between spotting pins.

The process scales spot intensities so that the normalised ratios provide an approximation of the ratio of gene expression between the two samples. Lowess detects systematic deviations and corrects them by carrying out a local weighted linear regression, as a function of \log_{10} intensity and subtracting the best-fit average \log_2 ratio from the observed ratios, for each data point. The normalisation used was a print-tip variant, i.e. the lowess normalisation was applied to each group of array elements deposited by a single spotting pin, to correct for spatial variation in the array. A balanced distribution of expression ratios independent of intensity was produced, which can be seen in Fig.1. For a general comparison between the ratio-intensity plots before and after application of the normalisation method, see Theoretical Appendix.

Data from microarrays are usually pre-processed in order to detect and remove noise and unreliable measurements. With any filtering method there is a trade-off between optimising measurement reliability and avoiding loss of information. As mentioned earlier, the reduced the data set of approximately 2400 genes was used for further analysis using SAM and other statistical methods. For further information on general data analysis see Theoretical appendix.

In order to find interesting genes, i.e. genes with expressing profiles which discriminated between tumour subtypes, SAM analysis was performed as a first step. A search for potentially interesting genes was also carried out by calculating the mean fold change within each class and investigating the result looking for the largest differences, i.e. a separation of classes.

A number of genes ranked fairly high with SAM and possessing a relatively high fold change difference were chosen for further analysis. Real-time PCR using Taqman was performed on four candidates. The results can be seen in Fig.2 (a-d).

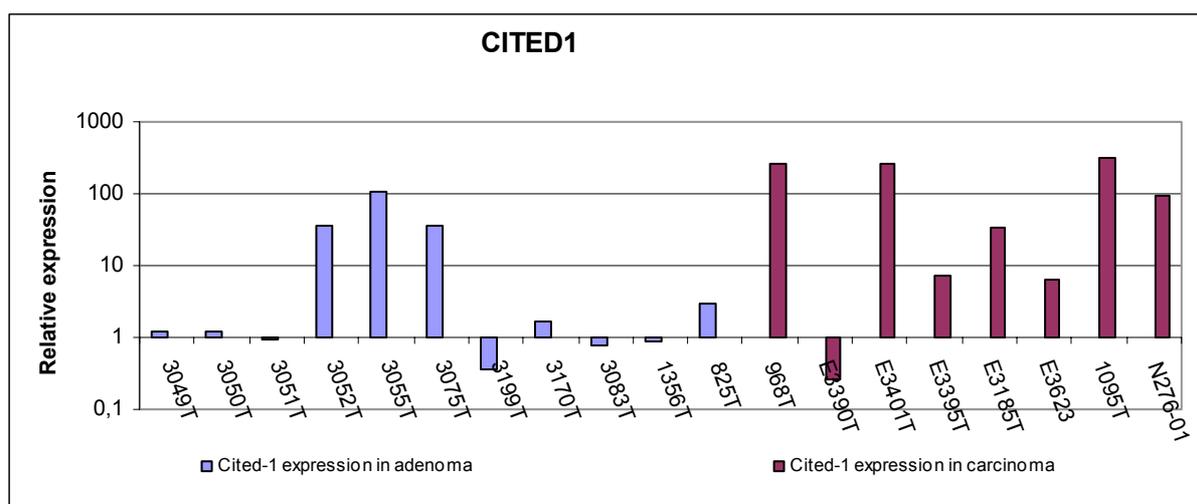


Fig.2 (a) shows relative expression of CITED1 in follicular adenomas (blue) and carcinomas (red).

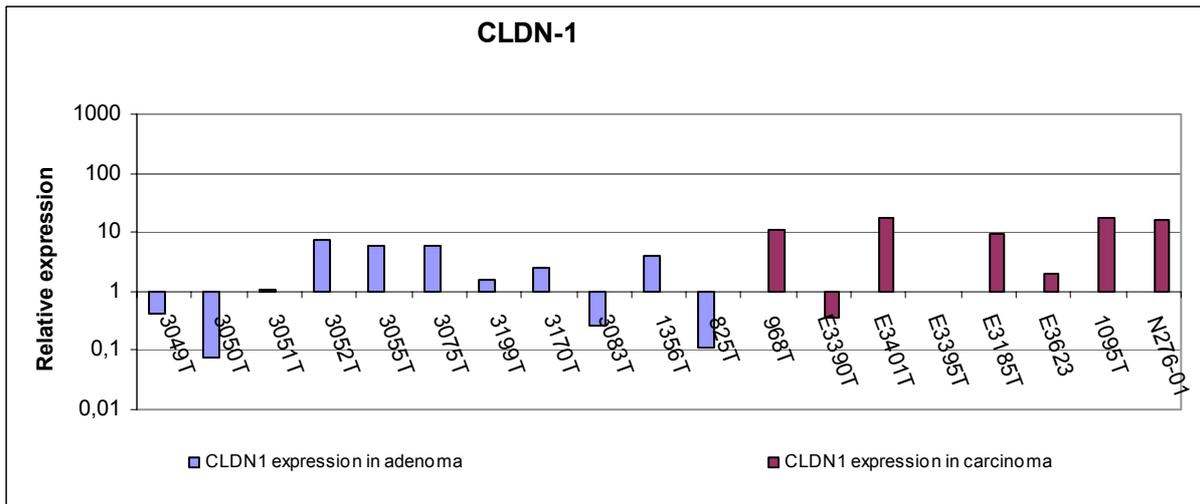


Fig.2 (b) shows relative expression of CLDN-1 in follicular adenomas (blue) and carcinomas (red).

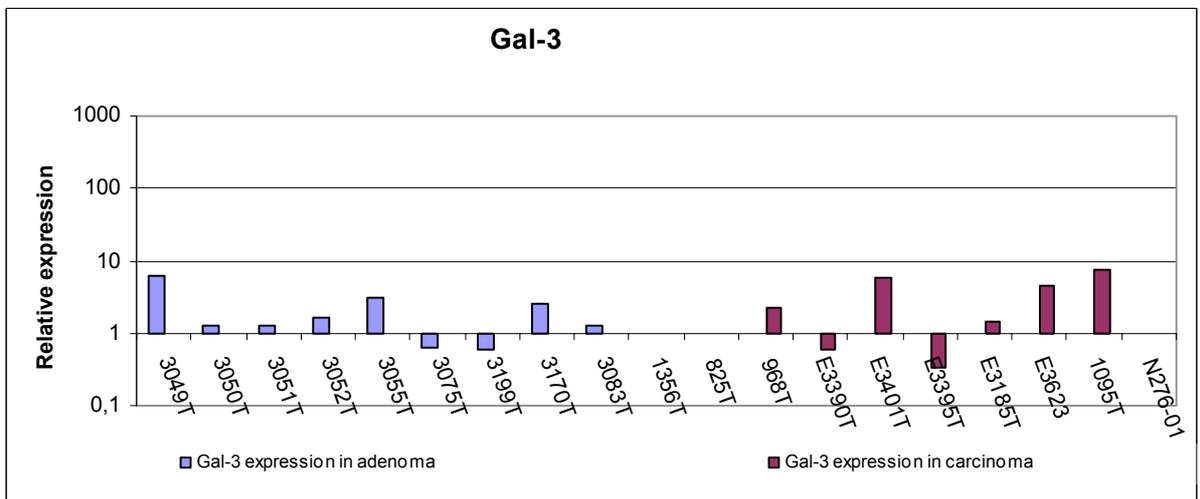


Fig.2 (c) shows relative expression of gal-3 in follicular adenomas (blue) and carcinomas (red).

