A Meta-Analysis of Thyroid Cancer Gene Expression Profiling Studies Identifies Important Diagnostic Biomarkers

Obi L Griffith[1], Adrienne Melck[2], Sam M Wiseman[2,3], and Steven JM Jones[1]

1. Canada's Michael Smith Genome Sciences Centre, British Columbia Cancer Agency; 2. Department of Surgery, University of British Columbia; 3. Genetic Pathology Evaluation Center, Prostate Research Center of Vancouver General Hospital & British Columbia Cancer Agency

British Columbia Cancer Agency
Vancouver, British Columbia, Canada

Canada's Michael Smith Genome Sciences Centre
www.bcgsc.ca

1. Abstract

Introduction: An estimated 4-7% of the population will develop a clinically significant thyroid nodule during their lifetime. In as much as one third of these cases pre-operative diagnoses by needle biopsy are inconclusive. In many cases, a patient will undergo a diagnostic surgery for what ultimately proves to be a benign lesion. Thus, there is a clear need for improved diagnostic tests to distinguish malignant from benign thyroid tumours. The recent development of high throughput molecular analytic techniques should allow the rapid evaluation of new diagnostic markers. However, researchers are faced with an overwhelming number of potential markers from numerous expression profiling studies. To address this challenge, we have carried out a systematic and comprehensive meta-analysis of potential thyroid cancer biomarkers from 21 published studies.

Methods: For each of the 21 studies, the following information was recorded wherever possible: Unique identifier (probe/tag/accession); gene name/description; gene symbol; comparison conditions; sample numbers for each condition; fold change; direction of change; and Pubmed ID. Clone accessions, probe ids or SAGE tags were mapped to a common gene identifier (Entrez gene) using the DAVID annotation tool, Affymetrix annotation files, and the DiscoverySpace SAGE tag mapping tool respectively. A heuristic ranking system was devised that considered the number of comparisons in agreement, total number of samples, average fold change and direction of change. Significance was assessed by random permutation tests. An analysis using gene lists produced from re-analyzed raw image files (ensuring standard methods) for a subset of the studies was performed to assess our method.

Results: In all overlap analysis groups considered except for one, we identified genes that were reported in multiple studies at a significant level ($p<0.05$). Considering the 'cancer versus non-cancer' group as an example, a total of 755 genes were reported from 21 comparisons and of these, 107 genes were reported more than once with a consistent fold-change direction. This result was highly significant ($p<0.0001$). Comparison to a subset analysis of microarrays re-analyzed directly from raw image files found some differences but a highly significant concordance with our method (p-value = 6.47E-68).

Conclusions: A common criticism of molecular profiling studies is a lack of agreement between studies. However, looking at a larger number of published studies, we find that the same genes are repeatedly reported and with a consistent direction of change. These genes may represent real biologic participants that through repeated efforts have overcome the issues of noise and error typically associated with such expression experiments. In some cases these markers have already undergone extensive validation and become important thyroid cancer markers. But, other high-ranking genes have not been investigated at the protein level. A comparison of our meta-review method (using published gene lists) to a meta-analysis of a smaller subset of studies (for which raw data were available) showed a strong level of concordance. Thus, we believe our approach represents a useful alternative for identifying consistent gene expression markers when raw data is unavailable (as is generally the case). Furthermore, we believe that this meta-analysis, and the candidate genes we have identified, may facilitate the development of a clinically relevant diagnostic marker panel.

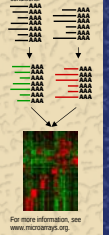
SAGE

Serial analysis of gene expression (SAGE) is a method of large-scale gene expression analysis that involves sequencing small segments of expressed transcripts ("SAGE tags") in such a way that the number of times a SAGE tag sequence is observed is directly proportional to the abundance of the transcript from which it is derived.

A description of the protocol and other references can be found at www.sagenet.org.

cDNA Microarrays

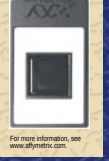
cDNA Microarrays simultaneously measure expression of large numbers of genes based on hybridization to cDNAs attached to a solid surface. Measures of expression are relative between two conditions.



For more information, see www.microarrays.org.

Affy Oligo Arrays

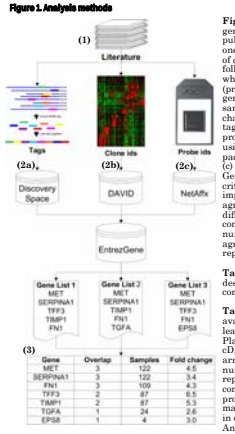
Affymetrix oligonucleotide arrays make use of tens of thousands of carefully designed oligos to measure the expression level of thousands of genes at once. A single labeled sample is hybridized at a time and an intensity value reported. Values are the based on numerous different probes for each gene or transcript to control for non-specific binding and chip inconsistencies.



