BIOLOGICAL MARKERS OF PROGNOSTIC IMPORTANCE IN BREAST PHYLLODES TUMOURS

TAN WAI JIN
B.Sc. (Hons), NUS

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARMENT OF ANATOMY
YONG LOO LIN SCHOOL OF MEDICINE
NATIONAL UNIVERSITY OF SINGAPORE

2015

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Tan Wai Jin

7th January 2015

ACKNOWLEDGEMENTS

This thesis will not be possible without the guidance, support and help from numerous people during the last four years.

First and foremost, I would like to thank my supervisor **Professor Tan Puay Hoon** for her continuous support and guidance throughout my undergraduate and graduate studies, providing me invaluable insights into pathology and translational research, and giving me countless opportunities to learn and improve. It has truly been a great privilege and honour to work under her mentorship.

I would also like to express my gratitude to my co-supervisor **Professor Bay Boon Huat** for his unwavering support and timely advice, which has helped me through critical stages in my studies. I sincerely thank him for all the guidance and valuable comments given.

I am grateful to **Dr Min-Han Tan** for his invaluable suggestions on many aspects of the projects, and to carry out some of the work in his laboratory in the Institute of Bioengineering and Nanotechnology, A*STAR. Many thanks also to his laboratory members **Dr Johnathan Lai**, **Dr Igor Cima**, **Dr Yukti Choudhury,** and **Dr Jamie Mong,** for generous assistance and guidance given in experiments setups and data analyses.

A special thanks to **Professor Teh Bin Tean** from the National Cancer Centre and his laboratory members for their constructive comments and suggestions for the molecular aspect of the project.

I would also like to thank the doctors from the breast research team in the Department of Pathology, Singapore General Hospital - **Dr Aye Aye Thike** for her expertise in pathology and kind assistance in immunoscoring and analysis, **Dr Jabed Iqbal** and **Dr Syed Salahuddin Ahmed** for sharing their clinical expertise in breast pathology. My heartfelt gratitude also goes to all staff of Department of Pathology, Singapore General Hospital for all the help and assistance rendered, with a special mention for **Associate Professor Alvin Lim** and **Ms Lim Tse Hui** for the FISH work.

The journey would not be as enriching and enjoyable without my current and ex-labmates/friends, for sharing many ups and downs together in troubleshooting experiments, and talking about everything under the sun. **Dr Yvonne Teng,** who taught me about molecular pathology and shared with me about life of a graduate student when I first joined. **Ms Cheok Poh Yian** who has taught me many aspects of the histotechnology work and generously

shared her experience in tissue microarrays and immunohistochemistry. **Mr Sai Sakktee Krisna**, for all the jokes and laughter, and for sharing the first few molecular pathology and technical experiences together. **Mr Jeffrey Lim,** for always taking care of the laboratory with his exceptional organizational skills, and for all the assistance given. **Ms Valerie Koh**, for her countless assistance rendered in time of needs and for sharing lots of stories outside of lab together. **Ms Jane Tan,** for her excellent technical skills and expertise on histology and her help on the special stain work. **Mr Ho Soo Keng, Mr Luke Chong,** and **Dr Emarene Mationg Kalaw**, for the enjoyable experience working together. Thanks for all the great memories we had together.

Many thanks also to colleagues and friends from Department of Anatomy, Yong Loo Lin School of Medicine, for sharing knowledge together and helping me in many instances, especially **Dr Ng Cheng Teng** and **Dr Chua Pei Jou**.

I am grateful to the **National University of Singapore** for the NUS Research Scholarship awarded, without which this journey will not be made possible.

On a personal note, I would like to thank my wonderful housemates and friends for the companionship and for taking care of me in times of need. **Vin Yee**, for sharing lab jokes and silly stuff together. **Huey Fen, Ee Ling** and **Qian Lyn** for always hearing me out. **Jessilyn**, for dragging me out for breaks and reminding me to eat.

Lastly, I would like to dedicate this work to my loved ones. My parents – **Tan Yit Huat** and **Ong Wee Chu**, for showering me with endless love, care, and motivation in bringing me this far. My sisters – **Wai Pei** and **Wai Ann**, for the love and joy of siblings. And **Yean Chert**, for the support, patience and love all these years.

TABLE OF CONTENTS

SUMMARY

Phyllodes tumours are uncommon fibroepithelial neoplasms of the breast accounting for less than 1% of primary breast tumours, comprising 6.9% of breast cancers reported in Singapore. They are classified into benign, borderline and malignant tumours by WHO, based on five histological parameters – stromal cellularity, stromal atypia, stromal mitoses, presence of stromal overgrowth and nature of the tumour border. This diagnostic framework remains challenging due to interpretive subjectivity and variable combinational permutations of histological criteria, leading to difficulties in accurate prediction of biological behaviour.

The first section of this work sought to address the combinational permutations of the parameters and to improve prediction of clinical outcome for phyllodes tumour patients. A total of 605 cases diagnosed in the Department of Pathology, Singapore General Hospital from January 1992 to December 2010 were examined retrospectively. Multivariate analysis showed that stromal atypia, stromal mitosis, stromal overgrowth and surgical margins were significant independent predictors for recurrence-free survival. A predictive nomogram which was constructed based on these criteria can be potentially used for patient counseling and clinical management of women diagnosed with phyllodes tumours.

Subsequently, 20 cases stratified into prognostically distinct categories were subjected to Affymeterix OncoScan™ FFPE microarray analysis of genome-wide copy number changes and mutational status. It was

found that cases which recurred/metastasized had higher chromosomal aberrations, with median events of 19 compared to 3.5 in non-recurrent cases. Also, high-level amplifications and homozygous deletions were detected exclusively in the former group. Regions of high-level amplifications included chromosome 1q32.1, 5p13.3, 3p25, and 7p12. Amplification of *EGFR* at chromosome 7p12 was confirmed on FISH and accompanied by intense EGFR immunostaining. Regions of homozygous deletion included *CDKN2A* (9p21) and *MACROD2* (20p12.1). Loss of protein expression was observed in cases with homozygous deletion at chromosome 9p21 involving *CDKN2A.* No mutations were identified in these cases.

The *KIT* gene (also known as CD117) was selected for further investigation. Activating mutations of the *KIT* gene was not observed and mutational status was not correlated with protein expression. However, CD117 immunoexpression was observed in 10% of 272 cases on tissue microarrays, with significant associations with borderline/malignant tumours and unfavourable pathological parameters including larger tumour size, increased number of mitoses, and permeative tumour margins. Also, patients with CD117-positive tumours had a poorer prognosis compared to patients with CD117-negative tumours.

Prognostic importance of Six1 and Pax3 was evaluated based on protein expression by immunohistochemistry on tissue microarrays to validate previous findings of increased expression observed at the transcript level in borderline/malignant phyllodes tumours. Six1 and Pax3 stromal

cytoplasmic expression was observed to be associated with higher tumour grade and adverse histological parameters. Patients having tumours with high Six1 expression experienced a shorter time to recurrence while patients having tumours with high Pax3 expression had a worse overall survival.

Overall, these findings provide an improved understanding in the biology of phyllodes tumours and warrant further investigations to elucidate mechanisms and pathways that are exploited by phyllodes tumours for malignant progression and aggressive behaviour.

LIST OF PUBLICATIONS

Journals

1. **Tan WJ**, Lai JC, Thike AA, Lim JCT, Tan SY, Koh VCY, Lim TH, Bay BH, Tan MH, Tan PH (2014). Novel genetic aberrations in breast phyllodes tumours: comparison between prognostically distinct groups. Breast Cancer Res Treat *145*, 635-645. *Impact factor: 4.198*

2. **Tan WJ**, Thike AA, Bay BH, Tan PH (2014). Immunohistochemical expression of homeoproteins Six1 and Pax3 in breast phyllodes tumours correlates with histological grade and clinical outcome. Histopathology *64*, 807-817. *Impact factor: 3.301*

3. **Tan WJ**, Thike AA, Tan SY, Tse GM, Tan MH, Bay BH, Tan PH (2015). CD117 expression in breast phyllodes tumours correlates with adverse pathologic parameters and reduced survival. Mod Pathol *28*, 352-358. *Impact factor: 6.364*

4. Tan PH, Thike AA, **Tan WJ**, Thu MMM, Busmanis I, Li H, Chay WY, Tan MH, Phyllodes Tumour Network Singapore (2012). Predicting clinical behaviour of breast phyllodes tumours: a nomogram based on histological criteria and surgical margins. J Clin Pathol *65,* 69-76. *Impact factor: 2.551*

The following articles are not included in the main body of the thesis:

5. Chong LYZ, Cheok PY, **Tan WJ**, Thike AA, Allen G, Ang MK, Ooi AS, Tan P, Teh BT, Tan PH (2012). Keratin 15, transcobalamin I and homeobox gene Hox-B13 expression in breast phyllodes tumors: novel markers in biological classification. Breast Cancer Res Treat *132*, 143-151. *Impact factor: 4.198*

6. Teng YHF, **Tan WJ**, Thike AA, Cheok PY, Tse GMK, Wong NS, Yip GWC, Bay BH, Tan PH (2011). Mutations in the epidermal growth factor receptor

(EGFR) gene in triple negative breast cancer: possible implications for targeted therapy. Breast Cancer Res *13*, R35. *Impact factor: 5.881*

Conference Proceedings

1. **Tan WJ**, Thike AA, Bay BH, Tan MH, Tan PH (2013). High throughput genomic screening reveals p16 loss in paraffin-embedded archival breast phyllodes tumours. Annual Scientific Congress of the Malaysian Oncological Society. Kuala Lumpur, Malaysia. ***Best Poster Prize***

2. **Tan WJ**, Thike AA, Lai JC, Tan MH, Bay BH, Tan PH (2013). Protein Expression and Mutation Status of C-kit in Phyllodes Tumours of the Breast. Proceedings of the Microscopy Society (Singapore) Annual General and Scientific Meeting, Singapore.

3. **Tan WJ**, Thike AA, Tan PH (2013). Clinicopathological Significance of Six1 Expression in Breast Phyllodes Tumours in SGH 20th Annual Scientific Meeting, Singapore.

4. **Tan WJ**, Thike AA, Tan PH (2012). CD117 expression in Breast Phyllodes Tumour: Correlation with Histological Parameters in Singhealth Duke-NUS Scientific Congress, Singapore.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	Akaike information criterion
BLAST	Basic local alignment search tool
CDK18	Cyclin-dependent kinase 18
CDKN2A	Cyclin-dependent kinase inhibitor 2A
CGH	Comparative genomic hybridization
Chr	Chromosome
CI	Confidence interval
CN	Copy number
Ct	Citrate buffer (pH6.0)
DAB	3, 3'-diaminobenzidine
DAPI	4', 6-diamidino-2-phenylindole
DFS	Disease free survival
DNA	Deoxyribonucleic acid
DR	Distant recurrences
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
ER	Estrogen receptor
ER2	Epitope retrieval solution 2 (pH9.0)
FA	Fibroadenoma
FFPE	Formalin-fixed, paraffin embedded
FGA	Fraction of genome altered
GISTs	Gastrointestinal stromal tumours
H&E	Haematoxylin and eosin-stained
HPF	High power field
HRP	Horse radish peroxidase
IHC	Immunohistochemistry
LOH	Loss of heterozygosity
LR	Local recurrences
MAPD	Median absolute pairwise difference
mRNA	Messenger ribonucleic acid
NCBI	National Center for Biotechnology Information
NCCN	National Comprehensive Cancer Network
OS	Overall survival
PCR	Polymerase chain reaction
PT	Phyllodes tumour
PR	Progesterone receptor
RFS	Recurrence-free survival
RT	Room temperature
SPF	S-phase fraction
SSC	Saline-sodium citrate
TE	Tris-EDTA buffer (pH9.0)
TMAs	Tissue microarrays

1.0 INTRODUCTION

1.1 Breast phyllodes tumours

Breast phyllodes tumours represent an unusual form of breast neoplasm. Sketchy reports of phyllodes tumours were documented in the literature as early as 1774 [Fiks, 1981] but the tumour was formally characterized only in 1838 by Johannes Müller (see Figure 1. for timeline). Johannes Müller, a prominent German anatomist and physiologist, described the tumour as 'Cystosarcoma phyllodes' in his book "*Über den feinern Bau und die Formen der krankhaften Geschwülste*" (On the Structural Details of Pathological Tumours). The gross description is worthy of quote in detail [Lee and Pack, 1931]:

> "The tumour forms a large firm mass, with a more or less uneven surface. The fibrous substance which constitutes the greater part of it is of a greyish white colour, extremely hard, and as firm as fibro-cartilage. Large portions of the tumour are made up entirely of this mass, but in some parts are cavities or clefts not lined with a distinct membrane. These cavities contain but little fluid; for either their parietes, which are hard like fibro-cartilage, and finely polished, lie in close apposition with each other, or a number of firm, irregular laminae sprout from the mass, and form the walls of the fissures; or excrescences of a foliated or wart-like form sprout from the bottom of the cavities and fill up their interior. These excrescences are perfectly smooth on their surface, and never contain cysts or cells. The laminae

> lie very irregularly, and project into the cavities and fissures like the folds of the psalterium in the interior of the third stomach of ruminant animals. In one instance the author saw these laminae here and there regularly notched or crenated like a cock's comb. Sometimes the laminae are but small, and the warty excrescences from the cysts very large, while in other instances both are greatly developed. Occasionally these warty excrescences are broad, sessile, and much indented; others have a more slender base, and somewhat resemble cauliflower condylomata."

Since the first detailed gross description by Müller, numerous terms were used to describe the same tumour such as intracanalicular fibroma and pseudosarcoma. However, the termed 'cystosarcoma' was gradually avoided as most of these tumours do not possess sarcomatous stroma. It was only in 1960 that a simpler term was coined by Lomonaco - 'tumour phyllodes'. 'Phyllodes tumour' has been the preferred term since 1982, as recommended in the second edition of the WHO Histological Classification of Breast Tumours and is now classified under the umbrella of fibroepithelial tumours in the latest edition.

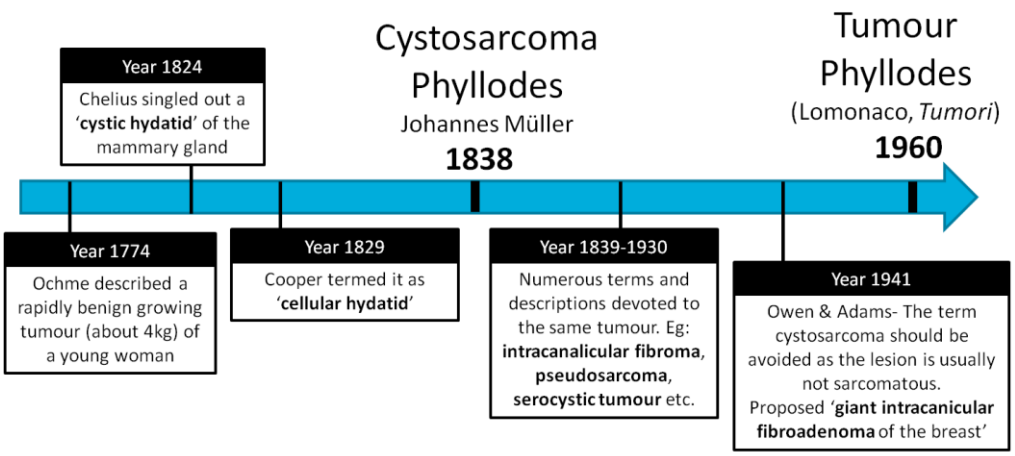

Figure 1.1 Timeline of the discovery and development of phyllodes tumours

1.1.1 Epidemiology

Phyllodes tumour is an uncommon neoplasm, constituting less than 1% of primary breast tumours in the United States and the incidence was reported to be different among various ethnic groups, with higher rates occurring in Asians/Pacific Islanders and Hispanics [Spitaleri et al., 2013]. However, the clinical outcome was not significantly different in Hispanic patients despite having larger tumours and higher mitotic rates as compared to Caucasian and Black patients [Pimiento et al., 2011]. In Singapore, the incidence of phyllodes tumour is reported to be 6.9% when compared to the frequency of breast cancers, possibly reflecting a true higher occurrence among Asian women [Tan et al., 2005a]. However, no significant incidence difference was observed between local ethnicities encompassing Chinese, Malay, Indian, and other groups [Tan et al., 2005a].

1.1.2 Clinical presentation

Phyllodes tumours present clinically as firm, painless palpable masses. In the 1930s Lee and Pack described phyllodes tumour as a rapidly growing mass which had a long initial period of quiescent growth. Rapid growth corresponded to a sudden in size without adhering to the skin [Lee and Pack, 1931]. Treves later emphasized that not all phyllodes tumours manifested enormous sizes or rapidity of growth. Nonetheless, the skin over large tumours may have a bluish discoloration [Pietruszka and Barnes, 1978], with some very rare occurrence of nipple retraction, cutaneous ulceration, fixation of skin, and nipple discharge [Norris and Taylor, 1967].

The average size of phyllodes tumours diagnosed is becoming smaller with the advent of imaging techniques such as mammography and ultrasound. Mammographically, phyllodes tumours typically appear as lobulated and benign opacities, with rare occasions of calcification [Stebbing and Nash, 1995]. Ultrasound findings usually include smooth contours, heterogenous pattern of internal echoes, absence of posterior shadowing and intramural cyst [Chao et al., 2000].

1.1.3 Clinical management

1.1.3.1 Surgery

Surgery is the commonest form of treatment for patients with phyllodes tumours and has remained the standard treatment to date. While most have moved away from mastectomy to breast conserving surgery, there is no clear consensus concerning the type of surgery to be performed. Some authors reported a better disease-free survival with mastectomy [Belkacémi et al., 2008; Pezner et al., 2008; Chen et al., 2005] while some found no impact of surgery type on recurrence rate [Asoglu et al., 2004; Macdonald et al., 2006]. Nonetheless, establishing a clear surgical margin is important as many reports have shown an association of positive margins with recurrence. [Asoglu et al., 2004; Tan et al., 2005a; Chen et al., 2005; Kapiris et al., 2001; Pandey et al., 2001]

The current NCCN (National Comprehensive Cancer Network) guidelines recommend wide local excision with at least 1cm margin regardless of grade [Carlson et al., 2010]. Sampling of lymph nodes is generally not advised as phyllodes tumours infrequently spread to lymph nodes [Gullett et al., 2009; Guillot et al., 2011]. However, given the suspicious histology and large size at clinical presentation especially in higher grade tumours, phyllodes tumours are often subjected to aggressive treatment including excessive sampling of lymph nodes which usually is a procedure reserved for breast carcinomas [Gullett et al., 2009].

1.1.3.2 Adjuvant therapy

Given the low incidence rate of phyllodes tumours, the effectiveness of adjuvant therapy especially for higher grade tumours is unknown. Scant literature of small studies has reported mixed results regarding the benefits of adjuvant therapy.

a) Radiotherapy

Cohn-Cedermark *et al* reported no improvement in survival and a few other authors were reserved in recommending of adjuvant radiotherapy [Chen et al., 2005; Chaney et al., 2000]. On the contrary, a few reports suggested potential benefits of adjuvant radiotherapy. Pandey *et al.* found slight improvement in disease free survival for patients although this was not significant statistically. Belkacémi *et al.* reported an improved local control rate but no impact on survival for borderline and malignant tumours. More recently, a relatively larger prospective study with a median follow up of 56 months concluded that adjuvant radiotherapy is a very effective treatment as none of the 46 patients developed local recurrence during the study period. A drawback of the study was that no control group was included [Barth et al., 2009].

b) Chemotherapy

Adjuvant chemotherapy is not well defined in the literature. A systematic trial comparing patients with and without chemotherapy revealed no significant differences in recurrence-free survival. The regimen was selected based on agents used for sarcomas - doxorubicin and dacarbazine

[Morales-Vásquez et al., 2007]. Other case reports showed efficacy in some women employing cisplatin, etoposide and ifosfamide-based chemotherapy [Hawkins et al., 1992; Burton et al., 1989; Turalba et al., 1986].

c) Endocrine therapy

Despite no proven role of endocrine therapy for phyllodes tumours in the literature, expression of estrogen and progesterone receptors has been previously evaluated in large series of phyllodes tumours in a retrospective manner. These hormonal receptors are generally positive in the epithelial component but rarely expressed in the stromal component [Tse et al., 2002; Sapino et al., 2006]. The limited potential in employing endocrine therapy for phyllodes tumours is largely due to the negative expression found in stromal cells, the principal neoplastic cell population of phyllodes tumours. Nonetheless, there has been documentation of positive ERβ (one of the two types of estrogen receptor apart from ERα) being expressed in phyllodes tumour stroma [Tse et al., 2002; Sapino et al., 2006]. However, the therapeutic potential of ERβ is still at a preliminary stage of research [Gallo et al., 2012].

1.1.4 Histopathology of breast phyllodes tumours

Histologically, it is characterized by a double-layered epithelial component arranged in cleft-like structures surrounded by a hypercellular spindle-cell stromal component (Figure 1.2). The combination of both components mimics a leafy architecture, which gives rise to its name phyllodes (*fýllo* in Greek means leaf).

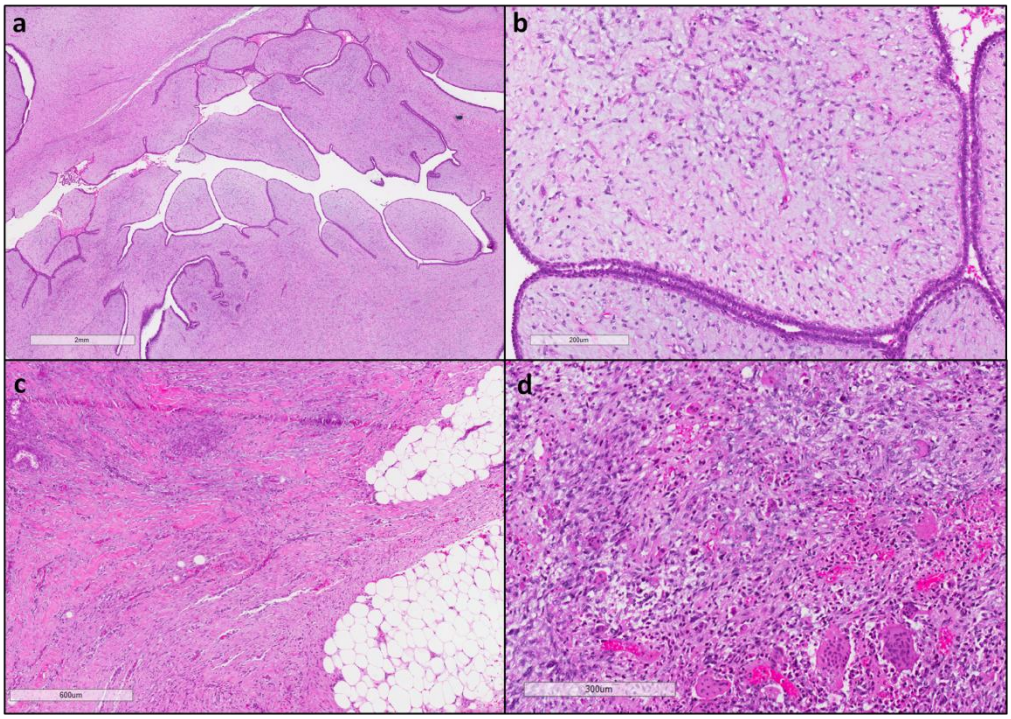

Figure 1.2 Histology of a phyllodes tumour. (a) Low magnification of a benign phyllodes tumour illustrating the cleft-like formation of the epithelial lining. Scale bar represents 2mm. (b) High magnification of the tumour showing hypercellular spindle-cell stroma. Scale bar represents 200µm. (c) An example of a borderline phyllodes tumour with a permeative border. Scale bar: 600µm. (d) An example of a malignant phyllodes tumour. Scale bar: 300µm.

Phyllodes tumours were noted as a perfectly benign disease by Johannes Müller in 1838. However, this was challenged subsequently by other researchers who reported recurrences and metastasis upon follow-up of patients with phyllodes tumours [Lee and Pack, 1931; Cooper and Ackerman, 1943]. White in 1940 proposed the existence of a malignant variant of the tumour after noting a patient who died from the disease with recurrences and metastasis prior to death, 19 months after she was first diagnosed with the tumour [White, 1940]. The malignant variant of phyllodes tumours was well recognized since then. In 1951, when Treves and Sunderland were not able to fit 18 of the 77 phyllodes tumours in their series

into either benign or malignant categories, and proposed a 'borderline' category in addition to benign and malignant categories [Treves and Sunderland, 1951].

Over the years, several grading systems have been suggested based on histological assessment. The criteria used for assessment were vastly similar although there was variation and subjectivity in terms of interpretation. Most supported a three-tiered grading scheme (commonly termed as benign, borderline, malignant proposed by Treves and Sunderland) while some advocated a two-tiered grading method (low grade and high grade) [Karim et al., 2009]. The development of molecular techniques has also driven researchers to better classify phyllodes tumours in recent years. Wang *et al.* supported a two-tiered grading (low/intermediate and high grade), observing no differences between the benign and borderline tumours in LOH (loss of heterozygosity) studies [Wang et al., 2006]. Laé *et al.* concurred with a two tiered grading system using CGH (comparative genomic hybridization) studies but the authors grouped borderline and malignant tumours as one category [Laé et al., 2007]. Jones *et al.* who performed array-CGH also agreed with a two-tiered grading based on cluster analysis of array-CGH data [Jones et al., 2008a]. However, more recently Ang *et al.* supported the three tiered grading based on expression profiles of 21 phyllodes tumours [Ang et al., 2011]. Amidst all these molecular classifications is a lack of consensus and hence histological assessment remains as the gold standard in grading phyllodes tumours.

In the 4th edition of the WHO classification of tumours of the breast, phyllodes tumours are classified into benign, borderline, and malignant groups. The three-tiered grading is preferred because it gives a greater certainty at the ends of the spectrums of these tumours [Tan et al., 2012]. Grading of the tumours is based on five histological parameters – stromal cellularity, stromal atypia, stromal mitosis, presence of stromal overgrowth and nature of the tumour border which are described in detail in the next page.

The stromal component is primarily assessed as it is that which undergoes malignant progression and drives clinical behaviour. Nonetheless, emerging evidence in the literature has shown that the epithelial component may play a role in disease progression by interacting with the stroma (Figure 1.3). Its innocence as a 'bystander' has been challenged with reports of abnormality observed in the epithelial component such as harbouring an independent set of genetic alterations apart from the stroma [Sawyer et al., 2000] and an increased frequency of hyperplasia and metaplasia compared to normal breast tissue [Karim et al., 2009b]. Besides, the subepithelial stromal region demonstrated different changes as compared to stroma further away from the epithelium. A higher mitotic activity was observed [Treves and Sunderland, 1951; Sawhney et al., 1992] and additional genetic changes were noted in peri-epithelial stromal cells [Jones et al., 2008a], suggesting interactions between epithelium and stroma.

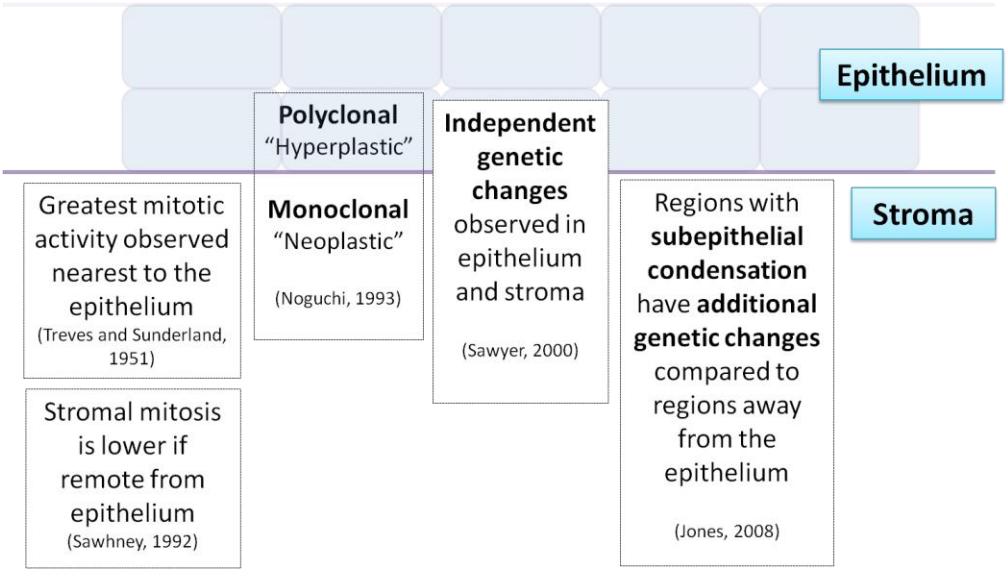

Figure 1.3 Role of epithelium in phyllodes tumours

The five histological parameters used to grade phyllodes tumours are:

1) Stromal cellularity (three tiered) - mild, moderate and marked cellularity

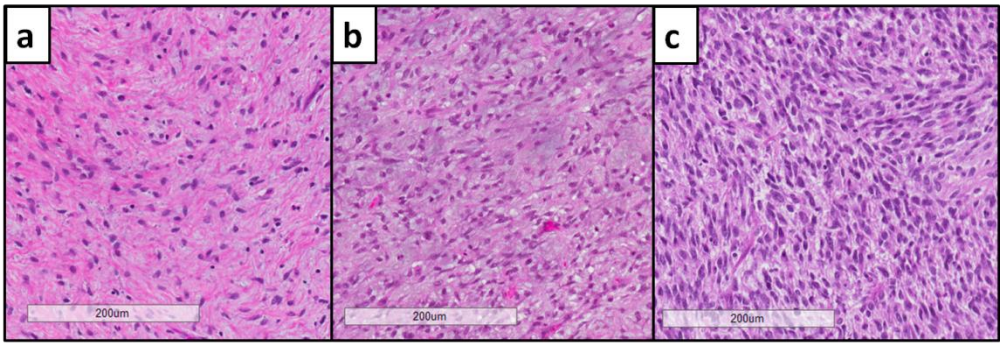

Figure 1.4 Example of (a) mild cellularity (b) moderate cellularity and (c) marked cellularity of the stromal component of phyllodes tumours. All three images were taken at the same magnification power. Scale bars: 200µm

2) Stromal atypia (three tiered) - none/mild, moderate, marked

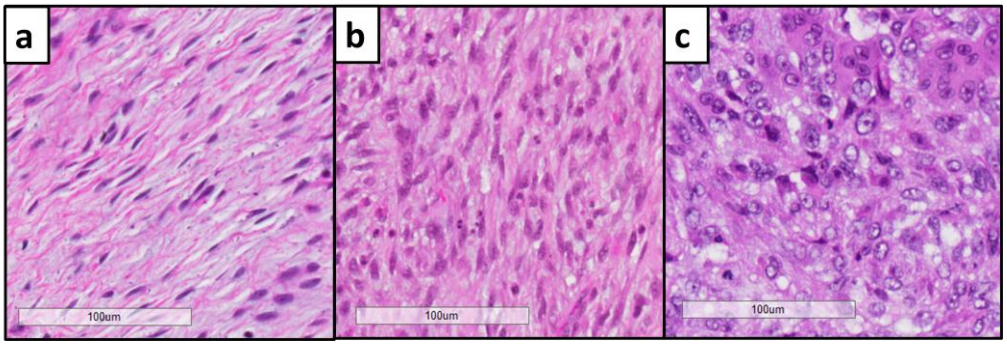

Figure 1.5 Example of (a) mild stromal atypia (b) moderate stromal atypia and (c) marked atypia of phyllodes tumours. Scale bars: 100µm

3) Stromal mitosis – assessing number of mitoses in ten high power fields (x40 objective and x10 eyepiece, 0.196mm^2)

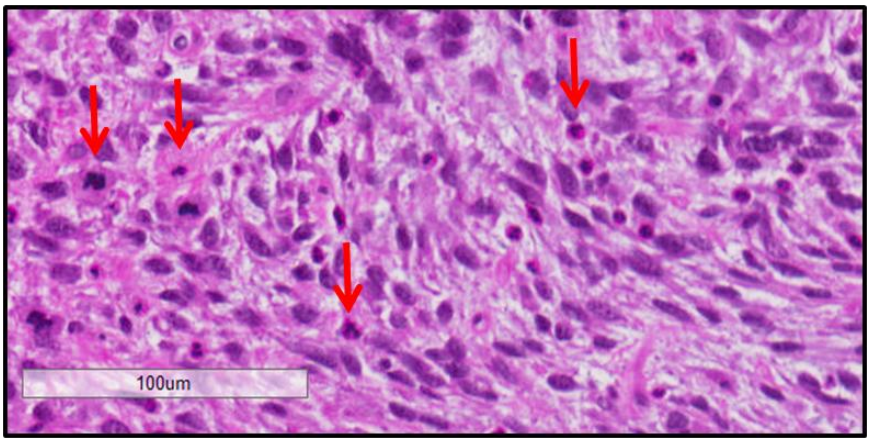

Figure 1.6 Example of mitotic activity (indicated by red arrows). There are more than 10 mitoses per one high power field (HPF) in this tumour sample. Scale bar: 100µm

4) Presence of stromal overgrowth – stromal proliferation to the degree where the epithelial component is absent in at least one low power field (x4 objective and x10 eyepiece, 22.9mm^2)

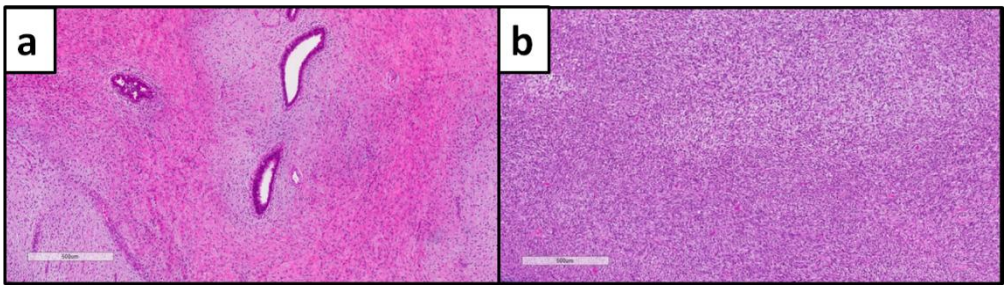

Figure 1.7 Presence of epithelial element in a low power field defines no stromal overgrowth (a) while absence of epithelial component defines stromal overgrowth (b). Scale bars: 100µm

5) Nature of the microscopic border, whether it is well circumscribed or permeative

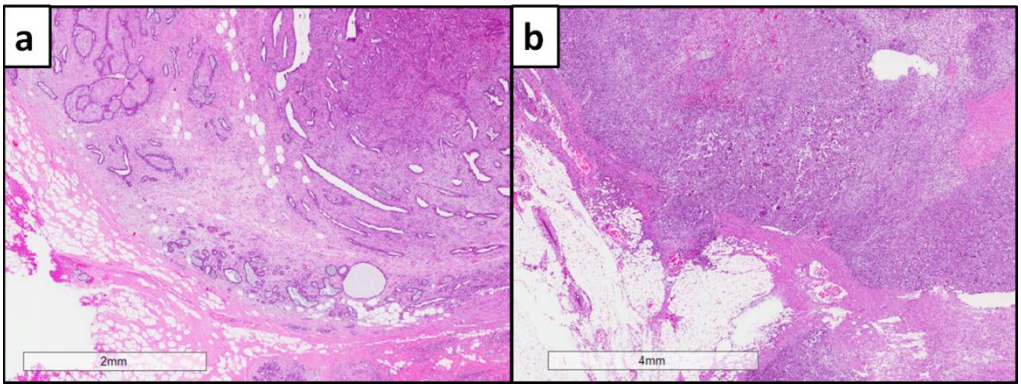

Figure 1.8 (a) An example of a benign phyllodes tumour illustrates a well-circumscribed tumour border. Scale bar: 2mm. (b) An example of a permeative border where stromal cells creep into the adjacent fat tissue. Scale bar: 4mm

A phyllodes tumour is diagnosed as benign when the following criteria were met: mild stromal cellularity, mild or no stromal atypia, occasional mitoses of not more than four per 10 high power field (HPF), absence of stromal overgrowth, and well-circumscribed margins. Conversely, a malignant phyllodes tumour is diagnosed with the combination of marked stromal cellularity and atypia, high mitotic activity of usually 10 or more per 10 HPF, presence of stromal overgrowth and permeative margins. Malignant phyllodes tumour is also diagnosed with the finding of a malignant heterologous element even without the presence of other histological features. A borderline phyllodes tumour shows some, but not all, of the characteristics observed in malignant lesions. See Table 1.1 for summarized features.

Table 1.1 Histological features of benign, borderline and malignant phyllodes tumours

Histological Feature	Benign	Borderline	Malignant*
Stromal Cellularity	Usually mild, may be non-uniform or diffuse	Usually moderate, may be non-uniform or diffuse	Usually marked and diffuse
Stromal Atypia	Mild or none	Mild or moderate	Marked
Mitotic Activity	Usually few (<5per 10HPF)	Usually frequent (5-9 per 10HPF)	Usually abundant (≥10 per 10HPF)
Stromal Overgrowth	Absent	Absent, or very focal	Present
Tumour Border	Well-circumscribed	Well-circumscribed, may be focally permeative	Permeative

*Presence of malignant heterologous element qualifies designation as a malignant phyllodes tumour, without the presence of other histological features

While the parameters for grading of phyllodes tumours appear well defined and clear, there is suboptimal standardization of assessing these parameters resulting in inherent observer variability. Also, there are variable ways of combining the parameters to derive the grade of phyllodes tumours. It is not determined which parameter has higher importance over another, and not all criteria could be fulfilled in assigning a grade for phyllodes tumours.

1.1.5 Molecular studies on breast phyllodes tumours

Molecular studies have been carried out to elucidate the underlying biological mechanisms of progression and behaviour of this unpredictable tumour, including copy number assessment, expression profiling, and screening for mutations. Copy number assessment is currently the most studied in phyllodes tumours while expression profiling and mutational analyses are relatively fewer as indicated in the literature review shown in Table 1.2.

