

RESEARCH ARTICLE

Open Access



A five-gene reverse transcription-PCR assay for pre-operative classification of breast fibroepithelial lesions

Wai Jin Tan¹, Igor Cima¹, Yukti Choudhury¹, Xiaona Wei¹, Jeffrey Chun Tatt Lim², Aye Aye Thike², Min-Han Tan^{1*} and Puay Hoon Tan^{2,3*}

Abstract

Background: Breast fibroepithelial lesions are biphasic tumors and include fibroadenomas and phyllodes tumors. Preoperative distinction between fibroadenomas and phyllodes tumors is pivotal to clinical management. Fibroadenomas are clinically benign while phyllodes tumors are more unpredictable in biological behavior, with potential for recurrence. Differentiating the tumors may be challenging when they have overlapping clinical and histological features especially on core biopsies. Current molecular and immunohistochemical techniques have a limited role in the diagnosis of breast fibroepithelial lesions. We aimed to develop a practical molecular test to aid in distinguishing fibroadenomas from phyllodes tumors in the pre-operative setting.

Methods: We profiled the transcriptome of a training set of 48 formalin-fixed, paraffin-embedded fibroadenomas and phyllodes tumors and further designed 43 quantitative polymerase chain reaction (qPCR) assays to verify differentially expressed genes. Using machine learning to build predictive regression models, we selected a five-gene transcript set (*ABCA8*, *APOD*, *CCL19*, *FN1*, and *PRAME*) to discriminate between fibroadenomas and phyllodes tumors. We validated our assay in an independent cohort of 230 core biopsies obtained pre-operatively.

Results: Overall, the assay accurately classified 92.6 % of the samples (AUC = 0.948, 95 % CI 0.913–0.983, $p = 2.51E-19$), with a sensitivity of 82.9 % and specificity of 94.7 %.

Conclusions: We provide a robust assay for classifying breast fibroepithelial lesions into fibroadenomas and phyllodes tumors, which could be a valuable tool in assisting pathologists in differential diagnosis of breast fibroepithelial lesions.

Keywords: Phyllodes tumors, Fibroadenomas, Fibroepithelial lesions, Prediction, Multigene assay, Core biopsy, FFPE

Background

Fibroadenomas and phyllodes tumors are fibroepithelial lesions of the breast, characterized by proliferation of both epithelial and stromal components. Fibroadenomas are more commonly encountered on core biopsies than the rarer phyllodes tumors (approximately 20 % and <1 % of breast core needle biopsies respectively) [1, 2]. The preoperative distinction between the two lesions has significant impact on subsequent treatment. The current

recommended management for phyllodes tumor diagnosed on core biopsy is wide excision without axillary staging regardless of grade [3]. Conversely, fibroadenomas are observed conservatively or, if tumors are larger than 2 cm, may be simply excised without achieving negative surgical margins [3]. This approach is due to the indolent behavior of fibroadenomas, despite sporadic reports of recurrences [4, 5], while phyllodes tumors have unpredictable outcomes with malignant tumors potentially progressing to metastasis and mortality [6–10]. It has been challenging separating cellular fibroadenoma from benign phyllodes tumor due to overlapping histological features, and this is particularly problematic on limited material of core biopsies, which may lead to over- or under-treatment for some patients, resulting in unnecessary anxiety and cost.

* Correspondence: mhtan@ibn.a-star.edu.sg; tan.puay.hoon@singhealth.com.sg

¹Division of Biodevices and Diagnostics, Institute of Bioengineering and Nanotechnology, 31 Biopolis Way, The Nanos, #04-01, Singapore 138669, Republic of Singapore

²Department of Pathology, Singapore General Hospital, 20 College Road, Academia, Level 7, Diagnostics Tower, Singapore 169856, Republic of Singapore
Full list of author information is available at the end of the article

Several studies have proposed differentiating histological features such as stromal cellularity, stromal overgrowth, fragmentation, subepithelial condensation and presence of adipose tissue within stroma on core biopsies being indicative of phyllodes tumor [11–13]. However, interpretation of these parameters is subjective, with interobserver variation and only moderate reproducibility between pathologists [11, 14]. Varied reports of immunohistochemical markers used in distinguishing phyllodes tumors from fibroadenomas suggest a lack of consensus and objectivity in assessing the expression of these biomarkers. Some authors reported Ki-67 expression to be helpful in diagnosing phyllodes tumors [15–17] but there are reports to the contrary [18, 19]. Lin et al. suggested a combination immunoscore of p16-INK4a and retinoblastoma-associated protein (pRB) [20] while Maity et al. reported expression of collagen I, III and CD105-positive microvessel density as parameters to differentiate the two lesions [21]. The vast majority of these studies were not conducted using pre-operative biopsies, which is where key management decision is required.

We set out to identify a useful molecular signature to help differentiate fibroadenomas from phyllodes tumors using pre-operative core biopsies to improve prediction of the final diagnosis.

Methods

Training set for assay development

The study received approval from the Centralized Institutional Review Board (CIRB 2005/002/F). As this was a retrospective study with anonymized cases, no specific patient consent was individually required. Forty-eight samples (24 fibroadenomas and 24 phyllodes tumors) were first employed as the training set for assay development. These included 10 paired core biopsies and surgical samples (20 samples), and 28 independent core and excisional samples from 38 patients (Table 1 and Additional file 1: Table S1). These formalin-fixed, paraffin-embedded (FFPE) samples were randomly selected from cases diagnosed at the Department of Pathology, Singapore General Hospital from 2008 to 2012. Hematoxylin and eosin (H&E)-stained slides were retrieved and reviewed. Phyllodes tumor was defined when there were well-developed fronds accompanied by increased stromal cellularity as opposed to fibroadenomas in which epithelial and stromal components were arranged in either intracanalicular or pericanalicular patterns without fronds or stromal hypercellularity. Differences in clinical features between fibroadenomas and phyllodes tumors were assessed with Mann–Whitney *U* test and Fisher's exact test.