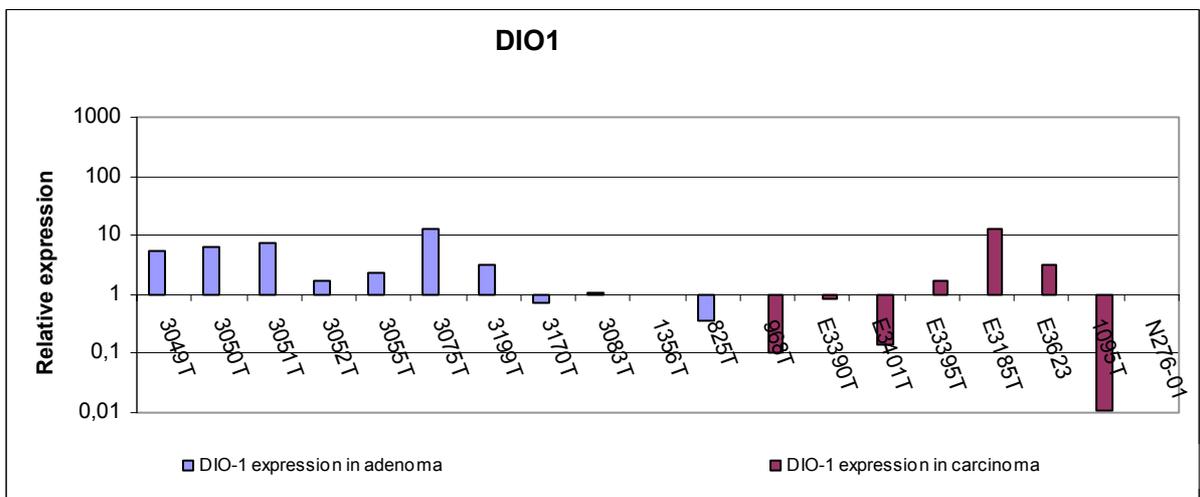


Fig.2 (d) shows relative expression of DIO1 in follicular adenomas and carcinomas.

Two of these genes, CITED1 and CLDN-1, had been previously selected when studying only a few of the tumours. They proved to still be interesting after extending the number of tumours in the study, and the RT-PCR results are also in agreement with array results for a convincing majority of the samples. Viewed on the whole the genes are over-expressed in carcinomas compared to adenomas.

CLDN-1 belongs to a family of genes that has been associated with various cancers [23, 24]. It is a tight-junction protein controlling cell-to-cell adhesion. Tight-junction proteins are specific for epithelial cells and seal neighbouring cells together to prevent leakage.

CITED1 is a CBP/p300-binding protein that does not bind directly to DNA, but is a transcriptional co-activator, also known as melanocyte-specific gene-1(MSG-1) [25]. It has previously been reported to be associated with other cancers [26] and even shown to be over-expressed in papillary thyroid carcinomas [27].

Gal-3 has been associated with neoplastic processes in various tissues [2, 28, 29], suggesting it could be a potential marker of malignancy in thyroid neoplasms, but the confirmation of these results in the case of follicular carcinomas and adenomas has not been seen. In fact, the accuracy has been reported to be fairly low, since benign lesions also expressed gal-3 [30] in a study aimed to evaluate gal-3 expression as a molecular marker. Our results show that the two classes of follicular adenomas and carcinomas can not be separated using this marker.

The fourth potential candidate, DIO1, is a gene with a protein product that belongs to the iodothyronine deiodinase family. It activates thyroid hormone by converting the prohormone thyroxine (T_4) by outer ring deiodination to bioactive triiodothyronine (T_3), and also degrades both hormones by inner ring deiodination. The RT-PCR results reproduced our earlier array results, where the gene is under-expressed in carcinomas, verifying that the gene is an interesting candidate with a biological function specifically connected to thyroid tissue.

To sum up, our real-time PCR results verified the majority of our array results. However, a difference worth mentioning could be the commonly seen larger fold changes obtained when using real-time PCR. This explains why the actual fold change in the expression profiles can differ from array results even though the profiles on the whole are consistent.

The statistical methods for further analysing the dataset fall into two classes, unsupervised and supervised. For general information on these statistical methods see Theoretical appendix. Unsupervised methods do not use any external information to organise the data. Supervised methods benefit from using some external information, usually the disease status of the samples, and it also involves the division of the entire data set into a training set and a test set. Classifiers are constructed by assigning predefined classes to expression profiles during training. Evaluation with the test set is performed to estimate the performance of the classifier, before it is applied to data with unknown classification.

Feature selection is a fundamental issue in gene expression-based tumour classification. The models to be built will use some features of the examples. One problem with gene expression data is that each example has too many features, and many of them are noisy and irrelevant for the learning problem. This is a common problem in pattern recognition and a number of approaches have been proposed to select a subset of the features to be used. One approach is to consider each feature individually, but methods considering pairs of genes that distinguish

well between classes have also been proposed [21]. Feature selection serves the purposes of reducing dimensionality, and thereby improves classification accuracy, and to identify genes that are relevant to the cause of disease or can be used as biomarkers for diagnosis of disease.

Two kinds of feature selection methods, filter (non-wrapper) methods and wrapper methods, have generally been studied [31, 32]. The essential difference between these methods is that a wrapper method makes use of the algorithm that will be used to build the final classifier, while a filter method does not. Because a learning algorithm is employed to evaluate the set of features considered, wrappers are time-expensive to run.

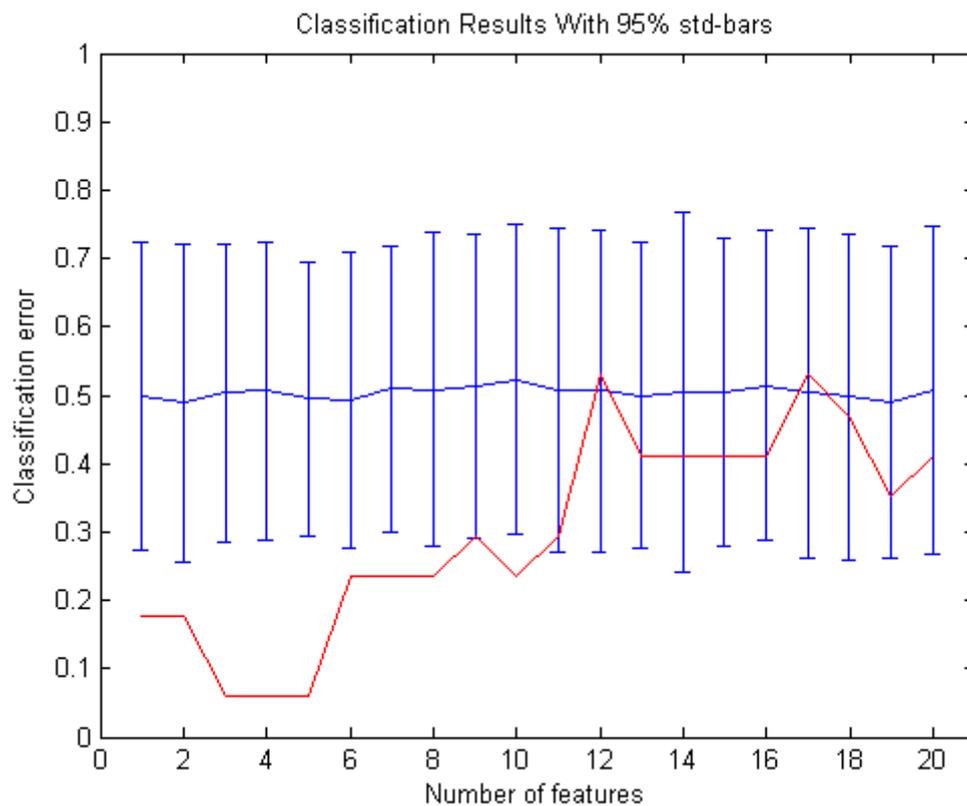


Fig.3 shows the performance obtained for leave-one-out validation (red) when using a number of features, up to 20, for the dataset of 2406 genes. A wrapper approach was used for feature selection, including Greedy pairs and Fisher's linear discriminant function. The result of a permutation test (blue) is also presented with 95%CI, showing that the results with 3, 4 and 5 genes could not have been obtained by chance. For details on methods and algorithms used see Material and methods or Theoretical appendix.