A few key areas investigated previously were:

1. classification of phyllodes tumours based on molecular differences **between tumour grades**;
2. identifying distinct changes **between epithelium and stroma** of phyllodes tumours which are biphasic in nature; and
3. discriminating **phyllodes tumours from fibroadenomas** which may be similar morphologically but are divergent in clinical behaviour.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours

Author	No. of cases	Type of materials	Type of Study	Key findings
Noguchi *et al.* (Cancer Res, 1993)	5 PTs; 10 FAs	Fresh frozen with stromal and epithelial components microdissected	Clonal analysis using *PGK* gene polymorphism	1) **PT vs FA**: PT consisted of both monoclonal and polyclonal components. FAs were polyclonal in origin. 2) **Stroma vs epithelium**: Stromal cells were monoclonal. Epithelial cells were polyclonal.
Noguchi *et al.* (Cancer, 1995)	11 PTs; 3 matched FAs	FFPE	Clonal analysis using *AR* gene polymorphism	1) **FA vs recurrent PT developed on the same site**: All were monoclonal in origin. The same allele of *AR* was inactivated in FA and PT. 2) Suggested that **monoclonal FAs** which could **progress to PT** should be differentiated from polyclonal FAs for management.
Kuijper *et al.* (J Pathol, 2002)	12 PTS; 25 FAs	FFPE with stromal and epithelial component microdissected	Clonal analysis using *AR* gene polymorphism	1) **PT:** Mostly, stroma was monoclonal and epithelium polyclonal. However, two cases were found to have monoclonal epithelium while three other cases of polyclonal stroma. 2) **FA**: Normal-appearing and hyperplastic epithelium of FA were polyclonal but carcinoma *in situ* from FA was monoclonal. Stroma of FA was polyclonal too.
Dietrich *et al.* (Human Pathol, 1997)	6	Short term cultures from excised tumours	Karyotyping	**Benign vs malignant PT:** Benign PTs had simple chromosomal changes. Malignant PT had complex karyotype.
Ladesich *et al.* (Cancer Genet Cytogenet, 2002)	1 low grade PT	Short term culture	Karyotyping	Complex karyotype can be found in low-grade PT and is not necessarily a sign of extreme malignancy.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours (continued)

Author	No. of cases	Type of materials	Type of Study	Key findings
Barbosa *et al.* (Cancer Genet Cytogenet, 2004)	1 PT; 3 FAs	Short term culture	Karyotyping	**PT vs FA**: i. PT was hyperdiploid. FA showed diploid or near-diploid modal chromosome number. ii. FA presented with mostly abnormalities on chr 16, 18 and 21. PT had numerous additional abnormalities.
Lu *et al.* (Genes, Chr & Cancer, 1997)	19	Fresh frozen	CGH	1) Gain of **1q** associated with stromal overgrowth. 2) Gain of 1q or loss of X chromosome associated with recurrence. 3) No evidence for gene amplification.
Sawyer *et al.* (Am J Pathol, 2000)	47	FFPE with microdissection of epithelial, stromal and adjacent normal components.	Allelic imbalance (AI) analysis using microsatellites on **chr 1q** and **3p**	1) 24% and 30% of PTs showed AI on chr 3p and chr 1 respectively. 2) **Stroma vs epithelium**: i. AI observed in both epithelium and stroma, sometimes as an independent genetic event. ii. 6% of PT showed microsatellite instability (MSI) in epithelium but not stroma 3) No association was observed between epithelial hyperplasia/atypia with epithelial AI. 4) Mapping chr 3p to *FHIT* (tumour suppressor) gene but no further experiment was done.
Jee *et al.* (Anal Cell Path,2003)	22	FFPE	CGH; LOH on chr **3p** only	1) Most common abnormality was **1q gain**. 2) No allelic imbalance found for LOH analysis. 3) No high level of amplifications seen.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours (continued)

Author	No. of cases	Type of materials	Type of Study	Key findings
Wang *et al.* (Breast Cancer Res Treat, 2006)	15 PTs; 11 FAs	FFPE stromal microdissection; 2 samples with epithelial microdissection - one benign epithelial and the other a carcinoma	LOH analysis	1) Supported **two**-tiered (low/intermediate and malignant) grading system : i. Allelic loss was significantly higher in high-grade PT vs low- and intermediate-grade PT (19.2% vs 4.7% and 4.7% respectively). ii. Loss of **7p12** was more frequently observed in high-grade PT while LOH at 3p24 was more frequently observed in low-/intermediate-grade PTs. 2) **Primary vs matched recurrent**: shared similar LOH with additional regions of LOH observed in recurrent cases. 3) **Epithelial vs stromal**: Epithelial component had lower allelic loss. Shared LOH regions include 4p15, 13p12, 13q32 and 17p. Stromal had additional LOH on chr 9 while epithelial had LOH on 17q, 18 and 22. 4) **PT vs FA**: FA had either lacked LOH or very low allelic loss rate (average 0.4%) vs PT (average 9.4%). 5) Mapping **7p12** (more observed in high-grade) to *HUS1*, a checkpoint homolog involved in DNA repair. No further experiment was carried out.
Laé *et al.* (Mod Pathology 2007)	30	Fresh frozen	CGH and FISH	1) Genetically segregate PT into **two** groups (benign, and borderline/malignant) 2) **1q gain** and **13q loss** were significantly associated with the borderline/malignant group. 3) *MDM2* (8-10 copies) and *MYC* (6 copies) were amplified in one case each.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours (continued)

Author	No. of cases	Type of materials	Type of Study	Key findings
Lv *et al.* (Breast Cancer Res Treat, 2008)	36	FFPE stromal microdissection	CGH	1) Genetically segregate PT into **two** groups (benign, and borderline/malignant) 2) Benign PTs have increased number of chromosomal **gains** especially **4q12** 3) Mapping **4q12** to *KIT, VEGFR* and *AFP* (alpha-fetoprotein) but no further experiment was carried out.
Jones *et al.* (J Pathol, 2008)	126 PTs; 3 FAs	Fresh frozen and FFPE	Array CGH/ Illumina Goldengate assay, mRNA expression profiling, mutational analysis and methylation analysis	1) Supported **two**-tiered grading system (benign/borderline and malignant) based on cluster analysis of array-CGH data: malignant group was characterized by regions of gains on **chr 7**. 2) Commonest change was +1q. In addition, +5p, +7, +8, -6, -9p (**9p21** involving $p16^{INK4a}$), -10p, -13 were associated with borderline/malignant PTs. 3) Mapping homozygous deletions of 9p21.3 to **$p16^{INK4a}$** : i. **methylation** observed in 2 malignant cases ii. **Mutation** (exons 1-3): A single missense P48L found in a malignant PT iii. **IHC**: loss of p16 protein expression associated with 9p deletion 4) **mRNA expression profiling:** Six gene sets (at 1q25, 2q36, 18q22, 5q23, 5q31 and 7p11) were associated with borderline/malignant group. **1q25** had most significant association with over-expression of genes of that region. 5) **With vs without sub-epithelial stroma condensation**: additional genetic changes were found, including deletion in **17p** and 8p12. 6) **PT vs FA:** FA showed no chromosomal-scale changes. 7) Mapping **17p** to **p53**: mutation not found, but protein expression was found in stroma of only 3/24 borderline/malignant tumours 8) Mapping 1p deletion to ***RBBP4, FABP3*** **and** ***HDAC1***: found no pathogenic mutations.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours (continued)

Author	No. of cases	Type of materials	Type of Study	Key findings
Kuijper *et al.* (Cell Oncol, 2009)	11 PTs; 3 FAs	Fresh frozen	CGH	1) **PT vs FA**: 91% of PTs showed chromosomal changes with mean number of 5.5 per case. No copy number change is detected in FA. 2) Copy number changes not correlated with tumour grade - copy number changes found in all PTs. 3) Loss of **16q** most frequently found. Mapping loss of **16q** to members of cadherin family (E-cadherin, M cadherin) 4) Amplifications (log_2 ratio ≥ 1) were seen infrequently - only on chromosome **5q** and **22q**. Mapping amplification regions (**5q** and **22q**) to *MAP3K1*, *MAPK1*, *PIK3R1*, and *PIK4CA.*
Ang *et al.* (Breast Cancer Res Treat, 2010)	21	Fresh frozen	Expression profiling, comparative genomic microarray analysis, IHC	1) Supported **three**-tiered pathology grading with a 29 gene list: mean chromosomal changes was 2.7, 4.2, and 9 for benign, borderline and malignant respectively. 2) **1q** gain most commonly associated with borderline and malignant tumours. 3) ***HOXB13*** was upregulated in malignant PTs compared with borderline and benign groups (~2.8-fold change)
Lee *et al.* (Cancer Genet Cytogenet, 2010)	1 mal PT	FFPE	Oligo-array CGH	Found gain of 1q, loss of **1p36** and **17q11.2.** Mapping the region of **1p36** to *SDHB* and **17q11.2** to *NF1* gene but no further experiment carried out.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours (continued)

Author	No. of cases	Type of materials	Type of Study	Key findings
Kim *et al.* (Virchows Archiv, 2009)	87	FFPE	Methylation analysis of 5 genes and IHC on *GSTP1*	1) Supported **two**-tiered (benign, borderline/malignant) grading system: mean number of methylated genes were 1.63 for benign, 3.13 and 3.70 respectively for borderline and malignant PT 2) Increasing methylation frequency observed with increasing grade of PTs. (5 genes - *RASSF1A*, *Twist*, *RARβ*, *GSTP1 and HIN-1)*
Huang *et al.* (Breast Cancer Res Treat, 2010)	86 PTs; 26 FAs	FFPE	Methylation profiling on 11 genes	1) **PT vs FA**: Elevated *RASSF1A* and *TWIST1* were significantly associated with PT vs FA. 2) *TWIST1* significantly associated with increasing malignancy of PTs. 3) Using ***RASSF1A*** (involved in cell cycle control and apoptotic signalling) and ***TWIST1*** (transcription factor involved in cell differentiation and survival) methylation to distinguish PT from FA. High positive predictive value of 0.83. (11 genes - *RASSF1A, TWIST1, RARβ, APC, WIF1, MGMT, MAL, CDKN2A, CDH1, TP73* and *MLH1*)
Sawyer *et al.* (J Pathol, 2002)	119	FFPE	Mutational analysis on ***APC*** *and* ***β-catenin***; ISH on Wnts; IHC on ***β-catenin*** and cyclin D1	1) **IHC**: β-catenin nuclear staining was observed in stroma but not in epithelial. Its expression was also associated with stromal cyclin D1 staining. 2) **Mutational analysis**: i. only one benign PT showed LOH (inference to mutation) on **APC** gene. ii. no *β-catenin* mutations (exon 3) were found. 3) **ISH**: Epithelial Wnt5a mRNA expression was associated with strong nuclear staining of β-catenin in stroma.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours (continued)

Author	No. of cases	Type of materials	Type of Study	Key findings
Jones *et al.* (J Pathol, 2008)	23	Fresh frozen, FFPE, 6 short term fibroblast and one established malignant cell line	Expression profiling, validation with ISH/IHC/FISH and functional studies on cell cultures	1) A group of 162 genes were over-expressed in borderline/malignant group, with a fold change of >2 2) Observation of over-expression of *PAX3*, *SIX1*, *TGFB2* and *HMGA2* in borderline/malignant group i. **ISH**: Overexpression of *PAX3*, *SIX1* and *TGFB2* observed in borderline/malignant PTs but not in benign PTs. ii. **Amplification**: No amplification of all four genes. iii. **Mutations**: No activating mutations found in *PAX3*, *SIX1* and *TGFB2*. *HMGA2* was not screened. iv. **M-FISH** on cell culture: *PAX3* translocation (of unclear significance) found in borderline/malignant vs benign. No translocation of *HMGA2* was observed. v. **IHC:** Nuclear staining of *HMGA2* was associated with borderline/malignant PTs. vi. **Functional studies:** Knockdown of *SIX1* and *HMGA2* had significant decreased proliferation and increased apoptosis observed in malignant but not borderline cell cultures. Conversely, such trend was observed in borderline but not malignant cell cultures with knockdown of *TGFB2*. Decreased cell proliferation was observed in both borderline and malignant cell cultures with *PAX3* knockdown.
Sawyer *et al.* (J Pathol, 2003)	30	FFPE	ISH, FISH, mutational analysis, IHC on **c-myc** and **c-kit**	1) **c-myc** expression was observed in both stroma and epithelium on IHC. Amplification of myc gene on FISH was observed in one malignant tumour which had weak immunostaining. 2) **c-kit** IHC staining was associated with malignant PTs. However, no activating mutations were found in exons 8-15 and 17.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours (continued)

Author	No. of cases	Type of materials	Type of Study	Key findings
Bose *et al.* (Anticancer Res, 2010)	17	FFPE	Mutational analysis and IHC on **c-kit**	1) c-kit IHC staining was not associated with tumour grade. 2) No activating mutations were found for exons 9, 11, 13 and 17. 3) A germline mutation was found at exon 17, codon 798 – IHC negative
Chen *et al.* (J Surg Res, 2000)	19	FFPE	IHC and mutational analysis of e11 on **c-kit** on 2 mal cases	1) Q556X (nonsense mutation) found in one sample and N564S (missense mutation) in another – strongly stained on IHC. 2) CD117 IHC expression was associated with increasing grade of PT.
Carvalho *et al.* (J Clin Pathol, 2004)	19	FFPE	IHC and mutational analysis of **c-kit** and **PDGFRA**	1) **IHC**: i. c-kit stromal positivity was associated with malignant PT. ii. No association was observed between PDGFRA stromal staining and tumour grade. 2) **Mutational analysis:** i. c-kit: A silent mutation at codon 798 was observed in 2 benign cases ii. PDGFRA: Intronic insertion IVS17-50insT and G956X (nonsense mutation) observed in one malignant and two benign cases.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours (continued)

Author	No. of cases	Type of materials	Type of Study	Key findings
Kersting *et al.* (Lab Invest, 2006)	58 PTs; 167 FAs	FFPE	Amplification analysis on ***EGFR*** using FISH, ISH and IHC	1) **Amplification of a CA repeat** within intron 1 of *EGFR* was observed in 41.8% with average copy number of 5.2 and was associated positively with IHC expression. 2) *EGFR* **whole gene amplification** (>4 copies per nucleus) was observed in 15.8% of all cases, two of which were high-level amplification (>10 copies per nucleus) and associated positively with IHC expression. 3) **PT vs FA**: None of the FAs showed EGFR expression on IHC and only 2.3% of FAs harboured amplification of the CA repeat in intron 1 of *EGFR*.
Agelopoulus *et al.* (Cell Oncol, 2007)	10 PTs	Fresh frozen	Expression profiling and IHC on PTs with known status of ***EGFR*** CA-SSR amplification	1) ***EGFR* CA-SSR amplified vs non-amplified cases**: Presence of *EGFR* CA-SSR amplification was associated with 230 differentially expressed genes. Many of these genes were involved in nucleic acid metabolism, signal transduction and receptor tyrosine kinase signalling pathway. 2) *EGFR* CA-SSR amplification was associated with **caveolin-1** and **esp15** protein expression on IHC.
Tse *et al.* (Breast Cancer Res Treat, 2009)	453	FFPE	FISH and IHC analysis of ***EGFR***	1) 24% of PTs expressed stromal EGFR on IHC. Staining was associated with tumour grade. 2) EGFR gene amplification (>4 copies per nucleus) was present in 8% of 12 strong IHC positive stain.

Table 1.2 Literature review on the molecular studies performed on phyllodes tumours (continued)

Author	No. of cases	Type of materials	Type of Study	Key findings
Korcheva *et al.*(Appl Immunohistochem Mol Morphol, 2011)	26	FFPE	Mutational analysis using mass spectroscopy	1) Identified a **S8R** substitution in ***FBX4*** gene present in 11.5% of tumours. 2) No other mutations were found in a panel of 321 mutations across 30 genes, including *EGFR*, *KIT*, *KRAS* and *TP53*.
Do *et al.* (Tumor Biol, 2012)	179	FFPE	Promoter methylation analysis on ***TWIST1***; IHC	TWIST1 promoter methylation was associated with tumour grade but not associated with protein expression on IHC.
Kuenen-Boumeester *et al.* (J Pathol, 1999)	19	Fresh frozen	Mutational analysis and IHC on ***TP53***	P128S and R273C missense mutation was observed in one case each. Only R273C mutation was associated with overexpression of p53 protein on IHC.
Woolley *et al.* (Mol Diagn, 2000)	1 mal PT	Cell culture from tumour	Mutational analysis on ***TP53***	No mutation was found in exons 5-8 of *TP53.*
Gatalica *et al.* (Pathol Res Pract, 2001)	1 mal PT	FFPE	Mutational analysis on ***TP53***	p53 protein expression was correlated with malignant PT. R248W mutation in exon 7 was found in one sample.

Abbreviations:
PT - phyllodes tumour; FA – fibroadenoma; FFPE - formalin-fixed, paraffin-embedded tissue; mal – malignant; chr – chromosome; CGH - comparative genomic hybridization; LOH - loss of heterozygosity; FISH - fluorescence *in situ* hybridization; IHC - immunohistochemistry

Four main molecular study types could be identified from Table 1.2 - chromosomal investigations, methylation, expression profiling and mutational analysis. Largely, there are no common observations of recurrent aberrations except for gain of chromosome 1q reported by a few groups [Jee et al., 2003; Lu et al., 1997; Jones et al., 2008; Laé et al., 2007; Lee et al., 2010; Ang et al., 2010]. Although it is widely agreed that higher grade tumours harbour increased molecular aberrations such as chromosomal changes and methylation events, none of the studies has correlated their findings with clinical outcome.

1.2 Prognostic factors

1.2.1 Prognosis of phyllodes tumours

The majority of phyllodes tumours behave in a benign fashion, with reported overall local recurrence rates of about 19% ranging from 0% to 67% in a review of retrospective series from 1951 to 2012 [Spitaleri et al., 2013]. Deaths from phyllodes tumours are rare, often preceded by distant metastases [Tan et al., 2005a]. The common sites for metastasis include lung and bone. The rates of recurrences and deaths of large series previously published are shown in Table 1.3.

Grade is a good indicator of biological behaviour of the tumour. Majority of those in the benign grade have a good prognosis while the course of malignant grade tumours is usually unfavourable. However, recurrences could occur regardless of grade, even in benign ones. While local recurrences are more commonly reported in the benign grades and distant metastases occurring almost exclusively in the malignant group, some benign tumours could recur with increased grade and advance with an aggressive behaviour. Such unpredictable behaviour of phyllodes tumours drives the search for prognostic factor(s) and/or marker(s) which could identify tumours with potential aggressive behaviour at an early stage so that they could be managed accordingly.

1.2.2 Clinical and histological prognostic factors

Age does not seem to be an important factor [Pietruszka and Barnes, 1978; Tan et al., 2005; Guillot et al., 2011; Parker and Harries, 2001] other

than sporadic reports of age being a significant factor for survival [Macdonald et al., 2006; Pezner et al., 2008]. The role of tumour size is not clear with some reporting larger tumour size having unfavourable outcomes [Belkacémi et al., 2008; Chaney et al., 2000; Roa et al., 2006] while others found no impact of tumour size on outcome [Pietruszka and Barnes, 1978; Tan et al., 2005a, Chen et al., 2005].

Although many histological prognostic factors have been investigated, findings between studies are not always consistent. Tan *et al.* reported that stromal atypia, cellularity and tumour borders were associated with local recurrence on univariate analysis [Tan et al., 2005a]. Multivariate analysis revealed surgical margin status and pseudoangiomatous stromal hyperplasia to be independent predictors for recurrence. Cohn-Cedermark *et al.* reported tumour necrosis and presence of heterologous stromal elements were independent prognostic factors on multivariate analysis [Cohn-Cedermark et al., 1991]. Chaney reported stromal overgrowth to be the only independent predictor of distance metastasis in a study of 101 patients [Chaney et al., 2000]. Nonetheless, surgical margins appear to be a rather consistent determinant for local recurrences [Pandey et al., 2001; Asoglu et al., 2004; Kapiris et al., 2001; Belkacémi et al., 2008] and distant metastasis [Pandey et al., 2001].

Table 1.3 Descriptive data of past large retrospective series conducted for phyllodes tumours

Author	Journal	Year	Period	N	Country	Age	Size	Grade (%)			Prognosis (%)		
								Ben	Bor	Mal	LR	DR	Death
Treves	Cancer	1951	1930–1949	77	USA	41[a]	10	54	23	23	67	11.6	10.4
Norris	Cancer	1967	Before1967	94	USA	45	6.4	Not graded			35	NR	17
Hajdu	Cancer	1976	1932–1976	199	USA	NR	4	75	NG	25	16	1.5	2
Chua	Aust N Z J Surg	1988	1978–1984	106	Singapore	30[a]	5	91	6	3	19	0.9	1
Cohn-Cedermark	Cancer	1991	1958–1986	77	Sweden	50	5	55[b]	NG	45[c]	19	21	21
Zurrida	Eur J Cancer	1992	1970–1989	216	Italy	NR	NR	65	14	21	12.5	NR	NR
Reinfuss	Cancer	1996	1952–1988	170	Poland	52[a]	7	54	11	35	8	16	16
Zissis	Breast Cancer Res Treat	1998	1981–1995	84	Greece	NR	6	65	17	18	2.3	1	1
Chaney	Cancer	2000	1944–1998	101	USA	41	6	58	12	30	4	8	16
Niezabitowski	Breast Cancer Res Treat	2001	1952–1998	118	Poland	49	NR	44	20	36	8	9	11
Chen	J Surg Oncol	2005	1985–2003	172	Taiwan	37[a]	5.8	76	7	17	11	1.7	2

Table 1.3 Descriptive data of past large retrospective series conducted for phyllodes tumours (continued)

Author	Journal	Year	Period	N	Country	Age	Size	Grade (%)			Prognosis (%)		
								Ben	Bor	Mal	LR	DR	Death
Tan	Am J Clin Pathol	2005	1992–2002	335	Singapore	42	5.4	75	16	9	12.8	2	2
Abdalla	J Egypt Natl Canc Inst	2006	1988–2003	79	Egypt	42	11	39	34	27	20	14	13
Ben hassouna	Am J Surg	2006	1986–2001	106	Tunisia	40[a]	8.3	59	25	26	12.2	7.5	8
Barrio	Am Surg Oncol	2007	1954–2005	293	USA	42	6	69	NG	31	8.5	1.7	2
Belkacémi	Int J Radiat Oncol Biol Phys	2008	1971-2003	443	France	40	4.6	64	18	18	19	3.4	NR
Guillot	Breast J	2011	1994–2008	165	France	44	3	69	22	9	10	1.2	1
Pimiento	J Am Coll Surg	2011	1999–2010	124	USA	45	4.5	49	35	16	6.5	1.6	NR
Tan	J Clin Path	2012	1992–2010	605	Singapore	43	5.2	73	18	9	11.2	1.1	2
Spitaleri	Crit Rev Oncol/Hematol	2013	1999-2010	172	Italy	44	NR	40	24	36	10.5	1.2	2
Kim	Breast Cancer Res Treat	2013	2000-2010	193	Korea	41[a]	4	75	17	8	9.3	4.1	2.1

[a]mean age, otherwise median; [b]benign and borderline; [c]probably or clearly malignant;
Ben=Benign; Bor=Borderline; Mal=Malignant; LR=local recurrence; DR=distant recurrence; N=number of cases; NG=not graded; NR=not reported

1.2.3 Biological markers as prognostic factors

The increasingly important role of immunohistochemistry as an adjunct tool in diagnostic surgical pathology has directed the attention of researchers from histological parameters to biological markers over the years. Extensive studies have been performed over the last decade in search of suitable biological markers to predict the behaviour of phyllodes tumours. An excellent review performed by Karim *et al.* which was published in Journal of Clinical Pathology in two separate articles detailed the summary of the findings for these studies [Karim et al., 2009b; Karim et al., 2013].

However, studies focusing on the prognostic impact of biomarkers are limited by the events of recurrences and deaths in each series. Table 1.4 shows the biological markers studied in relation to clinical outcome in terms of local recurrence, distant recurrence and/or deaths. Broadly, these markers can be categorized into cell cycle markers, angiogenesis markers, epithelial-mesenchymal transition (EMT) markers and receptor tyrosine kinases (Figure 1.9). Many recurring biological markers were investigated but consistent findings were scanty. For example, some authors reported p53 to be correlated with disease free recurrence [Niezabitowski et al., 2001; Kuijper et al., 2005a; Yonemori et al., 2006] but some observed no correlation [Kleer et al., 2001; Tan et al., 2005b; Esposito et al., 2006]. Ki67 and CD117 are other examples of well studied markers in phyllodes tumours but with mixed results in association with clinical outcome.

While immunohistochemistry is an excellent supplement to classic morphology [Ramos-Vara and Miller, 2014], many aspects of the technique are applied differently between laboratories. These include types and duration of fixation used, methods of antigen retrieval, staining protocol, and evaluation of staining results. Without standardization of these aspects, it is unlikely that the results are comparable between laboratories.

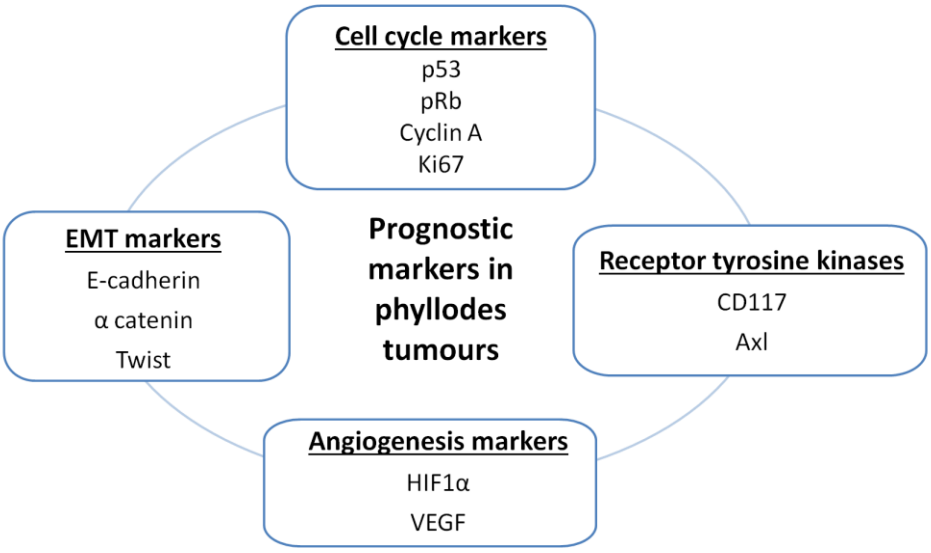

Figure 1.9 Prognostic markers studied in phyllodes tumours

Table 1.4 Studies on biological markers of prognostic importance

First Author, *Journal*	Year	Number of cases	Prognosis	Test	Marker	Summary
Kleer *Mod Pathol*	2001	20 7 benign (35%) 13 malignant (65%)	3 LR (15%)	IHC	Ki67, p53	Neither Ki67 or p53 can predict recurrence
Niezabitowski *Breast Cancer Res Treat*	2001	118 52 benign (44%) 24 borderline (20%) 42 malignant (36%)	17 recur (14.4%) -6 LR, 7 DR, 4 LR+DR 13 deaths (11%)	Flow cytometry; IHC	S-phase fraction (SPF) Ki67, p53	-SPF and p53 independent prognostic marker for DFS -Ki67 independently influenced OS.
Shpitz B *J Surg Oncol*	2002	23 16 benign (70%) 4 borderline (17%) 3 malignant (13%)	4/22 LR (17%) 1/22 death (4.5%)	IHC	HER2, Ki67, p53	No correlation with recurrence
Tse *Mod Pathol*	2003	186 106 benign (57%) 51 borderline (27%) 29 malignant (16%)	24/180 recur (13%) -21 LR, 3 DR	IHC	CD31, p53	No correlation with recurrence
Tse *Mod Pathol*	2004	179 101 benign (56%) 50 borderline (28%) 28 malignant (16%)	9 recur (5%) -7 LR, 2 DR	IHC	CD117	No relationship with recurrence
Kuijper *Am J Clin Path*	2005a	40 21 benign (53%) 8 borderline (20%) 11 malignant (27%)	10/37 recur (27%)	IHC	bcl-2, Cyclin A, Cyclin D1, Ki-67, Rb, p53, p16, and p21	Inverse relation between DFS and expression of cyclin A, Ki-67, pRb, p53

Table 1.4 Studies on biological markers of prognostic importance (continued)

First Author, *Journal*	Year	Number of cases	Prognosis	Test	Marker	Summary
Kuijper *Breast Cancer Res*	2005b	37 18 benign (49%) 8 borderline (21%) 11 malignant (30%)	10/30 recur (33%) -9 LR,1 DR	IHC	CAIX, CD31, HIF-1α, Ki67, p53, VEGF	HIF-1α predictive of DFS
Tan PH *Mod Pathol*	2005	335 250 benign (75%) 54 borderline (16%) 31 malignant (9%)	43 recur (13%) 7 deaths (2%)	IHC	CD117, p53	CD117 predicted recurrence
Esposito *Arch Pathol Lab Med*	2006	30 16 benign (53%) 8 borderline (27%) 6 malignant (20%)	4/25 recur (16%) 1 death	IHC	CD117, Endothelin-1, Ki67 p16, p21, p53	No correlation with recurrence
Yonemori *Pathol Res Prac*	2006	41 20 benign (49%) 5 borderline (12%) 16 malignant (39%)	11 recur (27%) -2 LR, 7 DR, 2 LR+DR Deaths not mentioned	IHC	CD117, EGFR, HER2, Ki67	Ki67 and p53 significantly correlated with DFS and OS.
Karim *J Clin Path*	2009	65 34 benign (52%) 23 borderline (35%) 8 malignant (12%)	9/62 recur (15%)	IHC	β-catenin, Wnt1, Wnt5a, SFRP4 and E-cadherin	DFS significantly decreased with positive epithelial E-cadherin staining
Karim *Histopathology*	2010	65 34 benign (52%) 23 borderline (35%) 8 malignant (12%)	9/62 recur (15%)	IHC	Cyclin D1, Ki67, p16, Rb	High epithelial expression of Rb associated with reduced DFS

Table 1.4 Studies on biological markers of prognostic importance (continued)

First Author, *Journal*	Year	Number of cases	Prognosis	Test	Marker	Summary
Jung *J Breast Cancer*	2010	67 39 benign (58%) 16 borderline (24%) 12 malignant (18%)	11 recur (16.4%) 3 deaths (4%)	IHC; mutation analysis	CD117, PDGFRA	Stromal CD117, PDGFRA and Ki67 expression is predictive of recurrence
Al-Masri *Ann Surg Oncol*	2012	43 16 benign (37%) 10 borderline (23%) 17 malignant (40%)	6 DR (14%)	IHC	CD10	CD10 significantly correlated with the occurrence of distant metastasis.
Tsang *Histopathology*	2012a	158 92 benign (58%) 43 borderline (27%) 23 malignant (15%)	22 recur (14%)	IHC	α-catenin, β-catenin, E-cadherin	α-catenin showed a significant correlation with recurrence
Tsang *Hum Pathol*	2012b	155 92 benign (59%) 42 borderline (27%) 21 malignant (14%)	33/152 recur (22%)	IHC	E-cadherin	Epithelial expression correlates with reduced DFS
Kwon *Tumor Biol*	2012	207 157 benign (76%) 34 borderline (16%) 16 malignant (8%)	26 recur (13%) -18 LR, 8 DR 8 deaths (4%)	IHC	CXCR4, Ezrin, HMGA2, N-cadherin, TGFβ, Twist, S100A4, SDF1	High Twist expression correlated with shorter DFS and OS.

Table 1.4 Studies on biological markers of prognostic importance (continued)

First Author, *Journal*	Year	Number of cases	Prognosis	Test	Marker	Summary
Do *Tumor Biol*	2012	179 103 benign (58%) 38 borderline (21%) 38 malignant (21%)	29 recur (16%) -24 LR, 5 DR 4 deaths (2.2%)	Methylation; IHC	Twist (methylation and IHC) NF-κB, SIP1, Snail, Slug (IHC only)	No relationship with recurrences
Ho *Histopathology*	2013	185 120 benign (65%) 48 borderline (26%) 17 malignant (9%)	11 recur (6%) 3 deaths (2%)	IHC	β-catenin ,CD34, VEGF	VEGF expression associated with poorer OS
Ren *Tumor Biol*	2014	140 80 benign (57%) 30 borderline (21.5%) 30 malignant (21.5%)	30 recur (21%) -20 LR, 10 DR	Real-time PCR, western blot, IHC	Axl and ST6GalNAcII	Axl and ST6GalNAcII was significantly correlated with distance metastasis by IHC

Abbreviations: LR- local recurrences; DR- distant recurrences; IHC- immunohistochemistry; PCR – polymerase chain reaction; OS- overall survival; DFS – disease free survival; SPF – S-phase fraction

1.2.3.1 CD117 as a prognostic marker in phyllodes tumours

CD117, or known as c-kit, is a type III transmembrane receptor tyrosine kinase encoded by the *KIT* gene located at chromosome 4q12. The receptor is characterized by an extracellular domain comprising five immunoglobulin-like repeats (encoded by exons 2-9), a transmembrane domain (exon 10), a juxtamembrane domain (exon 11), and two intracellular tyrosine kinase domains (exons 12-20) as shown in Figure 1.10. The extracellular domain binds to ligands for receptor activation. The juxtamembrane domain has an autoinhibitory mechanism maintaining the dormant state of an inactive receptor and tyrosine kinase domains act as docking sites for signalling molecules.

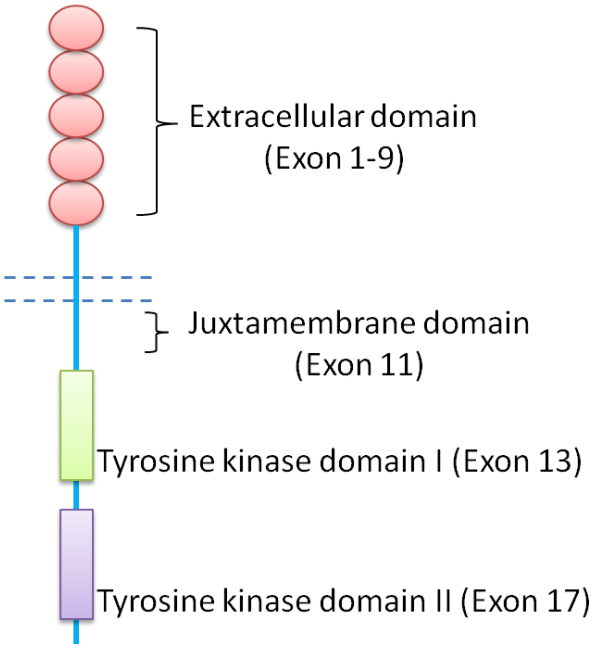

Figure 1.10 Schematic illustration of a CD117 receptor, made up of an extracellular domain comprising five immunoglobulin-like repeats, a juxtamembrane domain encoded by exon 11, and tyrosine kinase domains I and II encompassing exons 13 and 17.

CD117 exists as monomer in its inactive state. Upon binding of stem cell factor (KIT ligand), the extracellular domain undergoes conformational changes and allows two monomers to form a dimer, facilitating activation of the receptors via transphosphorylation (Figure 1.11). Activation of CD117 initiates various downstream signalling pathways such as the phosphatidylinositol 3'-kinase (PI3'-kinase), Src family kinases, mitogen-activated protein kinase (MAP kinase) signalling pathways, mediating cell survival, migration, and proliferation [Lennartsson and Rönnstrand, 2012].

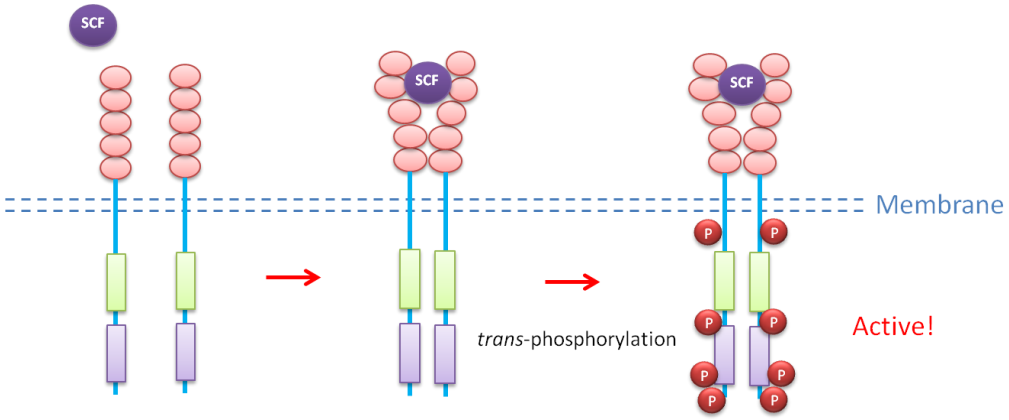

Figure 1.11 Activation of CD117. CD117 receptors exist as monomers in the inactive state. Upon binding of the stem cell factor (SCF), the extracellular domain undergoes conformational changes, allowing two monomers to be in close proximity, thereby forming a dimer which facilitates activation of the receptors via transphosphorylation, starting from the juxtamembrane domain, before proceeding to the tyrosine kinase domains.

CD117 has been implicated in various malignancies, including gastrointestinal stromal tumours (GISTs) [Corless et al., 2002], colon cancer [Gavert et al., 2013] and small cell lung carcinoma [Fischer et al., 2007]. Mutations of CD117 are commonly seen in these tumours especially those at exons 11 (juxtamembrane domain), 13 and 17 (kinase domains) and to a lesser extent in exon 9 (extracellular domain). Patients with GISTs harbouring overexpression or mutation of CD117 are responsive to tyrosine kinase inhibitors which can block the activity of an activated CD117 receptor. There has been previous suggestion that the stromal component of phyllodes tumours bear some similarities with GIST such as the spindled nature and spectrum of behaviour from benign to malignant [Carvalho et al., 2004].

Investigation of CD117 mutational status in phyllodes tumours started in year 2000 by Chen *et al.* where the authors found two point mutations Q556X and N564S at exon 11 [Chen et al., 2000]. Mutations in this domain release the autoinhibitory mechanism of the receptor, enabling the receptor to be activated even without the binding of a ligand. Following this, many other researchers have performed similar investigations, assessing both mutational status and protein expression of CD117 in phyllodes tumours (Table 1.5). However, reports vary across studies and data on prognosis of CD117-expressing tumours is scanty.