For more information, see www.affymetrix.com.

2. Methods

Figure 1. Analysis methods



Gene	Overlap	Samples	Fold change
MET	3	122	4.5
SERPINA1	3	122	3.4
FN1	3	109	4.3
TFF3	2	87	6.5
TIMP1	2	87	5.3
TGFA	1	24	2.6
EPS8	1	4	3.0

Fig 1: (1) Lists of differentially expressed genes were collected and curated from published studies. Each study consists of one or more comparisons between pairs of conditions (e.g. PTC vs. norm). The following information was recorded wherever possible: Unique identifier (probe, tag, accession); gene description; gene symbol; comparison conditions; sample numbers for each condition; fold change; direction of change. **(2)** SAGE tags, cDNA clone ids and Affymetrix probe ids were mapped to Entrez Gene using: (a) the DiscoverySpace software package[1]; (b) the DAVID Resource[2]; (c) the Affymetrix annotation files[3]. **(3)** Genes are ranked according to several criteria in the following order of importance: (i) number of comparisons in agreement (ie. listing the same gene as differentially expressed and with a consistent direction of change); (ii) total number of samples for comparisons in agreement; and (iii) average fold change reported for comparisons in agreement.

Table 1: Lists all abbreviations used to describe the samples and conditions compared in the various studies.

Table 2: A total of 34 comparisons were available from 21 studies, utilizing at least 10 different expression platforms. Platforms can be generally grouped into cDNA arrays (blue), oligonucleotide arrays (purple) and SAGE (pink). The numbers of 'up-/down-regulated' genes reported are for condition 1 relative to condition 2 for each comparison as provided. Only genes that could be mapped to a common identifier were used in our subsequent overlap analyses (see Analysis methods).

Table 1. Abbreviations for sample descriptions

ACL	Anaplastic thyroid cancer cell line
APTN	Autonomously functioning thyroid nodules
ATC	Anaplastic thyroid cancer
CTN	Cold thyroid nodule
FA	Follicular adenoma
FCL	Follicular carcinoma cell line
FTC	Follicular thyroid carcinoma
FVPTC	Follicular variant papillary carcinoma
GT	Goiter
HCC	Hurthle cell carcinoma
HN	Hyperplastic nodule
M	Metastatic
MACL	Anaplastic thyroid cancer cell line with metastatic capacity
Norm	Normal
PCL	Papillary carcinoma cell line
PTC	Papillary thyroid carcinoma
TCVPTC	Tall-cell variant PTC
UCL	Undifferentiated carcinoma cell line

3. Thyroid cancer expression data

Table 2. Thyroid cancer profiling studies included in analysis

Study	Platform	Genes/ features	Comparison		Up-/down
			Condition 1 (No. samples)	Condition 2 (No. samples)	
Chen *et al.* 2001	Atlas cDNA (Clontech)	588	M (1)	FTC (1)	18/40
Arnaldi *et al.* 2005	Custom cDNA	1807	FCL(1)	Norm (1)	9/20
			PCL(1)	Norm (1)	1/8
			UCL(1)	Norm (1)	1/7
			FCL(1), PCL(1), UCL(1)	Norm (1)	3/6
Huang *et al.* 2001	Affymetrix HG-U95A	12558	PTC (8)	Norm (8)	24/27
Aldred *et al.* 2004	Affymetrix HG-U95A	12558	FTC (9)	PTC(6), Norm(13)	142/0
			PTC (6)	PTC(9), Norm(13)	0/68
Cerutti *et al.* 2004	SAGE	N/A	FA(1)	PTC(1), Norm(1)	5/0
			PTC(1)	FA(1), Norm(1)	12/0
Eszlinger *et al.* 2001	Atlas cDNA (Clontech)	588	AFTN(3), CTN(3)	Norm(6)	0/16
Finley *et al.* 2004	Affymetrix HG-U95A	12558	PTC(7), FVPTC(7)	FA(14), HN(7)	48/85
Zou *et al.* 2004	Atlas cancer array	1176	MACL(1)	ACL(1)	43/21
Weber *et al.* 2005	Affymetrix HG-U133A	22283	FA(12)	PTC(12)	12/84
Hawthorne *et al.* 2004	Affymetrix HG-U95A	12558	GT(6)	Norm(6)	1/7
			PTC(8)	GT(6)	10/28
			PTC(8)	Norm(8)	4/4
Onda *et al.* 2004	Amersham custom cDNA	27648	ACL(11), ATC(10)	Norm(10)	31/56
Wasenius *et al.* 2003	Atlas cancer cDNA	1176	PTC(18)	Norm(3)	12/9
Barden *et al.* 2003	Affymetrix HG-U95A	12558	PTC(9)	FA(10)	59/45
Yano *et al.* 2004	Amersham custom cDNA	3968	PTC(7)	Norm(7)	54/0
Chevillard *et al.* 2004	custom cDNA	5760	PTC(3)	FA(4)	12/31
			FVPTC(3)	PTC(2)	123/16
Mazzanti *et al.* 2004	Hs-UniGem2 cDNA	10000	PTC(17), FVPTC(15)	FA(16), HN(15)	5/41
Takano *et al.* 2000	SAGE	N/A	PTC(1)	ATC(1)	3/10
			PTC(1)	FA(1)	4/1
			Norm(1)	FA(1)	6/0
			PTC(1)	ATC(1)	2/11
			PTC(1)	FA(1)	7/0
			PTC(1)	FTC(1)	2/1
Finley *et al.* 2004	Affymetrix HG-U95A	12558	PTC(9), PTC(11), FVPTC(13)	FA(16), HN(10)	59/55
Pauws *et al.* 2004	SAGE	N/A	FVPTC(1)	Norm(1)	33/9
Jarzab *et al.* 2005	Affymetrix HG-U133A	22283	PTC(16)	Norm(16)	75/27
Giordano *et al.* 2005	Affymetrix HG-U133A	22283	PTC(51)	Norm(4)	90/151
21 studies	**10 platforms**		**34 comparisons (473 samples)**		**1785**