Supervised learning was performed with leave-one-out validation, using a wrapper approach for feature selection, including Greedy pairs and Fisher's linear discriminant function. The performance for an increasing number of genes used can be seen in Fig. 3, and with three, four, and five genes the error rate is as low as 5.8%, i.e. only one sample is misclassified. The result of a permutation test is also visualised, showing that with a 95% confidence interval (CI) there is no possibility of obtaining the classification results with three genes by chance.

6 Discussion

Before interpreting the results of the study, it is necessary to examine the reliability, i.e. the sources of error present at different stages. One of the problems encountered early in the analytical procedure is achieving good quality raw data. The results of any further analysis with statistical methods will depend on the input quality, emphasising the need to put a lot of effort into obtaining good quality data.

First, the spot morphology may cause difficulties when the shape and size of the spots vary, mainly due to printing. This leads to a situation where the software cannot automatically analyse all spots on the array. Manual interpretation and adjustment is required. This procedure is very time consuming and of course subjective, and thereby hard to reproduce. The result is a trade-off between accuracy and time-consumption.

Second, cases of poor quality raw data may due to unexpressed genes in the reference used, result in a loss of raw data for certain spots. The result is either an unreliable ratio or no ratio at all from the spot. This can of course also be due to the lack of cDNA in the spot, or not being able to use the value because of merged spots. Unfortunately these cases of lacking information are indistinguishable, making interpretation of results harder. As mentioned earlier, a suggestion to be kept in mind for future studies is including tumour samples in the reference pool, in order to achieve reliable signal in the reference for all genes expressed in the tumour samples.

The effort of obtaining good quality raw data is followed by using the knowledge about the sources of errors when deciding on filtering criteria. The aim of the process of low-level analysis is refining the quality of data through balancing between not losing the valuable information present and at the same time eliminating bad spots and unreliable data from the data set.

Since no standard filter procedures for microarray experiments are available, different research groups use their own judgements. This contributes to the difficulties in comparing and reproducing experiments, within a research group as well as between research groups. Since the field of microarray technology is still very young, further research will be required before standards or recommendations for filtering criteria can be suggested.

The SAM approach, which uses knowledge about class labels, was used for further analysis. It uses criteria as variation and consistency of expression as well as mean fold change within each class in order to rank differentially expressed genes. Inconsistency within each class in an expression profile, i.e. large variation, will push genes further down in the list even though they may have large fold changes. The top genes may on the other hand carry a consistent expression profile but low fold changes, which is not the best choice for genes to be used as diagnostic markers and thereby a reason for not only using the ranking as a cut-off for choosing genes.

A potential candidate gene should preferably be found close to the top of the list and at the same time have an expression profile with large fold changes, indicating the gene is associated with malignancy. The four genes chosen for further analyses were chosen based on a combination of gene rank and fold change. To determine whether the genes were truly

potential candidate genes, and also to investigate if and to which extent array results were reliable and could be readily reproduced, real-time PCR was performed. Our results were confirmed with real-time PCR, indicating that the genes may prove to be useful as candidate genes in the future.

Both the genes CLDN-1 and CITED-1 were up-regulated in the carcinomas. This suggests the genes to be associated with tumour progression. These genes have previously been reported to play a role in other cancers, supporting this hypothesis. All CITED proteins strongly activate transcription, maybe through interacting with DNA-binding proteins and function as transcriptional coactivators [26]. CITED1 is suggested to enhance general transcription mediated by the SMAD transcription factors in a manner dependent on CBP/p300 [1]. Increased expression of CLDN-1 in colorectal cancers has been reported by Miwa *et al.* [23]. Their results imply involvement in transcriptional activation in the beta-catenin-Tcf/LEF signalling pathway. This proposes a role in tumourigenesis for the gene, since accumulation of beta-catenin is frequently observed in tumours arising.

The gene DIO1 is involved in the normal function of the thyroid. The observed down-regulation for the carcinomas indicates some kind of general disturbance present in this tissue.

Using immunohistochemical techniques, galectin-3 has been used to distinguish between follicular and papillary adenomas and carcinomas [33]. However, the interpretation of immunohistochemical staining is very subjective. Thereby it is open to possible interpretation errors, and can be technically difficult with cytological samples. Moreover, it is hypothesised that quantitative measurement with RT-PCR could yield a more objective way to distinguish benign from malignant nodules [30]. Gal-3 expression was at the RNA level detected not only in malignant tissue, but in all tissue specimens, not supporting the use of this marker for distinguishing follicular adenoma from follicular carcinoma with RT-PCR [30]. After investigating the potential of the marker on RNA level with RT-PCR, we can only agree with the results reported by Martins [30], which showed that the gal-3 expression, as a molecular marker in diagnosis of follicular adenoma or carcinoma, was questionable, since similar expression levels were shown in both adenomas and carcinomas.

Most scoring methods, for example SAM, do not use classification accuracy to measure a gene's ability to discriminate between tissue samples. Therefore, genes that are ranked according to these scores may not achieve the highest classification accuracy among genes in the experiments. Instead the optimal combination of genes should preferably be calculated by a feature selection procedure where the classification accuracy is used as a criterion function [32], as in our supervised learning wrapper approach.

Our choice of a linear classifier tries to avoid the problem with over-fitting the model to the data set, which is a greater risk with small datasets as in our case. Comparing linear classifiers to non-linear, they are faster than more complex algorithms. Moreover, there is a trade-off between flexibility in design and time consumption for performance, i.e. the number of parameters that can be influenced increases with complexity. An unwanted effect of flexibility in classifiers is the increased risk of over-fitting the model, making them better suited for larger data sets.

Considering our results with the minimum of the error rate of classification found using only three genes, the performance is very good. Permuting the class labels without obtaining anywhere near as good results further suggests really promising genes, but the top three genes

are not easily distinguished. Each round of leave-one-out validation returns a unique set of features differing slightly from each other since they are based on different training sets. Usually a consensus gene list is made, but that does not fully represent the truth.

The next step towards a clinical use is translating these results into a practical test in order to verify and investigate their potential as diagnostic markers. While the level of accuracy is not sufficient to replace histological examination, these molecular markers may be useful as a complement to morphology-based diagnosis. Moreover, the low fold change level demands careful interpretation, and may be relatively difficult to verify with real-time PCR or another clinically approachable method. It is still unclear whether our candidates will make it all the way to a future diagnostic test. Nevertheless, the study provides evidence that the clinical behaviour of follicular thyroid cancer can be anticipated by the analysis of the gene expression profiles.

A future scenario could be initial basic research monitoring thousands of genes in hundreds of samples using microarrays as a screening tool in the research laboratory, to identify genes providing optimal classification accuracy. Clinical applications will then monitor only this small subset of genes, avoiding the cost and complexity involved in performing large-scale microarray experiments.

Acknowledgements

First of all, I would like to thank everyone at the WCN Microarray Platform for all the help and support in the lab and outside, as well as providing a positive and stimulating work environment, making me feel very welcome as a member of the group. I would like to thank my supervisor Anders Isaksson for support and encouragement and Mårten Fryknäs for always taking the time to answer my questions and helping me solve encountered problems. Also, thanks to Ulrika Wickenberg for help and guidance concerning the data analysis. Finally, I would like to thank my examiner Prof. Ulf Pettersson.