Table 1.5 Summary of investigations of CD117 expression and mutation status in phyllodes tumours

Authors, year	N	Antibody and dilution	Scoring Criteria	CD117 expression	Mutation Status	Follow-up
Chen *et al.* 2000	19	Chemicon clone K69 1:100	≥10%	Associated with malignant grade	Q556X (exon 11) and N564S (exon 11) in one case each	NA
Sawyer *et al.* 2003	30	Novocastra 1:20	Intensity≥1	Associated with malignant grade	L510M (exon 10) in one case	NA
Tse *et al.* 2004	179	Novocastra 1:40	≥20% stromal cells of moderate-to-strong staining	Correlated with increasing grade	NA	Not correlated with recurrence
Carvalho *et al.* 2004	19	Novocastra 1:60	≥25%	Associated with malignant grade	Silent mutation Isoleucine 798 (exon 17) in two cases	NA
Tan *et al.* 2005b	335	Dako A4502 1:250	Any unequivocal stain	Associated with tumour grade	NA	Expression correlated with recurrence
Esposito *et al.* 2006	16	Dako polyclonal 1:100	Combined immunoreactive score ≥1	Associated with tumour grade	NA	Not correlated with recurrence

Table 1.5 Summary of investigations of CD117 expression and mutation status in phyllodes tumours

Authors, year	N	Antibody and dilution	Scoring Criteria	CD117 expression	Mutation Status	Follow-up
Yonemori *et al.* 2006	41	Dako A4502 Dilution not specified	>10% of at least moderate intensity	None of the cases were positive	NA	No positive cases for correlation with six distant recurrences
Djordjevic *et al.* 2008	47	Zymed 1:400	>0%	No correlation	Not detected in two CD117 positive tumours	NA
Bose *et al.* 2010	17	Dako 1:50	≥5%	No correlation	Silent mutation Isoleucine 798 (exon 17) in one case	NA
Jung *et al.* 2010	67	Dako polyclonal 1:300	>10%	Correlated with increasing grade	None detected in subset of 28 samples	Expression associated with recurrence
Noronha *et al.* 2011	33	Biocare clone Y145 1:100	≥20%	Differentially expressed in malignant grade	NA	NA

NA- Experiment not performed/ Data not available

1.2.3.2 Homeobox proteins as prognostic markers in phyllodes tumours

Homeoproteins are transcribed from homeobox genes which constitute a large group of developmental regulators for embryogenesis. Homeobox was first identified as a sequence motif in the fruit fly *Drosophila* and is well conserved across species. It encodes a signature homeodomain which forms a helix-turn helix three dimensional structure that binds to specific DNA elements, primarily those that contain a TAAT core motif [Samuel and Naora, 2005], enabling them to act as transcriptional regulators, triggering activation or suppression of downstream target genes including those involved in cell growth and differentiation.

The process of development shares common critical pathways as neoplasia [Kelleher et al., 2006], and aberrant expression of homeoproteins has been implicated in many solid tumours [Samuel and Naora, 2005] including colorectal cancers, breast cancers and sarcomas. Six1 and Pax3 are among these homeoproteins which are reportedly deregulated in tumours. Overexpression of Six1 was associated with higher tumour stage in cervical cancer and hepatocellular carcinoma [Zheng et al., 2010; Ng et al., 2006] while aberrant expression of Pax3 was demonstrated in Ewing sarcoma and glioma.

We previously found that Six1 and Pax3 were overexpressed in borderline/malignant phyllodes tumours at the transcript level in a collaborative project [Jones et al., 2008b]. The comprehensive expression profiling of 23 fresh frozen tumours revealed a 5-fold difference between

benign and borderline/malignant tumours among 162 differentially expressed genes. Moreover, knockdown of Six1 and Pax3 resulted in significantly decreased cell proliferation in short term cultures established from borderline and malignant phyllodes tumours. However, protein expression of both Six1 and Pax3 was not investigated and their prognostic significance in clinical samples was not addressed. It would be of interest to expand the study in a larger cohort to corroborate these findings with expression at the protein level.

1.3 Scope of study

The hypothesis of this study is that there are distinct subgroups within phyllodes tumours, having different course of disease which harbours unique characteristics that set them apart from each other.

This study aims to investigate and evaluate the prognostic importance of clinical and molecular parameters of phyllodes tumours, to understand the biology of phyllodes tumours from these perspectives, of which the eventual desired outcome is to improve management of patients with phyllodes tumours, with a more accurate ability to predict biological behaviour.

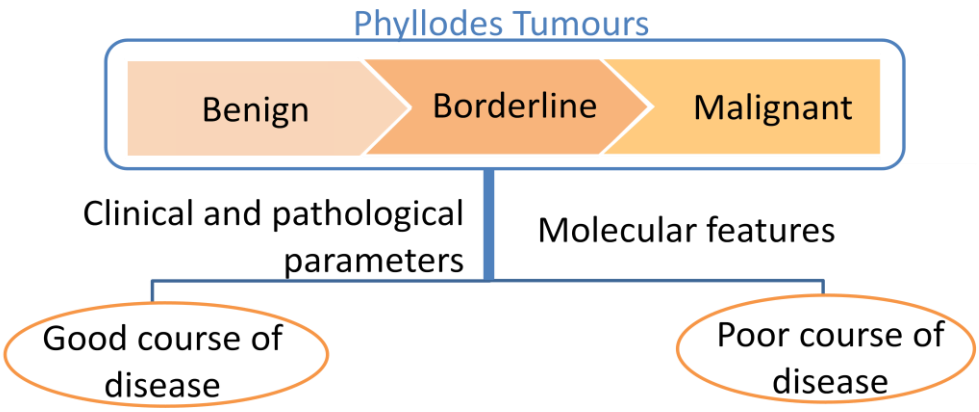

Figure 1.12 Diagram illustrating hypothesis and aims of the study

The scope of this study encompasses specific aims as follow:

1) To evaluate individual clinicopathological parameters in predicting recurrences in patients with phyllodes tumours and to propose a predictive model with these parameters
2) To screen for genetic aberrations in prognostically distinct groups of phyllodes tumours to understand the biological differences between tumours with good course of disease and tumours with poor course of disease
3) To investigate and validate the prognostic importance of receptor tyrosine kinase CD117, and homeobox proteins Six1 and Pax3 in phyllodes tumours

2.0 MATERIALS AND METHODS

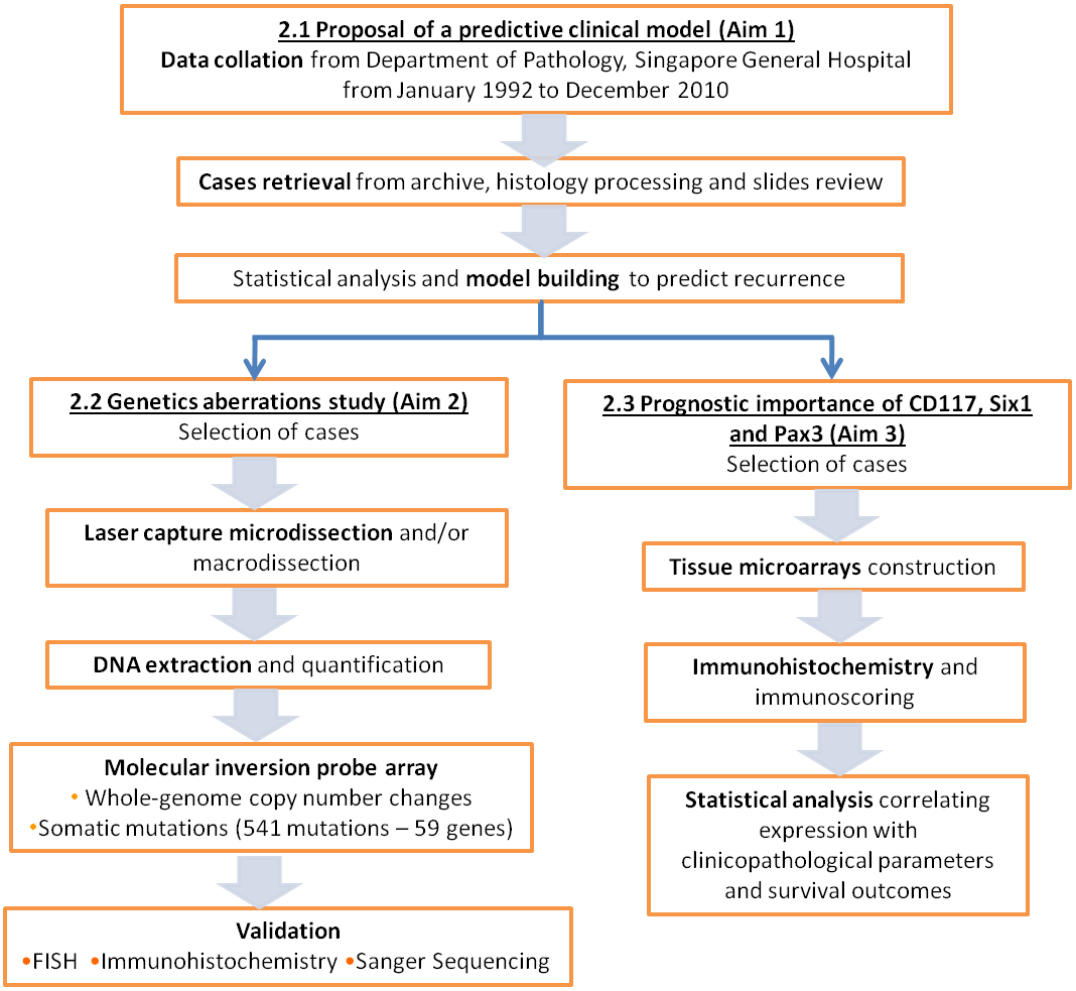

Figure 2.1 Overview of the project workflow

2.1 Proposal of a predictive clinical model

2.1.1 Study population and data collation

This study received approval from the SingHealth Centralized Institutional Review Board. Cases of breast phyllodes tumours diagnosed in the Department of Pathology, Singapore General Hospital between January 1992 and December 2010 were included in this study. Patient details and case information such as age, ethnicity, type of surgery performed, surgical margins of tumours, and follow-up data were retrieved from the information database.

Haematoxylin and eosin-stained (H&E) slides for all cases were retrieved from the archive room. Old slides were subjected to restaining and missing slides were replaced by re-sectioning of tissue from archival formalin-fixed, paraffin embedded (FFPE) tissue blocks.

2.1.2 Sample criteria

Slides were reviewed by a pathologist. The five histological parameters were reviewed and noted - stromal cellularity, stromal atypia, stromal mitotic activity, presence or absence of stromal overgrowth, and nature of microscopic margins.

Stromal cellularity and atypia were categorized as mild, moderate and marked. Stromal mitotic activity was quantified per 10 high-power microscopic fields (x40 objective and x10 eyepiece) over the most mitotically active stromal areas. Stromal overgrowth was defined as a low-power field (x4 objective and x10 eyepiece) that consisted of stroma only without

epithelial elements. Microscopic margins were classified as circumscribed or permeative if the margin between the tumour and surrounding breast tissue was well defined or permeative, respectively.

A phyllodes tumour was diagnosed as benign when the following criteria were fulfilled: mild stromal cellularity and atypia, occasional mitoses of not more than four per 10 high-power fields, no stromal overgrowth, and with circumscribed margins. Conversely, a malignant phyllodes tumour was diagnosed when the following features were present: marked stromal hypercellularity and atypia, brisk mitotic activity (10 or more per 10 high-power fields), presence of stromal overgrowth, and permeative margins. A borderline phyllodes tumour showed some, but not all of the characteristics observed in malignant phyllodes tumours.

2.1.3 Histology processing

2.1.3.1 Sectioning

Tissue was sectioned from the archival FFPE tissue block or a tissue microarray block using a manual rotary microtome (Leica RM2135, Leica Biosystems, Germany). Sections were placed on a floatation bath at room temperature and flattened using a pair of forceps. Sections were then transferred to a warm water bath with temperature at approximately 50°C using a glass slide before placing onto a glass slide, charged slide (Microsystems Plus Slides, Leica Biosystems, Germany), or membrane slides (1.0 mm PEN, Carl Zeiss, Germany) depending on downstream procedures.

To ensure tissue adherence, tissue on glass slides were left on a hot plate at 80°C for 3 minutes while charged slides were incubated in an oven overnight at 80°C. Membrane slides were left to air-dry at room temperature.

2.1.3.2 Haematoxylin and eosin (H&E) staining

Old slides were first incubated in xylene for ≥ 2 hours to remove cover slips and associated depex before being subjected to H&E re-staining. New slides of tissue sections dried on a hotplate were loaded directly into an automated H&E staining system on ST5010 Autostainer XL (Leica Biosystems, Germany). Briefly, tissues were dipped in two changes of xylene for 2 minutes each and rehydrated through graded alcohol. Subsequently, slides were washed in running water, incubated in two changes of haematoxylin and washed under running water again. Upon staining with eosin, tissues were dehydrated through graded alcohol and dipped in xylene before mounted with DPX mounting media and cover slips.

2.1.4 Model and nomogram building

All analyses were done using R V.2.13.0 (http://www.R-project.org), STATA version 11, and PASW Statistics for Windows version 18.0. Differences between benign, borderline and malignant phyllodes tumours were assessed with χ^2 test and Fisher's exact test.

Kaplan Meier survival curves were employed to estimate recurrence-free survival (RFS) and overall survival (OS), defined as time from date of surgery to date of first relapse and death, respectively, or to the date of last

follow-up for censored cases. Log-rank test was used to compare survival between groups. Univariate and multivariate Cox regression analyses were performed to evaluate the effect of potential features in predicting recurrences.

A backward elimination method was used applying in the Akaike information criterion (AIC) to select best performing features. Assumptions for proportional hazards were verified systematically. Multivariate Cox regression coefficients of the model were used to generate the nomogram. The nomogram performance was validated using 200 bootstrap set of resamples. Calibration plots were constructed to evaluate the ability of the nomogram in predicting RFS of individual patients at 1, 3, 5 and 10 years after surgery.

2.1.5 Comparison of models

A total histological score model was derived on the assumption that the potential features had a linear additive effect. Scores were assigned based on the criteria as listed in Table 2.1 for each category. The total histological score model was defined as the sum of the scores of the five criteria, which ranged from 5 to 13 points.

Table 2.1 Scores assigned for each category. Total score were derived by summing up the score from each criterion

Score / Criteria	1	2	3
Hypercellularity	Mild	Moderate	Marked
Atypia	Mild	Moderate	Marked
Mitotic rate	0-4	5-9	≥10
Overgrowth	Absent	Present	NA
Microscopic margin	Pushing/Circumscribed	Permeative	NA

NA: not applicable

The performance of the nomogram and the total histological score model was compared using Harrell's *c*-index and likelihood ratio. Harrell's *c*-index measures the predictive power of each model by evaluating concordance between predicted and observed response of the individual subject separately. For assessment using likelihood ratio test, the nomogram and total histological score model were first combined to give rise to a full model. Then the performance of a nested model comprising only the nomogram and total histological score respectively were evaluated with likelihood ratio test. Adequacy index was calculated to measure the percentage of variation explained by the respective nested model as compared to the full model.

2.2 Genetic aberrations study

2.2.1 Selection of cases for molecular investigation

Twenty cases were further selected from the cohort for molecular investigation based on clinical outcome. Cases were stratified into two prognostically distinct groups:

a) cases with recurrence, metastasis and/or death

b) cases without recurrence with a follow-up period of no less than two years, a duration supported by our own data of a median time to recurrence of 24.6 months.

2.2.2 Tissue dissection

2.2.2.1 Macrodissection

5-10 sections of tissue of 10µm thick on glass slides were deparaffinized with 3 changes of xylene for 2 minutes each and 2 changes of absolute alcohol for one minute each. Corresponding desired areas containing >95% lesional tissue pre-marked on a H&E slide by a histopathologist were macrodissected using a sterile scalpel blade for each case and subjected to DNA isolation downstream.

2.2.2.2 Laser capture microdissection

10-15 sections of 5µm thick tissues were sectioned on UV-screened membrane slides (1.0 mm PEN, Carl Zeiss, Germany) and stained with haematoxylin. The PALM® Microbeam system (Carl Zeiss, Germany) was employed. A tissue section was viewed through the imaging software of the system under the microscope and the desired tissue area was drawn. Then, a laser beam was used to cut and isolate the desired area onto the PALM® AdhesiveCaps (Carl Zeiss, Germany) for downstream analysis (Figure 2.2).

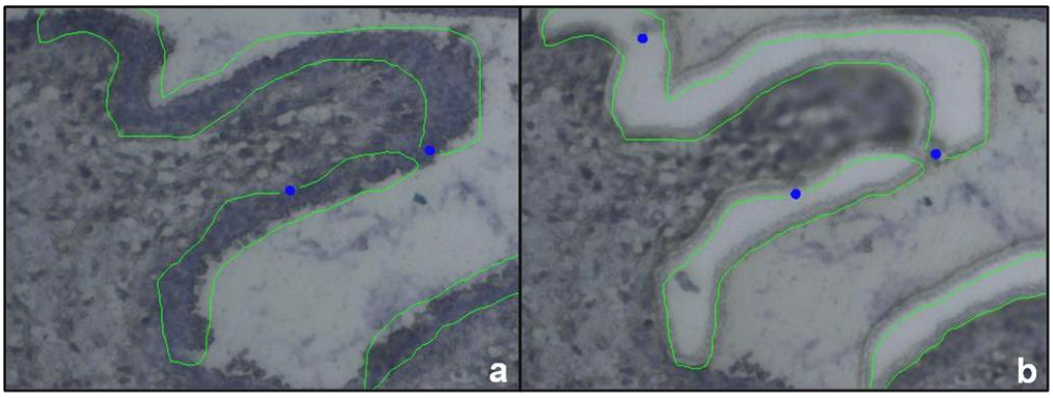

Figure 2.2 Procedure of laser capture microdissection. (a) Epithelium was drawn and highlighted before laser dissection. (b) After laser dissection, the epithelial component was isolated and captured for DNA extraction.

2.2.3 DNA extraction

DNA extraction from tissue dissected was performed using the Ambion® RecoverAll™ Total Nucleic Acid Isolation Kit (Life Technologies, USA) according to the manufacturer's instruction. Briefly, digestion buffer and protease were added to the tissue and incubated for 48 hours at 50°C. After the temperature had settled at room temperature, isolation additive and

ethanol mixture was added to the samples and mixed thoroughly. The mixture was then passed through a filter cartridge and washed several times. Nuclease digestion was performed with RNase mix prior to elution of DNA in nuclease-free water.

2.2.4 Genome wide molecular inversion probe array

Purified DNA was sent for processing via the utility of the OncoScan™ FFPE Express 2.0 Service (Affymetrix, Inc, CA). Samples were quantified with the PicoGreen® assay. Only samples with at least 75ng of DNA were passed and normalized. The OncoScan™ assay interrogates ~333,000 markers for copy number (CN) changes, and 541 somatic mutations from 59 genes specific for cancer (Supplementary Table 2). Technicalities and the design of probes were as previously described [Wang et al., 2012b]. Pre-processed data analysed using Nexus 6 Copy Number™ (Biodiscovery, CA) was delivered. The fraction of genome altered (FGA) for each tumour was calculated [Schiffman et al., 2010] as total lengths of gains and losses divided by the total genome length (3224.46Mb, Genome Reference Consortium Human Build 37).

Mutation score, a measurement of differences between the normal wild-type cluster and test sample, was provided for each mutation assay. A higher score denotes a larger difference from the wild-type cluster and hence a higher likelihood for the mutation to be real. A score of 4 was recommended for BRAF V600E assay to be valid while other mutation assays were recommended at a score of 9.

2.2.5 Validation

2.2.5.1 Sanger sequencing

Polymerase chain reaction (PCR) was performed for samples which harboured potential mutations detected on the OncoScan™ mutation panel. Primer sequences as shown in Table 2.3 were designed with Primer-BLAST. PCR was set up using Taq PCR Core Kit (Qiagen, Germany) in accordance with the manufacturer's protocol in a total volume of 50µl. The PCR cycling program was as follows: (1) 94°C for 5 minutes (1 cycle), (2) 94°C for 1 minute, 57°C for 1 minute, 72°C for 1 minute (40 cycles) and (3) 72°C for 10 minutes (1 cycle). Negative controls (non-template) were included in every PCR run. PCR products were analysed with 2% gel electrophoresis stained with GelRed™ and purified using QIAquick PCR Purification Kit (Qiagen, Germany). Purified PCR amplicons were sequenced by 1st BASE Pte Ltd (Singapore). DNA sequences were analysed using the NCBI database and BLAST software.

Table 2.2 Primer sequences designed using Primer-BLAST for PCR amplification and Sanger Sequencing

Mutation	NCBI Accession Number	Forward Sequence	Reverse Sequence	Size (bp)
ATM_p.T2666A_c.7996A<G	NM_000051.3	ATTCCAGCAGACCAGCCAAT	AGGGCTTGGGCAAAGGAAAT	237
BRCA1_c.134plus1G<T*	NG_005905.2	TGTCTCCACAAAGTGTGACCA	GGTGTTTCCTGGGTTATGAAGG	103
CDKN2A_p.W110X_c.329G<A	NM_000077.4	CCCTGGCTCTGACCATTCTG	ATGGTTACTGCCTCTGGTGC	309
CDKN2A_c.151minus1G<A*	NG_007485.1	CCCTGGCTCTGACCATTCTG	ATGGTTACTGCCTCTGGTGC	309
EGFR_p.G598V_c.1793G<T	NM_005228.3	TGTGCCCACTACATTGACGG	GTGCAGTTTGGATGGCACAG	131
EGFR_p.L858R_c.2573T<G	NM_005228.3	AGCAGGGTCTTCTCTGTTTCA	TGACCTAAAGCCACCTCCTT	200
KIT_p.V654A_c.1961T<C	NM_000222.2	GTTCCTGTATGGTACTGCATGCG	CAGTTTATAATCTAGCATTGCC	282
MEN1_c.654plus3A<G*	NG_008929.1	ATGCCTGGGTAGTGTTTGGG	ATGACAGCCAGGAAAAGGGG	263
NF2_p.V219M_c.655G<A	NM_000268.3	CAGTGTCTTCCGTTCTCCCC	TTTAGCAGTCTGGCCCTCAC	167
NOTCH1_p.Q2460X_c.7378C<T	NM_017617.3	CTGGCGGTGCACACTATTCT	TGTCCACAGGCGAGGAGTAG	136
NPM1_p.W288 frame shift_c.863insCCTG†	NM_002520.6	TCTCTGGCAGTGGAGGAAGT	CCATGTCTGACCACCGCTAC	289
RB1_p.E440X_c.1318G<T	NM_000321.2	TGCTAAAGCTGTGGGACAGG	ACGAACTGGAAAGATGCTGC	114
RET_p.A664D_c.1991C<A	NM_020630.4	TCTGCCTTCTGCATCCACTG	GAGGAGTAGCTGACCGGGAA	110
SMAD4_p.K507Q_c.1519A<C	NM_005359.5	GTCAGCTGCTGCTGGAATTG	GTTAAGGGCCCCAACGGTAA	234
TGFBR2_p.R497X_c.1489C<T	NM_001024847.2	CCCTGTGTTTGCTGGCTTTC	GTCGCCCTCGATCTCTCAAC	123
TP53_p.H179R_c.536A<G	NM_000546.5	CACTTGTGCCCTGACTTTCA	GGGCCAGACCTAAGAGCAAT	320
TP53_p.C124R_c.370T<C	NM_000546.5	TGAAGCTCCCAGAATGCCAG	AAGTCTCATGGAAGCCAGCC	229

Note: Mutations written in such format: Gene_p. amino acid position_c. DNA coding sequence position.

* Mutations at intronic regions

† Insertion of four bases (CCTG) between coding sequence position 863 and 864

2.2.5.2 Fluorescence *in situ* hybridization (FISH)

Tissue sections of 4µm thick on charged slide were deparaffinized and left to air-dry. Sections were treated in 0.2M hydrogen chloride for 20 minutes at room temperature, followed by washes in purified water and 2X saline-sodium citrate (SSC) buffer. The slides were then incubated in Pretreatment Solution (Abbott Molecular, USA) for 30 minutes at 80°C. Tissue was digested with Protease I Solution (Abbott Molecular), washed in 2X SSC buffer and dehydrated in graded alcohol.

Dual colour LSI EGFR SpectrumOrange™/ CEP 7 SpectrumGreen™ probes (Abbott Molecular) were applied and co-denatured. Probes were hybridized overnight at 37°C. Slide was washed in 2X SSC/0.3% Igepal® CA-630 (Sigma-Aldrich, USA) at 75°C then at room temperature for 2 minutes and 1 minute respectively. The tissue was counterstained with 4', 6-diamidino-2-phenylindole (DAPI) antifade solution (Vectashield, USA) and analyzed under an epifluorescence microscope. Signals from 60 non-overlapping nuclei were enumerated for copy numbers of EGFR and CEP 7. EGFR amplification was defined with the ratio of EGFR to CEP7 of 2 and above.

2.2.5.3 Immunohistochemistry

Immunohistochemistry was performed using CDK18, MDM4, RAF1, EGFR and p16 primary antibodies. Please refer to section 2.3.3 for further details.

2.3 Prognostic importance of CD117, Six1 and Pax3

2.3.1 Selection of cases

A subset of cases was used for this part of the study. All cases diagnosed from January 2003 to December 2010 were included.

2.3.2 Tissue microarrays (TMAs)

Tissue microarrays (TMAs) were constructed using the Manual Tissue Arrayer MTA-1 (Beecher Instruments Inc., Sun Prairie, USA). A brief summary is illustrated in Figure 2.3. First, three representative areas comprising high stromal cells density were marked on selected H&E slides for each case by a histopathologist. Then, the marked areas were identified on the corresponding FFPE blocks and punched with a 2.0mm diameter core. Before each core was transferred to the recipient block, a core wax from the recipient block was removed with the recipient puncher. The three cores were arrayed to three recipient blocks respectively. A maximum of 40 cores were arrayed in a single block, with a tonsil core as an orientation marker.

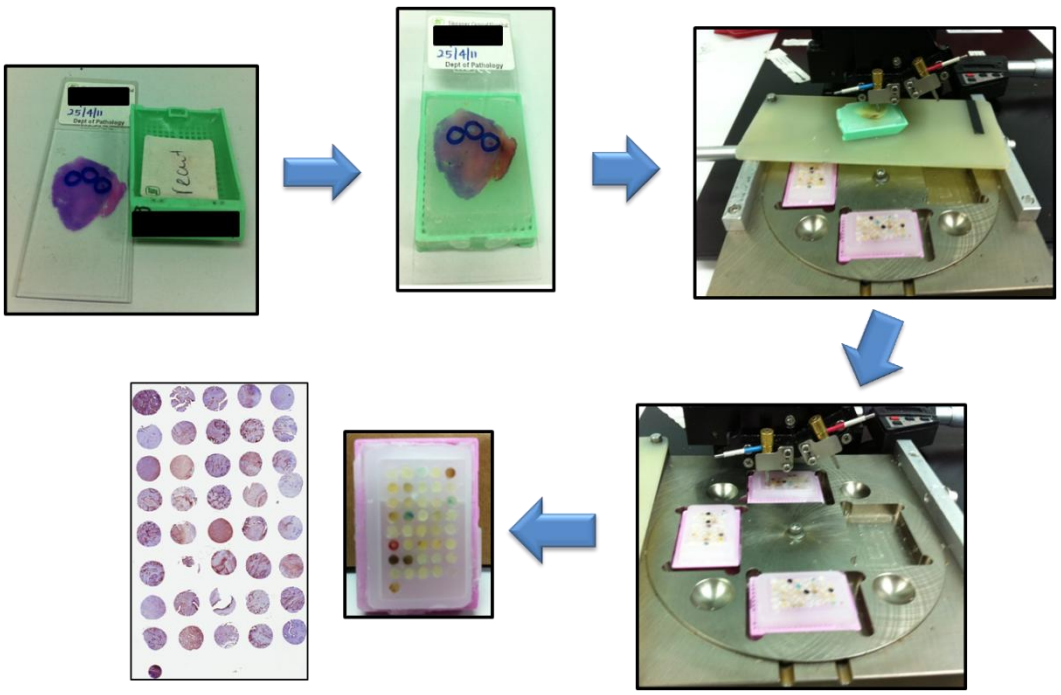

Figure 2.3 Construction of a tissue microarray. First, representative areas are marked by a histopathologist. Then the areas are identified on the corresponding green FFPE tissue block (donor block). The donor block is punched with a 2.0mm core (blue-coded puncher), and then transferred to a recipient wax block (purple as shown in figure) where a paraffin cylinder was previously removed by the recipient puncher (red-coded puncher). A maximum of 40 cores are arrayed in a single block with a tonsil core as an orientation marker (bottom left corner).

2.3.3 Immunohistochemistry

2.3.3.1 Immunohistochemical staining

The procedure of immunohistochemistry was optimized on three automated platforms for respective antibodies (please see Table 2.3). Antibody was diluted with 'Antibody Diluent with Background Reducing Components' (Dako, Glostrup, Denmark) to the optimized dilution. The three platforms used in this study are Autostainer Plus (Dako, Glostrup, Denmark), BenchMark ULTRA (Ventana Medical Systems, Arizona, USA), and BOND-MAX™ (Leica Biosystems, Germany).

The fundamental procedures for all three platforms are similar (Figure 2.4). First, 4µm thick tissue section on charged slides (Microsystems Plus Slides, Leica Biosystems, Germany) were deparaffinized and underwent pre-treatment for antigen retrieval. Endogenous peroxidase activity was blocked with hydrogen peroxide. Then primary antibody was added and incubated according to the optimized protocol. Secondary antibody conjugated with horse radish peroxidase (HRP) was applied after which 3, 3'-diaminobenzidine (DAB) was added. Counter-staining with haematoxylin was subsequently performed before sections were dehydrated and mounted with DPX and cover slips. All runs were performed together with controls as listed in Table 2.4.

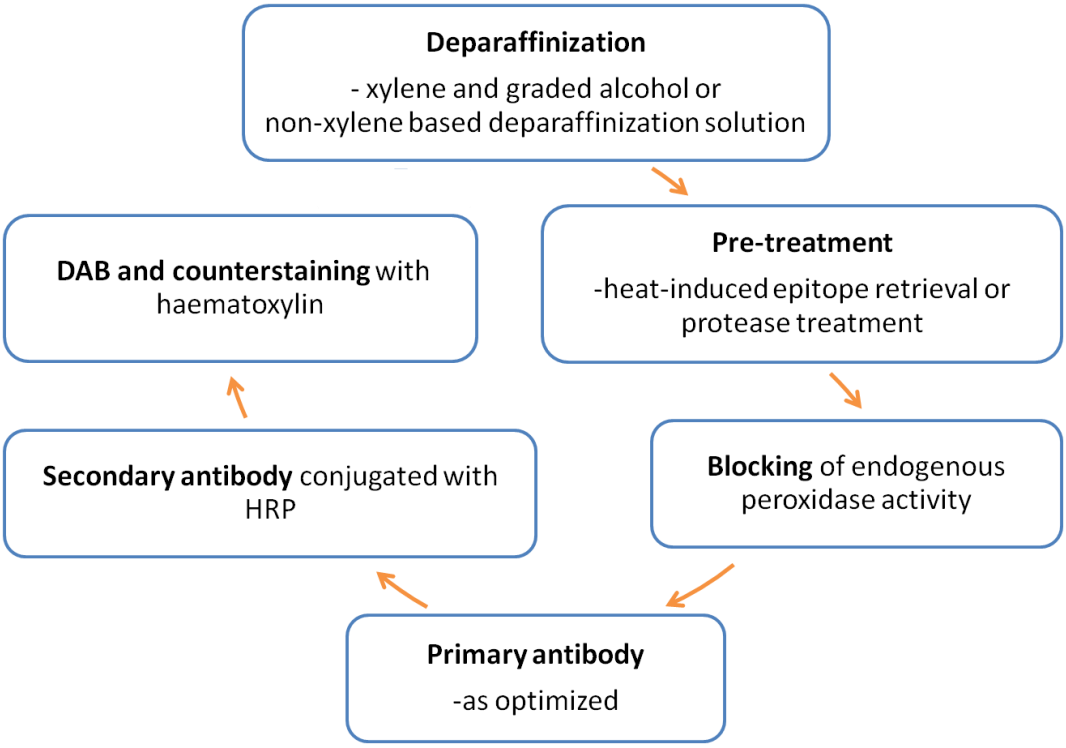

Figure 2.4 Procedure of immunohistochemistry staining

Table 2.3 Antibody details and optimized protocol for immunohistochemistry staining

Antibody	Source	Catalogue No.	Dilution	Time	Pre-treatment	Automation system	Secondary antibody and visualization
CD117	Dako	A4502	1:400	20 min	ER2 at 100°C 20min	Leica BOND-MAX™	Bond™ Polymer Refine Detection
CDK18	Sigma	HPA045429	1:100	60 min	Ct at 120°C 5 min	Dako Autostainer Plus	Dako EnVision® kit
EGFR	Dako	M7239	1:50	32 min	Protease 1 at RT 14 min	Ventana BenchMark ULTRA	Ventana ultraView Universal DAB Detection Kit
MDM4	Abcam	ab84393	1:100	30 min	Ct at 120°C 5 min	Dako Autostainer Plus	Dako EnVision® kit
p16	Santa Cruz	sc-56330	1:100	20 min	ER2 at 100°C 20 min	Leica BOND-MAX™	Bond™ Polymer Refine Detection
Pax3	Abcam	ab15717	1:150	30 min	TE at 98°C 15 min	Dako Autostainer Plus	Abcam anti-goat IgG 1:100 30min and DAB
Raf1	Sigma	HPA002640	1:50	30 min	Ct at 98°C 15 min	Dako Autostainer Plus	Dako EnVision® kit
Six 1	Sigma	HPA001893	1:100	60 min	Ct at 98°C 15 min	Dako Autostainer Plus	Dako EnVision® kit

Sigma: Sigma-Alrich; min: minutes; ER2: epitope retrieval solution 2 (pH9.0) ; TE: Tris-EDTA buffer (pH9.0); Ct: Citrate buffer (pH6.0); RT: room temperature; DAB: 3, 3'-diaminobenzidine

Table 2.4 Positive and negative controls used for each antibody and the cellular localization for the antibody expression

Antibody	Positive control	Negative control	Cellular localization	Threshold defined
CD117	Gastrointestinal stromal tumour	Lymphocytes	Cytoplasm and/or membrane	Positive: at least 1% or more stroma cells reactive
CDK18	Normal skin tissue	Omitting primary antibody	Cytoplasm	Nil
EGFR	Invasive ductal carcinoma	Lymphocytes	Cytoplasm and membrane	Nil
MDM4	Ovarian carcinoma	Omitting primary antibody	Nucleus	Nil
p16	Breast epithelium	Lymphocytes	Nucleus and cytoplasm	Nil
Pax3	Normal esophagus tissue	Lymphocytes	Nucleus and cytoplasm	High: Mean H-score of 14
Raf1	Normal kidney tissue	Omitting primary antibody	Nucleus or cytoplasm	Nil
Six 1	Normal post nasal space tissue	Lymphocytes	Nucleus and cytoplasm	High: Mean H-score of 63

2.3.3.2 Evaluation of staining

Slides were scanned using the ScanScope System (Aperio Technologies, Inc., USA) and viewed using the ImageScope software (Aperio). Staining was evaluated by eye-balling and scored by two blinded observers independently. Disagreements were reviewed jointly and resolved by consensus [Pinder et al., 2013].

Cellular localization of each stain evaluated is as listed in Table 2.4. Reactivity in each localization was assessed separately except for CD117, where cytoplasm and membrane were assessed together. Triplicates of cores were scored separately and the core with the highest expression was used for statistical analysis. Staining was evaluated in two aspects where appropriate – intensity of staining and percentage of tumour cells stained. Staining intensity was graded 0, 1+, 2+ and 3+ for nil, weak, moderate and strong intensity respectively. Percentage of tumour cells stained was estimated numerically. Semi-quantitative H-score [McCarty et al., 1985] ranging from 0 to 300 points was calculated as follows:

$$H - score = (1 \times \%\ of\ weakly\ stained\ tumour\ cells) + (2 \times \%\ of\ moderately\ stained\ tumour\ cells) + (3 \times \%\ of\ strongly\ stained\ tumour\ cells)$$

2.3.3.3 Toluidine blue staining

To further confirm that the CD117-positive cases were not confounded by mast cells, the sections were subjected to toluidine blue stain. Briefly, tissue sections of 4µm thick on glass slides were obtained. After deparaffinization with xylene and graded alcohols, tissue sections were incubated in 0.5% toluidine blue working solution for 6 minutes. Subsequently, slides were rinsed in changes of 95% alcohol immediately and dipped in 100% alcohol for one minute. Sections were then cleared with xylene and mounted with DPX mounting media. Bone marrow tissue containing mast cells were used as positive control. Presence of mast cells was identified by their metachromatic appearance (red-purple coloured) of the cytoplasm amidst the toluidine blue-stained background.

2.3.4 Statistical Analysis

Statistical analysis was performed with SPSS for Windows, version 18. Chi-square and Fisher's exact tests were used to analyze associations between categorical data. Means between groups were compared using analysis of variance (ANOVA). Kaplan-Meier survival curves were used to estimate recurrence-free survival and overall survival, defined as the time from date of surgery to date of first relapse and death from phyllodes tumour, respectively, or to the last follow-up date for censored cases. Log-rank test was performed to compare survival between groups. A p-value of less than 0.05 was considered a significant result.

3.0 RESULTS

3.1 Proposal of a predictive clinical model

3.1.1 Characteristics of the study population

A total of 605 cases of phyllodes tumours were diagnosed from January 1992 to December 2010 - 440 (72.7%) benign, 111 (18.4%) borderline and 54 (8.9%) malignant tumours. Patient ages ranged from 15 to 79, with a median age of 43 years. Clinicopathological characteristics stratified according to tumours are as shown in Table 3.1.1. Older patients and larger tumour sizes were observed more frequently with borderline and malignant grade tumours. Significant associations were observed with the five histological parameters and tumour grade, suggesting a good concordance with the grade classification. No significant association was observed with ethnicity and surgical margins with tumour grade.