Expression profiling by Whole-Genome DASL® High Throughput (HT) Assay

Representative tumor areas were identified of which three to seven sections of 10- μ m-thick sections from the

Table 1 Clinical features of the training cohort from 38 patients

Features	Fibroadenomas (n = 19)	Phyllodes tumors (n = 19)	<i>p</i> value
Age			
Median (range)	35 (17–80)	44 (18–64)	0.09
Size			
Median (range)	25 (15–50)	65 (25–220)	< 0.001
Ethnicity, n (%)			0.2
Chinese	13 (68.4)	11 (57.9)	
Malay	0 (0.0)	4 (21.0)	
Indian	2 (10.5)	1 (5.3)	
Others	4 (21.1)	3 (15.8)	
Histology			
Simple fibroadenoma	15 ^a		
Complex fibroadenoma	4 ^c		
Benign phyllodes tumor		13 ^b	
Borderline phyllodes tumor		3 ^c	
Malignant phyllodes tumor		3 ^c	

^aFour paired core biopsies and surgical excisions

^bThree paired core biopsies and surgical excisions

^cOne paired core biopsy and surgical excision

same FFPE tumor block were obtained, deparaffinized and macrodissected. RNA was extracted using the RNeasy FFPE kit (Qiagen, Hilden, Germany) and quantified by Nanodrop Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). A total of 100 ng was used for quality assessment by real-time amplification of the *RPL13A* gene (forward primer, 5'-CACTTGGGGA-CAGCATGAG-3', and reverse primer, 5'-GTAACCCCT TGGTTGTGCAT-3') using the Power SYBR® Green RNA-to-CT™ 1-Step Kit (Life Technologies, Carlsbad, CA, USA) on a CFX96™ Real-Time PCR instrument (Bio-Rad Laboratories, Hercules, CA, USA). Samples with threshold cycle (Ct) below 29 were further subjected to quality assessment on a bioanalyzer. Eligible samples were submitted for expression profiling on the Whole-Genome DASL® HT Assay (Illumina, Inc., San Diego, CA, USA) at the Biopolis Shared Facilities A*Star, Singapore. The assay interrogates 29,377 features using the HumanHT-12 v4 BeadChip (Illumina, Inc.). Quantile-normalized gene expression data pre-analyzed using GenomeStudio® (Illumina, Inc.) was delivered. Data are available through GEO [GEO: GSE78071].

Selection of normalization genes and differentiating genes

Normalization genes were selected based on the smallest value of coefficient of variation among all samples. Differentiating genes were selected using the Significance Analysis of Microarrays package [22] and filtered based on the following criteria: (1) *q*-value less than 0.05; (2) mean difference of expression above 500; (3) *R*-fold

above 1.5 (for genes highly expressed in phyllodes tumors) or less than 0.67 (for genes highly expressed in fibroadenomas).

Design of quantitative polymerase chain reaction (qPCR) assay

cDNA was synthesized from 1 μ g of RNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems®, Life Technologies, Carlsbad, CA, USA). Each qPCR assay consisted of 1X Power SYBR® Green PCR Master Mix (Life Technologies), 0.5 μ M of forward and reverse primer each and 1 μ l of 10-fold diluted cDNA as a template in a final total volume of 10 μ l. Primers were designed using Primer-BLAST (NCBI, Bethesda, MD, USA) [23] with accession number listed in Additional file 1: Table S2. Non-template control acted as a negative control. Specificity of the amplicons was verified by melt curve analysis.

Data quantification and model building

Delta Ct (Δ Ct) for each gene and sample was quantified as: $\Delta Ct = Ct(\text{gene, sample}) - \text{geomean}(Ct(\text{five normalization gene}))$. For comparison with expression of the Whole-Genome DASL® HT Assay, Δ Ct data was transformed to $2^{-\Delta Ct}$ as a positive linear scale and significance of correlation was analyzed with Pearson's correlation test. Similarly to Cima et al. and Kälén et al. [24, 25], we used Random Forest (RF) ensemble classifier [26] to rank the importance of gene transcripts differentiating fibroadenomas and phyllodes tumors on qPCR assays. The top seven performing genes were used to build predictive logistic regression models using exhaustive search for the best model. To this end we used the *glmulti*

package [27] with inclusion of the interaction terms. The best model was selected based on the lowest Akaike information criteria (AIC) value [28].

Validation cohort for model validation

The model of the multigene assay was tested on a separate set of 230 core biopsies with at least 2 years of follow-up. Hematoxylin and eosin (H&E)-stained slides were retrieved and reviewed. The outcome of the multigene assay was compared against the final diagnosis on the corresponding surgical excisions. Cases without subsequent surgical excisions were free from progression for at least 2 years and diagnosis made based on the initial core biopsy was used as the reference instead.

Results

Clinical features of training set

The clinical features and histology of the training set are shown in Table 1. Examples of the histological appearances of fibroadenoma and phyllodes tumor are shown in Fig. 1. Phyllodes tumors were significantly larger than fibroadenomas ($p < 0.001$). Median age of patients diagnosed with fibroadenomas and phyllodes tumors was 35 years and 44 years respectively ($p = 0.09$). No significant differences were observed for ethnicity distribution between the two groups of tumors.