References

1. M Schlumberger, F.P., *Thyroid tumors*. 1999, Paris: Éditions Nucléon.
2. Ringel, M.D., *Molecular diagnostic tests in the diagnosis and management of thyroid carcinoma*. Rev Endocr Metab Disord, 2000. **1**(3): p. 173-81.
3. Cvejic, D., et al., *Immunohistochemical localization of galectin-3 in malignant and benign human thyroid tissue*. Anticancer Res, 1998. **18**(4A): p. 2637-41.
4. Wynford-Thomas, D., B.M. Stringer, and E.D. Williams, *Dissociation of growth and function in the rat thyroid during prolonged goitrogen administration*. Acta Endocrinol (Copenh), 1982. **101**(2): p. 210-6.
5. Bartolazzi, A., et al., *Application of an immunodiagnostic method for improving preoperative diagnosis of nodular thyroid lesions*. Lancet, 2001. **357**(9269): p. 1644-50.
6. Shalon, D., S.J. Smith, and P.O. Brown, *A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization*. Genome Res, 1996. **6**(7): p. 639-45.
7. Solinas-Toldo, S., et al., *Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances*. Genes Chromosomes Cancer, 1997. **20**(4): p. 399-407.
8. Gingeras, T.R., et al., *Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic Mycobacterium DNA arrays*. Genome Res, 1998. **8**(5): p. 435-48.
9. MacBeath, G. and S.L. Schreiber, *Printing proteins as microarrays for high-throughput function determination*. Science, 2000. **289**(5485): p. 1760-3.
10. Kim, J.H., et al., *Osteopontin as a potential diagnostic biomarker for ovarian cancer*. Jama, 2002. **287**(13): p. 1671-9.
11. Heid, C.A., et al., *Real time quantitative PCR*. Genome Res, 1996. **6**(10): p. 986-94.
12. Chee, M., et al., *Accessing genetic information with high-density DNA arrays*. Science, 1996. **274**(5287): p. 610-4.
13. Yang, I.V., et al., *Within the fold: assessing differential expression measures and reproducibility in microarray assays*. Genome Biol, 2002. **3**(11): p. research0062.
14. Brazma, A., et al., *Minimum information about a microarray experiment (MIAME)-toward standards for microarray data*. Nat Genet, 2001. **29**(4): p. 365-71.
15. Geschwind, D. H., *Indirect Labeling and Detection of cDNA using tyramide signal amplification, Basic Protocol*.
<http://geschwindlab.medsch.ucla.edu/protocol/tyramide.htm>. (15 Jun. 2002)
16. Yang, Y.H., et al., *Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation*. Nucleic Acids Res, 2002. **30**(4): p. e15.
17. Dudoit, S., *Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments*, in *Technical report #578*. 2000.
18. Tran, P.H., et al., *Microarray optimizations: increasing spot accuracy and automated identification of true microarray signals*. Nucleic Acids Res, 2002. **30**(12): p. e54.
19. Tusher, V.G., R. Tibshirani, and G. Chu, *Significance analysis of microarrays applied to the ionizing radiation response*. Proc Natl Acad Sci U S A, 2001. **98**(9): p. 5116-21.
20. Xiong, M., et al., *Feature (gene) selection in gene expression-based tumor classification*. Mol Genet Metab, 2001. **73**(3): p. 239-47.
21. Bo, T. and I. Jonassen, *New feature subset selection procedures for classification of expression profiles*. Genome Biol, 2002. **3**(4): p. RESEARCH0017.

22. Golub, T.R., et al., *Molecular classification of cancer: class discovery and class prediction by gene expression monitoring*. Science, 1999. **286**(5439): p. 531-7.
23. Miwa, N., et al., *Involvement of claudin-1 in the beta-catenin/Tcf signaling pathway and its frequent upregulation in human colorectal cancers*. Oncol Res, 2000. **12**(11-12): p. 469-76.
24. Terris, B., et al., *Characterization of gene expression profiles in intraductal papillary-mucinous tumors of the pancreas*. Am J Pathol, 2002. **160**(5): p. 1745-54.
25. Shioda, T., et al., *Transcriptional activating activity of Smad4: roles of SMAD hetero-oligomerization and enhancement by an associating transactivator*. Proc Natl Acad Sci U S A, 1998. **95**(17): p. 9785-90.
26. Yahata, T., et al., *Selective coactivation of estrogen-dependent transcription by CITED1 CBP/p300-binding protein*. Genes Dev, 2001. **15**(19): p. 2598-612.
27. Huang, Y., et al., *Gene expression in papillary thyroid carcinoma reveals highly consistent profiles*. Proc Natl Acad Sci U S A, 2001. **98**(26): p. 15044-9.
28. Aratake, Y., et al., *Diagnostic utility of galectin-3 and CD26/DPPIV as preoperative diagnostic markers for thyroid nodules*. Diagn Cytopathol, 2002. **26**(6): p. 366-72.
29. Bernet, V.J., et al., *Determination of galectin-3 messenger ribonucleic Acid overexpression in papillary thyroid cancer by quantitative reverse transcription-polymerase chain reaction*. J Clin Endocrinol Metab, 2002. **87**(10): p. 4792-6.
30. Martins, L., et al., *Galectin-3 messenger ribonucleic acid and protein are expressed in benign thyroid tumors*. J Clin Endocrinol Metab, 2002. **87**(10): p. 4806-10.
31. Xing, E.P. and R.M. Karp, *CLIFF: clustering of high-dimensional microarray data via iterative feature filtering using normalized cuts*. Bioinformatics, 2001. **17 Suppl 1**: p. S306-15.
32. Xiong, M., X. Fang, and J. Zhao, *Biomarker identification by feature wrappers*. Genome Res, 2001. **11**(11): p. 1878-87.
33. Orlandi, F., et al., *Galectin-3 is a presurgical marker of human thyroid carcinoma*. Cancer Res, 1998. **58**(14): p. 3015-20.
34. Quackenbush, J., *Microarray data normalization and transformation*. Nat Genet, 2002. **32 Suppl**: p. 496-501.
35. Brody, J.P., et al., *Significance and statistical errors in the analysis of DNA microarray data*. Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12975-8.
36. Raffelsberger, W., et al., *Quality indicators increase the reliability of microarray data*. Genomics, 2002. **80**(4): p. 385-94.
37. Tu, Y., G. Stolovitzky, and U. Klein, *Quantitative noise analysis for gene expression microarray experiments*. Proc Natl Acad Sci U S A, 2002. **99**(22): p. 14031-6.
38. Eisen, M.B., et al., *Cluster analysis and display of genome-wide expression patterns*. Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14863-8.
39. Tamayo, P., et al., *Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation*. Proc Natl Acad Sci U S A, 1999. **96**(6): p. 2907-12.
40. Varela, J.C., et al., *Microarray analysis of gene expression patterns during healing of rat corneas after excimer laser photorefractive keratectomy*. Invest Ophthalmol Vis Sci, 2002. **43**(6): p. 1772-82.
41. Butte, A., *The use and analysis of microarray data*. Nat Rev Drug Discov, 2002. **1**(12): p. 951-60.
42. Furey, T.S., et al., *Support vector machine classification and validation of cancer tissue samples using microarray expression data*. Bioinformatics, 2000. **16**(10): p. 906-14.