Recurrences were documented in 80 (13.2%) cases. There was an increasing trend of recurrences observed with increasing grade - 48 (10.9%) benign grade tumours, 16 (14.4%) borderline grade tumours, and 16 (29.6%) malignant grade tumours. Mean and median times to recurrence were 29.8 and 24.6 months respectively. Of the 80 cases, 68 were local recurrences, 7 were distant recurrences, and 5 were local preceding distant recurrences (Table 3.1.1). Sites of distant recurrence which were histologically confirmed or radiologically detected include lung and pleura, liver, vertebra and soft tissue. Deaths were recorded in 19 women, 12 of which were due to

phyllodes tumours and the remaining seven from non-phyllodes tumour related causes.

Table 3.1.1 Clinicopathological features of 605 cases stratified according to tumour grade

Clinicopathological features	Benign (%) N=440	Borderline (%) N=111	Malignant (%) N=54	p-value
Age (years) (mean 42, median 43, range 15–79)				<0.001*
<mean age	227 (51.6)	36 (32.4)	16 (29.6)	
≥mean age	213 (48.4)	75 (67.6)	38 (70.4)	
Ethnicity				0.724
Chinese	307 (69.8)	81 (73.0)	37 (68.5)	
Malay	66 (15.0)	19 (17.1)	10 (18.5)	
Indian	31 (7.0)	3 (2.7)	0 (0)	
Others	36 (8.2)	8 (7.2)	7 (13.0)	
Size (mm) (mean 52, median 40, range 3–250)				<0.001*
<mean size	343 (78.0)	54 (48.6)	16 (29.6)	
≥mean size	97 (22.0)	57 (51.4)	38 (69.4)	
Stromal hypercellularity				<0.001*
Mild	302 (68.6)	21 (18.9)	3 (5.6)	
Moderate	131 (29.8)	75 (67.6)	25 (46.3)	
Marked	7 (1.6)	15 (13.5)	26 (48.1)	
Stromal atypia				<0.001[a]
Mild	412 (93.6)	54 (48.7)	0 (0)	
Moderate	28 (6.4)	54 (48.7)	30 (55.6)	
Marked	0 (0)	3 (2.6)	24 (44.4)	
Stromal overgrowth				<0.001[a]
Absent	440 (100)	86 (77.5)	14 (25.9)	
Present	0 (0)	25 (22.5)	40 (74.1)	
Stromal mitotic activity/10 hpf[b]				<0.001[a]
0–4	402 (91.4)	31 (27.9)	5 (9.2)	
5–9	34 (7.7)	51 (46.0)	15 (27.8)	
≥10	4 (0.9)	29 (26.1)	34 (63.0)	
Microscopic borders				<0.001[a]
Circumscribed	283 (64.3)	49 (44.1)	20 (37.0)	
Permeative	157 (35.7)	62 (55.9)	34 (63.0)	

Table 3.1.1 Clinicopathological features of 605 cases stratified according to tumour grade (continued)

Clinicopathological features	Benign (%) N=440	Borderline (%) N=111	Malignant (%) N=54	p-value
Surgical margins				0.511
Complete	257 (58.4)	58 (52.3)	34 (63.0)	
Focal involvement	168 (38.2)	44 (39.6)	16 (29.6)	
Diffuse involvement	15 (3.4)	9 (8.1)	4 (7.4)	
Recurrences				<0.001[a]
No	392 (89.1)	95 (85.6)	38 (70.4)	
Yes	48 (10.9)	16 (14.4)	16 (29.6)	
Local	48	16	4	
Distant	0	0	7	
Local and distant	0	0	5	

[a] statistical significance; [b] high power field

3.1.2 Factors affecting recurrences

3.1.2.1 Univariate analysis

A total of 552 patients were included in this analysis after discounting 19 non-local residents where follow-up data was not available and cases with follow-up period less than 3 months. Results of unadjusted univariate Cox analysis are shown in Table 3.1.2. Patients with borderline and malignant tumours had an increased risk of recurrence as compared to patients with benign tumours, with a hazard ratio of 1.67 and 3.83 respectively. Cases of involved surgical margins were more likely to encounter a recurrence as compared to cases without surgical margins involved, with a hazard ratio of 7.14. Stromal hypercellularity, stromal atypia, stromal overgrowth and nature of microscopic borders were also significantly associated with recurrence-free survival.

Table 3.1.2 Univariate analysis of features affecting recurrence free survival

Features	No. of patients	No. of events	Hazard ratio (95% CI)[c]	p-value
All	552	85		
Age	552	85	1.00 (0.98 to 1.02)	0.7006
Ethnicity				0.2095
Chinese	394	60	Reference	
Malay	89	17	1.29 (0.75 to 2.22)	
Indian	34	1	0.19 (0.03 to 1.39)	
Others	35	4	1.04 (0.38 to 2.87)	
Size	552	85	1.00 (1.00 to 1.01)	0.1725
Mitoses per 10 hpf[b]	552	85	1.05 (1.03 to 1.07)	<0.0001[a]
Grade				<0.0001[a]
Benign	399	46	Reference	
Borderline	103	17	1.67 (0.96 to 2.92)	
Malignant	50	19	3.83 (2.24 to 6.53)	
Stromal hypercellularity				0.0005[a]
Mild	302	31	Reference	
Moderate	208	41	2.22 (1.39 to 3.54)	
Marked	42	10	2.83 (1.39 to 5.78)	
Stromal atypia				<0.0001[a]
Mild	424	48	Reference	
Moderate	101	23	2.16 (1.316 to 3.55)	
Marked	27	11	4.42 (2.29 to 8.50)	
Stromal overgrowth				0.0003[a]
Absent	482	63	Reference	
Present	70	19	2.49 (1.49 to 4.15)	
Microscopic borders				0.0047[a]
Circumscribed	314	33	Reference	
Permeative	238	49	1.87 (1.20 to 2.91)	
Surgical margin				<0.0001[a]
Negative	314	15	Reference	
Positive	228	70	7.14 (4.07 to 12.52)	

[a] statistical significance; [b] high power field; [c] 95% confidence interval

3.1.2.2 Multivariate analysis

All features significantly associated with recurrence-free survival were incorporated for multivariate analysis (Table 3.1.3) except tumour grade due to issue of multicollinearity as tumour grade was derived from histological parameters. The multivariate analysis showed that stromal atypia, overgrowth and surgical margins affected recurrence-free survival significantly upon adjustment for interaction while stromal mitotic activity was close to statistical significance (p=0.058). Stromal hypercellularity and nature of microscopic borders were excluded in the final model as they did not affect recurrence-free survival significantly upon adjustment for interaction. Further, log-likelihood ratio test showed that adding the respective features into the model did not significantly improve the model (p=0.065 and p=0.329 respectively).

Table 3.1.3 Multivariate analysis of features affecting recurrence-free survival

Features	No. of patients	No. of events	Hazard Ratio (95% CI)[c]	p-value
Mitoses per 10 hpf[b]	552	82	1.03 (1.00 to 1.06)	0.0580
Surgical margin				
Negative	314	15	Reference	
Positive	238	67	8.37 (4.71 to 14.90)	<0.0001[a]
Stromal atypia				
Mild	424	48	Reference	
Moderate	101	23	1.79 (1.01 to 3.16)	0.0446[a]
Marked	27	11	3.28 (1.48 to 7.23)	0.0033[a]
Stromal overgrowth				
Absent	482	63	Reference	
Present	70	19	2.28 (1.19 to 4.36)	0.0126[a]

[a] statistical significance; [b] high power field; [c] 95% confidence interval

3.1.3 Nomogram

3.1.3.1 Constructing the nomogram

The multivariate Cox regression coefficients were used to derive the nomogram as shown in Figure 3.1.1a. Each feature corresponded to a weighted contribution in terms of points. For example, moderate atypia corresponds to 11 points, 10 mitoses/10hpf corresponds to 3 points, absent of overgrowth corresponds to zero point and positive surgical margin corresponds to 40 points. The sum of these points (Figure 3.1.1b) determined the likelihood of recurrence-free survival at 1 year, 3 years, 5 years, and 10 years. A case with accumulated total points of 54 will have a probability of recurrence-free survival of just above 0.9 at 1 year, 0.7 at 3 years, 0.58 at 5 years, and 0.5 at 10 years. The calculations can be automated through computerised programming for practical usage and are accessible at http://mobile.sgh.com.sg/ptrra/.

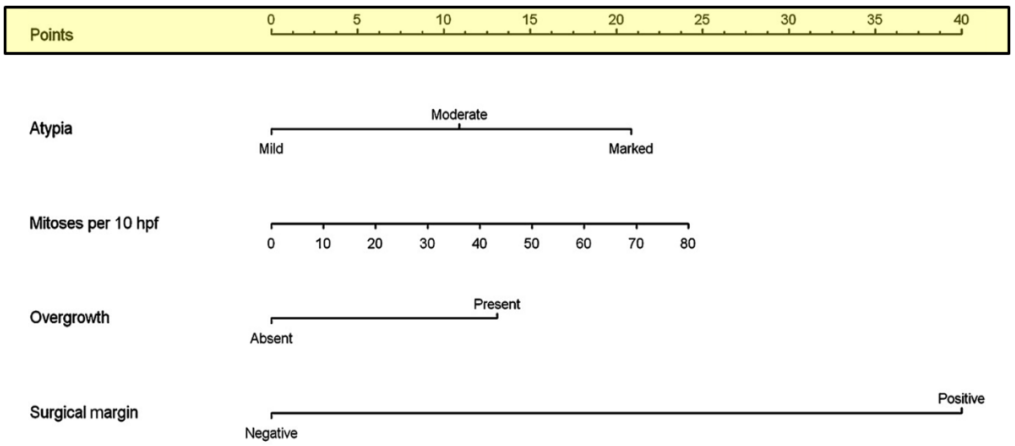

Figure 3.1.1a Nomogram for predicting recurrence-free survival of patients with phyllodes tumours. Each feature (atypia, mitoses per 10hpf, overgrowth and surgical margins) corresponds to a weighted contribution in terms of points (yellow bar). To use the nomogram, a line is drawn to the points bar for each feature and total points are calculated to be used for Figure 3.1.1b. For example, moderate atypia corresponds to 11 points, 10 mitoses/10hpf corresponds to 3 points, absence of overgrowth corresponds to zero point and positive surgical margin corresponds to 40 points. The total points from this case is 54.

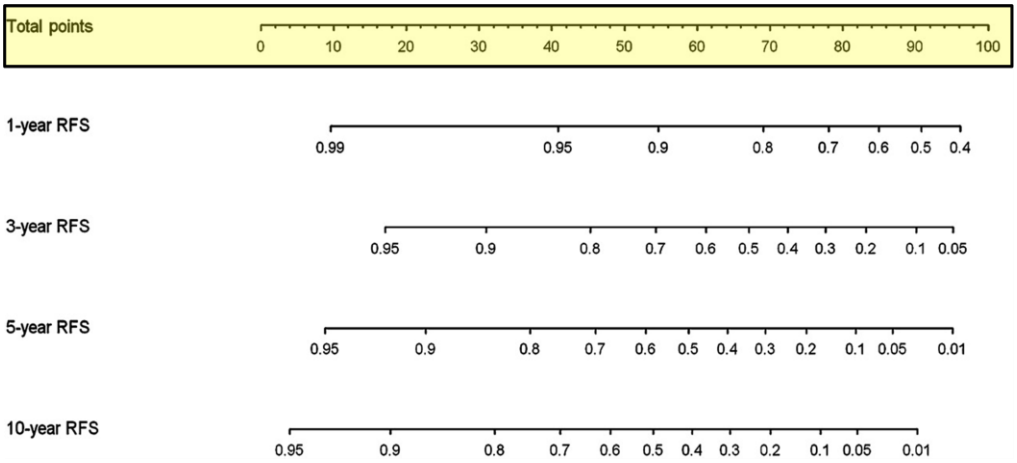

Figure 3.1.1b The total points accumulated from Figure 3.1.1a corresponds to the likelihood of recurrence-free survival at 1 year, 3 years, 5 years and 10 years. Draw a line from the total points bar (yellow) down for the respectively likelihood. For example the case which had total points of 54 will have a probability of recurrence-free survival of just above 0.9 at 1 year, 0.7 at 3 years, 0.58 at 5 years, and 0.5 at 10 years.

3.1.3.2 Validation of the nomogram

The nomogram's performance was evaluated through bootstrapping and the calibration plots of 200 bootstrap resamples are as shown in Figure 3.1.2. The mean and maximum deviations between observed and corrected outcomes are 0.03 and 0.02 at 1 year, 0.04 and 0.05 at 3 years, 0.04 and 0.07 at 5 years, and 0.04 and 0.08 at 10 years. The predictive accuracy of the nomogram measured by concordance index was 0.79.

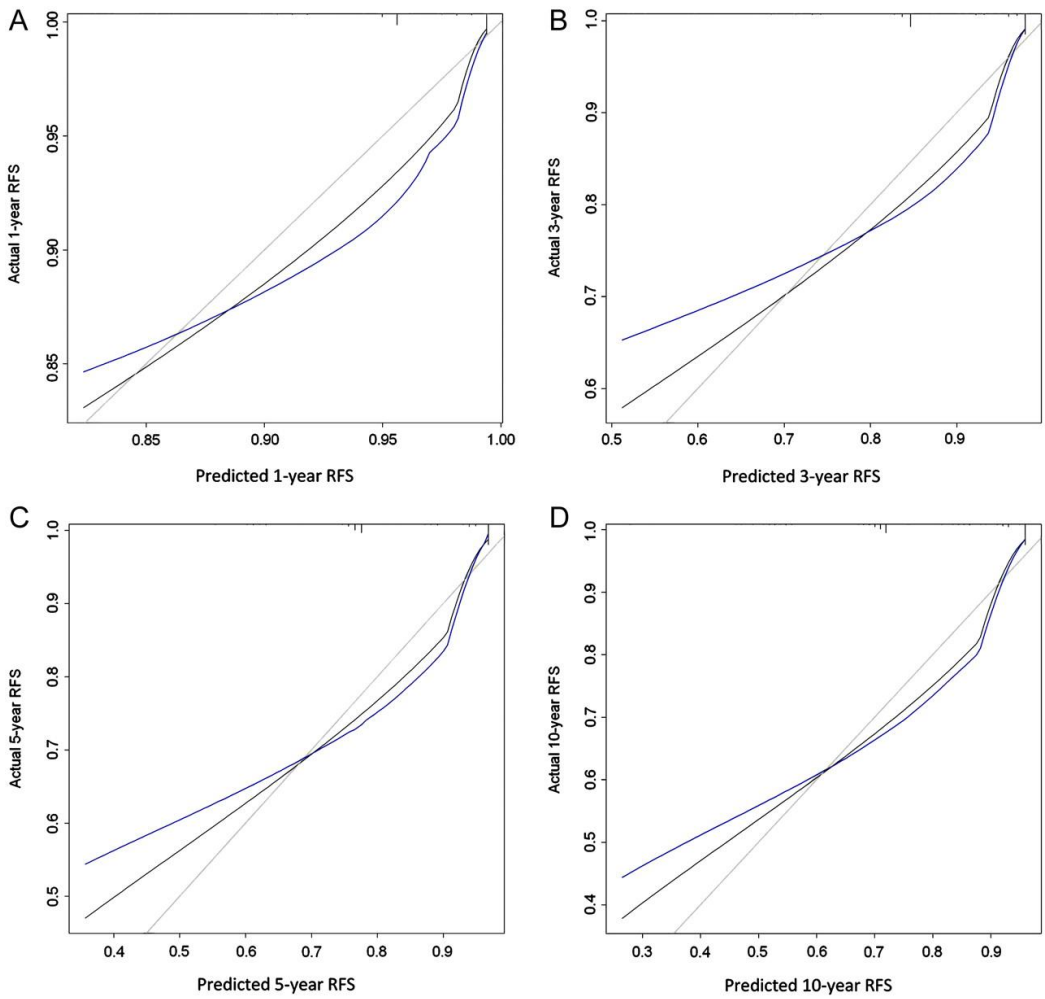

Figure 3.1.2 Calibration plots for the nomogram at 1 year (A), 3 years (B), 5 years (C), and 10 years (D). The ideal outcome, observed outcome, and the optimism corrected outcome are depicted as grey, black and blue lines. Mean and maximum deviations between observed and corrected outcomes are 0.03 and 0.02 at 1 year, 0.04 and 0.05 at 3 years, 0.04 and 0.07 at 5 years, and 0.04 and 0.08 at 10 years.

3.1.3.3 Comparison of nomogram with total histological score model

Both the nomogram and the total histological score model were significant as predictors of recurrence-free survival with hazard ratios of 1.05 (95% CI 1.04 to 1.06) and 1.27 (95% CI 1.15 to 1.40) respectively (Table 3.1.4). This means that for every point increase in the nomogram and total histological score model, there is an increase of risk of recurrence by 5% and 27% respectively. However, the nomogram has a range of scores between 0 and 100 while the total histological score model has a range of score between 5 and 13, accounting for the apparent higher hazard ratio in the total histological score model.

Table 3.1.4 Comparison of nomogram and histological score by hazard ratio and concordance index

	Hazard ratio[a] (95% CI)	**p-value**	**Concordance index**
Nomogram	1.05 (1.04 to 1.06)	<0.001	0.79
Total histological score	1.27 (1.15 to 1.40)	<0.001	0.65

[a]For every unit increase in score. It should be noted that the nomogram has a score range between 0 and 100, but the total histological score has a narrower range between 5 and 13, hence accounting for the apparently higher hazard ratio.

The Harrell's c-index of the nomogram was 0.79 while that of the total histological score model was 0.65, indicating a higher prediction accuracy of the nomogram as compared to the total histological score model. The likelihood ratio test also demonstrated a superior performance of the nomogram as compared to the total histological score model (Table 3.1.5). The individual nested nomogram and total histological score model, when compared to the full model, had an adequacy index of 99.9% and 22.7% respectively (Figure 3.1.3). Moreover, inclusion of the nomogram to the total histological score model significantly ($p<0.001$) improved the accuracy prediction but inclusion of the total histological score model to the nomogram did not improve accuracy prediction ($p=0.7026$).

Table 3.1.5 Comparison of nomogram and histological score by likelihood ratio test and adequacy index

Likelihood			**p-value**		**Adequacy index**	
Full model	Nomogram	Total score	Nomogram[a]	Total score[b]	Nomogram[c]	Total score[d]
94.15	94.01	21.38	<0.0001	0.7026	99.9%	22.7%

[a]Inclusion of nomogram significantly improves the prediction accuracy compared to a nested model of total score only
[b]Inclusion of the total score does not significantly improve the prediction accuracy compared to a nested model of nomogram only
[c]Percentage of variation explained by nomogram compared to the full model
[d]Percentage of variation explained by total score compared to the full model

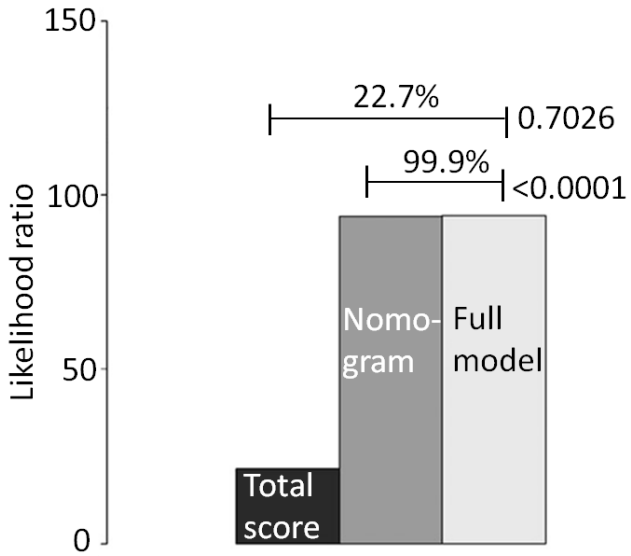

Figure 3.1.3 Comparison of the nomogram and the total histological score model. The predictive value of each model is represented by the likelihood ratio value (y-axis) and the adequacy index (%). 99.9% of the full model is attributable to the nomogram while only 22.7% of the full model is attributable to the total histological score model.

3.2 Genome-wide copy number and mutational analysis

3.2.1 Characteristics of the study population

Details of sample selected are as shown in Table 3.2.1. Tumour grade was significantly associated with recurrence/death. However, no significant differences were observed between the two groups in terms of pathological features except for mitotic activity (Table 3.2.2).

19 (95%) passed the quality control with median absolute pairwise difference (MAPD) ≤0.6 [Wang et al., 2007] and were further analysed for copy number changes and mutations. None of the microdissected tumours passed the required quantum for the assay (75ng) and hence not further analyzed.

Table 3.2.1 Features of 20 samples selected for genome-wide copy number and mutational analysis

No.	Age	Size (mm)	Ethnicity	Diagnosis	Prognosis
1	23	55	Chinese	Benign	No recurrence
2	50	100	Chinese	Benign	No recurrence
3	26	55	Malay	Benign	No recurrence
4	18	35	Malay	Benign	No recurrence
5	24	35	Indian	Benign	No recurrence
6	15	55	Chinese	Borderline	No recurrence
7	34	190	Chinese	Borderline	No recurrence
8	37	24	Chinese	Borderline	No recurrence
9	45	60	Malay	Borderline	No recurrence
10	55	250	Malay	Borderline	No recurrence
11	58	150	Chinese	Malignant	No recurrence
12	43	112	Chinese	Benign	Local Recurrence[a]
13	46	55	Chinese	Benign	Local Recurrence[a]
14	40	35	Chinese	Borderline	Local Recurrence[b]
15	44	210	Malay	Borderline	Metastasis preceded death
16	40	190	Chinese	Malignant	Metastasis preceded death
17	54	60	Chinese	Malignant	Death without recurrence
18	49	100	Others	Malignant	Lung metastasis
19	20	65	Others	Malignant	Lung metastasis
20	57	95	Malay	Malignant	Lung metastasis

[a] Recurred with grade progressed to borderline

[b] Recurred with increased size

Table 3.2.2 Clinicopathological characteristics of 19 phyllodes tumours stratified according to clinical behaviour

Features	Cases without recurrence (n=10)	Cases with recurrence/death (n=9)	p-value
Age (mean 38 years, median 40 years, range 15-58 years)			
≤median age	7 (70%)	3 (33.3%)	
>median age	3 (30%)	6 (66.7%)	0.179
Size (mean 94mm, median 60mm, range 24mm-250mm)			
≤median size	7 (70%)	3 (33.3%)	
>median size	3 (30%)	6 (66.7%)	0.179
Tumour grade			
Benign	5 (50%)	2 (22.2%)	
Borderline	5 (50%)	2 (22.2%)	
Malignant	0 (0%)	5 (55.6%)	0.023*
Stromal Hypercellularity			
Mild	4 (40%)	1 (11.1%)	
Moderate	5 (50%)	6 (66.7%)	
Marked	1 (10%)	2 (22.2%)	0.337
Stromal Atypia			
Mild	8 (80%)	4 (44.5%)	
Moderate	2 (20%)	3 (33.3%)	
Marked	0 (0%)	2 (22.2%)	0.175
Stromal Overgrowth			
Absent	9 (90%)	4 (44.4%)	
Present	1 (10%)	5 (55.6%)	0.057
Stromal Mitotic Activity/10 hpf			
0-4	8 (80%)	2 (22.2%)	
5-9	2 (20%)	4 (44.5%)	
≥10	0 (0%)	3 (33.3%)	0.027*
Microscopic Border			
Circumscribed	8 (80%)	3 (33.3%)	
Permeative	2 (20%)	6 (66.6%)	0.070
Haemorrhage			
Absent	6 (60%)	2 (22.2%)	
Present	4 (40%)	7 (77.8%)	0.170
Necrosis			
Absent	9 (90%)	5 (55.6%)	
Present	1 (10%)	4 (44.4%)	0.141
Surgical Margin			
Negative	6 (60%)	3 (33.3%)	
Positive	4 (40%)	6 (66.7%)	0.370

* significant statistically

3.2.2 Copy number analysis

Copy number changes were noted in all samples ranging from 1 to 83 events across tumours. An event was defined with two or more copy gains, or loss of single or both copies. Cases with recurrence and death harboured more aberrations with median events of 19 (range 0 – 72) compared to 3.5 (range 0 – 49) in the recurrence-free group. Fraction of genome altered (FGA) was also higher in the former group with median FGA of 6% as compared to 1.35% in the latter group. A summary of the alterations is shown in Table 3.2.3 and the details of aberrations for each sample are listed in Supplemental Table 1.

The most frequently observed loci of aberrations were 1q21.1, 2p11.1, 7q21.3, 7q33, 8p22, and 15q11.1 (Table 3.2.4). Gain of 1q21.1 and loss of 2p11.1 were more frequently observed in cases with recurrences/deaths. The function of genes within the 1q21.1 and 2p11.1 regions is largely unknown. However, 1q21.1 copy number variation was associated with neuroblastoma [Diskin et al., 2009] and hepatocellular carcinoma [Chen et al., 2010]. On the contrary, 7q21.3 gain was more frequently observed in the non-recurrent tumours. Though genes are also largely unknown in this region, a recent study revealed *SGCE* (sarcoglycan, epsilon) and *DYNC1I1* (dynein, cytoplasmic 1, intermediate chain 1) as a probable target gene at the 7q21.3 locus [Dong et al., 2011] in hepatocellular carcinoma.

Table 3.2.3 Summary of patient cohort and genomic alterations observed in 20 phyllodes tumours

No.	Diagnosis	Details	FGA[a]	Events		Amplification/Homozygous deletion (Candidate Gene)
				Gain	Loss	
Cases without recurrence			Med: 0.0135	Med: 3.5		
1	Benign	No recurrence	0.046	3	40	None
2	Benign	No recurrence	0.046	1	2	None
3	Benign	No recurrence	0.000	0	1	None
4	Benign	No recurrence	0.000	0	1	None
5	Benign	No recurrence	0.055	19	49	None
6	Borderline	No recurrence	0.007	2	1	None
7	Borderline	No recurrence	0.000	0	1	None
8	Borderline	No recurrence	0.000	2	2	None
9	Borderline	No recurrence	0.106	25	37	None
10	Borderline	No recurrence	0.020	0	4	None
11	Malignant	No recurrence			Failed quality control	
Cases with recurrence/metastasis/death			Med: 0.06	Med: 19		
12	Benign	Local Recurrence[b]	0.013	14	14	[A] 5p13.3 (*PDZD2*)
13	Benign	Local Recurrence[b]	0.021	0	4	[HD] 9p21 (*CDKN2A*)
14	Borderline	Local Recurrence[c]	0.077	0	2	None
15	Borderline	Metastasis preceded death	0.032	0	1	None
16	Malignant	Metastasis preceded death	0.320	22	24	[A] 7p12 (*EGFR*)
17	Malignant	Death without recurrence	0.240	11	72	[A] 3p25 (*RAF1*)
18	Malignant	Lung metastasis	0.060	15	4	[A] 1q32.1 (*MDM4*); [HD] 9p21 (*CDKN2A*)

Table 3.2.3 Summary of patient cohort and genomic alterations observed in 20 phyllodes tumours (continued)

No.	Diagnosis	Details	FGA[a]	Events		Amplification/Homozygous deletion (Candidate Gene)
				Gain	Loss	
Cases with recurrence/metastasis/death			Med: 0.06	Med: 19		
19	Malignant	Lung metastasis	0.138	5	39	[HD] 20p12.1 (*MACROD2*)
20	Malignant	Lung metastasis	0.003	1	2	None

[a] Fraction of Genome Altered (FGA) calculated as total lengths of gains and losses divided by the total genome length (3224.46Mb)

[b] Recurred with grade progressed to borderline [c]Recurred with increased size

[A] High-level amplification(>10copies gain); [HD] Homozygous deletion

Table 3.2.4 Six most frequently occurring cytogenetic alterations in 19 phyllodes tumours

Specific loci aberrations	Cases without recurrence (n=10)	Cases with recurrence/death (n=9)	Total (n=19)
1q21.1 gain	5 (50.0%)	8 (88.9%)	13 (68.4%)
2p11.1 loss	3 (30.0%)	7 (77.8%)	10 (52.6%)
7q21.3 gain	7 (70.0%)	4 (44.4%)	11 (57.9%)
7q33 gain	5 (50.0%)	4 (44.4%)	9 (47.4%)
8p22 gain	5 (50.0%)	5 (55.6%)	10 (52.6%)
15q11.1 loss	6 (60.0%)	5 (55.6%)	11 (57.9%)

3.2.2.1 High-level amplifications and candidate genes

High-level amplifications, defined as presence of more than 10 copies, were observed only in cases which recurred/died. Regions of amplification included chromosome 1q32.1, 3p25, 5p13.3, and 7p12 in four separate samples (Figure 3.2.1).

Amplification of chromosome 1q32.1 was observed in tumour sample #18, a malignant grade tumour which subsequently metastasized. A total of 71 genes were covered under the amplified region (Table 3.2.5), of which *MDM4* (MDM4, p53 regulator) and *CDK18* (cyclin-dependent kinase 18) were selected for further testing on immunohistochemistry (Figure 3.2.2). However, no overexpression of *MDM4* and *CDK18* on immunohistochemistry was observed, suggesting these genes were not overexpressed at the protein level.

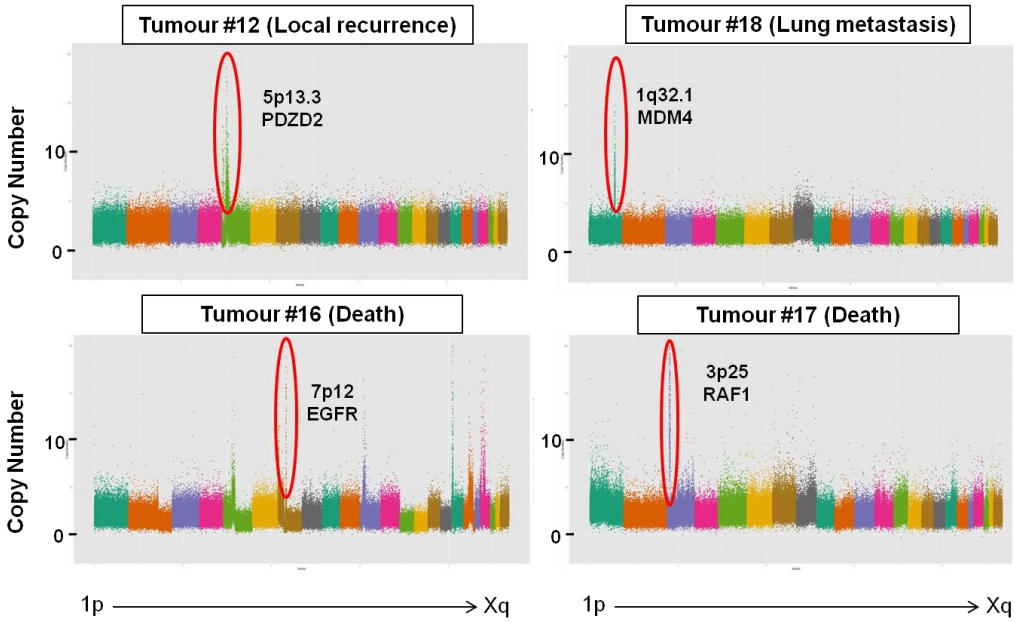

Figure 3.2.1 Manhattan plots showing copy number mapped to chromosomal location on the x-axis for four different samples. Each colour represents a chromosome on x-axis. Region of high level amplifications with candidate genes shown are highlighted in red circles.

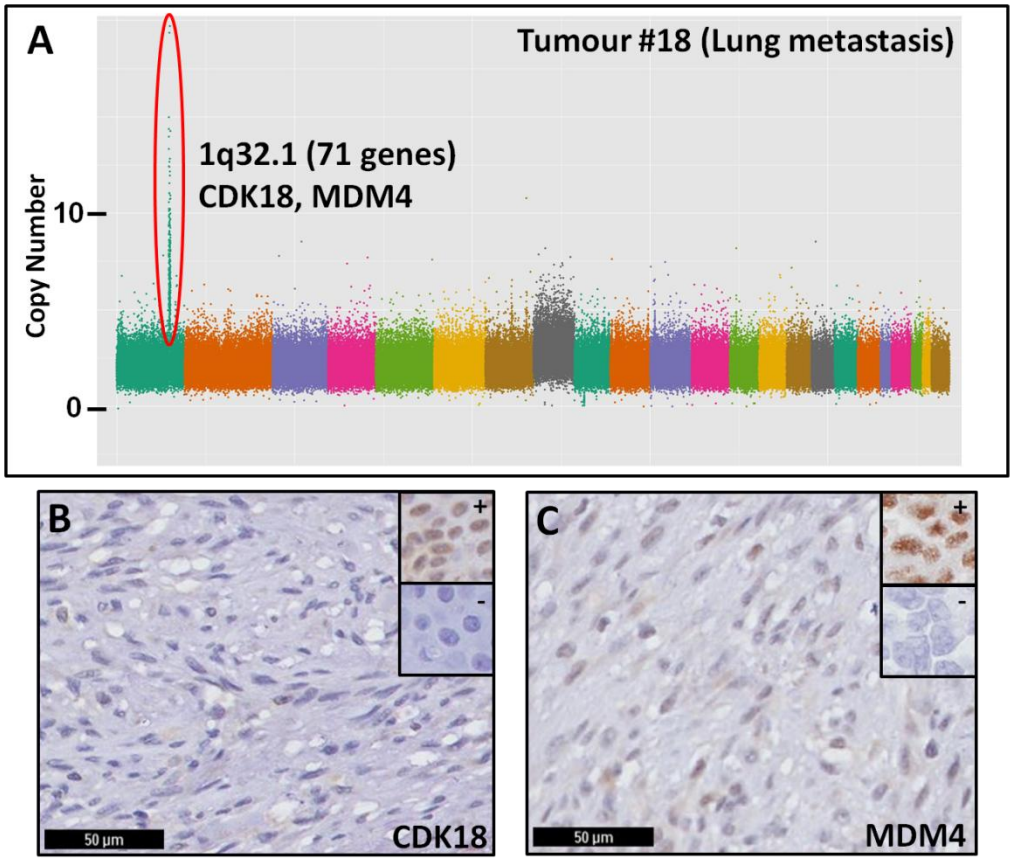

Figure 3.2.2 (A) Amplification of chromosome 1q32.1 was observed in sample #18 which subsequently metastasized. No overexpression of protein levels by candidate genes CDK18 (B) and MDM4 (C) were observed on immunohistochemistry, suggesting these genes were not amplified in this sample. Positive and negative controls are shown as insets. Scale bars: 50µm

Chromosome 3p25 was amplified in a case which the patient passed away without tumour recurrence. *RAF1* (Raf-1 proto-oncogene, serine/threonine kinase) was selected among the 53 candidate genes within the amplified region of 3p25 (Table 3.2.6) for immunohistochemical assessment. Patchy staining was observed on immunohistochemistry, indicating a low RAF1 expression in this sample (Figure 3.2.3).

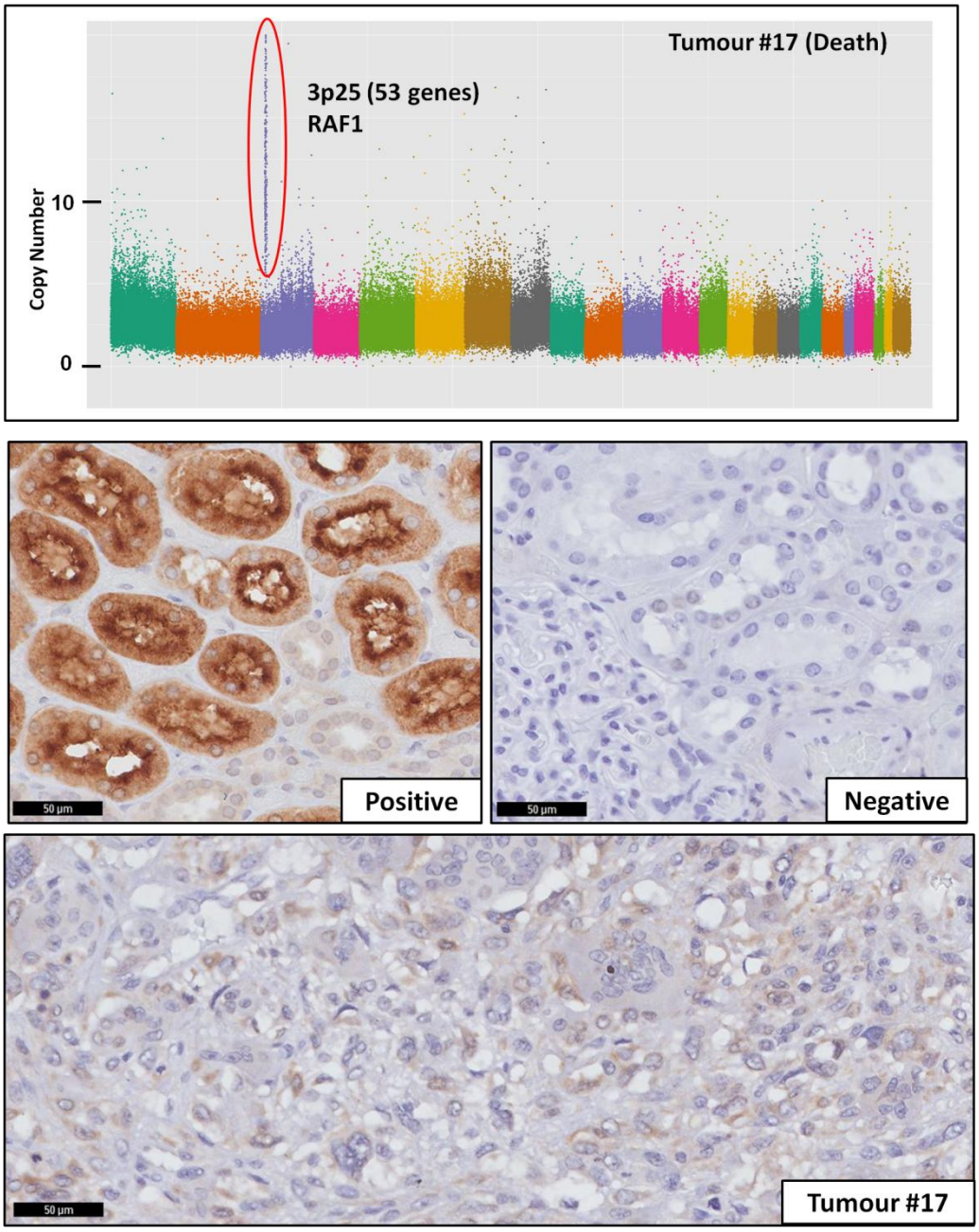

Figure 3.2.3 Amplification of chromosome 3p25 (top panel) was observed in sample #17 of which the patient died from the disease. Candidate gene RAF1 was however, patchy in this sample (bottom panel) and has a much weaker expression as compared to the positive control. Negative control is as shown on the right. Scale bar: 50µm.