Expression profiling and correlation with qPCR assays

Forty-seven samples (97.9 %) from 37 patients passed the quality control and were profiled successfully. Genes discriminating fibroadenomas and phyllodes tumors are listed in Additional file 1: Table S3. We designed and validated qPCR assays on 43 selected genes. Concordance

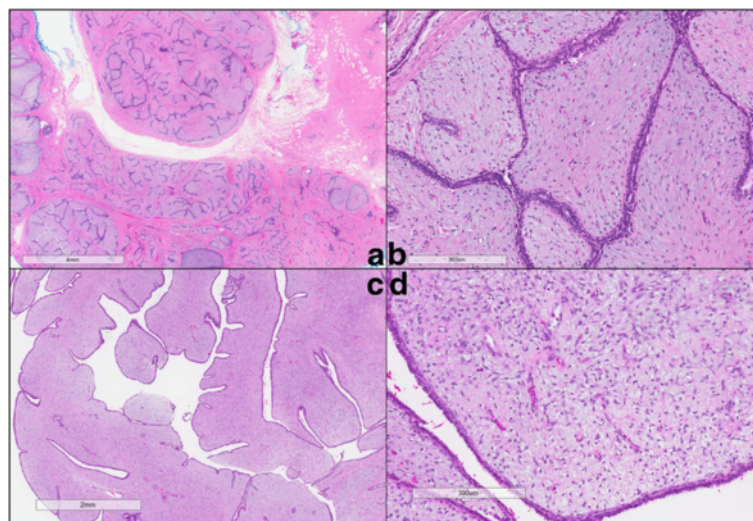


Fig. 1 Histology of fibroadenoma and phyllodes tumor on surgical excisions. **a, b** An example of a fibroadenoma at low-power (**a**) and high-power (**b**) magnifications. **c, d** An example of a benign phyllodes tumor at low-power (**c**) and high-power (**d**) magnifications

between expression profiling and the qPCR assays was assessed based on a pilot run on six representative samples (Additional file 1: Table S4). Twenty-three assays with Pearson's *r* of above 0.6 were further tested on the remaining 40 samples. One case was excluded due to insufficient material after expression profiling.

Development of a multigene qPCR panel

The results of ΔCt for all 23 qPCR assays were ranked using variable importance feature of the Random Forest classifier (Fig. 2). The seven most important genes in separating fibroadenomas and phyllodes tumors were *TRIM29*, *FN1*, *CCL19*, *ABCA8*, *NPTX2*, *APOD* and *PRAME*. A total of 268,435,456 candidate models were identified by *glmulti* based on these seven genes. We employed the genetic algorithm approach in the package to perform automated screening for the best model based on AIC value. A final five-gene model encompassing *APOD*, *ABCA8*, *PRAME*, *FN1*, and *CCL19* with AIC of 14.2 was returned with coefficients as listed in Table 2.

Performance of the model

Patient demographics of the validation cohort of 230 core biopsies are shown in Table 3. Overall the assay was able to predict 213 (92.6 %) cases accurately. The prediction accuracy rates for fibroadenomas and phyllodes tumors were 94.7 % (179/189) and 82.9 % (34/41) respectively (Table 4), with positive (PPV) and negative (NPV) predictive values of 77.3 % and 96.2 %. A receiver operating characteristics curve with an area under the curve (AUC) of 0.948 (95 % confidence interval (CI) 0.913–0.983, *p* = 2.51E-19) was generated, indicating a large effect size in expected diagnostic performance for the five-gene assay (Fig. 3).

Of the 17 discordant cases (Table 5), seven were diagnosed as phyllodes tumors on pathological reports but

Table 2 Coefficients of the best model in predicting diagnosis in the 46-sample set

Genes	Coefficients
<i>APOD</i>	2.95575
<i>APOD:ABCA8</i>	-0.11934
<i>PRAME:FN1</i>	-0.43165
<i>PRAME:CCL19</i>	0.08326

were predicted as fibroadenomas on our assay. Upon review of these seven cases, two cases contained high epithelium content (Fig. 4), two were heterogeneous on histology with focal areas suggestive of fibroadenomas, while three other cases were confirmed as phyllodes tumors on review. The other ten of the 17 discordant cases were diagnosed as fibroadenomas on pathological reports but were predicted as phyllodes tumors on our assay. Among these ten cases, six cases had post-operative excisional material available as reference while the remaining four cases were benchmarked against the pre-operative pathological diagnosis. Of the six with excisional material, four were unequivocally fibroadenomas on histology, one was a cellular fibroadenoma without prominent fronds, and one was a fibroadenoma with sclerosing adenosis. Of the four pre-operative biopsies, one was unequivocally fibroadenoma, two cases contained features in keeping with fibroadenoma with hyalinized leafy fronds noted albeit without stromal cellularity, and one was an indeterminate case with focal areas of hemorrhage and high cellularity, which could not be definitively concluded on review.

Table 3 Patient demographics of the validation cohort of 230 core biopsies

Characteristics	<i>N</i>	Percentage (%)
Age		
Median	46 years	
Range	15–75 years	
Ethnicity		
Chinese	164	71.3
Malay	23	10.0
Indian	11	4.8
Others	32	13.9
Diagnosis		
Fibroadenoma	189	82.2
with subsequent surgical excisions	58	25.2
without subsequent surgical excisions	131	57
Phyllodes tumors	41	17.8
Benign	22	9.6
Borderline	16	6.9
Malignant	3	1.3