4. Overlap analysis results

Table 3. Comparison groups analyzed for overlap

Overlap analysis group	Condition set 1	Condition set 2	# comps	# genes (multi-study)	p-value
Cancer vs. non-cancer	ACL, ATC, FCL, PTC, FVPTC, HCC, M, MACL, PCL, PTC, TCVPTC, UCL	AFTN, CTN, FA, GT, HN, Norm	21	755 (107)	<0.0001
Cancer vs. normal	ACL, ATC, FCL, PTC, FVPTC, HCC, M, MACL, PCL, PTC, TCVPTC, UCL	Norm	12	478 (53)	<0.0001
Cancer vs. benign	ACL, ATC, FCL, PTC, FVPTC, HCC, M, MACL, PCL, PTC, TCVPTC, UCL	AFTN, CTN, FA, GT, HN	8	332 (38)	<0.0001
Normal vs. benign	Norm	AFTN, CTN, FA, GT, HN	3	19 (1)	0.0113
PTC vs. non-cancer	FVPTC, PCL, PTC, TCVPTC	AFTN, CTN, FA, GT, HN, Norm	12	503 (82)	<0.0001
PTC vs. normal	FVPTC, PCL, PTC, TCVPTC	Norm	8	369 (49)	<0.0001
PTC vs. benign	FVPTC, PCL, PTC, TCVPTC	AFTN, CTN, FA, GT, HN	4	183 (13)	<0.0001
PTC vs. other	FVPTC, PCL, PTC, TCVPTC	Any other	15	528 (107)	<0.0001
PTC vs. FA	PTC	FA	6	222 (3)	0.0455
PTC vs. other	PTC, FCL	Any other	10	403 (15)	0.0003
Aggressive cancer vs. other	ACL, ATC, M, MACL	Any other	4	145 (4)	0.0402
ATC vs. other	ACL, ATC, MACL	Any other	3	91 (6)	<0.0001
Affy re-processed	PTC, FTC	Norm, FA	5	1317 (179)	<0.0001

Figure 2. Gene overlap for cancer vs. non-cancer analysis

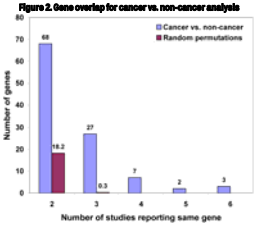


Table 3: Each overlap analysis group defines an artificial group of comparisons for which gene overlap was analyzed. In all groups considered except for one, we identified one or more genes that were reported in two or more studies. For example, the "cancer vs. non-cancer" group (highlighted) includes all comparisons between what we would consider 'cancer' (as in condition set 1) and 'non-cancer' (as in condition set 2). In this case, 21 comparisons met the criteria and produced a list of 755 potential cancer markers, 107 of which were identified in multiple studies. These 'multi-study cancer versus non-cancer markers' are summarized further in figure 2 and table 4.

Fig. 2: 107 genes were found in multiple studies for the cancer versus non-cancer analysis with overlap of two to six, much more than expected by chance.