Table 3.2.5 Candidate genes covered under the region of chromosome 1q32.1 for sample #18

No.	Gene Symbol	Start position	End position	Length	Name
1	**MIR5191**	201688635	201688755	121	microRNA 5191
2	**NAV1**	201617449	201796102	178654	neuron navigator 1
3	**RNU6-79**	201736888	201767778	30891	RNA, U6 small nuclear 79
4	**MIR1231**	201777738	201777830	93	microRNA 1231
5	**IPO9**	201798287	201853422	55136	importin 9
6	**SHISA4**	201857796	201861715	3920	shisa homolog 4 (Xenopus laevis)
7	**LMOD1**	201865583	201915716	50134	leiomodin 1 (smooth muscle)
8	**TIMM17A**	201924618	201939789	15172	translocase of inner mitochondrial membrane 17 homolog A (yeast)
9	**RNPEP**	201951765	201975275	23511	arginyl aminopeptidase (aminopeptidase B)
10	**ELF3**	201979689	201986315	6627	E74-like factor 3 (ets domain transcription factor, epithelial-specific)
11	**GPR37L1**	202092028	202098634	6607	G-protein coupled receptor 37 like 1
12	**ARL8A**	202102531	202113871	11341	ADP-ribosylation factor-like 8A
13	**PTPN7**	202116140	202130716	14577	protein tyrosine phosphatase, non-receptor type 7
14	**PTPRVP**	202137178	202158577	21400	protein tyrosine phosphatase, receptor type, V, pseudogene
15	**LGR6**	202163117	202288889	125773	leucine-rich repeat containing G protein-coupled receptor 6
16	**UBE2T**	202300784	202311094	10311	ubiquitin-conjugating enzyme E2T (putative)
17	**PPP1R12B**	202317829	202557697	239869	protein phosphatase 1, regulatory subunit 12B
18	**SYT2**	202559724	202679551	119828	synaptotagmin II
19	**KDM5B**	202696531	202777549	81019	lysine (K)-specific demethylase 5B
20	**KDM5B-AS1**	202780073	202781041	969	KDM5B antisense RNA 1

Table 3.2.5 Candidate genes covered under the region of chromosome 1q32.1 for sample #18 (continued)

No.	Gene Symbol	Start position	End position	Length	Name
21	**LOC641515**	202794328	202795421	1094	uncharacterized LOC641515
22	**LOC148709**	202830881	202844369	13489	actin pseudogene
23	**RABIF**	202847409	202858385	10977	RAB interacting factor
24	**KLHL12**	202860229	202896371	36143	kelch-like 12 (Drosophila)
25	**ADIPOR1**	202909959	202927700	17742	adiponectin receptor 1
26	**CYB5R1**	202931000	202936404	5405	cytochrome b5 reductase 1
27	**LOC401980**	202955579	202976393	20815	4933406M09Rik pseudogene
28	**TMEM183B**	202976535	202992668	16134	transmembrane protein 183B
29	**TMEM183A**	202976533	202993197	16665	transmembrane protein 183A
30	**PPFIA4**	203020310	203047864	27555	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 4
31	**MYOG**	203052256	203055166	2911	myogenin (myogenic factor 4)
32	**ADORA1**	203096835	203136533	39699	adenosine A1 receptor
33	**MYBPH**	203136938	203144942	8005	myosin binding protein H
34	**CHI3L1**	203148058	203155922	7865	chitinase 3-like 1 (cartilage glycoprotein-39)
35	**CHIT1**	203185206	203198860	13655	chitinase 1 (chitotriosidase)
36	**LOC730227**	203267885	203274453	6569	uncharacterized LOC730227
37	**BTG2**	203274663	203278729	4067	BTG family, member 2
38	**FMOD**	203309751	203320289	10539	fibromodulin
39	**PRELP**	203444882	203460479	15598	proline/arginine-rich end leucine-rich repeat protein
40	**OPTC**	203463270	203478077	14808	opticin
41	**ATP2B4**	203595914	203713209	117296	ATPase, Ca++ transporting, plasma membrane 4

Table 3.2.5 Candidate genes covered under the region of chromosome 1q32.1 for sample #18 (continued)

No.	Gene Symbol	Start position	End position	Length	Name
42	**SNORA77**	203698708	203698833	126	small nucleolar RNA, H/ACA box 77
43	**LINC00260**	203699704	203700979	1276	long intergenic non-protein coding RNA 260
44	**LAX1**	203734283	203745480	11198	lymphocyte transmembrane adaptor 1
45	**ZBED6**	203766650	203769590	2941	zinc finger, BED-type containing 6
46	**ZC3H11A**	203764750	203823256	58507	zinc finger CCCH-type containing 11A
47	**SNRPE**	203830739	203840280	9542	small nuclear ribonucleoprotein polypeptide E
48	**LINC00303**	204001574	204010392	8819	long intergenic non-protein coding RNA 303
49	**SOX13**	204042245	204096871	54627	SRY (sex determining region Y)-box 13
50	**ETNK2**	204100189	204121307	21119	ethanolamine kinase 2
51	**REN**	204123943	204135465	11523	renin
52	**KISS1**	204159468	204165619	6152	KiSS-1 metastasis-suppressor
53	**GOLT1A**	204167287	204183220	15934	golgi transport 1A
54	**PLEKHA6**	204187978	204329057	141080	pleckstrin homology domain containing, family A member 6
55	**LINC00628**	204337557	204338847	1291	long intergenic non-protein coding RNA 628
56	**PPP1R15B**	204372491	204380944	8454	protein phosphatase 1, regulatory subunit 15B
57	**PIK3C2B**	204391757	204459474	67718	phosphatidylinositol-4-phosphate 3-kinase, catalytic subunit type 2 beta
58	**MDM4**	204485506	204527248	41743	Mdm4 p53 binding protein homolog (mouse)
59	**LRRN2**	204586302	204654597	68296	leucine rich repeat neuronal 2
60	**NFASC**	204797781	204991950	194170	neurofascin
61	**CNTN2**	205012339	205047171	34833	contactin 2 (axonal)
62	**TMEM81**	205052256	205053588	1333	transmembrane protein 81

Table 3.2.5 Candidate genes covered under the region of chromosome 1q32.1 for sample #18 (continued)

No.	Gene Symbol	Start position	End position	Length	Name
63	**RBBP5**	205055269	205091150	35882	retinoblastoma binding protein 5
64	**DSTYK**	205111630	205180727	69098	dual serine/threonine and tyrosine protein kinase
65	**TMCC2**	205197037	205242471	45435	transmembrane and coiled-coil domain family 2
66	**NUAK2**	205271190	205290883	19694	NUAK family, SNF1-like kinase, 2
67	**KLHDC8A**	205305647	205326039	20393	kelch domain containing 8A
68	**LEMD1-AS1**	205342379	205356568	14190	LEMD1 antisense RNA 1
69	**LEMD1**	205350505	205391214	40710	LEM domain containing 1
70	**MIR135B**	205417429	205417526	98	microRNA 135b
71	**CDK18**	205473683	205501921	28239	cyclin-dependent kinase 18

Table 3.2.6 Candidate genes covered under the region of chromosome 3p25 for sample #17

No.	Gene Symbol	Start position	End position	Length	Name
1	**FGD5**	14860468	14976072	115605	FYVE, RhoGEF and PH domain containing 5
2	**FGD5-AS1**	14984285	14989948	5664	FGD5 antisense RNA 1
3	**NR2C2**	14989235	15090780	101546	nuclear receptor subfamily 2, group C, member 2
4	**MRPS25**	15090018	15106816	16799	mitochondrial ribosomal protein S25
5	**ZFYVE20**	15111579	15140655	29077	zinc finger, FYVE domain containing 20
6	**COL6A4P1**	15206868	15247466	40599	collagen, type VI, alpha 4 pseudogene 1
7	**CAPN7**	15247732	15294423	46692	calpain 7
8	**LOC100505696**	15295690	15306005	10316	uncharacterized LOC100505696
9	**SH3BP5**	15295862	15382901	87040	SH3-domain binding protein 5 (BTK-associated)
10	**IRAK2**	10206562	10285427	78866	interleukin-1 receptor-associated kinase 2
11	**TATDN2**	10290176	10322906	32731	TatD DNase domain containing 2
12	**GHRLOS2**	10326102	10327430	1329	ghrelin opposite strand RNA 2 (non-protein coding)
13	**GHRL**	10327433	10334631	7199	ghrelin/obestatin prepropeptide
14	**GHRLOS**	10327437	10335133	7697	ghrelin opposite strand/antisense RNA
15	**SEC13**	10342614	10362858	20245	SEC13 homolog (S. cerevisiae)
16	**MIR885**	10436172	10436246	75	microRNA 885
17	**ATP2B2**	10365706	10547268	181563	ATPase, Ca++ transporting, plasma membrane 2
18	**LINC00606**	10801168	10805877	4710	long intergenic non-protein coding RNA 606
19	**SLC6A11**	10857916	10980146	122231	solute carrier family 6 (neurotransmitter transporter, GABA), member 11
20	**SLC6A1-AS1**	11047783	11060910	13128	SLC6A1 antisense RNA 1

Table 3.2.6 Candidate genes covered under the region of chromosome 3p25 for sample #17 (continued)

No.	Gene Symbol	Start position	End position	Length	Name
21	**SLC6A1**	11034419	11080935	46517	solute carrier family 6 (neurotransmitter transporter, GABA), member 1
22	**HRH1**	11178778	11304939	126162	histamine receptor H1
23	**ATG7**	11314009	11599139	285131	autophagy related 7
24	**VGLL4**	11597543	11762220	164678	vestigial like 4 (Drosophila)
25	**TAMM41**	11831918	11888352	56435	TAM41, mitochondrial translocator assembly and maintenance protein, homolog (S. cerevisiae)
26	**SYN2**	12045861	12233532	187672	synapsin II
27	**TIMP4**	12194567	12200851	6285	TIMP metallopeptidase inhibitor 4
28	**PPARG**	12329348	12475855	146508	peroxisome proliferator-activated receptor gamma
29	**TSEN2**	12525930	12574820	48891	tRNA splicing endonuclease 2 homolog (S. cerevisiae)
30	**LOC100129480**	12581279	12586963	5685	uncharacterized LOC100129480
31	**MKRN2**	12598593	12625210	26618	makorin ring finger protein 2
32	**RAF1**	12625099	12705700	80602	v-raf-1 murine leukemia viral oncogene homolog 1
33	**THUMPD3**	9404716	9428475	23760	THUMP domain containing 3
34	**LOC440944**	9430536	9439174	8639	uncharacterized LOC440944
35	**SETD5**	9439402	9519838	80437	SET domain containing 5
36	**LHFPL4**	9540044	9595486	55443	lipoma HMGIC fusion partner-like 4
37	**MTMR14**	9691116	9744078	52963	myotubularin related protein 14
38	**CPNE9**	9745509	9771592	26084	copine family member IX
39	**BRPF1**	9773433	9789699	16267	bromodomain and PHD finger containing, 1
40	**OGG1**	9791627	9808353	16727	8-oxoguanine DNA glycosylase

Table 3.2.6 Candidate genes covered under the region of chromosome 3p25 for sample #17 (continued)

No.	Gene Symbol	Start position	End position	Length	Name
41	**CAMK1**	9799028	9811668	12641	calcium/calmodulin-dependent protein kinase I
42	**TADA3**	9821653	9834420	12768	transcriptional adaptor 3
43	**ARPC4**	9834178	9848789	14612	actin related protein 2/3 complex, subunit 4, 20kDa
44	**ARPC4-TTLL3**	9834231	9878040	43810	ARPC4-TTLL3 readthrough
45	**TTLL3**	9851643	9878040	26398	tubulin tyrosine ligase-like family, member 3
46	**RPUSD3**	9879532	9885702	6171	RNA pseudouridylate synthase domain containing 3
47	**CIDEC**	9908393	9921938	13546	cell death-inducing DFFA-like effector c
48	**JAGN1**	9932270	9936031	3762	jagunal homolog 1 (Drosophila)
49	**IL17RE**	9944295	9958084	13790	interleukin 17 receptor E
50	**IL17RC**	9958757	9975305	16549	interleukin 17 receptor C
51	**CRELD1**	9975523	9987097	11575	cysteine-rich with EGF-like domains 1
52	**PRRT3**	9987225	9994078	6854	proline-rich transmembrane protein 3
53	**PRRT3-AS1**	9989087	9996471	7385	PRRT3 antisense RNA 1

Sample #12 which recurred twice and progressed to a borderline grade harboured an amplification of chromosome 5p13.3. There were 19 candidate genes within the amplified region of 5p13.3 (Table 3.2.7). None were selected for further experimentation as no satisfactory antibody was available commercially.

Lastly, chromosome 7p12 was amplified in a sample of a patient who experienced metastasis and subsequently died from disease (sample #16). Nine genes were within the amplified region (Table 3.2.8) and *EGFR* (epidermal growth factor receptor) was selected for further testing on fluorescence *in situ* hybridization (FISH) and immunohistochemistry. *EGFR* amplification was confirmed on FISH with mean copy number of 25 per nucleus. The ratio of EGFR/CEP 7 signals enumerated from 60 nuclei was 8.3 (Figure 3.2.4). The sample exhibited strong EGFR staining on immunohistochemistry, indicating an overexpression of EGFR at the protein level.

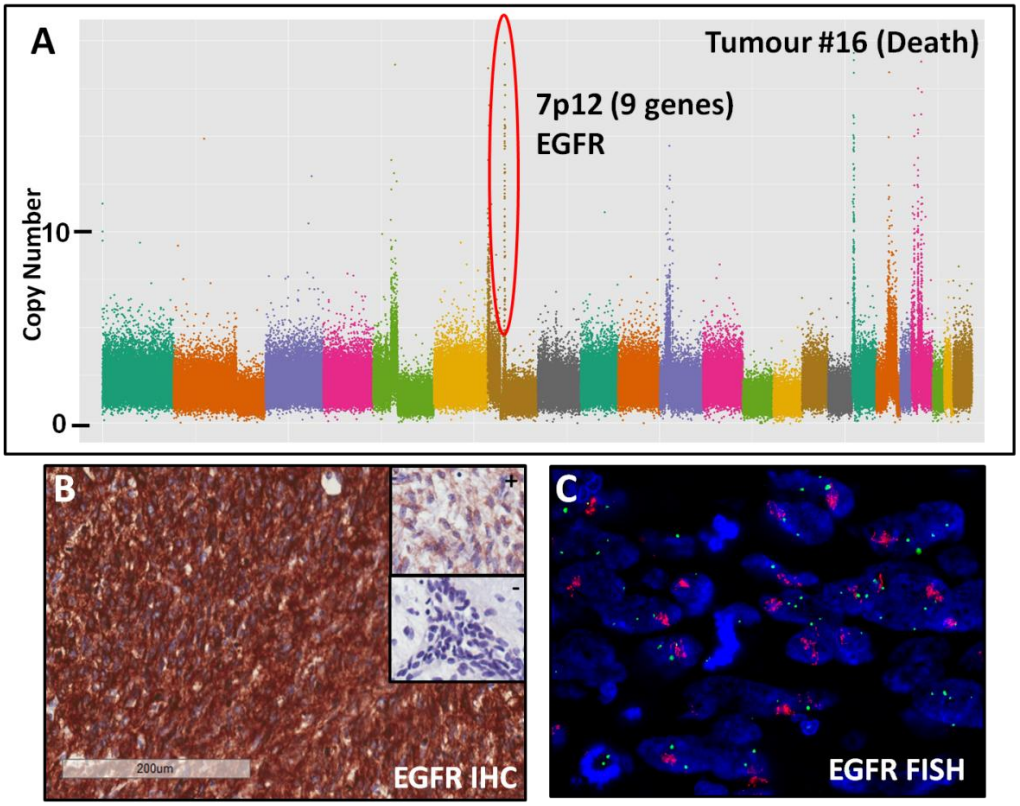

Figure 3.2.4 (A) Amplification of chromosome 7p12 was among other amplifications observed in sample #16 where the patient died from disease. (B) EGFR was overexpressed on immunohistochemistry with strong intensity as compared to an example of a positive case (inset). Scale bar: 200µm. (C) Fluorescence image showing amplification of *EGFR* with ratio of *EGFR* (red signals) to chromosome 7 (green signals) of 8.3 enumerated from 60 nuclei.

Table 3.2.7 Candidate genes covered under the region of chromosome 5p13.3 for sample #12

No.	Gene Symbol	Start position	End position	Length	Name
1	**CDH6**	31193761	31329253	135493	cadherin 6, type 2, K-cadherin (fetal kidney)
2	**DROSHA**	31400601	31532282	131682	drosha, ribonuclease type III
3	**C5orf22**	31532372	31555165	22794	chromosome 5 open reading frame 22
4	**MIR4279**	31936207	31936265	59	microRNA 4279
5	**PDZD2**	31799030	32111038	312009	PDZ domain containing 2
6	**GOLPH3**	32124823	32174425	49603	golgi phosphoprotein 3 (coat-protein)
7	**MTMR12**	32227110	32313114	86005	myotubularin related protein 12
8	**ZFR**	32354455	32444844	90390	zinc finger RNA binding protein
9	**SUB1**	32585604	32604185	18582	SUB1 homolog (S. cerevisiae)
10	**NPR3**	32710742	32791830	81089	natriuretic peptide receptor C/guanylate cyclase C (atrionatriuretic peptide receptor C)
11	**LOC340113**	32947548	32962573	15026	uncharacterized LOC340113
12	**TARS**	33440801	33468196	27396	threonyl-tRNA synthetase
13	**ADAMTS12**	33527286	33892124	364839	ADAM metallopeptidase with thrombospondin type 1 motif, 12
14	**RXFP3**	33936490	33939023	2534	relaxin/insulin-like family peptide receptor 3
15	**SLC45A2**	33944720	33984780	40061	solute carrier family 45, member 2
16	**AMACR**	33987090	34008220	21131	alpha-methylacyl-CoA racemase
17	**C1QTNF3**	34017962	34043371	25410	C1q and tumour necrosis factor related protein 3
18	**C1QTNF3-AMACR**	33987090	34124633	137544	C1QTNF3-AMACR readthrough
19	**RAI14**	34656432	34832717	176286	retinoic acid induced 14

Table 3.2.8 Candidate genes covered under the region of chromosome 7p12 for sample #16

No.	Gene Symbol	Start position	End position	Length	Name
1	**POM121L12**	53103348	53104618	1271	POM121 transmembrane nucleoporin-like 12
2	**FLJ45974**	53723201	53879624	156424	uncharacterized LOC401337
3	**HPVC1**	54268916	54270114	1199	human papillomavirus (type 18) E5 central sequence-like 1
4	**VSTM2A**	54610018	54636948	26931	V-set and transmembrane domain containing 2A
5	**LOC285878**	54624662	54639419	14758	uncharacterized LOC285878
6	**SEC61G**	54819939	54826939	7001	Sec61 gamma subunit
7	**EGFR**	55086724	55275031	188308	epidermal growth factor receptor
8	**LOC100507500**	55247442	55256642	9201	uncharacterized LOC100507500
9	**LANCL2**	55433140	55501435	68296	LanC lantibiotic synthetase component C-like 2 (bacterial)

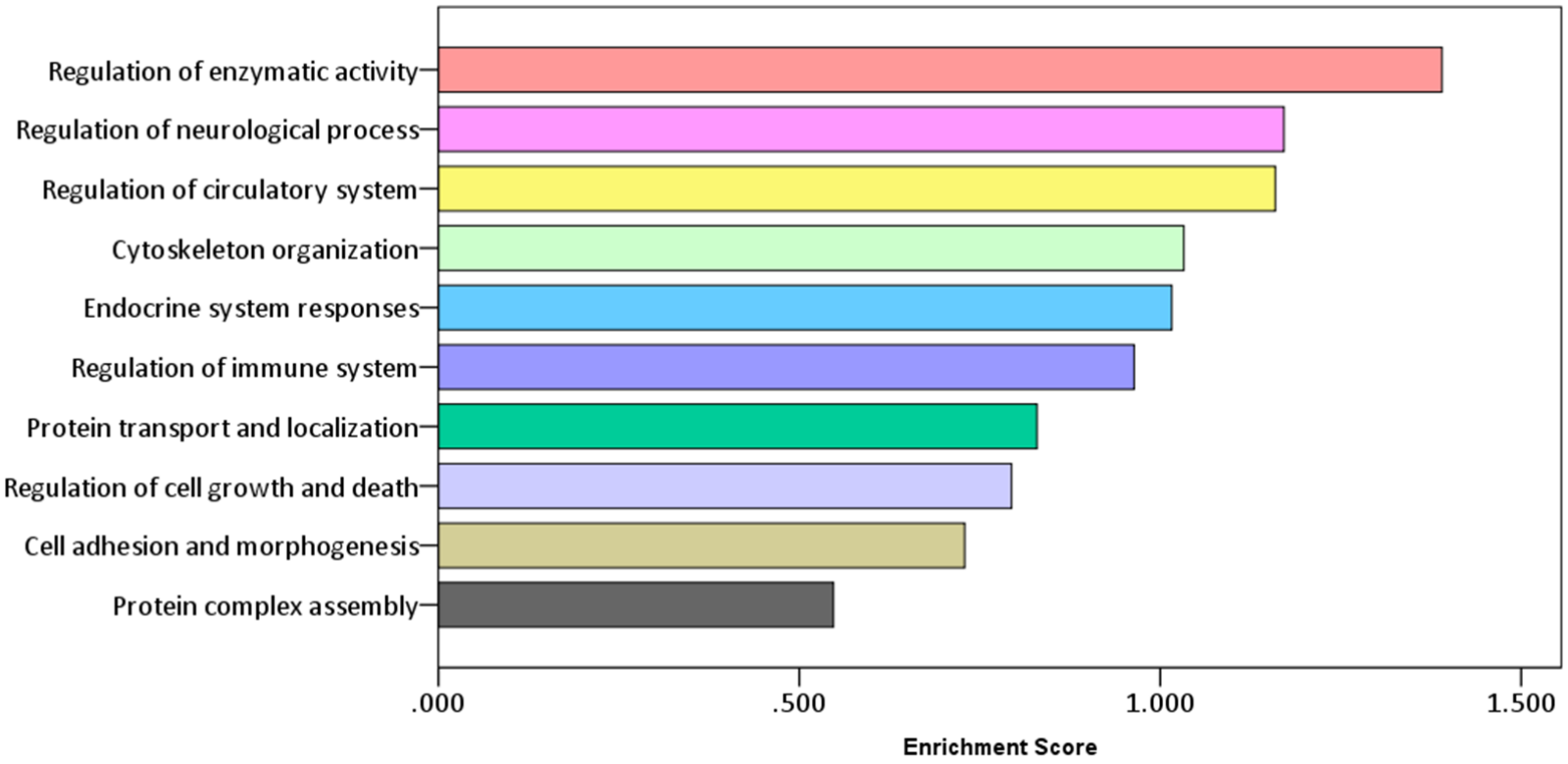

Figure 3.2.5 Functional classification of the 152 candidate genes under the amplified regions with the respective enrichment score

The 152 candidate genes under the amplified regions were classified into different functional groups using the DAVID functional annotation bioinformatics tool and were found to be highly involved in regulation of enzymatic activity, neurological processes and circulatory system (Figure 3.2.5).

3.2.2.2 Homozygous deletion

Two regions of homozygous deletions were observed – chromosome 9p21 involving *CDKN2A* (cyclin-dependent kinase inhibitor 2A) and chromosome 20p12.1. Chromosome 9p21 deletion was observed in two samples - one with a local recurrence of an increased grade and the other with metastasis. Both samples were negative when tested with CDKN2A (or more commonly known as p16) antibody on immunohistochemistry (Figure 3.2.6).

Homozygous deletion at chromosome 20p12.1 was observed in a case which had lung metastasis. Deletion at this region was frequently reported even in non-cancer patients and could be a result of genome instability [Bradley et al., 2010].

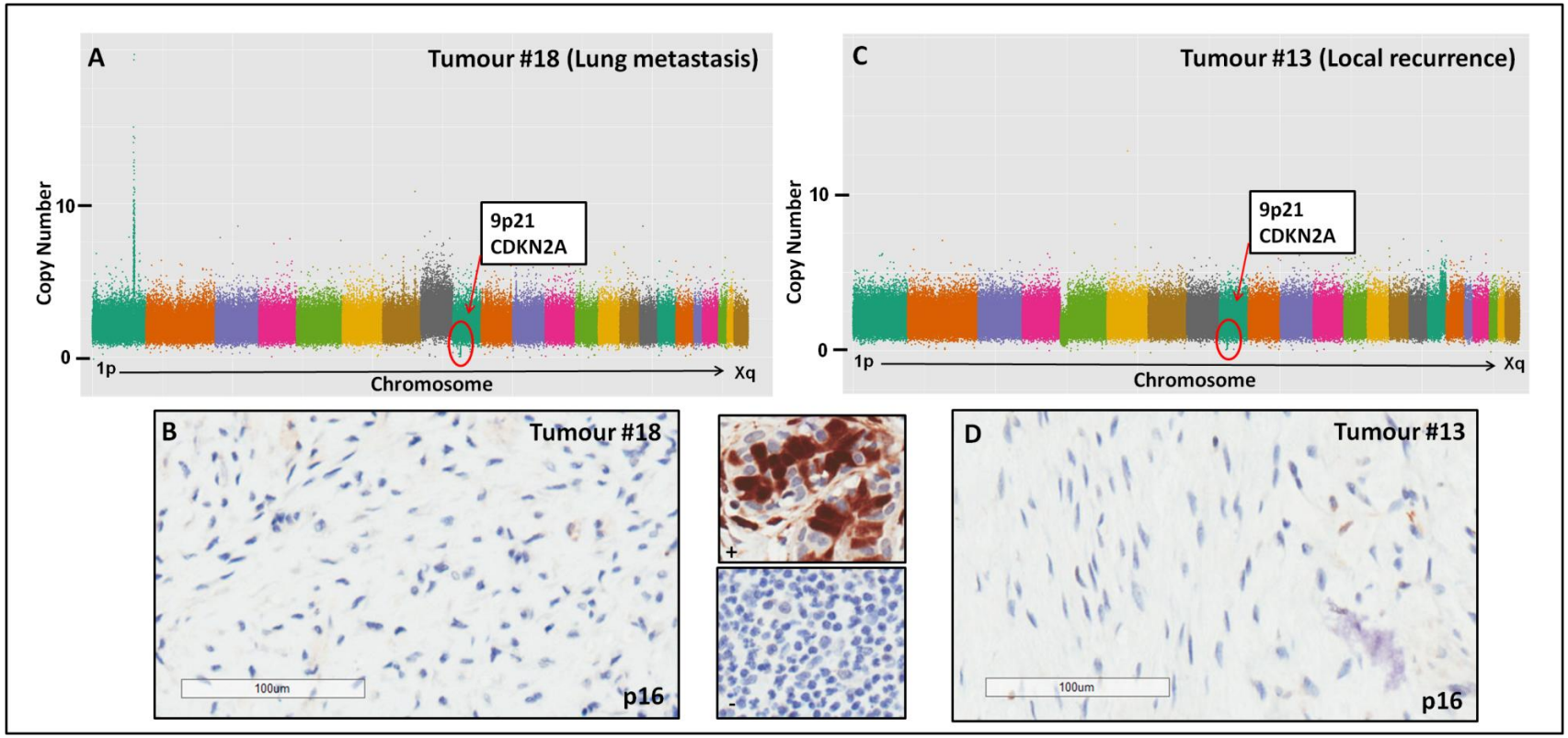

Figure 3.2.6 Homozygous deletion of chromosome 9p21 covering *CDKN2A* gene was observed in tumour #18 (A) and tumour #13 (C). Corresponding protein staining showed loss of expression in both tumours (B, D) Positive and negative controls as shown in between (B) and (D). Scale bars: 100µm.

3.2.3 Mutations and validation with Sanger Sequencing

None of the 541 mutation assays (Supplemental Table 2) for the 19 samples scored above 9 points, the recommended threshold for a mutation to be valid. However, the threshold for BRAF V600E assay was recommended at 4 points. Hence, to validate the absence of mutations in other assays, Sanger sequencing was performed on selected assays which had score 5 and above (Table 3.2.9). Results from Sanger sequencing corroborated the negative findings on OncoScan™ mutation assay. No mutations were identified. Examples of gel electrophoresis and chromatograms are shown in Figure 3.2.7.

Table 3.2.9 Mutation score (threshold>5) of each mutation detected on the OncoScan™ mutation panel

Tumour ID, prognosis / Mutations	#4 No recurrence	#6 No recurrence	#14 Local recurrence	#17 Died	#18 Metastasis	#19 Metastasis	#20 Metastasis	Total
ABL1_p.M351T_c.1052T<C	-	-	-	-	-	-	5.29[c]	1
APC_p.S1341R_c.4023T<G	-	-	-	-	-	-	6.11[c]	1
ATM_p.Q2442P_c.7325A<C	-	-	-	-	-	-	5.95[c]	1
ATM_p.T2666A_c7996A_G	-	-	-	5.74	-	-	-	1
BRCA1_c.134plus1G<T[a]	-	-	-	-	-	5.22	5.38	2
CDKN2A_pW110X_c329G_A	-	-	-	-	6.26	5.33	-	2
CDKN2A_c.151minus1G<A[a]	-	-	-	-	-	5.51	-	1
CTNNB1_p.G34E_c.101G<A	-	-	-	-	-	-	5.53[c]	1
EGFR_p.G598V_c.1793G<T	5.02	-	-	-	-	-	6.52	2
EGFR_p.L858R_c.2573T<G	-	-	-	-	-	-	6.09	1
FGFR3_p.K650Q_c.1948A<C	-	-	-	-	-	-	5.52[c]	1
KIT_p.V654A_c.1961T<C	-	-	-	-	-	-	6.13	1
MEN1_c.654plus3A<G	-	-	5.71	-	-	-	-	1
NF2_p.V219M_c.655G<A	-	5.49	-	-	-	-	-	1
NOTCH1_p.Q2460X_c.7378C<T	5.23	5.30	-	-	-	-	-	2
NPM1_p.W288 frame shift12_c.863insCCTG[b]	-	-	-	5.67	-	-	-	1
PTEN_p.I101T_c.302T<C	-	-	-	-	-	-	5.66[c]	1
PTEN_c.1027minus2A<G[a]	-	-	-	-	-	-	5.12[c]	1
RB1_p.E440X_c.1318G<T	-	-	-	5.30	-	-	-	1

Table 3.2.9 Mutation score (threshold>5) of each mutation detected on the OncoScan™ mutation panel (continued)

Tumour ID, prognosis / Mutations	#4 No recurrence	#6 No recurrence	#14 Local recurrence	#17 Died	#18 Metastasis	#19 Metastasis	#20 Metastasis	Total
RET_p.A664D_c.1991C<A	5.59	-	-	-	-	-	-	1
SMAD4_p.E330A_c.989A<C	-	-	-	-	-	-	6.13[c]	1
SMAD4_p.K507Q_c.1519A<C	-	-	-	-	-	-	7.50	1
TGFBR2_p.R497X_c.1489C<T	-	5.83	-	-	-	-	-	1
TP53_p.H179R_c.536A<G	-	-	5.33	-	-	-	5.19	2
TP53_p.C124R_c.370T<C	-	-	-	-	-	-	5.47	1
Total mutations	3	3	2	3	1	3	15	30

[a] Mutations at intronic regions

[b] Insertion of four bases (CCTG) between coding sequence position 863 and 864. No good sequencing data was obtained.

[c] Sanger sequencing not performed

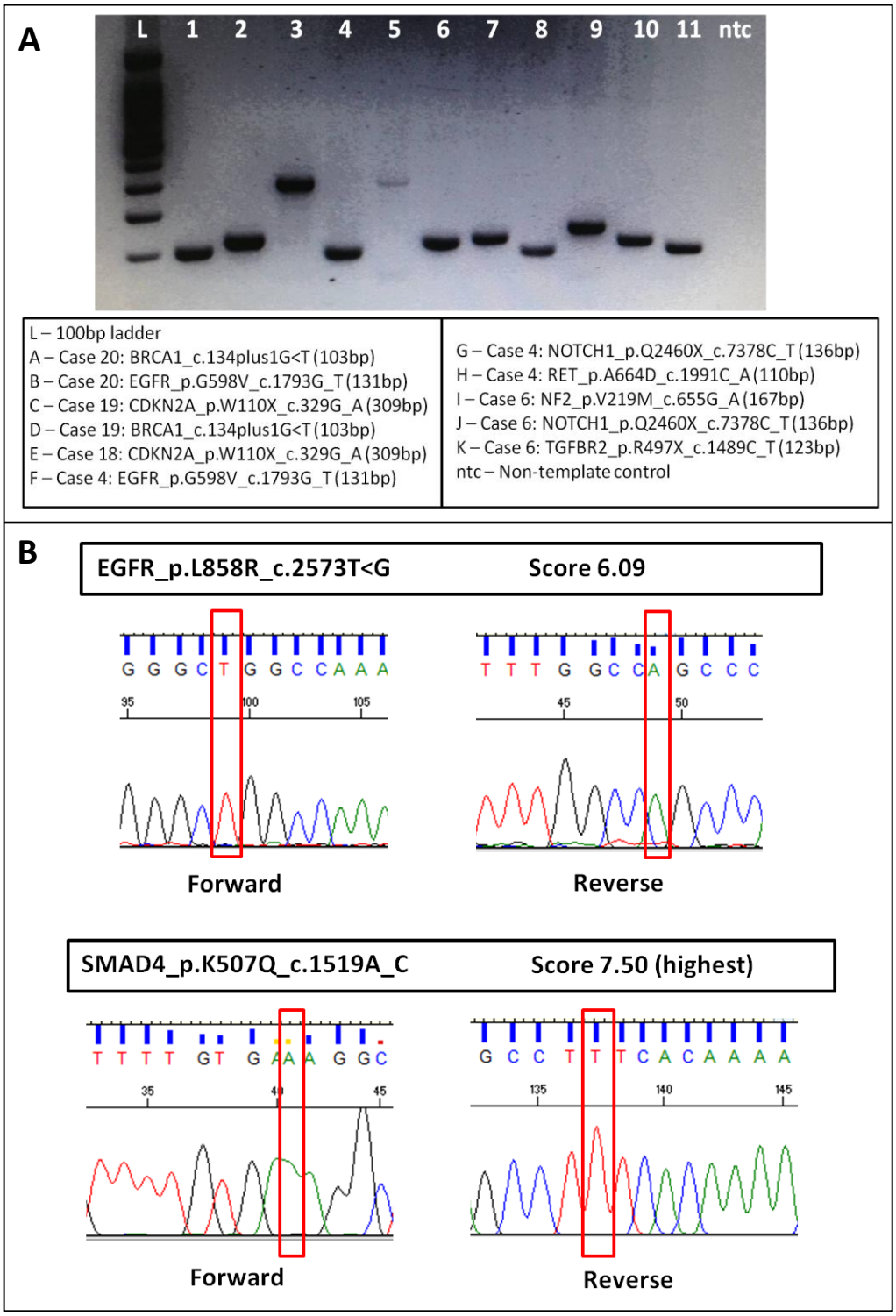

Figure 3.2.7 Polymerase chain reaction (PCR) was designed targeting the specific mutation region. Amplified products were sequenced and screened for presence of mutation. (A) Gel electrophoresis for the PCR products. (B) Examples of two assays with chromatograms showing targeted regions in red boxes. No mutation was observed in both assays.

3.3 CD117 protein expression and mutation status in phyllodes tumours

3.3.1 Characteristic of the study population

A total of 272 cases were screened for protein expression of CD117 on immunohistochemistry using tissue microarray sections. These cases constituted a series of consecutive cases diagnosed in the Department of Pathology, Singapore General Hospital from January 2003 to December 2010. Of the 272 cases, 189 (69.5%) were benign, 60 (22.1%) were borderline, and 23 (8.4%) were malignant. Median age of this cohort was 43 years, ranging from 15 to 79 years. Tumour size range was 8-250mm with a median of 38mm. Ethnic distribution of patients comprised 187 (68.8%) Chinese, 42 (15.4%) Malay, 15 (5.5%) Indian and 28 (10.3%) of other ethnic origins.

Data from 248 patients were available for follow-up analysis, after discounting cases of loss to follow-up and those in whom follow-up was less than three months. There were 24 (10%) recurrences in this series, of which 18 cases were local recurrences and 6 were distant recurrences. Five deaths were documented- one benign, two borderline, and two malignant cases. The benign case recurred with progression to borderline grade tumour and metastasis to the lung before death. One patient with borderline tumour died from acute pancolitis and the other patient with borderline tumour had lung metastasis leading to death. One malignant case had a lung metastasis preceding death, the other passed away without documented recurrence.

3.3.2 Immunohistochemistry results

Among the 272 cases, 28 cases (10%) were positive for CD117 of which 9 were benign, 14 were borderline and 5 were malignant tumours. Example of a positive case is shown in Figure 3.3.1c. As mast cells are also positively indicated by CD117 (Figure 3.3.1e), toluidine blue was performed to identify these cells which could be a confounding factor in assessing the staining status of stromal cells. The percentage of positively stained stromal cells ranged from 1% to 5% with a mean and median percentage of 1.4% and 1% respectively.

CD117 positivity was significantly associated with borderline/malignant tumours ($p<0.001$). Also, a significant association was observed with larger tumour size, increased cellularity, atypia, number of mitoses, permeative tumour margins, and presence of haemorrhage (Table 3.3.1). A trend of shorter time to recurrence was observed in tumours which expressed CD117 compared to tumours which did not express CD117 (Figure 3.3.2). In addition, patients with CD117-positive tumours had a worse overall survival compared to patients with CD117-negative tumours (Figure 3.3.3).