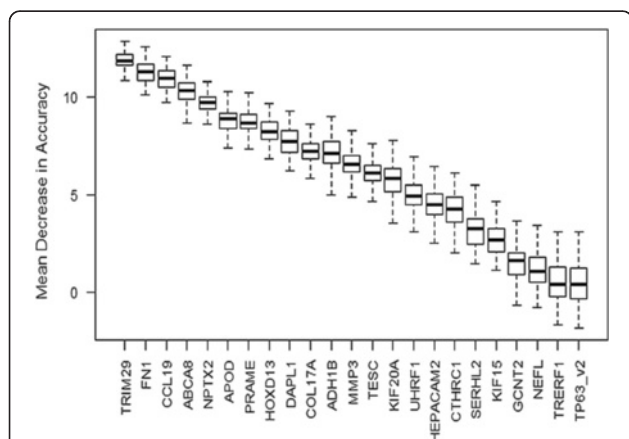


Fig. 2 Box plot of mean decrease in accuracy for 100 Random Forest trees generated for each gene. Mean decrease in accuracy measures the importance of each gene to the classification

Table 4 Performance of the five-gene model in predicting diagnosis in the independent validation cohort of 230 core biopsies

Predicted outcome from the five-gene assay	Diagnosis from pathological report		
	Fibroadenomas	Phyllodes tumors	
Fibroadenomas	179	7	NPV = 0.962
Phyllodes tumors	10	34	PPV = 0.773
	Spec = 0.947	Sen = 0.829	

The five-gene assay has an overall accuracy of 92.6 %, with a sensitivity (sen) of 82.9 % and specificity (spec) of 94.7 %. The positive predictive value (PPV) and negative predictive value (NPV) are 77.3 % and 96.2 % respectively

Among the 230 core biopsies, the pre-operative pathological diagnoses were inconclusive for 22 cases where the term ‘fibroepithelial lesion’ was assigned, and there were three cases where the preoperative diagnoses were incongruous with the post-operative outcome (Table 6). Of these 25 cases, the five-gene assay was 80 % (20/25) accurate in classification with a PPV of 94.7 %.

Discussion

Classification of breast fibroepithelial tumors based on differentiating morphological and immunohistochemical features on pre-operative material has been challenging with variable findings across different groups (see Table 7 for summary). Jacobs et al. and Lee et al. first described individual pathological parameters which might help to differentiate fibroadenomas and phyllodes tumors in these limited samples [11, 16]. Jara-Lazaro proposed a combination of histological and immunohistochemical markers to indicate phyllodes tumors on core biopsies [15] but did not weigh the relative importance of each parameter in predicting phyllodes tumors. Morgan addressed this question by proposing a predictive tool including coefficient factors for each parameter to distinguish between fibroadenomas and phyllodes tumors but this has yet to be validated in an independent

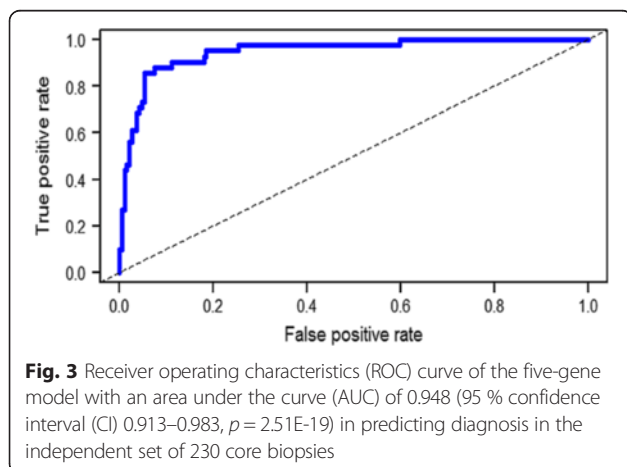


Table 5 Seventeen cases with discordant outcomes between the five-gene assay and pathological diagnosis in the validation cohort. Post-operative diagnoses were used as benchmark reference unless otherwise stated

Sample ID	Five-gene assay	Pathological diagnosis
CB22	FA	Benign PT ^a
CB116	FA	Benign PT
CB95	FA	Benign PT
CB28	FA	Benign PT ^a
CB68	FA	Benign PT ^a
CB120	FA	Borderline PT ^a
CB130	FA	Borderline PT
CB26	PT	Cellular FA
CB77	PT	FA
CB24	PT	FA
CB29	PT	FA
CB82	PT	FA
CB126	PT	FA
CB144	PT	FA ^{b,c}
CB184	PT	FA ^{b,c}
CB251	PT	FA ^b
CB258	PT	FA ^{b,d}

FA fibroadenoma, PT phyllodes tumor

^aPre-operative pathological diagnoses were inconclusive or discordant with post-operative pathological diagnoses (see Table 6 asterisked cases)

^bCases of core biopsies without subsequent surgical excisions. Outcome of the five-gene assay was benchmarked against the pre-operative pathological diagnosis

^cFeatures in keeping with fibroadenoma with hyalinized leafy fronds noted albeit without stromal cellularity

^dFocal areas of hemorrhage and high cellularity, diagnosis could not be definitively concluded on review

series of core biopsies [12]. Our study is the first to investigate differentiating features of fibroepithelial lesions on pre-operative material at the molecular level. We have developed a five-gene assay using a systematic approach based on genome-wide expression profiling data and validated the assay in an independent cohort of 230 pre-operative core biopsies of breast fibroepithelial lesions, the largest cohort reported so far. The pre-operative core biopsies were FFPE tissue containing low-quality RNA. Accordingly, our assay has been developed using RNA extracted from limited FFPE materials from core biopsies and thus is expected to perform on such material in the clinical setting.