43. Radmacher, M.D., L.M. McShane, and R. Simon, *A paradigm for class prediction using gene expression profiles*. J Comput Biol, 2002. **9**(3): p. 505-11.
44. Rajeevan, M.S., et al., *Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR*. J Mol Diagn, 2001. **3**(1): p. 26-31.
45. Rajeevan, M.S., et al., *Use of real-time quantitative PCR to validate the results of cDNA array and differential display PCR technologies*. Methods, 2001. **25**(4): p. 443-51.

Theoretical appendix

Normal function of the thyroid

The production of the thyroid hormones T_3 and T_4 starts with iodine and tyrosine uptake from the blood stream. Tyrosine is then built into thyroglobulin, and after enzymatic coupling of iodine the final hormones are stored in the cavity of the follicles until secretion, which takes place by the releasing of T_4 and T_3 from thyroglobulin and through diffusion through the membrane into the blood stream. The hormones are distinguished by the fact that T_4 contains four iodine atoms and T_3 only three. Normally the follicle cells produce 20-30 times more T_4 than T_3 .

The thyroid is considered “the master gland of metabolism” and the effects of thyroid hormones in target cells are carried out through binding to receptors in the nucleus, which leads to a series of events resulting in a change of transcriptional activity, and thereby a change in metabolism.

The production of thyroid hormones is regulated by TSH (thyroid stimulating hormone), a hormone from the pituitary gland, which in turn is regulated by TRH (TSH-regulating hormone), produced by the hypothalamus. TSH secretion is also influenced by T_3 and T_4 concentrations in the blood through negative feedback. While thyroxine is produced directly in the follicle cells, the main part of triiodothyronine exists as a prohormone, turning into the final hormone through the release of one iodine atom from thyroxine. Triiodothyronine has a biological effect about five times higher than thyroxine and mainly contributes to the effects of thyroid hormones. Several effects of T_3 and T_4 can be seen:

- they stimulate heat production
- they are necessary for normal growth
- they are necessary for normal development of the central nervous system
- they activate of the sympathetic nervous system
- they help oxygen get into cells
- lack of thyroid hormones leads to slower reflexes and reduced physical performance

Calcitonin has several effects, e.g. inhibition of the bone marrow suppression, increases the uptake of calcium in bone marrow and inhibits the reabsorption of calcium in the kidneys, all leading to a decrease in the calcium concentration in the blood.

Data analysis

Low-level analysis

A normalisation process is performed to adjust the individual hybridisation intensities to balance them appropriately so that meaningful comparisons can be made. Reasons for normalisation include unequal amounts of RNA used, differences in efficiency of labelling and detection between the dyes used, all in order to remove systematic errors and allow comparison within and between slides. The choice of a robust and adequate method is crucial for the quality of the data obtained. Non-linear algorithms such as the curve-smoother locally weighted linear regression (lowess) are nowadays preferred over global normalisation methods, since they can remove intensity-dependent dye-specific effects[34]. Spatial effects on fluorescence intensities can also be corrected for by normalisation, since most normalisation methods can be applied either globally (to the entire dataset) or locally (to physical subsets of the dataset). Local normalisation has the advantage that it can correct for systematic spatial variation in the array, inconsistencies among the spotting pins used, variability in slide surface, and slight local conditions in hybridisation conditions. A balanced distribution of expression ratios independent of intensity is produced, which can be seen when comparing the R-I (ratio-intensity) plots of Fig. 1 (a) and (b)[34], describing the distribution of data before and after normalisation.

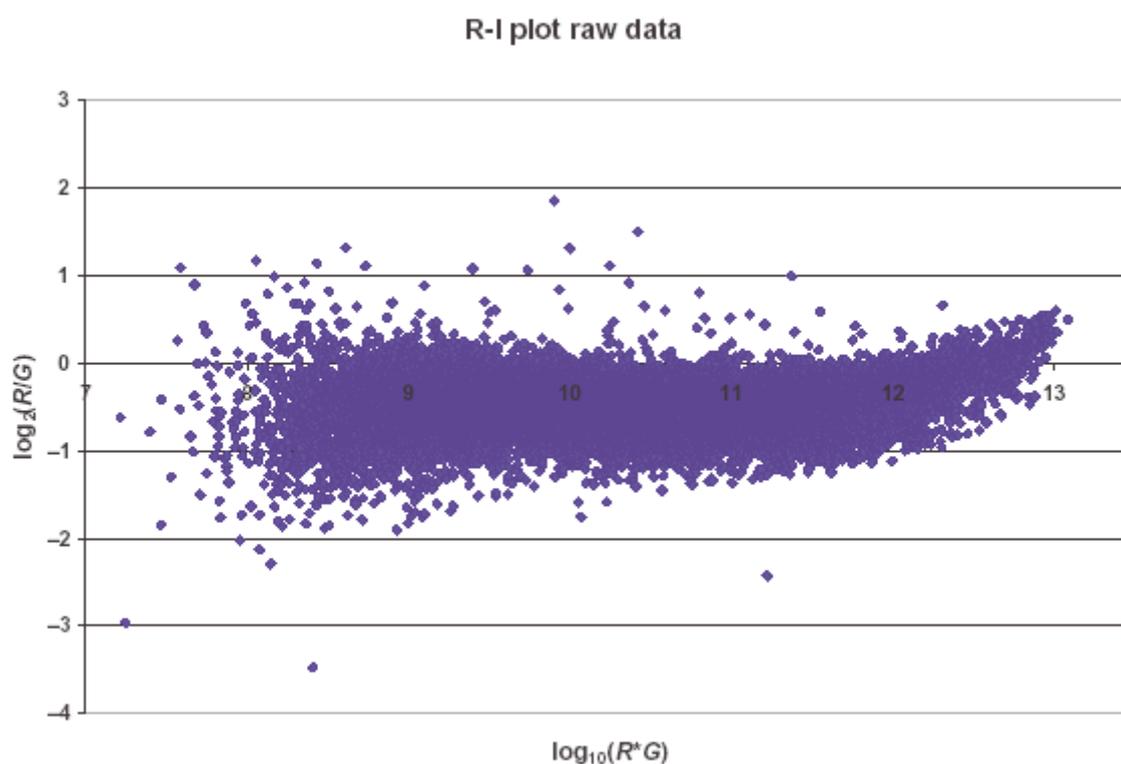


Fig. 1 (a) The R-I (ratio-intensity) plot reveals intensity-specific artefacts in the \log_2 (ratio) measurements.