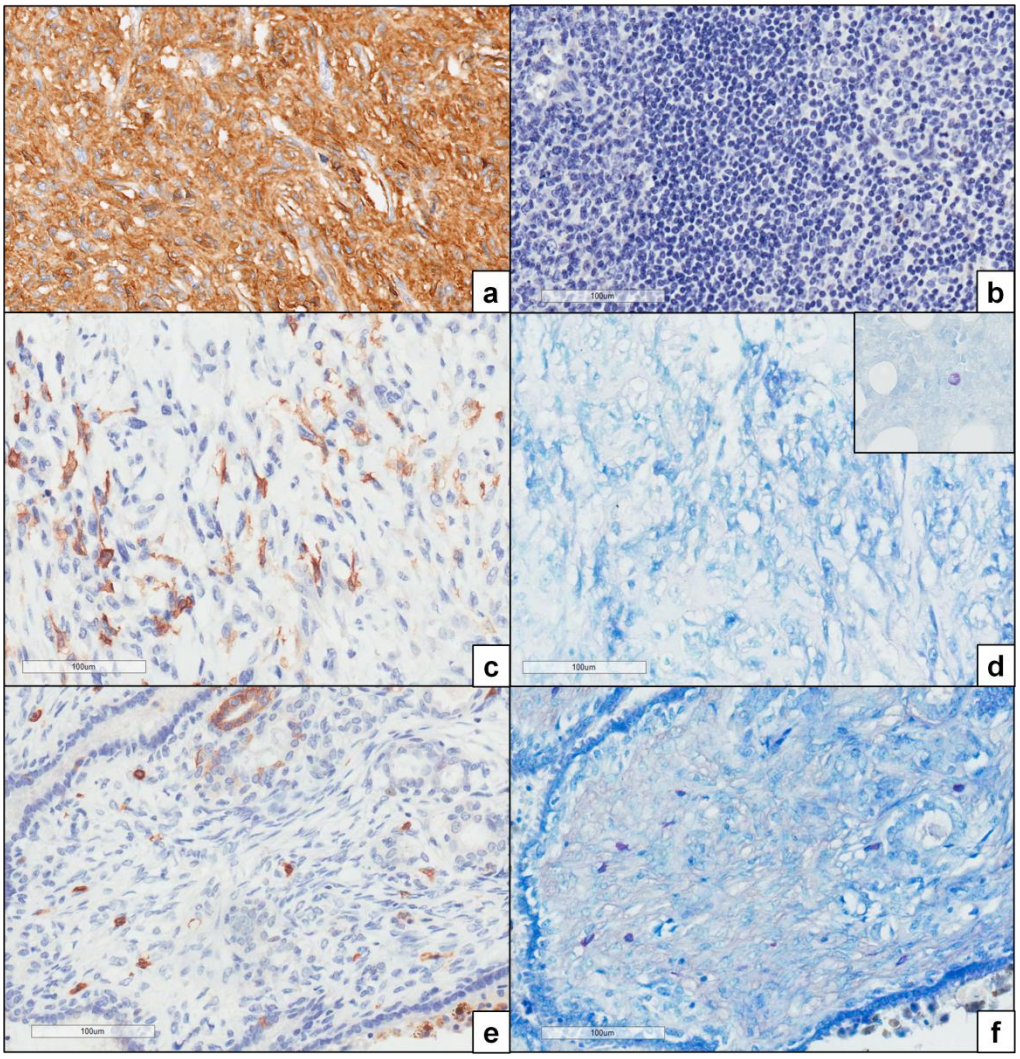

Figure 3.3.1 CD117 immunostaining and toluidine blue staining. (a) CD117 positive control. (b) CD117 negative control. (c, d) Example of a CD117-positive phyllodes tumour with the corresponding negative toluidine blue staining. Positive control with a purple metachromatic appearance of a mast cell is shown on the top right inset. (e, f) Mast cells which are positive for both CD117 and toluidine blue staining are excluded from the assessment in the stromal cells. Scale bars: 100µm

Table 3.3.1 Clinicopathological characteristics of phyllodes tumours in association with CD117 stromal positivity

Clinicopathological parameters	CD117 Negative (%)	CD117 Positive (%)	*p*-value
Age *(mean 43 years, median 43 years, range 15–79)*			
≤43 years	124 (90.5)	13 (9.5)	
> 43 years	120 (88.9)	15 (11.1)	0.694
Tumour Size *(mean 51mm, median 38mm, range 8–250mm)*			
≤51 mm	175 (93.1)	13 (6.9)	
>51 mm	69 (82.1)	15 (17.9)	0.009[a]
Tumour grade			
Benign	180 (95.2)	9 (4.8)	
Borderline	46 (76.7)	14 (23.3)	
Malignant	18 (78.3)	5 (21.7)	<0.001[a]
Stromal Hypercellularity			
Mild	142 (95.3)	7 (4.7)	
Moderate	87 (82.1)	19 (17.9)	
Marked	15 (88.2)	2 (11.8)	0.003[a]
Stromal Atypia			
Mild	207 (92.0)	18 (8.0)	
Moderate	31 (81.6)	7 (18.4)	
Marked	6 (66.7)	3 (33.3)	0.01[a]
Stromal Mitosis			
0-4	181 (94.8)	10 (5.2)	
5-9	39 (84.8)	7 (15.2)	
>9	24 (68.6)	11 (31.4)	<0.001[a]
Stromal Overgrowth			
No	218 (91.2)	21 (8.8)	0.059
Yes	26 (78.8)	7 (21.2)	
Microscopic Margin			
Circumscribed	192 (93.2)	14 (6.8)	
Permeative	52 (78.8)	14 (21.2)	0.002[a]
Necrosis			
Absent	217 (91.2)	21 (8.8)	
Present	27 (79.4)	7 (20.6)	0.062
Haemorrhage			
Absent	152 (95.0)	8 (5.0)	
Present	92 (82.1)	20 (17.9)	0.001[a]

[a] denotes statistically significant results

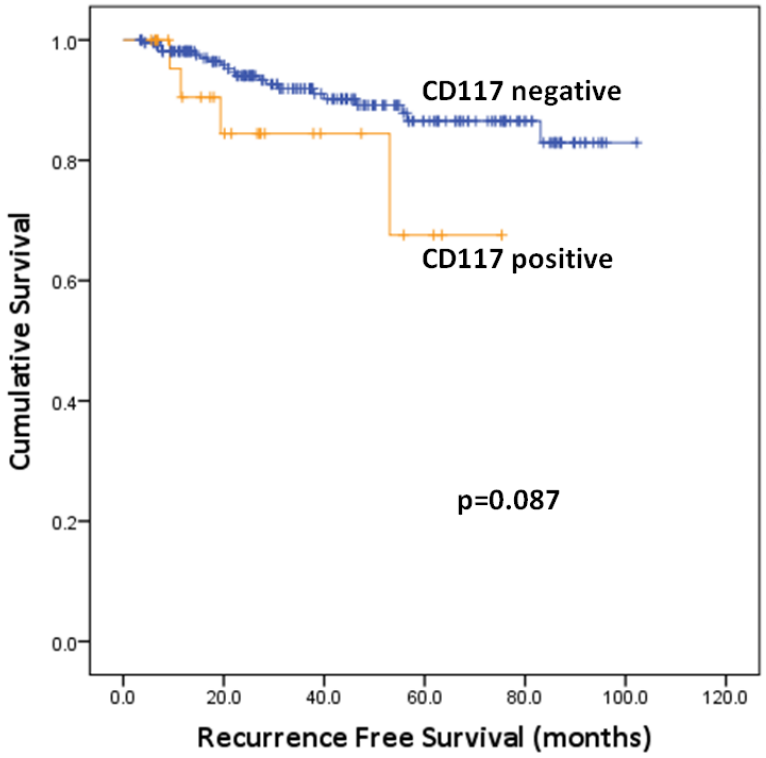

Figure 3.3.2 A shorter time to recurrence was observed in patients with CD117-positive tumours

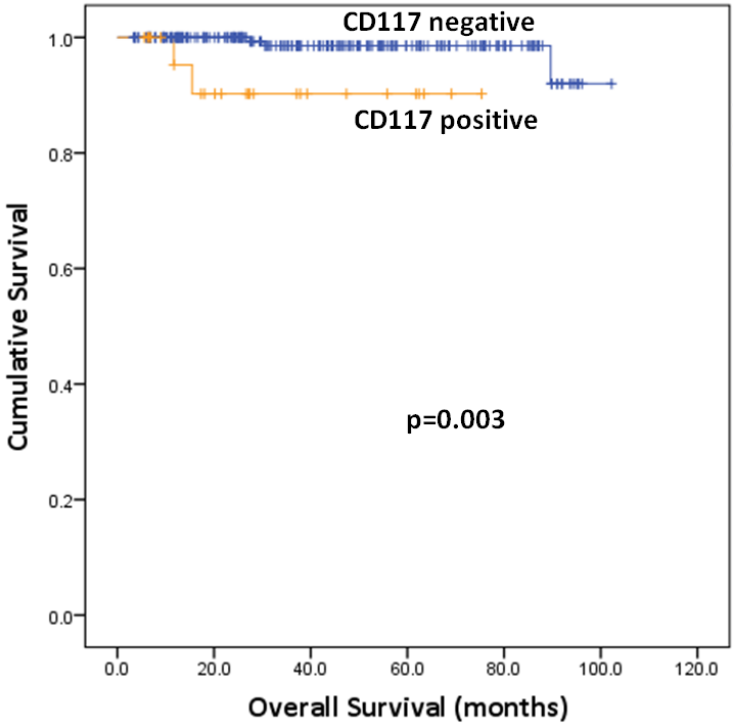

Figure 3.3.3 Patients with CD117-positive tumours had a worse overall survival compared to patients with CD117-negative tumours

3.3.3 Mutational analysis for CD117/KIT

Results from section 3.2.3 were analyzed further with a focus on the *KIT* gene. A total of 17 commonly reported mutations including the tyrosine kinase domains of exons 11, 13 and 17 were found in the panel of OncoScan™ FFPE assay (Table 3.3.2). The corresponding immunohistochemistry statuses of the 19 cases which did not harbour these mutations are shown in Table 3.3.3.

Table 3.3.2 17 commonly reported *KIT* mutations included in the OncoScan™ assay

No.	*KIT* mutations	Exon number
1	D52N	2
2	Y503_F504ins	9
3	W557R	11
4	V559A	11
5	V560D	11
6	L576P	11
7	F584S	11
8	P585P	11
9	K642E	13
10	V654A	13
11	T670I	13
12	I798I	13
13	D816Y	17
14	N822K	17
15	Y823D	17
16	V825A	17
17	E839K	17

Table 3.3.3 Staining status of KIT protein in the 19 cases interrogated on the OncoScan™ FFPE assay

Tumour	Diagnosis	Epithelial-stromal percentage	KIT protein staining status
1	Benign	40-60	Negative
2	Benign	15-85	Positive
3	Benign	35-75	Negative
4	Benign	40-60	Negative
5	Benign	30-70	Negative
6	Borderline	15-85	Positive
7	Borderline	20-80	Negative
8	Borderline	50-50	Positive
9	Borderline	30-70	Positive
10	Borderline	10-90	Negative
12	Benign	20-80	Negative
13	Benign	20-80	Positive
14	Borderline	20-80	Negative
15	Borderline	10-90	Positive
16	Malignant	0-100	Negative
17	Malignant	1-99	Positive
18	Malignant	5-95	Negative
19	Malignant	0-100	Negative
20	Malignant	30-70	Negative

3.4 Immunohistochemical expression of homeoproteins Six1 and Pax3 in phyllodes tumours

The rationale for investigating protein expression of Six1 and Pax3 in phyllodes tumours was to extend the findings of a collaborative project which found that Six1 and Pax3 was overexpressed in borderline/malignant phyllodes tumours at the transcript level [Jones et al., 2008b]. It is of interest to corroborate the findings with expression at the protein level. The study population as described in Section 3.3 and the corresponding tissue microarrays were employed for this study to investigate protein expression of Six1 and Pax3 by immunohistochemistry. As these are transcription factors, expression in both nucleus and cytoplasm was assessed. Also, epithelial and stromal expression was determined separately to identify distinct expression between the two components.

3.4.1 Six1 expression

Analysis was available in 270 cases after discounting loss of cores due to processing. Nine cases were without epithelial component and excluded from the epithelial analysis. Mean expression quantified as H-score stratified according to tumour grades and localization are shown in Figure 3.4.1. In the stromal component, Six1 expression was significantly associated with increasing tumour grade but this trend was observed only with cytoplasmic expression as illustrated in Figure 3.4.1. Conversely in the epithelial component, an inverse trend was observed where a decreasing mean H-score

was associated with increasing tumour grade and this was exhibited in the nucleus instead of cytoplasm as seen in the stromal component.

For analysis with clinicopathological parameters, H-score was classified into low and high expression using the overall mean, 63 points, as threshold. High expression was defined as H-score exceeding 63 points while low expression was defined with 63 points and below. In the stromal component, high cytoplasmic expression was associated with larger tumour size, higher cellularity, atypia and mitoses, stromal overgrowth, permeative microscopic margins and haemorrhage (Table 3.4.1). Stromal nuclear expression was not significantly associated with any of the parameters.

Table 3.4.1 Associations between clinicopathological parameters and stromal expression of Six1

Clinicopathologica parameters	Six1 expression in stromal nuclei			Six1 expression in stromal cytoplasm		
	H-score ≤ mean	H-score > mean	*p*-value	H-score ≤ mean	H-score > mean	*p*-value
Age (mean 43 years, median 43 years, range 15–79)						
≤43 years	75 (50.3)	62 (51.2)		62 (48.1)	62 (44.0)	
>43 years	74 (49.7)	59 (48.8)	0.903	67 (51.9)	79 (56.0)	0.542
Tumour Size (mean 51mm, median 38mm, range 8–250mm)						
≤51 mm	100 (67.1)	87 (71.9)		98 (76.0)	89 (63.1)	
>51 mm	49 (32.9)	34 (28.1)	0.428	31 (24.0)	52 (36.9)	0.025[a]
Tumour Grade						
Benign	103 (69.1)	85 (70.2)		104 (80.6)	84 (59.6)	
Borderline	29 (19.5)	30 (24.8)		22 (17.1)	37 (26.2)	
Malignant	17 (11.4)	6 (5.0)	0.126	2 (2.3)	20 (14.2)	<0.001[a]
Stromal Hypercellularity						
Mild	84 (56.4)	65 (53.7)		83 (64.3)	66 (46.8)	
Moderate	55 (36.9)	50 (41.3)		39 (30.2)	66 (46.8)	
Marked	10 (6.7)	6 (5.0)	0.682	7 (5.4)	9 (6.4)	0.013[a]
Stromal Atypia						
Mild	121 (81.2)	102 (84.3)		118 (91.5)	105 (74.5)	
Moderate	20 (13.4)	18 (14.9)		8 (6.2)	30 (21.3)	
Marked	8 (5.4)	1 (0.8)	0.116	3 (2.3)	6 (4.3)	0.001[a]
Stromal Mitosis						
0-4	103 (69.1)	87 (71.9)		105 (81.4)	85 (60.3)	
5-9	25 (16.8)	20 (16.5)		16 (12.4)	29 (20.6)	
>9	21 (14.1)	14 (11.6)	0.817	8 (6.2)	27 (19.1)	<0.001[a]
Stromal Overgrowth						
Absent	129 (86.6	108 (89.3)		123 (95.3)	114 (80.9)	
Present	20 (13.4)	13 (10.7)	0.577	6 (4.7)	27 (19.1)	<0.001[a]
Microscopic Margins						
Circumscribed	113 (75.8)	92 (76.0)		109 (84.5)	96 (68.1)	
Permeative	36 (24.2)	29 (24.0)	1.000	20 (15.5)	45 (31.9)	0.002[a]
Necrosis						
Absent	127 (85.2)	109 (90.1)		117 (90.7)	119 (84.4)	
Present	22 (14.8)	12 (9.9)	0.271	12 (9.3)	22 (15.6)	0.143
Haemorrhage						
Absent	88 (59.1)	72 (59.5)		87 (67.4)	73 (51.8)	
Present	61 (40.9)	49 (40.5)	1.000	42 (32.6)	68 (48.2)	0.009[a]

Note: Mean H-score for Six1 is 63. [a]Statistically significant results

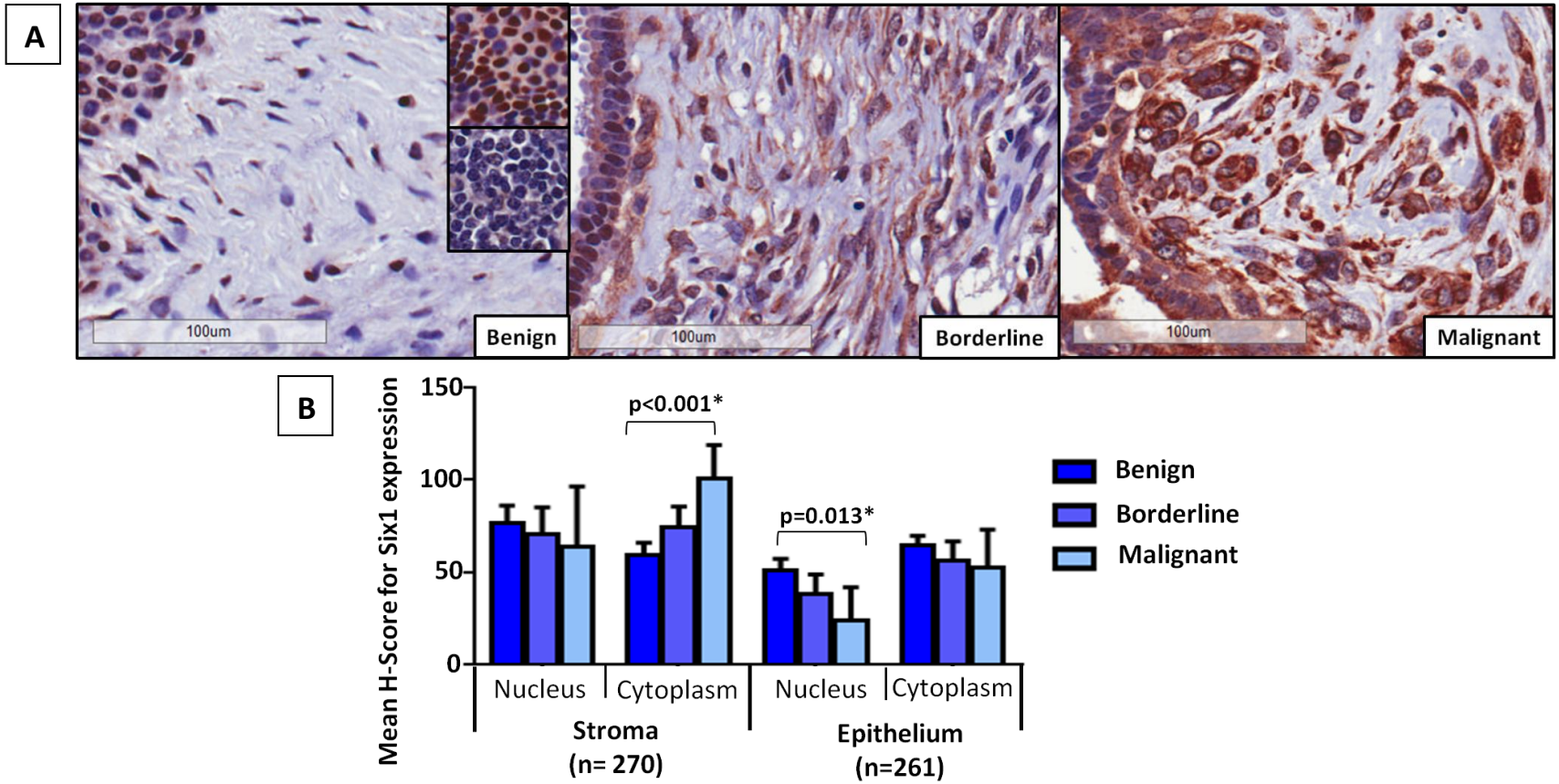

Figure 3.4.1 Six1 expression in the stroma and epithelium of phyllodes tumours. (A) Expression of Six1 in the stromal cytoplasm increases with tumour grade. Scale bar: 100µm. (B) Bar charts of mean H-score stratified according to localization of staining. Error bars represent 95% confidence interval. Expression in stromal cytoplasm and epithelial nuclei was significantly associated with tumour grade

In the epithelial component, high nuclear expression was associated with smaller tumour size and circumscribed tumour margin while high cytoplasmic expression was associated with circumscribed tumour margin only (Table 3.4.2).

Table 3.4.2 Associations between clinicopathological parameters and epithelial expression of Six1

Clinicopathological parameters	Six1 expression in epithelial nuclei			Six1 expression in epithelial cytoplasm		
	H-score ≤ mean	H-score > mean	*p*-value	H-score ≤ mean	H-score > mean	*p*-value
Age (mean 43 years, median 43 years, range 15–79)						
≤43 years	87 (47.3)	34 (44.2)		69 (53.9)	65 (48.9)	
>43 years	97 (52.7)	43 (55.8)	0.684	59 (46.1)	68 (51.1)	0.458
Tumour Size (mean 51mm, median 38mm, range 8–250mm)						
≤51mm	120 (65.2)	65 (84.4)		89 (69.5)	96 (72.2)	
>51mm	64 (34.8)	12 (15.6)	0.002[a]	39 (30.5)	37 (27.8)	0.684
Tumour Grade						
Benign	125 (67.9)	63 (81.8)		85 (66.4)	103 (77.4)	
Borderline	45 (24.5)	13 (16.9)		34 (26.6)	24 (18.0)	
Malignant	14 (7.6)	1 (1.3)	0.037[a]	9 (7.0)	6 (4.5)	0.139
Stromal Hypercellularity						
Mild	101 (54.9)	48 (62.3)		71 (55.5)	78 (58.6)	
Moderate	72 (39.1)	27 (35.1)		52 (40.6)	47 (35.3)	
Marked	11 (6.0)	2 (2.6)	0.366	5 (3.9)	8 (6.0)	0.555
Stromal Atypia						
Mild	152 (82.6)	69 (89.6)		104 (81.3)	117 (88.0)	
Moderate	28 (15.2)	8 (10.4)		22 (17.2)	14 (10.5)	
Marked	4 (2.2)	0 (0)	0.234	2 (1.6)	2 (1.5)	0.294
Stromal Mitosis						
0-4	130 (70.7)	60 (77.9)		87 (68.0)	103 (77.4)	
5-9	35 (19.0)	10 (13.0)		26 (20.3)	19 (14.3)	
>9	19 (10.3)	7 (9.1)	0.444	15 (11.7)	11 (8.3)	0.228
Stromal Overgrowth						
Absent	166 (90.2)	71 (92.2)		115 (89.8)	122 (91.7)	
Present	18 (9.8)	6 (7.8)	0.815	13 (10.2)	11 (8.3)	0.671
Microscopic Margins						
Circumscribed	136 (73.9)	68 (88.3)		93 (72.7)	111 (83.5)	
Permeative	48 (26.1)	9 (11.7)	0.013[a]	35 (27.3)	22 (16.5)	0.037[a]
Necrosis						
Absent	162 (88.0)	71 (92.2)		110 (85.9)	123 (92.5)	
Present	22 (12.0)	6 (7.8)	0.386	18 (14.1)	10 (7.5)	0.110
Haemorrhage						
Absent	105 (57.1)	54 (70.1)		77 (60.2)	82 (61.7)	
Present	79 (42.9)	23 (29.9)	0.053	51 (39.8)	51 (38.3)	0.899

Note: Mean H-score for Six1 is 63. [a]Statistically significant results

Among the 18 cases with local recurrences, 11 (61.1%) expressed high Six1 stromal cytoplasmic expression. All 6 cases with distant recurrences also expressed high Six1 stromal cytoplasmic expression. Patients with tumours expressing high stromal cytoplasmic Six1 or low epithelial nuclear Six1 expression was found to have a shorter time to recurrence with Kaplan Meier survival analysis, although only the latter was significant statistically (Figure 3.4.2). However, tumours with both high stromal and low epithelial expression did not exhibit significant differences in recurrence-free survival as compared with those which had either high stromal expression or low epithelial expression (p=0.136, results not shown). Analysis of overall survival returned no significant differences in prognoses between patients with high Six1 versus low Six1 expressing tumours.

3.4.2 Pax3 expression

Analysis was available for 271 cases with 268 cases containing an epithelial component. Overall mean and median H-scores were 14 and 10 respectively, ranging from 0 to 80 points. Similar to what was observed for Six1 expression, tumour grade was positively associated with expression of Pax3 in the stromal cytoplasm but not in the stromal nucleus as illustrated in Figure 3.4.3. Also, the expression in the epithelial nucleus was inversely correlated with tumour grade, a similar trend observed in Six1 expression.

For analysis of Pax3 expression in association with clinicopathological parameters, mean H-score of 14 was employed as a threshold to define low (≤14 points) and high expression (>14 points) respectively.

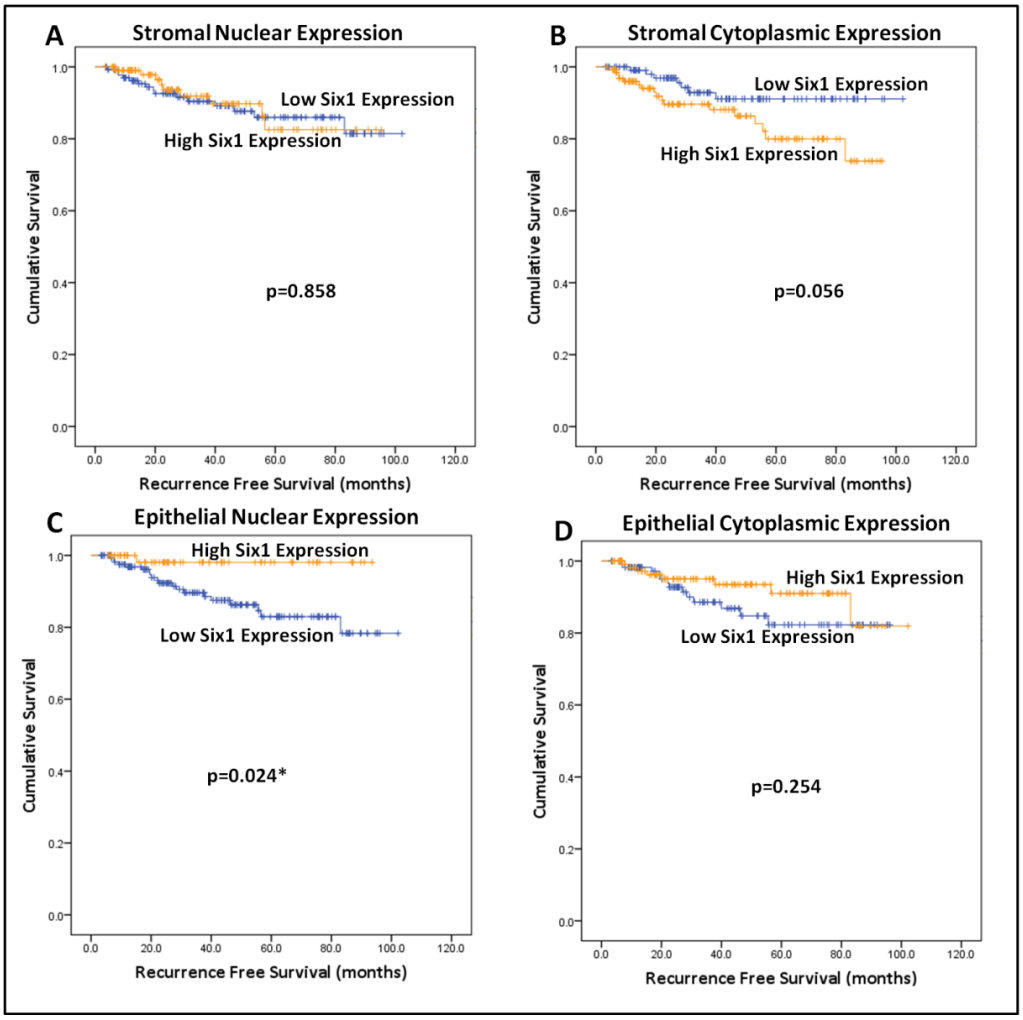

Figure 3.4.2 Kaplan Meier survival curves demonstrating differences in recurrence-free survival between tumours expressing high Six1 and low Six1 in (A) stromal nucleus; (B) stromal cytoplasm; (C) epithelial nucleus; (D) epithelial cytoplasm. A trend of shorter recurrence-free survival was observed in (B) and a significant difference was observed in (C).

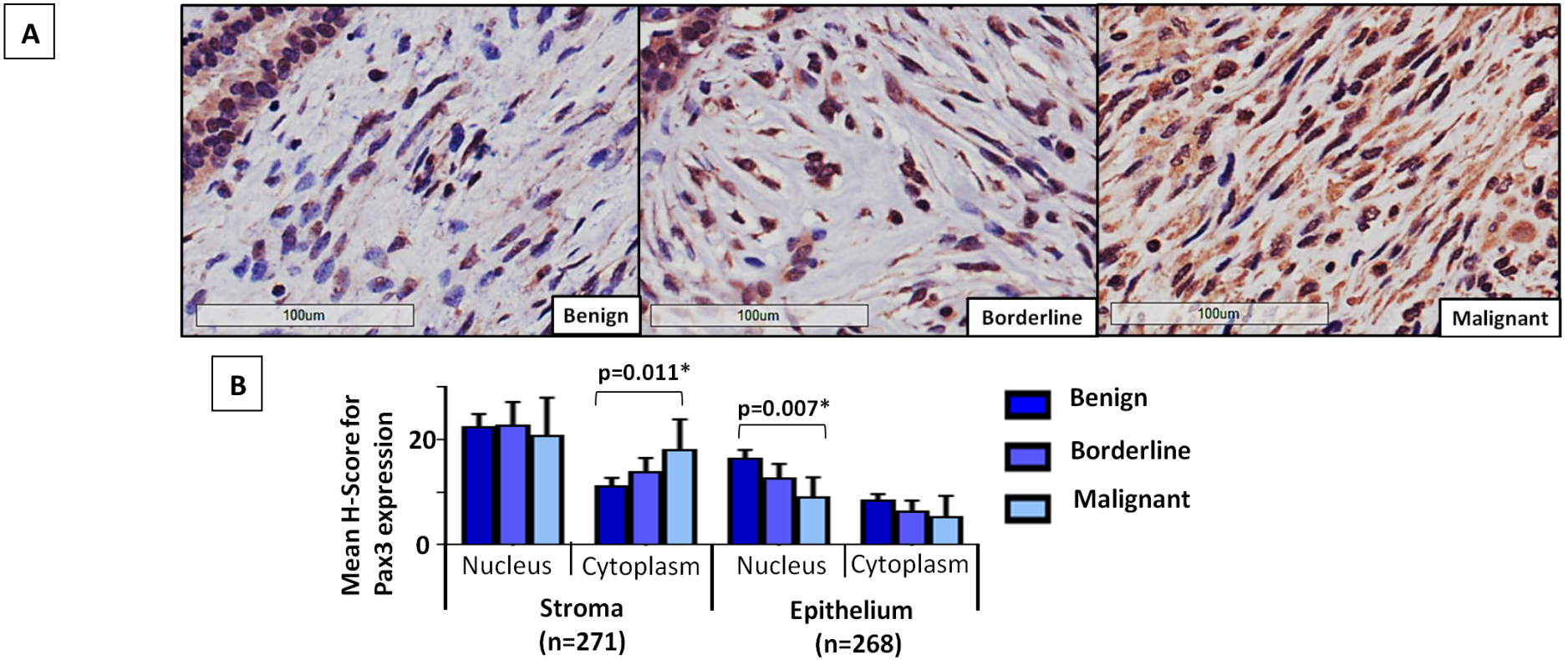

Figure 3.4.3 Pax3 expression in the stroma and epithelium of phyllodes tumours. (A) Expression of Pax3 in the stromal cytoplasm increases across tumour grade but the trend is not seen in the stromal nucleus. Scale bar: 100µm. (B) Bar charts of mean H-score stratified according to localization of staining for Pax3 expression. Error bars represent 95% confidence interval. Expression in stromal cytoplasm and epithelial nuclei was significantly associated with tumour grade.

In the stromal component, high cytoplasmic expression was associated with stromal overgrowth, permeative microscopic margins, presence of necrosis and haemorrhage. Expression in the nucleus was not significantly associated with any clinicopathological parameters (Table 3.4.3).

In the epithelium, nuclear expression of Pax3 was significantly associated with stromal overgrowth only (Table 3.4.4). However, when H-score was analyzed as a continuous variable using ANOVA test, higher epithelial nuclear expression was associated with mild stromal atypia ($p=0.008$), lower stromal mitoses ($p=0.03$), absence of necrosis ($p=0.011$) and haemorrhage ($p=0.025$). No significant association was observed with epithelial cytoplasmic expression.

Pax3 expression was not associated with recurrence-free survival regardless of localization. Nonetheless, patients with tumours exhibiting high stromal cytoplasmic Pax3 expression experienced a poorer overall survival (Figure 3.4.4).

Table 3.4.3 Associations between clinicopathological parameters and stromal expression of Pax3

Clinicopathological parameters	Pax3 expression in stromal nucleus			Pax3 expression in stromal cytoplasm		
	H-score ≤ mean	H-score > mean	*p*-value	H-score ≤ mean	H-score > mean	*p*-value
Age (mean 43 years, median 43 years, range 15–79)						
≤43 years	43 (46.2)	94 (52.8)		85 (47.2)	39 (42.9)	
>43 years	50 (53.8)	84 (47.2)	0.310	95 (52.8)	52 (57.1)	0.521
Tumour Size (mean 51mm, median 38mm, range 8–250mm)						
≤51 mm	61 (65.6)	127 (71.3)		130 (72.2)	58 (63.7)	
>51 mm	32 (34.4)	51 (28.7)	0.335	50 (27.8)	33 (36.3)	0.165
Tumour Grade						
Benign	65 (69.9)	124 (69.7)		136 (75.6)	53 (58.2)	
Borderline	20 (21.5)	39 (21.9)		33 (18.3)	26 (28.6)	
Malignant	8 (8.6)	12 (8.4)	0.996	11 (6.1)	12 (13.2)	0.011[a]
Stromal Hypercellularity						
Mild	51 (54.8)	98 (55.1)		101 (56.1)	48 (52.7)	
Moderate	34 (36.6)	72 (40.4)		69 (38.3)	37 (40.7)	
Marked	8 (8.6)	8 (4.5)	0.371	10 (5.6)	6 (6.6)	0.853
Stromal Atypia						
Mild	79 (84.9)	145 (81.5)		151(91.5)	73 (80.2)	
Moderate	10 (10.8)	28 (15.7)		24 (13.3)	14 (15.4)	
Marked	4 (4.3)	5 (2.8)	0.454	5 (2.8)	4 (4.4)	0.685
Stromal Mitosis						
0-4	67 (72.0)	124 (69.7)		132 (73.3)	59 (64.8)	
5-9	16 (17.2)	29 (16.3)		28 (15.6)	17 (18.7)	
>9	10 (10.8)	25 (14.0)	0.744	20 (11.1)	15 (16.5)	0.313
Stromal Overgrowth						
Absent	82 (88.2)	156 (87.6)		168 (93.3)	70 (76.9)	
Present	11 (11.8)	22 (12.4)	1.000	12 (6.7)	21 (23.1)	<0.001[a]
Microscopic Margins						
Circumscribed	70 (75.3)	136 (76.4)		146 (81.1)	60 (65.9)	
Permeative	23 (24.7)	42 (23.6)	1.000	34 (18.9)	31 (34.1)	0.007[a]
Necrosis						
Absent	79 (84.9)	158 (88.8)		164 (91.1)	73 (80.2)	
Present	14 (15.1)	20 (11.2)	0.440	16 (8.9)	18 (19.8)	0.018[a]
Haemorrhage						
Absent	56 (60.2)	104 (58.4)		114 (63.3)	46 (50.5)	
Present	37 (39.8)	74 (41.6)	0.796	66 (36.7)	45 (49.5)	0.05[a]

Note: Mean H-score for Pax3 is 14. [a]statistically significant results

Table 3.4.4 Associations between clinicopathological parameters and epithelial expression of Pax3 localized to nuclei and cytoplasm respectively

Clinicopathological parameters	Pax3 expression in epithelial nuclei			Pax3 expression in epithelial cytoplasm		
	H-score ≤ mean	H-score > mean	*p*-value	H-score ≤ mean	H-score > mean	*p*-value
Age (mean 43 years, median 43 years, range 15–79)						
≤43 years	60 (42.6)	63 (49.6)		109 (49.3)	27 (57.4)	
>43 years	81 (57.4)	64 (50.4)	0.270	112 (50.7)	20 (42.6)	0.338
Tumour Size (mean 51mm, median 38mm, range 8–250mm)						
≤51mm	97 (68.8)	91 (71.7)		157 (71.0)	31 (66.0)	
>51mm	44 (31.2)	36 (28.3)	0.689	64 (29.0)	16 (34.0)	0.487
Tumour Grade						
Benign	91 (64.5)	98 (77.2)		155 (70.1)	34 (72.3)	
Borderline	37 (26.2)	22 (17.3)		49 (22.2)	10 (21.3)	
Malignant	13 (9.2)	7 (5.5)	0.076	17 (7.7)	3 (6.4)	0.937
Stromal Hypercellularity						
Mild	73 (51.8)	76 (59.8)		122 (55.2)	27 (57.4)	
Moderate	59 (41.8)	46 (36.2)		89 (40.3)	16 (34.0)	
Marked	9 (6.4)	5 (3.9)	0.352	10 (4.5)	4 (8.5)	0.450
Stromal Atypia						
Mild	111 (78.7)	113 (89.0)		182 (82.4)	42 (89.4)	
Moderate	25 (17.7)	12 (9.4)		33 (14.9)	4 (8.5)	
Marked	5 (3.5)	2 (1.6)	0.076	6 (2.7)	1 (2.1)	0.488
Stromal Mitosis						
0-4	93 (66.0)	98 (77.2)		156 (70.6)	35 (74.5)	
5-9	29 (20.6)	16 (12.6)		36 (16.3)	9 (19.1)	
>9	19 (13.5)	13 (10.2)	0.117	29 (13.1)	3 (6.4)	0.418
Stromal Overgrowth						
Absent	119 (84.4)	119 (93.7)		195 (88.2)	43 (91.5)	
Present	22 (15.6)	8 (6.3)	0.019[a]	26 (11.8)	4 (8.5)	0.620
Microscopic Margins						
Circumscribed	107 (75.9)	99 (78.0)		172 (77.8)	34 (72.3)	
Permeative	34 (24.1)	28 (22.0)	0.772	49 (22.2)	13 (27.7)	0.448
Necrosis						
Absent	120 (85.1)	117 (92.1)		196 (88.7)	41 (87.2)	
Present	21 (14.9)	10 (7.9)	0.086	25 (11.3)	6 (12.8)	0.802
Haemorrhage						
Absent	107 (75.9)	99 (78.0)		131 (59.3)	29 (61.7)	
Present	34 (24.1)	28 (22.0)	0.772	90 (40.7)	18 (38.3)	0.870

Note: Mean H-score for Pax3 is 14. [a]statistically significant results

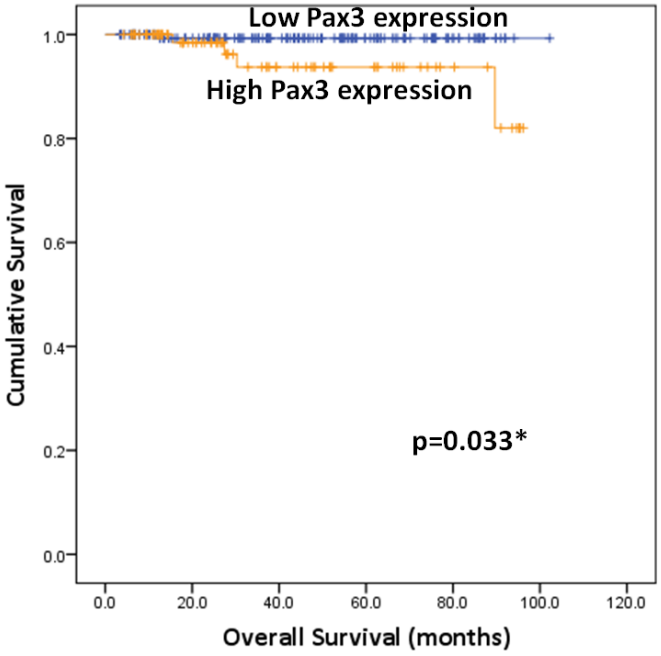

Figure 3.4.4 Kaplan Meier survival analysis showing patients with tumours expressing high stromal cytoplasmic Pax3 had a poorer overall survival.