Comparatively in surgical excisional materials, Huang et al. proposed a two-gene test derived from methylation profiling of an 11-gene panel in 86 samples [29], which described an elevated *RASSF1A* and/or *TWIST1* methylation observed in phyllodes tumors as compared to fibroadenomas. They further evaluated the test in a separate validation cohort of 19 samples and reported a sensitivity and specificity of 0.33 and 0.75 respectively,

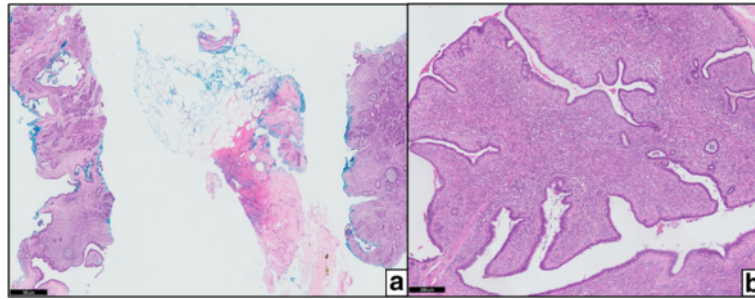


Fig. 4 Example of a discordant case containing high epithelium content. The five-gene assay predicted the core biopsies (a) as fibroadenoma but the final surgical excision (b) was diagnosed as phyllodes tumor on pathological reports

Table 6 Cases with inconclusive pre-operative pathological diagnoses (n = 22), and discordant pre-and post-operative pathological diagnoses (n = 3). Among these cases, the five-gene assay was 80 % (20/25) accurate in predicting the final post-operative outcome

Sample ID	Pathological reports		Outcome from five-gene assay
	Pre-operative	Post-operative	
CB22	FA	Benign PT	FA ^a
CB36	FA	Benign PT	PT
CB111	FA	Benign PT	PT
CB21	FEL	Benign PT	PT
CB31	FEL	Benign PT	PT
CB40	FEL	Benign PT	PT
CB43	FEL	Benign PT	PT
CB57	FEL	Benign PT	PT
CB59	FEL	Benign PT	PT
CB96	FEL	Benign PT	PT
CB99	FEL	Benign PT	PT
CB107	FEL	Benign PT	PT
CB114	FEL	Benign PT	PT
CB20	FEL	Borderline PT	PT
CB42	FEL	Borderline PT	PT
CB55	FEL	Borderline PT	PT
CB73	FEL	Borderline PT	PT
CB78	FEL	Borderline PT	PT
CB101	FEL	Borderline PT	PT
CB115	FEL	Borderline PT	PT
CB122	FEL	Borderline PT	PT
CB28	FEL	Benign PT	FA ^a
CB68	FEL	Benign PT	FA ^a
CB120	FEL	Borderline PT	FA ^a
CB77	FEL	FA	PT ^a

FA fibroadenoma, PT phyllodes tumor, FEL fibroepithelial lesion without definitive assignment into fibroadenoma or phyllodes tumor

^aInaccurate classification by the five-gene assay benchmarked against the post-operative pathological diagnosis

with a PPV and NPV of 0.83 and 0.23. However, development of the test from a pre-selected panel of 11 genes may not be representative and the sample size of the validation cohort was too small to be conclusive. In contrast, our assay has a better sensitivity and specificity at 0.83 and 0.95 despite a lower PPV of 0.77. In a separate study by Kuijper interrogating the transcriptome differences between five fibroadenomas and eight phyllodes tumors, *CTAG1/2*, *PRAME*, *HOXC13*, *ELF5* and *FABP7* were among 96 other transcripts found to be highly differentially expressed between fibroadenomas and phyllodes tumors [30]. More recently, Vidal et al. reported a cluster of 47 epithelial- and luminal-related genes was found to be more expressed in fibroadenomas than phyllodes tumors among 105 breast cancer-related genes studied [31]. Findings from these studies however, were not further deployed as a test to distinguish fibroadenomas from phyllodes tumors on pre-operative materials despite the significant differential expression observed.

The training cohort comprised a mixture of surgical excisions and core biopsies with varying classifications of fibroepithelial lesions, simulating a realistic clinical scenario. Phyllodes tumors comprise benign, borderline and malignant grades on a continuous spectrum [32]. It is important that the assay works across the spectrum although one may argue that the malignant grade of phyllodes tumors is rarely in the histologic differential diagnosis between fibroadenomas and phyllodes tumors, and hence the assay may have little utility in the separation of fibroadenomas from malignant phyllodes tumors. The proportion of malignant phyllodes tumors included in the training cohort concurs with the incidence of malignant phyllodes tumors reported in the literature [33]. Nevertheless, even with the exclusion of malignant phyllodes tumors in the training cohort, differences of expression between fibroadenomas and phyllodes tumors for genes selected for the assay still fall within our selection criteria (R-fold differences above 1.5 and mean differences above 500) and hence would not have altered the assay development outcome. Excluding the

Table 7 Summary of selected studies investigating differential features of fibroadenomas and phyllodes tumors on pre-operative materials

Author	N	Assessment approach	Key findings
Jacobs [16]	16 FAs 12 PTs	Histology Immunohistochemistry	Stromal cellularity, mitoses, expression of Ki-67 and topoisomerase IIa might help determining phyllodes tumors
Lee [11]	38 FAs 44 PTs	Histology	Features indicating phyllodes tumors: (1) increased stromal cellularity in at least 50 % of the specimen; (2) stromal overgrowth ; (3) fragmentation and; 4) presence of adipose tissue within stroma
Jara-Lazaro [15]	21 FAs 36 PTs	Histology Immunohistochemistry	(1) Marked stromal hypercellularity and nuclear atypia, stromal overgrowth, and ill-defined lesional margins exclusively predicted phyllodes tumor on excision. (2) Expression of Ki67 ≥ 5 % , topoisomerase IIa ≥ 5 % and reduced CD34 correlates significantly with phyllodes tumors.
Morgan [12]	91 ^a FAs 21 PTs	Histology	Proposed two putative predictive tools: (1) $Logit P(x) = -0.9014 (age) - 3.61 (mitosis) + 11.156$ (2) $Z = 0.8909(age) + 0.0347(\% stroma) + 0.5274 (mitoses/10HPF)$
Yasir [13]	37 FAs 27 PTs	Histology	Stromal mitoses and/or three or more histological features were helpful in predicting phyllodes tumors on excisions.