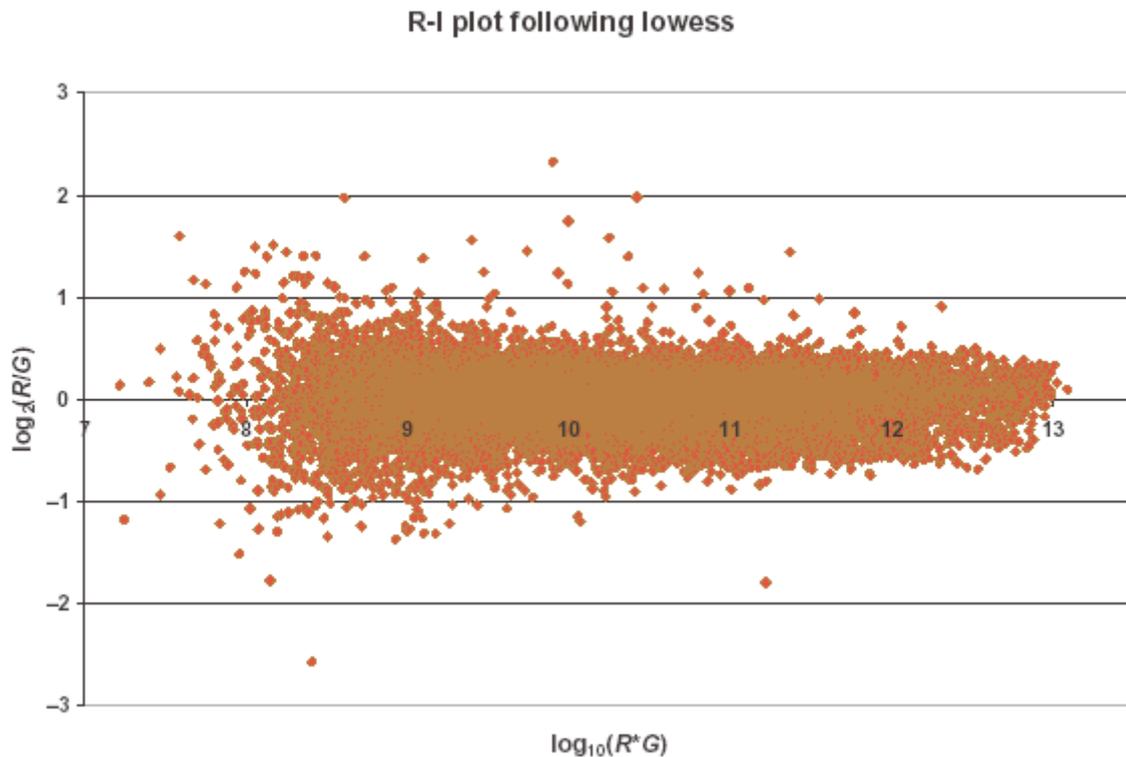


Fig. 1 (b) shows that application of print-tip lowess normalisation can correct for both systematic variation as a function of intensity and spatial variation between spotting pins.

Many of the processes involved in microarray experiments are highly non-linear, resulting in quite noisy measurements. Consequently, data from microarrays are usually pre-processed in order to detect and remove noise and unreliable measurements. An aspect common to all array techniques is the extent of reliability and variance in measurements. As with sequencing, best measures of reliability can be made only when large data sets containing repetitions and overlapping data are available. Yang *et al.*[13] propose a procedure for eliminating low-quality data for replicates within and between slides, underlining the importance of replicates in generating high-quality expression data. Several issues remain to be resolved in respect to filtering of microarray data and a central question is how to determine the best discriminating “quality-criterion” to use. A number of quality indicators and tests have been proposed[35-37]. With any filtering method there is a trade-off between optimising measurement reliability and avoiding loss of information.

High-level analysis

The statistical methods for furthering analysing the dataset fall into two classes, unsupervised and supervised.

Unsupervised methods do not use any external information to organise the data. Hierarchical clustering is one of the most widely used clustering techniques, where the main goal is to produce a tree in which the nodes represent subsets of a dataset[38]. Self-organising maps (SOM)[39] were among the first machine learning methods to be applied to classification of cancers. It is a clustering approach based on hypothetical neural structures called feature

maps, which are adapted by the effect of the input samples. Golub *et al.*[22] reported a prediction using such a model discovering the distinction between *acute myeloid leukaemia* (AML) and *acute lymphoblastic leukaemia* (ALL). Unsupervised classification applications have also used the k-means clustering algorithm, which as well as SOM require knowledge on the number of clusters representing the data, since samples are categorised into a fixed number of k clusters[40].

Supervised methods benefit from using some external information, usually the disease status of the samples, and it also involves the division of the entire data set into training and a test set. Classifiers are constructed by assigning predefined classes to expression profiles during training. Testing with the test set is performed to estimate the performance of the classifier, before it is applied to data with unknown classification. Supervised methods include k-nearest neighbour classification, support vector machines (SVM), and neural networks[41]. The support vector machine is a powerful supervised learning algorithm that has been applied to classify complex expression patterns. It can be seen as an optimisation problem, which finds mathematical structures to differentiate positive and negative training samples. Furey *et al.*[42] have implemented a system for the classification of normal and ovarian cancer tissues based on a SVM model.

A general framework for prediction based on gene expression profiles is proposed by Radmacher *et al.*[43] consisting of initially evaluating if it is appropriate to use class prediction for the given data set, selecting prediction method, performing cross-validation of the class prediction and finally assessing the significance of the results by permutation testing. A permutation method is used to calculate the probability of producing a cross-validated error rate as small as observed given no association between class membership and expression profiles, mimicked by randomly permuting the labels among the gene expression profiles.

Real-time PCR

cDNA microarrays is one of the leading technologies to identify differentially expressed genes and genome-wide transcription patterns. The reliability of these techniques to detect transcriptional differences representative of samples is affected by several factors. Microarray results are influenced by array production, RNA extraction, probe labelling, hybridisation conditions, and image analysis, so genes identified as differentially expressed must be validated with another method.

Real-time RT-PCR methods that monitor product accumulation are the current state-of-the-art for gene quantification[11], even though the level of expression differences measured by real-time PCR for many genes has been shown to differ when compared to results from the arrays[44, 45]. This disagreement may be due to the lack of specificity and thereby ability for DNA fragments to discriminate between splice variants on arrays. Until arrays are optimised for measurement of sequence-specific transcripts, it is possible for cDNA array results to be masked by cross-hybridisation. The strength of real-time PCR assays as a secondary validation procedure lies in its potential to quantify relative changes in expression of large numbers of known and novel genes with limited RNA in a rapid and precise manner.

The high through-put Taqman PCR (Applied Biosystems) method use the 5'-nuclease activity of Taq DNA polymerase in conjunction with DNA probes labelled with quencher and reporter dyes. A positive PCR reaction results in the removal of the reporter dye from the influence of the quencher dye, leading to an increase in measurable fluorescence, proportional to the amount of amplicon produced. The real-time reaction information allows quantification of target nucleic acid.

Appendix I

Total RNA isolation

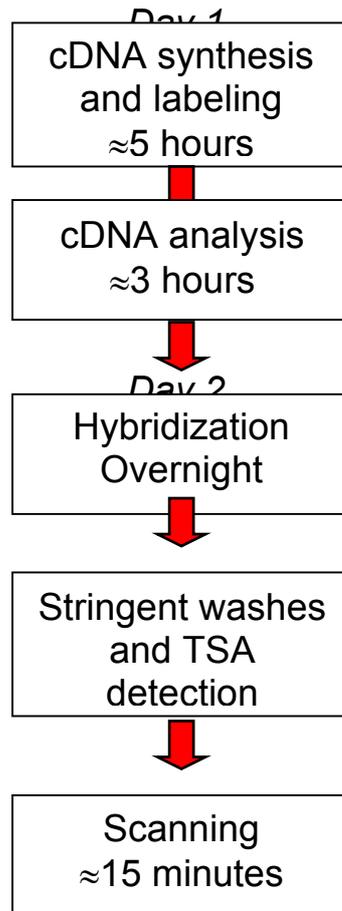
1. Tumors are surgically removed from patients and rapidly transported to pathology where a pathologist removes a piece(s) for microarray analysis. These tumor samples are rapidly frozen in Liquid Nitrogen and stored at -80°C until use.
2. The frozen tumor specimen is removed from the freezer and, using a scalpel, small pieces (approximately 50-100 milligrams each) are cut off of the larger tumor specimen while it is still frozen. As the small pieces are cut off, they are immediately placed into 10-12mls of **TRIZOL Reagent (GibcoBRL Life Technologies)** (http://www2.lifetech.com:80/catalog/techline/molecular_biology/Manuals_PPS/15596.pdf) that is contained within a 50ml Screw Cap Tube at RT. For a given tumor sample, up to 1 gram can be placed into this volume of TRIZOL, and typically 0.5-1 gram of tissue is used per isolation if possible. Note, that as little as 0.2 gram has been successfully used in this volume but it may be better to use 5-8mls TRIZOL for specimens of this size (the TRIZOL Reagent Protocol calls for 1ml TRIZOL per 50-100mg tissue processed).
3. The tumor sample in TRIZOL is homogenized using a PowerGen 125 Tissue Homogenizer (Fisher Scientific) starting at 5000 RPM and gradually going up to approximately 20,000 RPM over a period of 30-60 seconds at RT. The homogenization is performed until a homogeneous solution is obtained and very few visible tumor pieces can be seen.
4. Incubate at RT for 5-10 minutes after homogenization.
5. The TRIZOL/tumor homogenate is transferred to a 50ml Oak Ridge Centrifuge tube and centrifuged at 12,000g (10,000RPM on SS34) for 5-10 minutes at 4°C .
6. Remove the upper fat layer by using a Pasteur pipette hooked up to a vacuum flask (the upper fat layer, if present, has a yellowish appearance while the TRIZOL homogenate remains red).
7. NOTE: From here to Step 12, the protocol is exactly as written in the TRIZOL Reagent Protocol (see above for URL).
8. Add 0.2ml Chloroform per 1ml TRIZOL Reagent used in Step 2. Shake tube vigorously for 15-30 seconds by hand and incubate at RT for 5 minutes.
9. Spin at 12,000g for 15 minutes at 4°C . Carefully remove the upper aqueous phase, which contains the total RNA, and place this in a new 50ml Oak Ridge Centrifuge tube.
10. Precipitate the RNA by adding 0.5ml Isopropyl alcohol per 1ml TRIZOL Reagent used in Step 2. Incubate at RT for 10 minutes, then spin at 12,000g for 10 minutes at 4°C .
11. Carefully decant the supernatant and wash the pellet once using 75% ETOH by adding 1ml 75% ETOH per 1ml TRIZOL used in Step 2. Spin at 7,500g (7500 RPM in SS34) for 5 minutes at 4°C .
12. Carefully decant the supernatant and air dry the pellet for 10-20 minutes at RT.
13. Resuspend the RNA pellet in 200-300 μl DEPC-DH₂O or DEPC-TE, and take A260 if necessary.