3.4.3 Combinational analysis of Six1 and Pax3

For this purpose, only the expression in the stromal cytoplasm was analyzed as it associated significantly with tumour grade and clinicopathological parameters among all other localizations. Six1 expression was found to be associated positively with Pax3 (Spearman's ρ= 0.458, p-value<0.001), whereby a high Six1 stromal expression was observed with high stromal Pax3 expression (Figure 3.4.5)

When the two markers were analyzed in combination, tumours with low expression of both Six1 and Pax3 were found to have a better recurrence-free survival compared to tumours exhibiting either high expression of Six1/Pax3 alone or high expression of both Six1 and Pax3 (Figure 3.4.6).

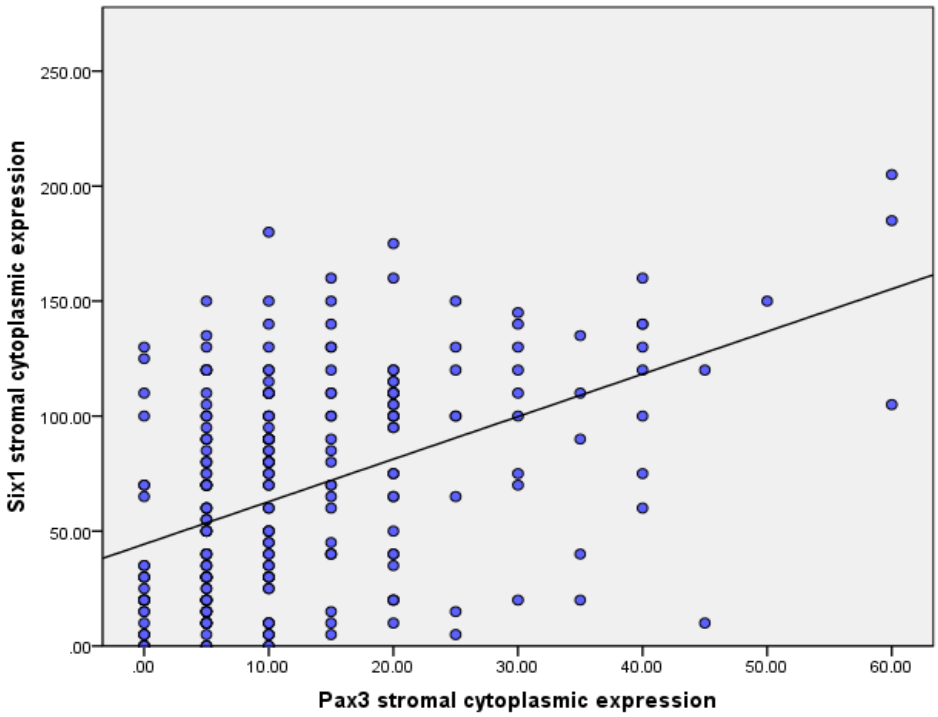

Figure 3.4.5 Scatterplot illustrating a positive correlation (Spearman's ρ= 0.458, p-value<0.001) between Six1 stromal cytoplasmic expression and Pax3 stromal cytoplasmic expression measured in H-score.

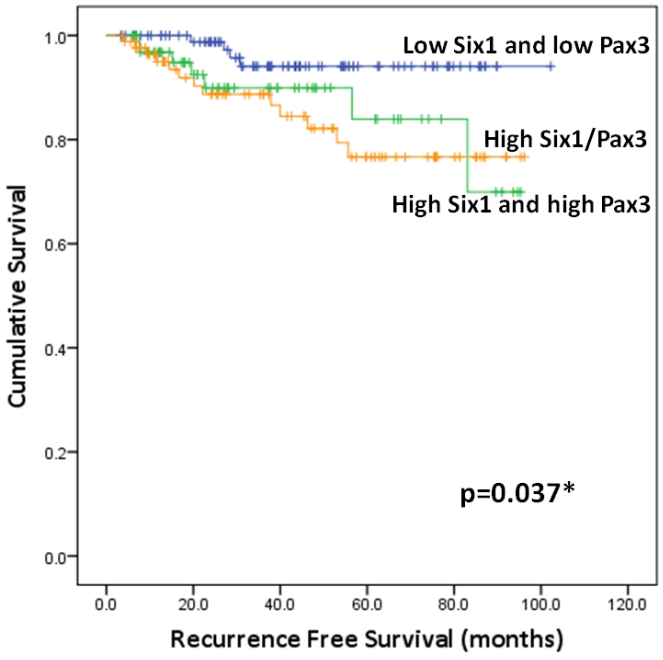

Figure 3.4.6 Kaplan Meier survival analysis illustrating patients with tumours exhibiting low expression of both Six1 and Pax3 had a better recurrence-free survival than patients with tumours exhibiting high expression of either marker or both markers.

4.0 DISCUSSION

4.1 Nomogram as a prognostic and predictive tool for phyllodes tumour patients

A nomogram is fundamentally a graphical calculation instrument which can be based on any type of statistical function such as linear regression or Cox regression models [Shariat et al., 2009]. In medicine, nomograms are typically used as clinical prognostic and predictive models. They currently represent the most accurate tool for predicting outcomes in cancer patients with reported superior performance over other predictive tools such as risk-grouping, look-up tables, tree analysis and artificial neural network [Shariat et al., 2009]. Nomograms are gaining popularity among clinicians especially those made available with a user-friendly interface and accessible via internet such as the MSKCC (Memorial Sloan Kettering Cancer Center) prostate cancer nomogram and the Karakiewicz nomogram for predicting survival in patients with renal cell carcinoma.

While many studies have described important prognostic factors in phyllodes tumour, none have translated the findings into such an application. As phyllodes tumours are uncommon, the lack of sufficient events and a small study cohort might have impeded the success of such a statistical modelling method. The large cohort size of this study made modelling viable with adequate events and long follow-up data. Moreover, the relatively homogenous patient cohort is managed in a uniform manner within a single institution, reducing possible bias. Also, the robust follow-up data comprising

institutional medical records and information from the National Registry of Births and Deaths reduces possible inaccuracy.

The development of this nomogram for patients with phyllodes tumours allows discussions of management strategies for the individual patient, eliminating generalization based solely on histological grade. The permutations of integrating the assessment of stromal atypia, mitoses, cellularity, overgrowth and nature of microscopic margins to derive a grade render variability, leading to difficulties in accurately predicting behaviour. Nonetheless, it is not advocated that the grading of phyllodes tumours should be replaced as histological grade reflects the biology of the tumour, as evidenced by its significant association with recurrence free survival. Moreover, utility of grade enables comparison across studies as it is the most widely used form of diagnostic assessment [Spitaleri, 2013].

The final nomogram incorporated three histological criteria (stromal atypia, mitoses and overgrowth) and surgical margins, simply referred to as the AMOS criteria. Stromal cellularity and nature of microscopic margins are excluded in the model because they are not independently associated with recurrence in the multivariate analysis. Even in cases with negative surgical margins, the nature of microscopic borders was not predictive of recurrence free survival. Surgical margins are frequently reported to be significantly associated with recurrence in previous studies [Pandey et al., 2001; Asoglu et al., 2004; Kapiris et al., 2001; Belkacémi et al., 2008], concurring with our

findings which showed that having a positive surgical margin greatly increases the risk of recurrence.

The performance of the nomogram is superior to that of the total histological score derived arbitrarily based on a linear additive effect. This reflects the complexity of the biological impact of each parameter affecting tumour behaviour, which could not be simplified with a linear additive scale. Each parameter has a certain weight as to how much it affects recurrence free survival and is accounted for in the nomogram. The nomogram has a good discriminating ability with a concordance index of 0.79 and high calibration accuracy with small deviation at 1 year, 3 years, 5 years and 10 years. However, a similar issue faced by all nomograms is whether the performance of the nomogram could be extrapolated to be applied in an external cohort and in a different population. Recently, a Japanese group validated the utility of this nomogram by demonstrating the ability of the nomogram in predicting recurrence-free survival for patients with phyllodes tumours albeit in a small cohort of 43 patients [Nishimura et al., 2014].

Compared to other nomograms developed, the relatively small number of events is a limitation of the study. The rare occurrence coupled with the innate biology of the tumour is a common challenge faced by many researchers investigating phyllodes tumours. Nonetheless, the recurrence rate of this series (13.2%) corroborates the reported rates of recurrence in the literature, largely ranging from 10% to 20% [Spitaleri et al., 2013]. Although some may argue that local recurrence may not truly reflect the end

point of the disease, previous reports have shown that local recurrences precede distant metastases which could cause demise except tumours which are primarily diagnosed as malignant grade tumours as they can metastasize or lead to death without local recurrence [Tan et al., 2005a]. Besides, recurrence often leads to subsequent treatment that could instil psychological distress and affect quality of life. Hence, probability of recurrences would be a valuable piece of information for patients to know even if it does not reflect the endpoint of the disease [Kattan, 2008].

4.2 Genome-wide copy number and mutational analysis

The aims and design of this study differed slightly from those investigated previously. Instead of identifying differences between tumour grades, this study aims to elucidate differences between prognostically distinct tumours. Hence, tumours were stratified according to clinical outcome instead of traditional tumour grade (benign, borderline and malignant) as recurrences could occur regardless of grade. Although such stratification leads to an imbalance in representation of histological grades in the two groups, as it was within expectation to find more benign tumours in the non-recurrent group and malignant tumours in the recurrent/metastasis group, it is not the aim of this present study to perform a molecular profile of tumours according to grade, but to identify apparent candidate gene(s) which could possibly drive malignant clinical behaviour. The other aspect of this study was to identify distinct changes between the epithelium and stroma of the tumours. However, the attempt of interrogating a specific compartment

was unsuccessful due to insufficient yield of genomic DNA from the microdissected samples. The whole section samples nonetheless, had 95% success rate (19/20 tumours) of being analyzed. High throughput screening on FFPE materials are challenging as the quality of DNA is often compromised. Hence, the high success rate of this technology is advantageous particularly for rare tumours such as the phyllodes tumours as fresh frozen tissues are difficult to obtain.

Local recurrences were grouped together with those experiencing metastasis and death in this study. Again, some may argue that local recurrences may not truly reflect the disease endpoint and may be confounded by inadequate surgical margins. However, local recurrences are clinically relevant as tumours which recur locally may progress in grade. Moreover, local recurrences usually precede distant metastases that eventually lead to demise [Tan et al., 2005a]. It is further exemplified in this study that the exclusive presence of high-level amplification and homozygous deletions in cases which recurred, including those that recurred locally, showed that these aberrations were not limited only to cases which metastasized but can also be discovered in benign tumours that recur locally. Although it is true that surgical margin involvement is a key factor for recurrences as shown in the previous section of this study, it is believed that the underlying biology also plays an important role, since there are many benign tumours do not recur locally despite involvement of surgical margins. In the 19 cases which were interrogated successfully in this study, surgical margin status was not significantly different between the two prognostically

distinct groups. It may be argued that it is uncertain if recurrence-free tumours will truly be free from recurrence given a longer follow-up, but the results from the previous section indicated that recurrences in breast phyllodes tumours occurred at a median duration of 24.6 months, a period encompassed in the follow-up of the cases evaluated.

One recurrent molecular aberration in phyllodes tumours is gain of chromosome 1q. Gain of 1q was reported to be associated with borderline/malignant grade [Laé et al., 2007; Jones et al., 2008a; Ang et al., 2011] and recurrence [Lu et al., 1997]. However, chromosome 1q is a large region encompassing numerous known and unknown genes. Jones *et al.* correlated findings of the 1q gain with expression data and found that genes set at region 1q25 showed the most significant association with copy number data [Jones et al., 2008a]. Also reported by the authors is the gain of 1q21.1, a similar region found in this study to be the most frequent aberration found among the 19 tumours. The function of the genes located within this region is largely unknown. However, it was previously reported that the chromosome region of 1q21.1 encompasses many of the *NBPF* (neuroblastoma breakpoint family) genes and its gain was associated with neuroblastoma [Diskin et al., 2009].

While chromosomal aberrations are important in giving an aerial view of alterations harboured by tumours, it is difficult to identify specific target gene(s) which confer(s) growth and survival advantage to the tumours within a large region of aberrations. Target genes are more easily identified in the

case of amplifications, homozygous deletions and chromosome translocation [Vogelstein et al., 2013]. Such findings reported in phyllodes tumours hitherto are limited. Some of the few specific target genes previously identified in phyllodes tumours include *EGFR*, *MYC, MDM2* and *CDKN2A* [Laé et al., 2007; Sawyer et al., 2003; Kersting et al., 2006; Tse et al., 2009]. In this present study, four regions of amplifications (1q32.1, 3p25, 5p13.3, 7p12) were identified as shown in section 3.2.2.1 with one specific target gene (*EGFR*) isolated and confirmed on FISH and immunohistochemistry. *EGFR* amplification was previously reported by Kersting *et al.* and Tse *et al.* in isolated cases. Agelopoulos *et al.* further demonstrated that *EGFR* amplification in phyllodes tumours is associated with 230 differentially expressed genes [Agelopoulos et al., 2007], notably with *CAV1* (Caveolin 1) and *EPS15* (epidermal growth factor receptor pathway substrate 15). Three other amplified regions in this study have no target genes isolated, however the functional annotation analysis revealed that genes covered within the amplified region are highly involved in regulation of enzymatic activity, suggesting the exploitation of such pathways for tumour survival and migration.

Two regions of homozygous deletions were observed in this study - chromosome 9p21 involving *CDKN2A* (cyclin-dependent kinase inhibitor 2A) and chromosome 20p12.1. Cases with homozygous deletion at 9p21 concurred with loss of protein expression of CDKN2A (better known as p16). This finding is in line with those of Jones *et al.* who reported that loss of p16

expression was associated with interstitial deletion of 9p21 involving the *CDKN2A* region in malignant tumours. Interestingly, even in malignant tumours which did not harbour this deletion, *CDKN2A* methylation was observed, suggesting the importance of its downregulation in malignant progression [Jones et al., 2008a]. On the contrary, Karim *et al.* observed that p16 protein expression was elevated with increasing tumour grade with no molecular information available [Karim et al., 2010]. The homozygous deletion at chromosome 20p12.1 was reported to be one of the hotspots of large rare deletions in human genome [Bradley et al., 2010]. It is found in non-cancer patients and could arise due to genome instability. It is however unclear whether this will result in susceptibility towards cancer as deletions at this region are also noted in colorectal cancer [Davison et al., 2005] and gastric cancer cell lines [Tada et al., 2010].

Yet another key genomic alteration in tumours is represented by somatic mutations, acquired and accumulated over the lifetime of the cancer patient [Stratton et al., 2009]. It is important to identify somatic mutations as they confer selective growth advantage to cancer cells, enabling them to outgrow normal cells and evade cell death. From a clinical point of view, knowing the specific mutations allows targeted therapy to be administered such as imatinib and gefitinib targeting selective c-kit [Heinrich et al., 2003] and EGFR-mutant cells respectively [Paez et al., 2004]. Mutational analyses previously performed in phyllodes tumours have presented with inconsistent findings across studies. Korcheva *et al.* identified a S8R substitution in the

FBX4 (an E3 ubiquitin ligase) gene to be present in 11.5% of tumours after screening 321 mutations across 30 genes [Korcheva et al., 2011]. Jones *et al*. found a missense mutation P48L in the *CDKN2A* gene in a malignant phyllodes tumour after screening 35 tumours [Jones et al., 2008a]. Some authors have reported the presence of *TP53* mutations in isolated cases of malignant tumours [Kuenen-Boumeester et al., 1999; Gatalica et al., 2001] but others have found none [Woolley et al., 2000; Korcheva et al., 2011]. In this current study, no mutation was found after screening for 541 mutation hotspots of 59 genes including oncogenes and tumour suppressor genes such as *ABL1*, *BRAF*, *BRCA1*, *CDKN2A*, *EGFR*, *ERBB2*, *KIT*, *KRAS*, *PTEN*, *RB1* and *TP53*. This implies that the genomic landscape of the phyllodes tumours may differ from that of the common carcinoma which usually harbours these mutations and need to be explored as such.

In summary for this section of the study, higher chromosomal aberrations were observed in phyllodes tumours which recurred/metastasized with amplifications and homozygous deletions identified exclusively in this group. Regions of amplifications at 1q32.1, 5p13.3, 3p25 and homozygous deletions at 20p12.1 were previously not reported in phyllodes tumours. Moreover, no mutation has been identified. Limitations of this study are the small sample size and its retrospective nature. Nonetheless, the findings warrant expansion to future investigations to identify driver genes which would be useful to elucidate possible pathways exploited by phyllodes tumours for malignant progression.

4.3 CD117 protein expression and mutation status in phyllodes tumours

The mutation panel in this study included frequently reported CD117 mutations of exons 9, 11, 13 and 17. As described in the previous section, no mutation was observed in these hotspots. This concurs with findings reported by Jung *et al.* who found no mutations in exons 9, 11, 13 and 17 in a subset of 28 samples [Jung et al., 2010]. Djordjevic and Hanna also found no mutations in two cases of CD117 positive tumours [Djordjevic and Hanna, 2008]. Other reports found no activating mutations but discovered a silent mutation of isoleucine 798 (exon 17) [Carvalho et al., 2004; Bose et al., 2010] and a point mutation of L510M (exon 10) of unknown significance [Sawyer et al., 2003]. It was also observed in this study that CD117 protein expression did not correlate with mutation status. Although none of the cases had mutations, 37% of the 19 tumours expressed CD117 at the protein level. The high percentage of CD117-positive tumours prompted further investigation on a larger series of 272 cases, of which the percentage of CD117-positive tumours reduces to 10%. This was rationalized when the findings of the larger series showed that the CD117 expression was associated with borderline/malignant tumours. The percentage of borderline/malignant tumours was relatively higher in the smaller series of 19 tumours (63%) compared to the larger series (30%).

In addition to association with borderline/malignant grade, CD117 expression was also associated with unfavourable outcome. Patients with CD117-positive tumours had a significantly poorer survival outcome. A trend of shorter recurrence free survival was also observed in CD117-positive cases

despite not significant statistically. This concurs with observations reported by Tan *et al.* and Jung *et al.* where CD117 expression was correlated with recurrence [Tan et al., 2008b; Jung et al., 2010]. On the contrary, Tse *et al.* and Esposito *et al.* observed no associations in their study of 179 and 16 cases respectively [Tse et al., 2004; Esposito et al., 2006]. In other CD117-expressing tumours, multivariate analyses showed that CD117 expression was an independent prognostic marker for oesophageal squamous cell carcinoma patients [Fan et al., 2013]. Multivariate analysis for this section of study is not feasible due to the low number of events documented. With the low number of events, the accuracy, precision and significance of the coefficients estimated by the analysis will become unreliable [Peduzzi et al., 1995]. Nonetheless, it is worthwhile to note the implication of a poorer clinical outcome of CD117 expressing tumours despite the limitation of a small number of events documented, suggesting an underlying aggressive nature of these tumours.

The inconsistent findings of CD117 protein expression from various groups could be attributed to usage of different antibodies, staining protocols and scoring criteria. Standardization of protocols between laboratories is challenging and there is no universal consensus as to which staining protocol and scoring criteria are the best. Hence, these protocols need to be optimized and validated accordingly in respective laboratories [Loughrey et al., 2006]. The antibody Dako A4502 employed in this study is validated for diagnostic use and was previously evaluated by other authors to have high

sensitivity and specificity across different tumours [Went, 2004; Lucas et al., 2003]. Furthermore, the staining protocol was complemented with toluidine blue staining to exclude possibly confounding contribution of mast cells, as previously suggested by Djodjervic and Hanna that the expression of CD117 in phyllodes tumours observed by other authors might have just been a mast cell phenomenon [Djordjevic and Hanna, 2008]. All cases initially defined to be CD117 positive were negative on toluidine blue, reinforcing the initial observation that these were not mast cells but stromal cells instead. A 1% cutoff for scoring criteria was used here in view that CD117 is not normally expressed in breast stromal cells [Kondi-Pafiti et al., 2010]. Hence, presence of the protein even in a low percentage could indicate an abnormal state. Although the low percentage could be partly contributed by the use of tissue microarrays which represent a small proportion of the entire tumour, it has been previously shown that tissue microarrays provide a reliable replication of phyllodes tumours in terms of biomarker expression [Ho et al., 2013].

Patients with GISTs harbouring overexpression or mutation of CD117 are responsive to tyrosine kinase inhibitors which can block the activity of an activated CD117 receptor. This largely motivates the investigation of CD117 expression and mutation status in phyllodes tumours as there has been previous suggestion that the stromal component of phyllodes tumours bear some similarities with GIST such as the spindled nature and spectrum of behaviour from benign to malignant [Carvalho et al., 2004]. However, recent insights into the roles of CD117 in cancer shed important light on the

different types of CD117 expressing tumours. CD117-expressing tumours are broadly classified into two groups [Pittoni et al., 2011]. The first group is usually characterized by gain-of-function (activating) CD117 mutations which play a central pathogenetic role in neoplasm initiation. These tumours are made up of cells that normally express CD117, such as GIST which arise from the oncogenic transformation of interstitial cells of Cajal that normally express high levels of CD117. On the contrary, the second group of tumours has rare occurrence of CD117 mutations and are composed of cells that do not normally express CD117. In this case, CD117 has a passive role and its expression is acquired during tumour progression. This may explain the lack of mutations found in phyllodes tumours as compared to GISTs. The stromal component of phyllodes tumour, which likely arises from breast mesenchymal tissue, does not usually express CD117 under normal circumstances.

The lack of activating mutations in phyllodes tumours suggests that the therapeutic option of using a tyrosine kinase inhibitor such as imatinib in patients with phyllodes tumours is unlikely to be effective. Nonetheless, the associations of CD117 protein expression with borderline/malignant phyllodes tumours, worse pathologic parameters and poorer prognosis are real phenomena that suggest a role of CD117 in the biological behaviour of breast phyllodes tumours.

4.4 Immunohistochemical expression of homeoproteins Six1 and Pax3 in phyllodes tumours

In a comprehensive expression profiling performed by Jones *et al.* which employed 23 fresh frozen phyllodes tumours, Six1 and Pax3 was found to have the highest fold change (5-fold difference) among 162 differentially expressed genes comparing borderline/malignant tumours to benign phyllodes tumours [Jones et al., 2008b]. The difference was predominantly observed in the stromal component. To validate the findings from the microarray data, the authors performed mRNA *in situ* hybridization in a separate set of 49 formalin-fixed paraffin embedded tissue and observed a similar association. These findings indicate that *Six1* and *Pax3* were highly transcribed but whether the protein was translated was not investigated. Results from this section filled in the gap, where protein expression of Six1 and Pax3 was assessed in phyllodes tumours by immunohistochemistry. A significant association of Six1 and Pax3 immunohistochemical expression was observed with tumour grade, corroborating what was reported by Jones *et al.* However, such a trend was only observed in the cytoplasm, but not in the nucleus, of the stromal cells. The findings of cytoplasmic expression associating with increasing tumour grade suggest a possible non-transcriptional effect of these homeoproteins in the cytoplasm, indicating that its role is not solely dependent on its DNA-binding activity which occurs in the nucleus. Cytoplasmic function of transcription factors playing a role in tumour progression was previously exemplified in STAT3 where it was found

to promote cell migration by binding to stathmin, thus stabilizing microtubule formation in the cytoplasm [Ng et al., 2006].

In contrast to the observations in the stroma, a negative association of homeoprotein expression with tumour grade was observed instead in the epithelial component. This resonates with previous reports that the epithelial component has its own distinct set of molecular changes as compared to the stromal component [Sawyer et al., 2000]. The significance and mechanism of elevated expression of homeoproteins in the epithelium of lower grade phyllodes tumours were unclear. However, Six1 expression was shown to stimulate cellular proliferation through induction of cyclin A1 in mammary cancer cell lines [Coletta et al., 2004]. This corroborates the previous reports of noting an actively proliferating epithelium in benign phyllodes tumours, and an inverse correlation of epithelial hyperplasia with tumour grade [Tan et al., 2005a; Pietruszka and Barnes, 1978]. It could be possible that reactivation of homeoprotein expression was used in the early stages of the tumourigenesis to trigger cellular proliferation in the epithelium, thus keeping the epithelial-stromal architecture in balance in lower grade tumours.

Despite a small number of distant recurrences documented in this series, it is nonetheless interesting to note that all tumours which subsequently metastasized exhibited high expression of Six1 in the primary tumours. Jones *et al.* also noted that metastatic phyllodes tumours had a particularly high *Six1* overexpression among the Six1-overexpressed malignant phyllodes tumour [Jones et al., 2008b]. This trend is not only

limited to phyllodes tumours but also observed in breast cancer, where patients with Six1 overexpressed tumours had a shortened time to relapse, metastasis and death [Micalizzi et al., 2009]. Six1 was demonstrated to mediate breast tumour cell metastasis by multiple means such activating TGF-β signalling through miR-106b-25 microRNA cluster [Smith et al., 2012], promoting VEGF-C production and lymphangiogenesis [Wang et al., 2012a], and inducing epithelial-mesenchymal transition [McCoy et al., 2009]. Further, association of Six1 overexpression with metastasis was also implicated in other solid tumours. Ng *et al.* found that increased Six1 protein expression in hepatocellular carcinoma patients was significantly correlated with pathologic tumour-node-metastasis stage and venous infiltration [Ng et al., 2006]. Yu *et al.* demonstrated using mouse model that pulmonary metastasis of rhabdomyosarcoma was inhibited when Six1 expression was reduced [Yu et al., 2004], suggesting a functional role of Six1 in tumour dissemination.

A high expression of stromal Pax3 was significantly associated with reduced overall survival despite the limitation of a small number of mortalities documented. Studies on Pax3 with follow-up clinical data are scant in the literature with a study reporting a high expression of Pax3 was associated with higher grade gliomas and poorer survival [Chen et al., 2012]. Aberrant Pax3 expression was often associated with rhabdomyosarcoma, where chromosomal rearrangement producing oncogenic fusion protein Pax3-Foxo1 (forkhead box O1) was observed. Jones *et al.* did not find any evidence of fusion transcript involving *Pax3* and *Foxo1* genes, nor presence of

activating mutations and amplifications of *Pax3* in their series of phyllodes tumours. However, the authors identified a hybrid transcript of unknown significance between *Pax3* and *FARSB*, a phenylalanyl-tRNA synthetase [Jones et al., 2008b]. In this study, the Pax3 antibody employed targets against the N-terminal amino acids 2-12 of Pax3, enabling expression of most isoforms of Pax3 protein to be detected, including possible fusion proteins which consist of Pax3 at its N-terminal.

The crosstalk between Six1 and Pax3 has been investigated in developmental studies, with some reporting regulation of one molecule by the other. Ridgeway and Skerjanc demonstrated that Pax3 expression was able to induce Six1 in skeletal myogenesis of mouse embryonic cell lines [Ridgeway and Skerjanc, 2001]. On the contrary, Grifone *et al*. reported that Six1 homeoprotein regulated Pax3 expression instead, during myogenesis in mouse embryos [Grifone et al., 2005]. Jones *et al.* demonstrated that knockdown of Pax3 resulted in downregulation of Six1 in phyllodes tumour cell lines, suggesting that Pax3 may be the upstream molecular of Six1 in phyllodes tumours [Jones et al., 2008a]. A positive association was observed in this study between Six1 and Pax3 stromal expression. However, an investigation into the potential regulation of Six 1 by Pax3 is out of the scope of the present study as this is a retrospective series on FFPE materials. Nonetheless, a combinatorial panel of the two markers showed that low expression of both augured a better recurrence free survival compared to tumours expressing highly in one marker alone. Moreover, no significant

difference in prognosis was observed between tumours with high expression of one marker alone and tumours with high expression of both markers, suggesting the two homeoproteins work in concert instead of having a cumulative effect.

Recently, there have been concerns regarding the reliability and representativeness of using tissue microarrays to assess biomarker expression of tumours as the relatively small amount of tissue is not directly comparable to the whole tumours [Pinder et al., 2013]. In this present study, three cores of 2mm for each tumour were employed and the highest yielding H-score among the three cores was selected, in view of a previous study where a high concordance was found between tissue microarray and whole sections using the highest H-scores among phyllodes tumours core replicates [Ho et al., 2013]. Moreover, use of tissue microarrays have been validated even in highly heterogeneous tumours such as ovarian and breast cancer [Camp et al., 2000; Rosen et al., 2004]

In summary, this section of the study showed that stromal cytoplasmic expression of homeoproteins Six1 and Pax3 in phyllodes tumours was associated with higher tumour grade, unfavourable clinicopathological parameters and poorer prognosis. In addition, the epithelium and stroma exhibited different staining patterns, suggesting a distinct behaviour of the two components in tumour progression.

4.5 Overall conclusions and future studies

The aim of the study was to investigate and evaluate the prognostic importance of clinical and molecular parameters of phyllodes tumours. It is demonstrated through the main body of this thesis that a few clinical and molecular parameters have contributed in tandem to the course of the disease. Histological parameters (atypia, mitoses, stromal overgrowth) and surgical margins were significant predictors for recurrence-free survival. High-level amplifications and homozygous deletions were detected exclusively in tumours which recur/ metastasize and expression of biological markers CD117, Six1 and Pax3 were found to be associated with unfavourable histological parameters and poorer prognoses.

From these findings, it is shown that the biology of phyllodes tumours is complex with multiple factors contributing to malignancy of the tumour, with some factors overlapping as evidenced in the superiority of the nomogram to a linear total histological score model. The heterogeneity in the behaviour of phyllodes tumours extends beyond its histological grade and molecular aberrations are not limited to malignant grade tumours, but can be discovered also in benign tumours which later recur. The common mutations found in carcinoma were not seen in phyllodes tumours, suggesting that tumourigenesis and exploitation of oncogenic pathways of phyllodes tumours are distinct from that of carcinoma and should be investigated separately as such. Nonetheless, the associations of tyrosine kinase receptor CD117 and

homeoproteins Six1 and Pax3 with poorer prognoses suggest that these markers may play a role in the biology of phyllodes tumour progression.

The eventual desired outcome of this work is to improve the management of patients with phyllodes tumours. The nomogram developed can be used for patient counseling and clinical management of women diagnosed with phyllodes tumours albeit requiring further validation from external larger cohorts. Although CD117 expression is associated with borderline/malignant tumour grade, a lack of activating mutations and overexpression suggest that the use of a tyrosine kinase inhibitor in phyllodes tumours as a therapeutic option is unlikely to be effective. The inclusion of molecular features (genetic aberrations, CD117, Six1 and Pax3) into the nomogram was not feasible as the molecular studies were based on a subset cohort with insufficient number of events for meaningful statistical evaluation. Nonetheless, it will be interesting to incorporate such molecular features in the nomogram in the future if there is a sufficient number of events.

Overall strengths of this study are the large study cohort and a comprehensive examination of phyllodes tumours ranging from molecular investigations to clinical application. The retrospective nature of the study and the single cohort of the study population are the overall limitations of the study. Potential future work includes investigation of the mutational status of genes not commonly reported in cancer using next generation sequencing, and performing functional studies using *in vivo* and *in vitro*

models such as patient derived xenograft mouse model and primary cell culture models.

The understanding of phyllodes tumours in comparison to breast carcinoma is significantly lacking, largely due to the rare occurrence of such tumours and unavailability of commercial phyllodes tumour cell lines. As such, functional investigations on markers or molecules which were observed to be associated with malignancy in phyllodes tumours could not be performed without availability of cellular models. There is a need for this area to be addressed, particularly in this part of the world where phyllodes tumours are relatively more frequent compared to the West. It is nonetheless positive that at this point in time, in addition to the functional work performed by Jones *et al.* 10 years ago, a group in Taiwan and two groups in China this year have published functional studies using primary cell cultures developed in their respectively laboratories, to study implication of their hypothesized molecules including stem cell markers, microRNAs and cell adhesion molecule in phyllodes tumours [Lin et al., 2014; Gong et al., 2014; Ren et al., 2014].

Finally, there are several other questions in the field of phyllodes tumours that remain to be addressed. The relationship between epithelium and stroma and how these two components affect disease progression is still uncertain. Laser capture microdissection could be performed using flash frozen materials where quality of DNA is better preserved. Also, whether phyllodes tumours belong to a spectrum of disease with benign and

malignant ends akin to fibroadenomas and sarcomas respectively is yet to be explored.

BIBLIOGRAPHY

Abdalla HM, and Sakr MA (2006). Predictive factors of local recurrence and survival following primary surgical treatment of phyllodes tumors of the breast. J Egypt Natl Canc Inst *18*, 125-133.

Agelopoulos K, Kersting C, Korsching E, Schmidt H, Kuijper A, August C, Wülfing P, Tio J, Boecker W, van Diest PJ, Brandt B, and Buerger H (2007). Egfr amplification specific gene expression in phyllodes tumours of the breast. Cell Oncol *29*, 443-451.

Al-Masri M, Darwazeh G, Sawalhi S, Mughrabi A, Sughayer M, and Al-Shatti M (2012). Phyllodes tumor of the breast: role of CD10 in predicting metastasis. Ann Surg Oncol *19*, 1181-1184.

Ang MK, Ooi AS, Thike AA, Tan P, Zhang Z, Dykema K, Furge K, Teh BT, and Tan PH (2011). Molecular classification of breast phyllodes tumors: validation of the histologic grading scheme and insights into malignant progression. Breast Cancer Res Treat *129*, 319-329.

Asoglu O, Ugurlu MM, Blanchard K, Grant CS, Reynolds C, Cha SS, and Donohue JH (2004). Risk factors for recurrence and death after primary surgical treatment of malignant phyllodes tumors. Ann Surg Oncol *11*, 1011-1017.

Barbosa ML, Ribeiro EM, Silva GF, Maciel ME, Lima RS, Cavalli LR, and Cavalli IJ (2004). Cytogenetic findings in phyllodes tumor and fibroadenomas of the breast. Cancer Genet Cytogenet *154*, 156-159.

Barrio AV, Clark BD, Goldberg JI, Hoque LW, Bernik SF, Flynn LW, Susnik B, Giri D, Polo K, Patil S, and Van Zee KJ (2007). Clinicopathologic features and long-term outcomes of 293 phyllodes tumors of the breast. Ann Surg Oncol *14*, 2961-2970.

Barth RJ (1999). Histologic features predict local recurrence after breast conserving therapy of phyllodes tumors. Breast Cancer Res Treat *57*, 291-295.

Belkacémi Y, Bousquet G, Marsiglia H, Ray-Coquard I, Magné N, Malard Y, Lacroix M, Gutierrez C, Senkus E, Christie D, Drumea K, Lagneau E, Kadish SP, Scandolaro L, Azria D, and Ozsahin M (2008). Phyllodes tumor of the breast. Int J Radiat Oncol Biol Phys *70*, 492-500.

Ben Hassouna J, Damak T, Gamoudi A, Chargui R, Khomsi F, Mahjoub S, Slimene M, Ben Dhiab T, Hechiche M, Boussen H, and Rahal K (2006). Phyllodes tumors of the breast: a case series of 106 patients. Am J Surg *192*, 141-147.

Bose P, Dunn ST, Yang J, Allen R, El-Khoury C, and Tfayli A (2010). c-Kit expression and mutations in phyllodes tumors of the breast. Anticancer Res *30*, 4731-4736.

Bradley WE, Raelson JV, Dubois DY, Godin E, Fournier H, Privé C, Allard R, Pinchuk V, Lapalme M, Paulussen RJ, and Belouchi A (2010). Hotspots of large rare deletions in the human genome. PLoS One *5*, 10.