FAs fibroadenomas, PTs phyllodes tumors

^aDiagnoses not confirmed on excisions

three malignant tumor cases in the validation cohort only slightly reduces the sensitivity and PPV from 0.829 and 0.773 to 0.816 and 0.756 respectively. It is not our aim to investigate the differential expression between the phyllodes tumor grades although there is a trend of differences observed between grades in the expression of these five genes (results not shown). The sample sizes of borderline and malignant phyllodes tumors would be too small for meaningful analysis.

Several underlying factors which potentially limit the performance of the assay resulting in 17 discordant outcomes between the assay and pathological diagnosis include tumor heterogeneity and the issue of sampling on core biopsies. These factors may also have contributed to the three discordant pathological diagnoses between pre-operative core and post-operative excision materials. Core biopsies offer insight into only part of a tumor, which may not truly represent its entirety. Also, it is not uncommon for phyllodes tumors to contain areas indistinguishable from fibroadenomas, as seen in two discordant phyllodes tumor cases incorporating focal areas suggestive of fibroadenomas. Two other discordant phyllodes tumors harbored high epithelium content. The contribution of the epithelial component to the performance of the assay has yet to be ascertained although previous studies have shown that mutations were found in the stromal but not epithelial component [34, 35].

The limitation of our validation cohort is that the sample size for phyllodes tumor is small but the test was validated on a larger number of fibroadenomas, which have higher incidence compared to phyllodes tumors. We incorporated fibroadenomas on core biopsies which were

not excised surgically although these may theoretically include uncertainty as the diagnoses are based solely on the core biopsy and not on the excised tumor. However, precluding fibroadenomas without subsequent excisions would result in a selection bias due to the exclusion of a large portion of representative cases. Moreover, the incidence of phyllodes tumor subsequent to a fibroadenoma diagnosis on core biopsy is very low [36], with an average duration of 12 months to the final correct diagnosis.

We do not advocate that the current diagnostic framework be replaced by the assay. Apart from the histological findings, clinical decision whether to proceed with surgical excision takes into account other factors such as radiological size and characteristics, as well as patient symptoms. For instance, a diagnosis of fibroadenoma on core biopsy may still be followed by excision if there is radiologic-pathologic discordance, or if the lesion is large or symptomatic. A diagnosis of phyllodes tumor on core biopsy however, warrants excision. Incorporating the results from our assay allows an additional tool that can be integrated into the decision-making process, enhancing precision especially when it affirms the pathological assessment on core biopsy. The gene assay is also helpful for pathologists in interpreting these lesions when the histological characteristics are indeterminate or ambiguous. This is exemplified by the 22 fibroepithelial lesions without a conclusive classification on core biopsy in the validation cohort. The multigene assay was able to classify 82 % of these cases accurately with a PPV of 94.7 %. The practicality and utility of the assay however, will need to be further validated in prospective studies.

The five-gene assay includes genes of various biological functions. *FNI* (fibronectin 1) encodes a major component of the extracellular matrix. *APOD* (apolipoprotein D) and *ABCA8* (ATP-binding cassette, sub-family A member 8) encode transporter proteins while *PRAME* (preferentially expressed antigen in melanoma) and *CCL19* (chemokine ligand 19) genes are involved in immunoregulatory processes. Some of these genes were reported to be useful in differential diagnosis of other forms of tumors such as *FNI* as a marker for renal cell carcinoma aggressiveness [37], *PRAME* as a marker for differentiating Müllerian carcinoma from malignant mesothelioma [38] and *ABCA8* as part of a multigene gene assay for classifying cancer types [39]. While the individual functional role of these genes has not been implicated in breast fibroepithelial lesions, we found that these markers work best in combination for differential diagnosis between fibroadenomas and phyllodes tumors, as derived from our model algorithm. Nonetheless, it would be of interest to investigate the functional roles of these genes in breast fibroepithelial lesions in future studies.

Conclusions

We have developed a practical molecular assay for fibroepithelial lesions, classifying fibroadenomas and phyllodes tumors in pre-operative core biopsies. This may serve as an adjunctive aid for accurate pathological diagnosis. Prospective real-world trials will be helpful to determine whether improved surgical decision-making, supported by more accurate histological diagnosis, will lead to better outcomes.

Additional file

Additional file 1: Supplementary tables. **Table S1** Details of 48 samples from 38 patients constituting the training cohort. **Table S2** Primers designed for potential differentiating genes and normalization genes. **Table S3** Significant genes differentially expressed between fibroadenomas (FAs) and phyllodes tumors (PTs). **Table S4** Correlation between expression profiling and qPCR assays based on a pilot run of six representative samples. Genes with good correlation value ($r \geq 0.6$) were subjected to testing on remaining 40 samples. (XLSX 59 kb)

Abbreviations

ABCA8: ATP-binding cassette, sub-family A member 8; AIC: Akaike information criteria; *APOD*: Apolipoprotein D; AUC: area under the curve; *CCL19*: chemokine ligand 19; CI: confidence interval; Ct: threshold cycle; FA: fibroadenoma; FEL: fibroepithelial lesion; FFPE: formalin-fixed, paraffin-embedded; *FNI*: fibronectin 1; H&E: hematoxylin and eosin; HT: high throughput; *NPTX2*: neuronal pentraxin II; NPV: negative predictive value; PPV: positive predictive value; *PRAME*: preferentially expressed antigen in melanoma; pRB: retinoblastoma-associated protein; PT: phyllodes tumor; qPCR: quantitative polymerase chain reaction; RF: Random Forest; *TRIM29*: tripartite motif containing 29.