Appendix II

INDIRECT LABELING AND DETECTION OF cDNA USING TYRAMIDE SIGNAL AMPLIFICATION (TSA)

BASIC PROTOCOL

Here we describe the TSA protocol originally developed by MICROMAX (NEN, PerkinElmer) with minor modifications. The use of TSA detection enables the researcher to use as little as 500 ng of total RNA, which is 50 to 100 times less than the amount of RNA used in direct cDNA labeling. TSA appears to provide a consistent and reproducible signal amplification method for cDNA microarray analysis, both across arrays and the genes tested (Karsten et al, 2001). Briefly, the experimental samples (total RNA) are converted into fluorescein (F1) and biotin (B) labeled cDNA during reverse transcription. Purified cDNA probes are mixed and simultaneously hybridized to the microarray. F1 and B labeled cDNAs are sequentially detected with conjugate reporter molecules after hybridization and stringent washes. The microarray is first incubated with Anti-F1-horseradish peroxidase (HRP) – antibody-enzyme conjugate, which specifically binds to the hybridized F1 labeled cDNA probe. The enzyme portion of the conjugate is horseradish peroxidase, which catalyzes the deposition of Cy3 labeled tyramide amplification reagent. The reaction results in the immediate deposition of numerous Cy3 labels adjacent to immobilized HRP. This step is an actual amplification of the signal – due to its enzymatic nature the amount of tyramide (Cy3) is greatly amplified relative to cDNA hapten (fluorescein). Prior to the second TSA step, the residual HRP is inactivated. Streptavidin-HRP binds to Biotin labeled cDNA and catalyzes the deposition of Cy5 labeled tyramide amplification reagent. This process is much more involved and tedious than direct methods, but allows the use of 50 to 100 times less starting RNA. The protocol provided below is minimally modified from manufacturer's instructions.



Materials

- Commercial or custom made microarray
- Sample total or polyA+ RNA
- MICROMAX TSA Labeling and Detection Kit (MPS521, MPS103, MPS522, NEN™ Life Science Products)
- 1 N Sodium Hydroxide (NaOH) (freshly prepared)
- 1M Tris-HCl, pH 7.5 (Life Technologies, Inc.)
- 70% ethanol diluted with DEPC-treated water, ice cold
- 100% Isopropanol
- 5 M Ammonium Acetate
- 0.5 M EDTA
- Dimethyl Sulfoxide (DMSO) (Sigma Chemicals)
- TN buffer
 - 0.1 M Tris-HCl, pH 7.5
 - 0.15 M NaCl
- TNT buffer
 - 0.1 M Tris-HCl, pH 7.5
 - 0.15 M NaCl
 - 0.05% Tween-20
- TNB buffer
 - 0.1 M Tris-HCl, pH 7.5
 - 0.15 M NaCl
 - 0.5% Blocking reagent (kit)

NOTE: *Blocking reagent should be added in small portions to TN buffer with continuous stirring. The solution should be heated gradually to 65°C with continuous stirring until the blocking reagent is completely dissolved. This may take several hours. The solution should be slightly opaque, without visible particles. Aliquot and store at -20°C.*

TNB-G buffer

0.1 M Tris-HCl, pH 7.5

0.15 M NaCl

0.5% Blocking reagent (kit)

10% Goat Serum (Life Technologies)

UV cross-linking apparatus (e.g. GS GeneLinker™ UV Chamber, BIO-RAD)

Day 1.

I. Perform cDNA synthesis and labeling reaction

1. From 0.5 to 10 µg of total RNA in a final volume of 20 µl mix with:

1 µl Reaction Mix concentrate (from the kit)

1 µl Biotin- or fluorescein-dCTP (from the kit)

NOTE: *The amount of each labeled probe needed for microarray hybridization (one slide) depends on the hybridization area of the slide. Typically, 1.5 µg of total RNA is sufficient for a 9000-element microarray covering the area of a 24 mm X 50 mm microscope cover glass slide.*

NOTE: *All operations involving fluorescein-dCTP, Cy3- and Cy5-tyramide, including post-hybridization washes and development should be conducted in a darkened room, as these reagents are light sensitive. Turning off the room lights is sufficient.*

2. Incubate the tubes for 10 min at 65°C to denature secondary RNA structures. Cool to room temperature for 5 minutes to anneal the primers to the RNA template.

3. Place the tube in a 42°C water bath and add 2.5 µl of reaction buffer and 2 µl of AMV reverse transcriptase/RNase inhibitor mix.

4. Incubate for 3 hours at 42°C.

NOTE: *Although NEN recommends incubation for 60 minutes, more consistent results can be obtained with incubation for at least 3 hours.*

II. Hydrolyze RNA template

5. Place the tube on ice for 5 minutes and add 2.5 µl of 0.5M EDTA (pH 8.0) and 2.5 µl 1N NaOH to stop the reaction and hydrolyze the RNA. Mix components well.

6. Incubate for 30 min at 65°C.

NOTE: *Do not exceed 30 minutes!*

7. Place the tube on ice for 5 minutes and neutralize the reaction by adding 6.5 µl of 1M Tris-HCl, pH 7.5

III. cDNA precipitation and purification

It is very important to purify labeled cDNA probes as unincorporated fluorescein-labeled and biotinylated nucleotides generate high background signals. Although column purification is often recommended, isopropanol precipitation can also efficiently remove unincorporated nucleotides to an acceptable level.