4. Overlap analysis results (cont'd)

Table 4. Cancer versus non-cancer genes identified in 4 or more independent studies

Gene	Description	Comp's (Up/Down)	N	Fold Change
MET	met proto-oncogene (hepatocyte growth factor receptor)	0/0	202	3.03
TFF3	trefoil factor 3 (intestinal)	0/6	196	-14.70
SERPINA1	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	6/0	192	15.84
EPS8	epidermal growth factor receptor pathway substrate 8	5/0	186	3.15
TIMP1	tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	5/0	142	5.38
TGFA	transforming growth factor, alpha	4/0	165	4.64
QPCT	glutaminyl-peptide cyclotransferase (glutaminyl cyclase)	4/0	153	7.31
PROS1	protein S (alpha)	4/0	149	4.32
CRABP1	cellular retinoic acid binding protein 1	0/4	146	-11.55
FN1	fibronectin 1	4/0	128	7.68
FCGBP	Fc fragment of IgG binding protein	0/4	108	-2.41
TPO	thyroid peroxidase	0/4	91	-4.69

Table 4: shows a partial list (genes identified in 4 or more comparisons) from the cancer vs. non-cancer analysis. A complete table for this group and all others are available as supplementary data (www.bcgsc.ca/bioinfo/gethyroid).

Figure 3. Affymetrix subset analysis

Fig 3. A comparison of genes with multi-study evidence based on published lists versus a smaller subset re-analysed from raw microarray data showed a highly significant level of agreement (p-value = 6.47E-68). The 107 cancer vs.non-cancer multi-study genes (overlap of two or more) showed a concordance of 0.177 (± 0.048, 95% C.I.) with the 179 multi-study genes identified from the re-analysed Affymetrix subset. In total, there were 43 genes identified by both methods.

Table 5. Utility of stained markers for distinguishing benign from tumour.

Marker	% Pos. Benign	% Pos Cancer	p-value	Variable importance
BCL2	78	34.4	0*	24.074
CCND1	45.7	89	0*	22.57
P16	5	44.8	0*	16.41
P21	38	75	0*	8.687
CCNE1	37.4	69.2	0*	5.227
KIT	18	2.2	0*	3.4
S100	1	13.8	0.001*	1.558
HER3	34.3	56.5	0.003	0.478
AMFR	20.4	38.7	0.005	1.653
KI67	5	17.2	0.007	1.451
HER4	85.9	72.7	0.026	0.06
HER1	60	73.6	0.046	4.189
SERPINA1	8	17.6	0.047	0.84
P27	50	62.5	0.078	3.507
P57	2	7.3	0.096	0.198
P63	2	7.3	0.096	0.156
TTF1	85	90.9	0.217	1.904
P53	7	4.2	0.389	0
TG	100	98.9	0.476	3.638
CDX2	0	1.1	0.476	0
ESR1	0	1.1	0.476	0
PR	0	1.1	0.482	0
HER2	1	2.2	0.604	0
WT1	1	0	1	0
TSH	0	0	N/A	0

Table 5: Twenty-five markers were stained, scored and analyzed on a tissue microarray consisting of 100 benign and 105 malignant tissue samples (6 follicular, 90 papillary, 3 Hurthle cell, and 6 medullar). Using Pearson Chi-Square or Fisher's Exact test (where appropriate) 13 markers were found to be significantly associated ($p<0.05$) with disease status (benign vs. cancer). After multiple testing correction (Bonferroni) seven markers were still significant. All 25 markers were submitted to the Random Forests classification algorithm with a target outcome of cancer versus benign. *A classifier was produced with an overall error rate of 0.189, sensitivity of 79.2% and specificity of 83%.*

5. Conclusions and Future work

Conclusions:
> A significant number of genes are consistently identified in the literature as differentially expressed between different thyroid tissue and tumour subtypes.
> Our approach represents a useful method for identifying consistent gene expression markers when raw data is unavailable (as is generally the case).
> Some markers have previously undergone extensive validation while others have not yet been investigated at the protein level.
> Preliminary immunohistochemistry analysis on a TMA of over 200 thyroid samples for 25 antibodies show promising results.
> The addition of candidate genes from the meta-analysis may facilitate the development of a clinically relevant diagnostic marker panel.

Future work:
> Continue validation of putative markers by immunohistochemistry on TMA.
> Development of a clinically useful classifier for thyroid tissue based on results of TMA.

6. Acknowledgments

funding | Natural Sciences and Engineering Council of Canada (OG); Michael Smith Foundation for Health Research (OG, SW, and SJ); Canadian Institutes of Health Research (OG); BC Cancer Foundation

references | 1. Varhol et al, unpublished, http://www.bcgsc.ca/discoveryspace/; 2. Dennis et al. 2003, http://david.abcc.ncifcrf.gov/; 3. Affymetrix, http://www.affymetrix.com/support/index.affx.