Competing interests

MHT, PHT, WJT, and IC have filed for patents in molecular diagnostics with this assay. All other authors have no other competing interests.

Authors' contributions

WJT performed RNA extraction, developed and validated the qPCR assays, analyzed the data, and drafted the manuscript. IC participated in model development, performed data analysis, and drafted the manuscript. YC designed the workflow of the study and participated in the revision of the manuscript. XW participated in the data analysis of WG-DASL and revision of the manuscript. JLCT collected FFPE samples, collated clinical data and helped to draft the manuscript. AAT participated in selection of cases and revision of the manuscript. MHT conceived and designed the study, coordinated all experiments, and critically revised the manuscript. PHT conceived and supervised the study, participated in case selection, and critically revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This study was supported by the Institute of Bioengineering and Nanotechnology (Biomedical Research Council, Agency for Science, Technology and Research, Singapore) and the Biomedical Research Council-Science and Engineering Research Council (BMRC-SERC) Joint Diagnostics Grant. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We would like to thank Dr Vivien Koh of Biopolis Shared Facilities, Agency of Science, Technology and Research (A*Star), Singapore, for her contribution to microarray work.

Author details

¹Division of Biodevices and Diagnostics, Institute of Bioengineering and Nanotechnology, 31 Biopolis Way, The Nanos, #04-01, Singapore 138669, Republic of Singapore. ²Department of Pathology, Singapore General Hospital, 20 College Road, Academia, Level 7, Diagnostics Tower, Singapore 169856, Republic of Singapore. ³Duke-NUS Graduate Medical School, 8 College Road, Singapore 169857, Republic of Singapore.

Received: 1 January 2016 Accepted: 25 February 2016

Published online: 09 March 2016

References

- Dillon MF, Hill ADK, Quinn CM, O'Doherty A, McDermott EW, O'Higgins N. The accuracy of ultrasound, stereotactic, and clinical core biopsies in the diagnosis of breast cancer, with an analysis of false-negative cases. *Ann Surg.* 2005;242(5):701–7.
- Komenaka IK, El-Tamer M, Pile-Spellman E, Hibshoosh H. Core needle biopsy as a diagnostic tool to differentiate phyllodes tumor from fibroadenoma. *Arch Surg.* 2003;138(9):987–90.
- National Comprehensive Cancer Network. Clinical practice guidelines in oncology. Breast cancer version 3.2014. http://www.nccn.org/professionals/physician_gls/pdf/breast.pdf. Accessed 2 September 2014.
- Grady I, Gorsuch H, Wilburn-Bailey S. Long-term outcome of benign fibroadenomas treated by ultrasound-guided percutaneous excision. *Breast J.* 2008;14(3):275–8.
- Wang ZC, Buraimoh A, Iglehart JD, Richardson AL. Genome-wide analysis for loss of heterozygosity in primary and recurrent phyllodes tumor and fibroadenoma of breast using single nucleotide polymorphism arrays. *Breast Cancer Res Treat.* 2006;97(3):301–9.
- Tan PH, Jayabaskar T, Chuah KL, Lee HY, Tan Y, Hilmy M, et al. Phyllodes tumors of the breast: the role of pathologic parameters. *Am J Clin Pathol.* 2005;123(4):529–40.
- Tan PH, Thike AA, Tan WJ, Thu MMM, Busmanis I, Li H, et al. Predicting clinical behaviour of breast phyllodes tumours: a nomogram based on histological criteria and surgical margins. *J Clin Pathol.* 2012;65(1):69–76.
- Parker SJ, Harries SA. Phyllodes tumours. *Postgrad Med J.* 2001;77(909):428–35.
- Belkacemi Y, Bousquet G, Marsiglia H, Ray-Coquard I, Magné N, Malard Y, et al. Phyllodes tumor of the breast. *Int J Radiat Oncol Biol Phys.* 2008;70(2):492–500.
- Moffat CJ, Pinder SE, Dixon AR, Elston CW, Blamey RW, Ellis IO. Phyllodes tumours of the breast: a clinicopathological review of thirty-two cases. *Histopathology.* 1995;27(3):205–18.
- Lee AHS, Hodi Z, Ellis IO, Elston CW. Histological features useful in the distinction of phyllodes tumour and fibroadenoma on needle core biopsy of the breast. *Histopathology.* 2007;51(3):336–44.
- Morgan JM, Douglas-Jones AG, Gupta SK. Analysis of histological features in needle core biopsy of breast useful in preoperative distinction between fibroadenoma and phyllodes tumour. *Histopathology.* 2010;56(4):489–500.