8. Add 2.7 µl of 5M Ammonium Acetate and 31 µl of isopropanol.

9. Vortex and incubate at 4°C for 1 hour.
10. Centrifuge 30 min at top speed (12,000 to 14,000 g), 4°C and carefully decant supernatant.
11. Wash twice with 70% ethanol, centrifuge, aspirate ethanol and quickly dry under vacuum.
NOTE: *Do not overdry the pellet.*
12. Resuspend in a suitable volume of hybridization buffer. For a large 24 mm X 50 mm microarray we resuspend each labeled pellet in 15 µl of Q buffer for a total hybridization volume of 30 µl.
13. Use 1µl (100 ng) for cDNA analysis. Store the remainder at -20°C in the dark for up to few days.

IV. Analyze cDNA labeling efficiency

NEN's TSA labeling and detection kit offers all the necessary reagents to check the quality of cDNA synthesis and efficiency of incorporation of labeled nucleotide prior to the hybridization step. The protocol uses a semi-quantitative, membrane-based colorimetric assay. The control of cDNA quality is a desirable step for every microarray protocol considering the costs of reagents and slides.

14. Make serial dilutions of biotin- and fluorescein control DNA:
 - 2 µl control cDNA (from the kit) + 18 µl 1X TE buffer
 - 5 µl of above + 5 µl 1 X TE
 - 5 µl of above + 5 µl 1 X TE
 - 5 µl of above + 5 µl 1 X TE
15. Make serial dilutions of the experimental cDNA:
 - 1 µl experimental cDNA + 9 µl 1 X TE
 - 5 µl of above + 5 µl 1 X TE
 - 5 µl of above + 5 µl 1 X TE
16. Apply 1 µl of both control and test dilutions in duplicate on a nylon membrane (supplied in the kit). Air dry.
17. Wet the membrane in 2 X SSC and UV cross-link in order to attach cDNA to the support covalently (1200 X 10² microjoules).
NOTE: *Handle membrane with plastic forceps and gloves.*
18. Block the membrane for 30 min at 3.5 ml TNB-G blocking buffer.
19. Remove blocking buffer.
20. Dilute, mix, and add the following reagents to the membrane.
 - 20 µl Streptavidin-HRP and 20 µl Anti-Fluorescein-HRP conjugate in 3.5 ml TNB-G blocking buffer.
 - Incubate for 30 min at room temperature.
21. Wash 4 times for 5 min each with TNT buffer and air-dry the filter.
22. Mix the following and add 4CN Plus working solution to the membrane.
 - 0.2 ml 4CN Plus reagent + 10 ml diluent working solution (1 ml BB diluent into 9 ml water)
23. Incubate in the dark for 30 minutes. You should see the distinct signals from all three serial dilutions. The 1 to 40 experimental sample dilutions should be of the same intensity as the 1 to 80 control dilutions. If all spots are visible proceed to the hybridization step.
NOTE: *If two or fewer experimental spots are visible, repeat the dot-blot analysis. If similar results are obtained perform another cDNA synthesis reaction.*

V. Set up hybridization

24. Add equal amounts of experimental cDNA from sample labeled with fluorescein and sample labeled with biotin, mix well and denature the secondary structures by heating for 2 min at 90°C.

25. Open the hybridization chamber, add 60 µl of 2 X SSC buffer into the chamber well and place the microarray slide in the chamber over the buffer well.

NOTE: *The purpose of the buffer is to insure chamber humidity during hybridization that the array does not dry out. Although the buffer comes into contact with the bottom of the slide, it should not interfere with the array hybridization surface. The amount of buffer may vary depending on the type of chamber used.*

26. Carefully apply the probe suspension in one drop onto the hybridization area of the slide. Insure that the solution covers the entire array surface (the region containing gridded cDNA clones). Try to avoid bubbles.

NOTE: *Be careful not to touch the array surface with a micropipet tip when applying hybridization mixture. To prevent bubbles make sure there is no debris or dust on the hybridization surface before applying cover slip or hybridization mixture. The slide can be dusted with compressed air mixture.*

27. Very carefully place the cover slip over the hybridization mixture using fingers or forceps. Start making contact between the edge of the cover slip and the hybridization solution applied then slowly place the cover slip down onto the hybridization area of the slide to avoid bubbles and provide an even distribution of hybridization mixture across the array.

NOTE: *There are several ways to apply the cover slip. Another common procedure is to gently overlay a cover slip onto the array hybridization area and apply the probe suspension (hybridization mixture) onto the edge of the cover slip. The mixture is drawn underneath the cover slip by capillary action. There is no preferred method, use the procedure that you feel most confident with. It is useful to practice this step several times with 30 µL of water and a blank microscope slide, before starting an experiment.*

28. Seal the chamber tightly and incubate it in a water bath at 65°C overnight, but not more than 16 hours. Make sure that array is positioned horizontally during transfer and incubation.

NOTE: *If using Boekel hybridization oven, carefully place the slide inside according to manufacturers instructions and seal the oven. Incubate at 65°C overnight.*

29. Incubate overnight at 65°C in a humid environment.

Day 2.

VI. Perform stringent washes and TSA detection

30. Vertically immerse the slide in 0.5 X SSC, 0.01% SDS letting the cover slip fall off. Place the slide in a new container with the same buffer and wash the microarray for 5 minutes with gentle agitation on an orbital shaker.

NOTE: *Perform all washing and detection steps quickly to minimize time between incubation and wash steps in order to prevent drying. Drying of the microarray results in a very high background signals. Perform all washing steps at room temperature in a covered container with gentle agitation.*

31. Wash in 0.06 X SSC, 0.01% SDS for 5 minutes.

32. Wash in 0.06 SSC for 2 minutes.

33. Incubate with 600 µl of TNB-10% Goat Serum buffer for 10 minutes.

NOTE: *Perform all incubation steps at room temperature in a covered container with wet blotting paper underneath the slide without agitation. It is important to keep the slide surface from drying out during incubation.*

34. Wash slide with TNT buffer for 1 minute and decant.

35. Incubate with 300 μ l of Anti-FI-HRP conjugate solution for 10 minutes:

4 μ l (Z) Anti-FI-HRP conjugate + 400 μ l TNB-10% Goat Serum buffer

NOTE: *Perform all incubation steps by applying carefully the specified solutions directly onto hybridization area of the slide. Incubate in a covered container out of direct light. Make sure that hybridization area is completely covered with a specified solution.*

NOTE: *Make sure after incubation or washes that the solution is removed by gentle tilting of the slide. Dry the sides of the slide with a dust-free KimWi paper by capillary action. Do not touch the hybridization area.*

36. Wash 3 X 1 minute in TNT buffer and decant.

NOTE: *Perform all washes with agitation at low speed.*

37. Incubate with 300 μ l of Cyanine-3 Tyramide solution for 10 minutes:

1 μ l (R) Cyanine-3 Tyramide (diluted in 20 μ l of DMSO) + 500 μ l (S) Amplification Diluent

38. Wash 3 X 5 minutes in TNT buffer and decant

39. Incubate with 300 μ l of HRP Inactivation Solution for 10 minutes:

10 μ l (U) 3 M Na Acetate + 290 μ l (T) HRP Inactivation Reagent

40. Wash 3 X 1 minute in TNT buffer and decant

41. Incubate with 300 μ l of Streptavidin-HRP solution for 10 minutes:

4 μ l (J) Streptavidin-HRP conjugate + 400 μ l TNB-10% Goat Serum buffer

42. Wash 3 X 1 minute in TNT buffer and decant

43. Incubate with 300 μ l of Cyanine-5 Tyramide solution for 10 minutes:

1 μ l (V) Cyanine-5 Tyramide (diluted in 20 μ l of DMSO) + 500 μ l (S) Amplification Diluent

43. Wash 3 X 5 minutes in TNT buffer and decant

44. Wash 1 minute in 0.06 X SSC and decant.

45. Spin the slide dry at room temperature, 500 g.

46. Scan the slide.