13. Yasir S, Gamez R, Jenkins S, Visscher DW, Nassar A. Significant histologic features differentiating cellular fibroadenoma from phyllodes tumor on core needle biopsy specimens. *Am J Clin Pathol*. 2014;142(3):362–9.
14. Lawton TJ, Acs G, Argani P, Farshid G, Gilcrease M, Goldstein N, et al. Interobserver variability by pathologists in the distinction between cellular fibroadenomas and phyllodes tumors. *Int J Surg Pathol*. 2014;22(8):695–8.
15. Jara-Lazaro AR, Akhilesh M, Thike AA, Lui PC-W, Tse GM-K, Tan PH. Predictors of phyllodes tumours on core biopsy specimens of fibroepithelial neoplasms. *Histopathology*. 2010;57(2):220–32.
16. Jacobs TW, Chen Y-Y, Guinee JDG, Holden JA, Cha I, Bauermeister DE, et al. Fibroepithelial lesions with cellular stroma on breast core needle biopsy: are there predictors of outcome on surgical excision? *Am J Clin Pathol*. 2005;124(3):342–54.
17. Yohe S, Yeh IT. “Missed” diagnoses of phyllodes tumor on breast biopsy: pathologic clues to its recognition. *Int J Surg Pathol*. 2008;16(2):137–42.
18. Kaya R, Pestereli HE, Erdogan G, Gülkesen KH, Karaveli S. Proliferating activity in differential diagnosis of benign phyllodes tumor and cellular fibroadenomas: is it helpful? *Pathol Oncol Res*. 2001;7(3):213–6.
19. Barwijuk-Machala M, Musiatowicz B, Cylwik J, Reszec J, Augustynowicz A. AgNOR, Ki-67 and PCNA expression in fibroepithelial tumours of the breast in correlation with morphological features. *Folia Morphol (Warsz)*. 2004;63(1):133–5.
20. Lin C-K, Tsai W-C, Lin Y-C, Yu J-C. Biomarkers distinguishing mammary fibroepithelial neoplasms: a tissue microarray study. *Appl Immunohistochem Mol Morphol*. 2014;22(6):433–41.
21. Maity PP, Chatterjee S, Das RK, Mukhopadhyay S, Maity A, Maulik D, et al. Finding an optimum immuno-histochemical feature set to distinguish benign phyllodes from fibroadenoma. *Micron*. 2013;48:34–41.
22. samr: SAM: significance analysis of microarrays. R package version 2.0. <http://CRAN.R-project.org/package=samr>. Accessed 13 October 2014.
23. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinforma*. 2012;13:134.
24. Kalin M, Cima I, Schiess R, Fankhauser N, Powles T, Wild P, et al. Novel prognostic markers in the serum of patients with castration-resistant prostate cancer derived from quantitative analysis of the pten conditional knockout mouse proteome. *Eur Urol*. 2011;60(6):1235–43.
25. Cima I, Schiess R, Wild P, Kaelin M, Schuffler P, Lange V, et al. Cancer genetics-guided discovery of serum biomarker signatures for diagnosis and prognosis of prostate cancer. *Proc Natl Acad Sci U S A*. 2011;108(8):3342–7.
26. Liaw A, Wiener M. Classification and Regression by randomForest. *R News*. 2002;2(3):18–22.
27. glmulti: Model selection and multimodel inference made easy. R package version 1.0.7. <http://CRAN.R-project.org/package=glmulti>. Accessed 13 October 2014.
28. Calcagno V, de Mazancourt C. glmulti: an R package for easy automated model selection with (generalized) linear models. *J Stat Softw*. 2010;34(12):1–29.
29. Huang KT, Dobrovic A, Yan M, Karim RZ, Lee CS, Lakhani SR, et al. DNA methylation profiling of phyllodes and fibroadenoma tumours of the breast. *Breast Cancer Res Treat*. 2010;124(2):555–65.
30. Kuijper A. Pathogenesis and progression of fibroepithelial breast tumors. The Netherlands: Utrecht University; 2006.
31. Vidal M, Peg V, Galvan P, Tres A, Cortes J, Ramon y Cajal S, et al. Gene expression-based classifications of fibroadenomas and phyllodes tumours of the breast. *Mol Oncol*. 2015;9(6):1081–90.
32. Tan BY, Acs G, Apple SK, Badve S, Bleiweiss IJ, Brogi E, et al. Phyllodes tumours of the breast: a consensus review. *Histopathology*. 2016;68(1):5–21.
33. Spitaleri G, Toesca A, Botteri E, Bottiglieri L, Rotmensz N, Boselli S, et al. Breast phyllodes tumor: a review of literature and a single center retrospective series analysis. *Crit Rev Oncol Hematol*. 2013;88(2):427–36.
34. Lim WK, Ong CK, Tan J, Thike AA, Ng CC, Rajasegaran V, et al. Exome sequencing identifies highly recurrent MED12 somatic mutations in breast fibroadenoma. *Nat Genet*. 2014;46(8):877–80.
35. Tan J, Ong CK, Lim WK, Ng CC, Thike AA, Ng LM, et al. Genomic landscapes of breast fibroepithelial tumors. *Nat Genet*. 2015;47(11):1341–5.
36. Jacobs TW, Chen YY, Guinee DG, Eby PR, Thike AA, Vohra P, Tan PH. Phyllodes tumor subsequent to a diagnosis of fibroadenoma on breast core needle biopsy: frequency and characteristics. In: USCAP 103rd annual meeting 2014: March 1–7 2014. San Diego, CA, USA: Modern Pathology; 2014. p. 56A.
37. Waalkes S, Atschekzei F, Kramer MW, Hennenlotter J, Vetter G, Becker JU, et al. Fibronectin 1 mRNA expression correlates with advanced disease in renal cancer. *BMC Cancer*. 2010;10:503.
38. Brenne K, Nymoen DA, Reich R, Davidson B. PRAME (preferentially expressed antigen of melanoma) is a novel marker for differentiating serous carcinoma from malignant mesothelioma. *Am J Clin Pathol*. 2012;137(2):240–7.
39. Xu K, Cui J, Olman V, Yang Q, Puett D, Xu Y. A comparative analysis of gene-expression data of multiple cancer types. *PLoS One*. 2010;5(10), e13696.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