Camp RL, Charette LA, and Rimm DL (2000). Validation of tissue microarray technology in breast carcinoma. Lab Invest *80*, 1943-1949.

Carlson RW, Allred DC, Anderson BO, Burstein HJ, Carter WB, Edge SB, Erban JK, Farrar WB, Forero A, Giordano SH, Goldstein LJ, Gradishar WJ, Hayes DF, Hudis CA, Ljung BM, Marcom PK, Mayer IA, McCormick B, Pierce LJ, Reed EC, Smith ML, Somlo G, Topham NS, Ward JH, Winer EP, and Wolff AC (2010). Breast cancer: noninvasive and special situations. J Natl Compr Canc Netw *8*, 1182-1207.

Carvalho S, Silva AO, Milanezi F, Ricardo S, Leitão D, Amendoeira I, and Schmitt FC (2004). c-KIT and PDGFRA in breast phyllodes tumours: overexpression without mutations? J Clin Pathol *57*, 1075-1079.

Chaney AW, Pollack A, McNeese MD, Zagars GK, Pisters PW, Pollock RE, and Hunt KK (2000). Primary treatment of cystosarcoma phyllodes of the breast. Cancer *89*, 1502-1511.

Chao TC, Lo YF, Chen SC, and Chen MF (2002). Sonographic features of phyllodes tumors of the breast. Ultrasound Obstet Gynecol *20*, 64-71.

Chen CM, Chen CJ, Chang CL, Shyu JS, Hsieh HF, and Harn HJ (2000). CD34, CD117, and actin expression in phyllodes tumor of the breast. J Surg Res *94*, 84-91.

Chen J, Xia L, Wu X, Xu L, Nie D, Shi J, Xu X, Ni L, Ju S, Wu X, Zhu H, and Shi W (2012). Clinical significance and prognostic value of PAX3 expression in human glioma. J Mol Neurosci *47*, 52-58.

Chen L, Chan THM, and Guan XY (2010). Chromosome 1q21 amplification and oncogenes in hepatocellular carcinoma. Acta Pharmacol Sin *31*, 1165-1171.

Chen WH, Cheng SP, Tzen CY, Yang TL, Jeng KS, Liu CL, and Liu TP (2005). Surgical treatment of phyllodes tumors of the breast: retrospective review of 172 cases. J Surg Oncol *91*, 185-194.

Chua CL, Thomas A, and Ng BK (1988). Cystosarcoma phyllodes-Asian variations. ANZ J Surg *58*, 301-305.

Cohn-Cedermark G, Rutqvist LE, Rosendahl I, and Silfverswärd C (1991). Prognostic factors in cystosarcoma phyllodes. A clinicopathologic study of 77 patients. Cancer *68*, 2017-2022.

Coletta RD, Christensen K, Reichenberger KJ, Lamb J, Micomonaco D, Huang L, Wolf DM, Müller-Tidow C, Golub TR, Kawakami K, and Ford HL (2004). The Six1 homeoprotein stimulates tumorigenesis by reactivation of cyclin A1. Proc Natl Acad Sci U S A *101*, 6478-6483.

Cooper WG, and Ackerman LV (1943). Cystosarcoma phyllodes: with a consideration of its malignant variant. Surg Gynecol Obstet *77*, 279–283.

Corless CL, McGreevey L, Haley A, Town A, and Heinrich MC (2002). KIT mutations are common in incidental gastrointestinal stromal tumors one centimeter or less in size. Am J Pathol *160*, 1567-1572.

Davison EJ, Tarpey PS, Fiegler H, Tomlinson IPM, and Carter NP (2005). Deletion at chromosome band 20p12.1 in colorectal cancer revealed by high resolution array comparative genomic hybridization. Genes Chromosomes Cancer *44*, 384-391.

Dietrich CU, Pandis N, Rizou H, Petersson C, Bardi G, Qvist H, Apostolikas N, Bøhler PJ, Andersen JA, Idvall I, Mitelman F, and Heim S (1997). Cytogenetic findings in phyllodes tumors of the breast: karyotypic complexity differentiates between malignant and benign tumors. Hum Pathol *28*, 1379-1382.

Diskin SJ, Hou C, Glessner JT, Attiyeh EF, Laudenslager M, Bosse K, Cole K, Mossé YP, Wood A, Lynch JE, Pecor K, Diamond M, Winter C, Wang K, Kim C, Geiger EA, McGrady PW, Blakemore AI, London WB, Shaikh TH, Bradfield J, Grant SF, Li H, Devoto M, Rappaport ER, Hakonarson H, and Maris JM (2009). Copy number variation at 1q21.1 associated with neuroblastoma. Nature *459*, 987-991.

Djordjevic B, and Hanna WM (2008). Expression of c-kit in fibroepithelial lesions of the breast is a mast cell phenomenon. Mod Pathol *21*, 1238-1245.

Do SI, Kim JY, Kang SY, Lee JJ, Lee JE, Nam SJ, and Cho EY (2013). Expression of TWIST1, Snail, Slug, and NF-κB and methylation of the TWIST1 promoter in mammary phyllodes tumor. Tumour Biol *34*, 445-453.

Dong H, Zhang H, Liang J, Yan H, Chen Y, Shen Y, Kong Y, Wang S, Zhao G, and Jin W (2011). Digital karyotyping reveals probable target genes at 7q21.3 locus in hepatocellular carcinoma. BMC Med Genomics *4*, 60.

Esposito NN, Mohan D, Brufsky A, Lin Y, Kapali M, and Dabbs DJ (2006). Phyllodes tumor: a clinicopathologic and immunohistochemical study of 30 cases. Arch Pathol Lab Med *130*, 1516-1521.

Fan H, Yuan Y, Wang J, Zhou F, Zhang M, Giercksky KE, and JMN, and Suo Z (2013). CD117 expression in operable oesophageal squamous cell carcinomas predicts worse clinical outcome. Histopathology *62*, 1028-1037.

Fiks A (1981). Cystosarcoma phyllodes of the mammary gland-Müller's tumor. For the 180th birthday of Johannes Müller. Virchows Arch *392*, 1-6.

Fischer B, Marinov M, and Arcaro A (2007). Targeting receptor tyrosine kinase signalling in small cell lung cancer (SCLC): what have we learned so far? Cancer Treat Rev *33*, 391-406.

Gatalica Z, Finkelstein S, Lucio E, Tawfik O, Palazzo J, Hightower B, and Eyzaguirre E (2001). p53 protein expression and gene mutation in phyllodes tumors of the breast. Pathol Res Pract *197*, 183-187.

Gavert N, Shvab A, Sheffer M, Ben-Shmuel A, Haase G, Bakos E, Domany E, and Ben-Ze'ev A (2013). c-Kit is suppressed in human colon cancer tissue and contributes to L1-mediated metastasis. Cancer Res *73*, 5754-5763.

Gong C, Nie Y, Qu S, Liao JY, Cui X, Yao H, Zeng Y, Su F, Song E, and Liu Q (2014). miR-21 induces myofibroblast differentiation and promotes the malignant progression of breast phyllodes tumors. Cancer Res *74*, 4341-4352.

Grifone R, Demignon J, Houbron C, Souil E, Niro C, Seller MJ, Hamard G, and Maire P (2005). Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. Development *132*, 2235-2249.

Guillot E, Couturaud B, Reyal F, Curnier A, Ravinet J, Laé M, Bollet M, Pierga JY, Salmon R, Fitoussi A, and Breast Cancer Study Group of the Institut C (2011). Management of phyllodes breast tumors. Breast J *17*, 129-137.

Gullett NP, Rizzo M, and Johnstone PAS (2009). National surgical patterns of care for primary surgery and axillary staging of phyllodes tumors. Breast J *15*, 41-44.

Hajdu SI, Espinosa MH, and Robbins GF (1976). Recurrent cystosarcoma phyllodes: a clinicopathologic study of 32 cases. Cancer *38*, 1402-1406.

Hawkins RE, Schofield JB, Wiltshaw E, Fisher C, and McKinna JA (1992). Ifosfamide is an active drug for chemotherapy of metastatic cystosarcoma phyllodes. Cancer *69*, 2271-2275.

Heinrich MC, Corless CL, Demetri GD, Blanke CD, von Mehren M, Joensuu H, McGreevey LS, Chen CJ, Van den Abbeele AD, Druker BJ, Kiese B, Eisenberg B, Roberts PJ, Singer S, Fletcher CDM, Silberman S, Dimitrijevic S, and Fletcher JA (2003). Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumor. J Clin Oncol *21*, 4342-4349.

Ho SK, Thike AA, Cheok PY, Tse GMK, and Tan PH (2013). Phyllodes tumours of the breast: the role of CD34, vascular endothelial growth factor and β-catenin in histological grading and clinical outcome. Histopathology *63*, 393-406.

Huang KT, Dobrovic A, Yan M, Karim RZ, Lee CS, Lakhani SR, and Fox SB (2010). DNA methylation profiling of phyllodes and fibroadenoma tumours of the breast. Breast Cancer Res Treat *124*, 555-565.

Jee KJ, Gong G, Ahn SH, Park JM, and Knuutila S (2003). Gain in 1q is a common abnormality in phyllodes tumours of the breast. Anal Cell Pathol (Amst) *25*, 89-93.

Jones AM, Mitter R, Poulsom R, Gillett C, Hanby AM, Tomlinson IPM, Sawyer EJ, and Phyllodes Tumour Consortium (2008b). mRNA expression profiling of phyllodes tumours of the breast: identification of genes important in the development of borderline and malignant phyllodes tumours. J Pathol *216*, 408-417.

Jones AM, Mitter R, Springall R, Graham T, Winter E, Gillett C, Hanby AM, Tomlinson IPM, Sawyer EJ, and Phyllodes Tumour Consortium (2008a). A comprehensive genetic profile of phyllodes tumours of the breast detects important mutations, intra-tumoral genetic heterogeneity and new genetic changes on recurrence. J Pathol *214*, 533-544.

Jung CW, Suh KS, Lee JS, Kim JR, Chang ES, Sul HJ, and Park MJ (2010). Mutation-free expression of c-Kit and PDGFRA in phyllodes tumors of the breast. J Breast Cancer *13*, 257.

Kapiris I, Nasiri N, A'Hern R, Healy V, and Gui GP (2001). Outcome and predictive factors of local recurrence and distant metastases following primary surgical treatment of high-grade malignant phyllodes tumours of the breast. Eur J Surg Oncol *27*, 723-730.

Karim RZ, Gerega SK, Yang YH, Horvath L, Spillane A, Carmalt H, Scolyer RA, and Lee CS (2009a). Proteins from the Wnt pathway are involved in the pathogenesis and progression of mammary phyllodes tumours. J Clin Pathol *62*, 1016-1020.

Karim RZ, Gerega SK, Yang YH, Spillane A, Carmalt H, Scolyer RA, and Lee CS (2010). p16 and pRb immunohistochemical expression increases with increasing tumour grade in mammary phyllodes tumours. Histopathology *56*, 868-875.

Karim RZ, O'Toole SA, Scolyer RA, Cooper CL, Chan B, Selinger C, Yu B, Carmalt H, Mak C, Tse GM, Tan PH, Putti TC, and Lee CS (2013). Recent insights into the molecular pathogenesis of mammary phyllodes tumours. J Clin Pathol *66*, 496-505.

Karim RZ, Scolyer RA, Tse GM, Tan PH, Putti TC, and Lee CS (2009b). Pathogenic mechanisms in the initiation and progression of mammary phyllodes tumours. Pathology *41*, 105-117.

Kattan MW (2008). Nomograms are difficult to beat. Eur Urol *53*, 671-672.

Kelleher FC, Fennelly D, and Rafferty M (2006). Common critical pathways in embryogenesis and cancer. Acta Oncol *45*, 375-388.

Kersting C, Kuijper A, Schmidt H, Packeisen J, Liedtke C, Tidow N, Gustmann C, Hinrichs B, Wülfing P, Tio J, Boecker W, van Diest P, Brandt B, and Buerger H (2006). Amplifications of the epidermal growth factor receptor gene (egfr) are common in phyllodes tumors of the breast and are associated with tumor progression. Lab Invest *86*, 54-61.

Kim JH, Choi YD, Lee JS, Lee JH, Nam JH, Choi C, Park MH, and Yoon JH (2009). Borderline and malignant phyllodes tumors display similar promoter methylation profiles. Virchows Arch *455*, 469-475.

Kim S, Kim JY, Kim DH, Jung WH, and Koo JS (2013). Analysis of phyllodes tumor recurrence according to the histologic grade. Breast Cancer Res Treat.

Kleer CG, Giordano TJ, Braun T, and Oberman HA (2001). Pathologic, immunohistochemical, and molecular features of benign and malignant phyllodes tumors of the breast. Mod Pathol *14*, 185-190.

Kondi-Pafiti A, Arkadopoulos N, Gennatas C, Michalaki V, Frangou-Plegmenou M, and Chatzipantelis P (2010). Expression of c-kit in common benign and malignant breast lesions. Tumori *96*, 978-984.

Korcheva VB, Levine J, Beadling C, Warrick A, Countryman G, Olson NR, Heinrich MC, Corless CL, and Troxell ML (2011). Immunohistochemical and molecular markers in breast phyllodes tumors. Appl Immunohistochem Mol Morphol *19*, 119-125.

Kuenen-Boumeester V, Henzen-Logmans SC, Timmermans MM, van Staveren IL, van Geel A, Peeterse HJ, Bonnema J, and Berns EM (1999). Altered expression of p53 and its regulated proteins in phyllodes tumours of the breast. J Pathol *189*, 169-175.

Kuijper A, Buerger H, Simon R, Schaefer KL, Croonen A, Boecker W, van der Wall E, and van Diest PJ (2002). Analysis of the progression of fibroepithelial tumours of the breast by PCR-based clonality assay. J Pathol *197*, 575-581.

Kuijper A, de Vos RAI, Lagendijk JH, van der Wall E, and van Diest PJ (2005a). Progressive deregulation of the cell cycle with higher tumor grade in the stroma of breast phyllodes tumors. Am J Clin Pathol *123*, 690-698.

Kuijper A, Snijders AM, Berns EMJJ, Kuenen-Boumeester V, van der Wall E, Albertson DG, and van Diest PJ (2009). Genomic profiling by array comparative genomic hybridization reveals novel DNA copy number changes in breast phyllodes tumours. Cell Oncol *31*, 31-39.

Kuijper A, van der Groep P, van der Wall E, and van Diest PJ (2005b). Expression of hypoxia-inducible factor 1 alpha and its downstream targets in fibroepithelial tumors of the breast. Breast Cancer Res *7*, R808-818.

Kwon JE, Jung WH, and Koo JS (2012). Molecules involved in epithelial-mesenchymal transition and epithelial-stromal interaction in phyllodes tumors: implications for histologic grade and prognosis. Tumour Biol *33*, 787-798.

Ladesich J, Damjanov I, Persons D, Jewell W, Arthur T, Rogana J, and Davoren B (2002). Complex karyotype in a low grade phyllodes tumor of the breast. Cancer Genet Cytogenet *132*, 149-151.

Laé M, Vincent-Salomon A, Savignoni A, Huon I, Fréneaux P, Sigal-Zafrani B, Aurias A, Sastre-Garau X, and Couturier J (2007). Phyllodes tumors of the breast segregate in two groups according to genetic criteria. Mod Pathol *20*, 435-444.

Lee BJ, and Pack GT (1931). Giant intracanalicular myxoma of the breast: the so-called cystosarcoma phyllodes mammae of Johannes Müller. Ann Surg *93*, 250-268.

Lee J, Wang J, Torbenson M, Lu Y, Liu QZ, and Li S (2010). Loss of SDHB and NF1 genes in a malignant phyllodes tumor of the breast as detected by oligo-array comparative genomic hybridization. Cancer Genet Cytogenet *196*, 179-183.

Lennartsson J, and Rönnstrand L (2012). Stem cell factor receptor/c-Kit: from basic science to clinical implications. Physiol Rev *92*, 1619-1649.

Lin JJ, Huang CS, Yu J, Liao GS, Lien HC, Hung JT, Lin RJ, Chou FP, Yeh KT, and Yu AL (2014). Malignant phyllodes tumors display mesenchymal stem cell features and aldehyde dehydrogenase/disialoganglioside identify their tumor stem cells. Breast Cancer Res *16*, R29.

Loughrey MB, Trivett M, Beshay V, Dobrovic A, Kovalenko S, Murray W, Lade S, Turner H, McArthur GA, and JZ, and Waring PM (2006). KIT immunohistochemistry and mutation status in gastrointestinal stromal tumours GISTs evaluated for treatment with imatinib. Histopathology *49*, 52-65.

Lu YJ, Birdsall S, Osin P, Gusterson B, and Shipley J (1997). Phyllodes tumors of the breast analyzed by comparative genomic hybridization and association

of increased 1q copy number with stromal overgrowth and recurrence. Genes Chromosomes Cancer *20*, 275-281.

Lucas DR, Al-Abbadi M, Tabaczka P, Hamre MR, Weaver DW, and Mott MJ (2003). c-Kit Expression in desmoid fibromatosis: comparative Immunohistochemical evaluation of two commercial antibodies. Am J Clin Pathol *119*, 339-345.

Lv S, Niu Y, Wei L, Liu Q, Wang X, and Chen Y (2008). Chromosomal aberrations and genetic relations in benign, borderline and malignant phyllodes tumors of the breast: a comparative genomic hybridization study. Breast Cancer Res Treat *112*, 411-418.

Macdonald OK, Lee CM, Tward JD, Chappel CD, and Gaffney DK (2006). Malignant phyllodes tumor of the female breast: association of primary therapy with cause-specific survival from the Surveillance, Epidemiology, and End Results (SEER) program. Cancer *107*, 2127-2133.

McCarty KS, Miller LS, Cox EB, and Konrath J (1985). Estrogen receptor analyses. Correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies. Arch Pathol Lab Med *109*, 716-721.

McCoy EL, Iwanaga R, Jedlicka P, Abbey NS, Chodosh LA, Heichman KA, Welm AL, and Ford HL (2009). Six1 expands the mouse mammary epithelial stem/progenitor cell pool and induces mammary tumors that undergo epithelial-mesenchymal transition. J Clin Invest *119*, 2663-2677.

Micalizzi DS, Christensen KL, Jedlicka P, Coletta RD, Barón AE, Harrell JC, Horwitz KB, Billheimer D, Heichman KA, Welm AL, Schiemann WP, and Ford HL (2009). The Six1 homeoprotein induces human mammary carcinoma cells to undergo epithelial-mesenchymal transition and metastasis in mice through increasing TGF-beta signaling. J Clin Invest *119*, 2678-2690.

Morales-Vásquez F, Gonzalez-Angulo AM, Broglio K, Lopez-Basave HN, Gallardo D, Hortobagyi GN, and De La Garza JG (2007). Adjuvant chemotherapy with doxorubicin and dacarbazine has no effect in recurrence-free survival of malignant phyllodes tumors of the breast. Breast J *13*, 551-556.

Ng DC, Lin BH, Lim CP, Huang G, Zhang T, Poli V, and Cao X (2006a). Stat3 regulates microtubules by antagonizing the depolymerization activity of stathmin. J Cell Biol *172*, 245-257.

Ng KT, Man K, Sun CK, Lee TK, Poon RT, Lo CM, and Fan ST (2006b). Clinicopathological significance of homeoprotein Six1 in hepatocellular carcinoma. Br J Cancer *95*, 1050-1055.

Niezabitowski A, Lackowska B, Rys J, Kruczak A, Kowalska T, Mitus J, Reinfuss M, and Markiewicz D (2001). Prognostic evaluation of proliferative activity

and DNA content in the phyllodes tumor of the breast: immunohistochemical and flow cytometric study of 118 cases. Breast Cancer Res Treat *65*, 77-85.

Nishimura R, Tan PH, Thike AA, Tan MH, Taira N, Li HH, and Ohsumi S (2014). Utility of the Singapore nomogram for predicting recurrence-free survival in Japanese women with breast phyllodes tumours. J Clin Pathol *67*, 748-750.

Noguchi S, Motomura K, Inaji H, Imaoka S, and Koyama H (1993). Clonal analysis of fibroadenoma and phyllodes tumor of the breast. Cancer Res *53*, 4071-4074.

Noguchi S, Yokouchi H, Aihara T, Motomura K, Inaji H, Imaoka S, and Koyama H (1995). Progression of fibroadenoma to phyllodes tumor demonstrated by clonal analysis. Cancer *76*, 1779-1785.

Noronha Y, Raza A, Hutchins B, Chase D, Garberoglio C, Chu P, Weiss L, and Wang J (2011). CD34, CD117, and Ki-67 expression in phyllodes tumor of the breast: an immunohistochemical study of 33 cases. Int J Surg Pathol *19*, 152-158.

Norris HJ, and Taylor HB (1967). Relationship of histologic features to behavior of cystosarcoma phyllodes. Analysis of ninety-four cases. Cancer *20*, 2090-2099.

Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, Naoki K, Sasaki H, Fujii Y, Eck MJ, Sellers WR, Johnson BE, and Meyerson M (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science *304*, 1497-1500.

Pandey M, Mathew A, Kattoor J, Abraham EK, Mathew BS, Rajan B, and Nair KM (2001). Malignant phyllodes tumor. Breast J *7*, 411-416.

Parker SJ, and Harries SA (2001). Phyllodes tumours. Postgrad Med J *77*, 428-435.

Peduzzi P, Concato J, Feinstein AR, and Holford TR (1995). Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol *48*, 1503-1510.

Pezner RD, Schultheiss TE, and Paz IB (2008). Malignant phyllodes tumor of the breast: local control rates with surgery alone. Int J Radiat Oncol Biol Phys *71*, 710-713.

Pietruszka M, and Barnes L (1978). Cystosarcoma phyllodes: a clinicopathologic analysis of 42 cases. Cancer *41*, 1974-1983.

Pimiento JM, Gadgil PV, Santillan AA, Lee MC, Esposito NN, Kiluk JV, Khakpour N, Hartley TL, Yeh IT, and Laronga C (2011). Phyllodes tumors: race-related differences. J Am Coll Surg *213*, 537-542.

Pinder SE, Brown JP, Gillett C, Purdie CA, Speirs V, Thompson AM, Shaaban AM, and Translational Subgroup of the NBCSG (2013). The manufacture and assessment of tissue microarrays: suggestions and criteria for analysis, with breast cancer as an example. J Clin Pathol *66*, 169-177.

Pittoni P, Piconese S, Tripodo C, and Colombo MP (2011). Tumor-intrinsic and -extrinsic roles of c-Kit: mast cells as the primary off-target of tyrosine kinase inhibitors. Oncogene *30*, 757-769.

Ramos-Vara JA, and Miller MA (2014). When tissue antigens and antibodies get along: revisiting the technical aspects of immunohistochemistry-the red, brown, and blue technique. Vet Pathol *51*, 42-87.

Reinfuss M, Mituś J, Duda K, Stelmach A, Ryś J, and Smolak K (1996). The treatment and prognosis of patients with phyllodes tumor of the breast: an analysis of 170 cases. Cancer *77*, 910-916.

Ren D, Li Y, Gong Y, Xu J, Miao X, Li X, Liu C, Jia L, and Zhao Y (2014). Phyllodes tumor of the breast: role of Axl and ST6GalNAcII in the development of mammary phyllodes tumors. Tumour Biol *35*, 9603-9612.

Ridgeway AG, and Skerjanc IS (2001). Pax3 is essential for skeletal myogenesis and the expression of Six1 and Eya2. J Biol Chem *276*, 19033-19039.

Roa JC, Tapia O, Carrasco P, Contreras E, Araya JC, Muñoz S, and Roa I (2006). Prognostic factors of phyllodes tumor of the breast. Pathol Int *56*, 309-314.

Rosen DG, Huang X, Deavers MT, Malpica A, Silva EG, and Liu J (2004). Validation of tissue microarray technology in ovarian carcinoma. Mod Pathol *17*, 790-797.

Samuel S, and Naora H (2005). Homeobox gene expression in cancer: insights from developmental regulation and deregulation. Eur J Cancer *41*, 2428-2437.

Sawhney N, Garrahan N, Douglas-Jones AG, and Williams ED (1992). Epithelial-stromal interactions in tumors. A morphologic study of fibroepithelial tumors of the breast. Cancer *70*, 2115-2120.

Sawyer EJ, Hanby AM, Ellis P, Lakhani SR, Ellis IO, Boyle S, and Tomlinson IP (2000). Molecular analysis of phyllodes tumors reveals distinct changes in the epithelial and stromal components. Am J Pathol *156*, 1093-1098.

Sawyer EJ, Hanby AM, Rowan AJ, Gillett CE, Thomas RE, Poulsom R, Lakhani SR, Ellis IO, Ellis P, and Tomlinson IPM (2002). The Wnt pathway, epithelial-

stromal interactions, and malignant progression in phyllodes tumours. J Pathol *196*, 437-444.

Sawyer EJ, Poulsom R, Hunt FT, Jeffery R, Elia G, Ellis IO, Ellis P, Tomlinson IP, and Hanby AM (2003). Malignant phyllodes tumours show stromal overexpression of c-myc and c-kit. J Pathol *200*, 59-64.

Schiffman JD, Hodgson JG, VandenBerg SR, Flaherty P, Polley MYC, Yu M, Fisher PG, Rowitch DH, Ford JM, Berger MS, Ji H, Gutmann DH, and James CD (2010). Oncogenic BRAF mutation with CDKN2A inactivation is characteristic of a subset of pediatric malignant astrocytomas. Cancer Res *70*, 512-519.

Shariat SF, Karakiewicz PI, Godoy G, and Lerner SP (2009). Use of nomograms for predictions of outcome in patients with advanced bladder cancer. Ther Adv Urol *1*, 13-26.

Shpitz B, Bomstein Y, Sternberg A, Klein E, Tiomkin V, Kaufman A, Groisman G, and Bernheim J (2002). Immunoreactivity of p53, Ki-67, and c-erbB-2 in phyllodes tumors of the breast in correlation with clinical and morphologic features. J Surg Oncol *79*, 86-92.

Smith AL, Iwanaga R, Drasin DJ, Micalizzi DS, Vartuli RL, Tan AC, and Ford HL (2012). The miR-106b-25 cluster targets Smad7, activates TGF-β signaling, and induces EMT and tumor initiating cell characteristics downstream of Six1 in human breast cancer. Oncogene *31*, 5162-5171.

Spitaleri G, Toesca A, Botteri E, Bottiglieri L, Rotmensz N, Boselli S, Sangalli C, Catania C, Toffalorio F, Noberasco C, Delmonte A, Luini A, Veronesi P, Colleoni M, Viale G, Zurrida S, Goldhirsch A, Veronesi U, and De Pas T (2013). Breast phyllodes tumor: A review of literature and a single center retrospective series analysis. Crit Rev Oncol Hematol *88*, 427-436.

Stebbing JF, and Nash AG (1995). Diagnosis and management of phyllodes tumour of the breast: experience of 33 cases at a specialist centre. Ann R Coll Surg Engl *77*, 181-184.

Stratton MR, Campbell PJ, and Futreal PA (2009). The cancer genome. Nature *458*, 719-724.

Tada M, Kanai F, Tanaka Y, Sanada M, Nannya Y, Tateishi K, Ohta M, Asaoka Y, Seto M, Imazeki F, Yoshida H, Ogawa S, Yokosuka O, and Omata M (2010). Prognostic significance of genetic alterations detected by high-density single nucleotide polymorphism array in gastric cancer. Cancer Sci *101*, 1261-1269.

Tan PH, Jayabaskar T, Chuah KL, Lee HY, Tan Y, Hilmy M, Hung H, Selvarajan S, and Bay BH (2005a). Phyllodes tumors of the breast: the role of pathologic parameters. Am J Clin Pathol *123*, 529-540.

Tan PH, Jayabaskar T, Yip G, Tan Y, Hilmy M, Selvarajan S, and Bay BH (2005b). p53 and c-kit (CD117) protein expression as prognostic indicators in breast phyllodes tumors: a tissue microarray study. Mod Pathol *18*, 1527-1534.

Tan PH, Tse G, Lee A, Simpson JF, and Hanby AM (2012). Fibroepithelial Tumours. In WHO Classification of Tumours of the Breast, Lakhani S, Ellis I, Schnitt S, Tan P, and van de Vijver M, eds. (Lyon: International Agency for Research on Cancer), pp. 142-147.

Treves N, and Sunderland DA (1951). Cystosarcoma phyllodes of the breast: a malignant and a benign tumor; a clinicopathological study of seventy-seven cases. Cancer *4*, 1286-1332.

Tsang JY, Mendoza P, Lam CC, Yu AM, Putti TC, Karim RZ, Scolyer RA, Lee CS, Tan PH, and Tse GM (2012a). Involvement of α- and β-catenins and E-cadherin in the development of mammary phyllodes tumours. Histopathology *61*, 667-674.

Tsang JY, Mendoza P, Putti TC, Karim RZ, Scolyer RA, Lee CS, Pang AL, and Tse GM (2012b). E-cadherin expression in the epithelial components of mammary phyllodes tumors. Hum Pathol *43*, 2117-2123.

Tse GMK, Lee CS, Kung FYL, Scolyer RA, Law BKB, Lau TS, and Putti TC (2002). Hormonal receptors expression in epithelial cells of mammary phyllodes tumors correlates with pathologic grade of the tumor: a multicenter study of 143 cases. Am J Clin Pathol *118*, 522-526.

Tse GMK, Lui PCW, Scolyer RA, Putti TC, Kung FYL, Law BKB, Lau TS, and Lee CS (2003). Tumour angiogenesis and p53 protein expression in mammary phyllodes tumors. Mod Pathol *16*, 1007-1013.

Tse GMK, Lui PCW, Vong JSL, Lau KM, Putti TC, Karim R, Scolyer RA, Lee CS, Yu AMC, Ng DCH, Tse AKY, and Tan PH (2009). Increased epidermal growth factor receptor (EGFR) expression in malignant mammary phyllodes tumors. Breast Cancer Res Treat *114*, 441-448.

Tse GMK, Putti TC, Lui PCW, Lo AWI, Scolyer RA, Law BKB, Karim R, and Lee CS (2004). Increased c-kit (CD117) expression in malignant mammary phyllodes tumors. Mod Pathol *17*, 827-831.

Turalba CI, el-Mahdi AM, and Ladaga L (1986). Fatal metastatic cystosarcoma phyllodes in an adolescent female: case report and review of treatment approaches. J Surg Oncol *33*, 176-181.

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, and Kinzler KW (2013). Cancer genome landscapes. Science *339*, 1546-1558.

Wang CA, Jedlicka P, Patrick AN, Micalizzi DS, Lemmer KC, Deitsch E, Casás-Selves M, Harrell JC, and Ford HL (2012a). SIX1 induces lymphangiogenesis

and metastasis via upregulation of VEGF-C in mouse models of breast cancer. J Clin Invest *122*, 1895-1906.

Wang Y, Cottman M, and Schiffman JD (2012b). Molecular inversion probes: a novel microarray technology and its application in cancer research. Cancer Genet *205*, 341-355.

Wang Y, Moorhead M, Karlin-Neumann G, Wang NJ, Ireland J, Lin S, Chen C, Heiser LM, Chin K, Esserman L, Gray JW, Spellman PT, and Faham M (2007). Analysis of molecular inversion probe performance for allele copy number determination. Genome Biol *8*, R246.

Wang ZC, Buraimoh A, Iglehart JD, and Richardson AL (2006). 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 *97*, 301-309.

Went PT (2004). Prevalence of KIT expression in human tumors. J Clin Oncol *22*, 4514-4522.

White JW (1940). Malignant variant of cystosarcoma phyllodes. Am J Cancer *40*, 458-464.

Woolley PV, Gollin SM, Riskalla W, Finkelstein S, Stefanik DF, Riskalla L, Swaney WP, Weisenthal L, and McKenna RJ (2000). Cytogenetics, immunostaining for fibroblast growth factors, p53 sequencing, and clinical features of two cases of cystosarcoma phyllodes. Mol Diagn *5*, 179-190.

Yonemori K, Hasegawa T, Shimizu C, Shibata T, Matsumoto K, Kouno T, Ando M, Katsumata N, and Fujiwara Y (2006). Correlation of p53 and MIB-1 expression with both the systemic recurrence and survival in cases of phyllodes tumors of the breast. Pathol Res Pract *202*, 705-712.

Yu Y, Khan J, Khanna C, Helman L, Meltzer PS, and Merlino G (2004). Expression profiling identifies the cytoskeletal organizer ezrin and the developmental homeoprotein Six-1 as key metastatic regulators. Nat Med *10*, 175-181.

Zheng XH, Liang PH, Guo JX, Zheng YR, Han J, Yu LL, Zhou YG, and Li L (2010). Expression and clinical implications of homeobox gene Six1 in cervical cancer cell lines and cervical epithelial tissues. Int J Gynecol Cancer *20*, 1587-1592.

Zissis C, Apostolikas N, Konstantinidou A, Griniatsos J, and Vassilopoulos PP (1998). The extent of surgery and prognosis of patients with phyllodes tumor of the breast. Breast Cancer Res Treat *48*, 205-210.

Zurrida S, Bartoli C, Galimberti V, Squicciarini P, Delledonne V, Veronesi P, Bono A, de Palo G, and Salvadori B (1992). Which therapy for unexpected phyllode tumour of the breast? Eur J Cancer *28*, 654-657.

APPENDIX

Supplemental Table 1 Details of chromosomal aberrations observed in 19 phyllodes tumours which passed quality control on the OncoScan® FFPE assay. Copy number (CN) gain is defined with gain of two copies of chromosomes and above. CN loss is defined with loss of one copy while homozygous copy loss is defined with loss of both copies.

Sample	Event	Chromosome Region	Cytoband	Length (bp)
1	CN Gain	chr1 0-1,222,808	1p36.33	1222809
1	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
1	CN Loss	chr6 43,695,136-46,893,364	6p21.1 - p12.3	3198229
1	CN Loss	chr6 47,280,591-58,742,393	6p12.3 - p11.1	11461803
1	CN Loss	chr6 61,000,000-65,425,900	6q11.1 - q12	4425901
1	CN Loss	chr6 66,891,799-67,311,468	6q12	419670
1	CN Loss	chr6 69,470,911-70,061,379	6q12 - q13	590469
1	CN Loss	chr6 70,511,448-71,744,348	6q13	1232901
1	CN Loss	chr6 72,354,086-78,485,370	6q13 - q14.1	6131285
1	CN Loss	chr6 80,410,916-84,593,162	6q14.1 - q14.2	4182247
1	CN Loss	chr6 85,152,217-90,311,887	6q14.3 - q15	5159671
1	CN Loss	chr6 91,295,356-91,746,966	6q15	451611
1	CN Loss	chr6 93,213,080-97,358,701	6q16.1	4145622
1	CN Loss	chr6 97,997,670-105,695,999	6q16.1 - q21	7698330
1	CN Loss	chr6 109,938,879-113,818,039	6q21	3879161
1	CN Loss	chr6 114,308,474-116,560,702	6q21 - q22.1	2252229
1	CN Loss	chr6 117,149,837-122,168,871	6q22.1 - q22.31	5019035
1	CN Loss	chr6 123,339,341-125,505,267	6q22.31	2165927
1	CN Loss	chr6 132,876,773-132,941,505	6q23.2	64733
1	CN Loss	chr6 133,228,189-134,141,984	6q23.2	913796
1	CN Loss	chr6 134,498,087-136,280,453	6q23.2 - q23.3	1782367
1	CN Loss	chr6 149,779,031-152,499,603	6q25.1	2720573
1	CN Loss	chr6 152,626,737-154,442,806	6q25.2	1816070
1	CN Loss	chr6 155,085,569-156,832,599	6q25.2 - q25.3	1747031
1	CN Loss	chr6 159,572,985-160,064,613	6q25.3	491629
1	CN Loss	chr6 160,446,357-171,115,067	6q25.3 - q27	10668711
1	CN Loss	chr8 6,788,988-6,911,399	8p23.1	122412
1	CN Gain	chr9 75,699,001-75,774,198	9q21.13	75198
1	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
1	CN Loss	chr15 22,796,596-38,019,749	15q11.2 - q14	15223154
1	CN Loss	chr15 44,098,711-48,523,894	15q15.3 - q21.1	4425184
1	CN Loss	chr15 51,291,741-52,238,170	15q21.2	946430
1	CN Loss	chr15 53,028,725-54,518,893	15q21.3	1490169
1	CN Loss	chr15 54,901,706-55,450,209	15q21.3	548504
1	CN Loss	chr15 58,066,967-58,611,769	15q21.3	544803
1	CN Loss	chr15 59,657,100-60,637,187	15q22.2	980088
1	CN Loss	chr15 60,987,346-61,189,584	15q22.2	202239
1	CN Loss	chr16 24,134,820-24,233,073	16p12.2 - p12.1	98254
1	CN Loss	chr17 2,948,968-3,751,282	17p13.3 - p13.2	802315

1	CN Loss	chr17 5,398,280-6,811,774	17p13.2 - p13.1	1413495
1	CN Gain	chr20 45,903,411-45,992,048	20q13.12	88638
1	CN Loss	chr21 45,297,385-47,523,589	21q22.3	2226205
1	CN Loss	chr22 48,870,985-50,030,785	22q13.32 - q13.33	1159801
2	CN Gain	chr10 92,137,961-92,295,085	10q23.31	157125
2	CN Loss	chrX 2,699,968-58,545,809	Xp22.33 - p11.1	55845842
2	CN Loss	chrX 61,944,035-155,270,560	Xq11.1 - q28	93326526
3	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
4	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
5	CN Loss	chr1 5,004,041-5,919,059	1p36.32 - p36.31	915019
5	CN Loss	chr1 148,915,961-149,761,266	1q21.2	845306
5	CN Gain	chr1 87,785,758-87,818,488	1p22.3	32731
5	CN Gain	chr1 145,419,163-145,527,032	1q21.1	107870
5	CN Loss	chr2 23,526,044-23,617,364	2p24.1	91321
5	CN Loss	chr3 10,440,707-11,019,981	3p25.3	579275
5	CN Loss	chr5 174,176,962-175,576,829	5q35.2	1399868
5	CN Gain	chr5 148,821,669-148,977,145	5q32	155477
5	CN Gain	chr6 25,992,859-26,064,704	6p22.2	71846
5	CN Gain	chr6 26,146,304-26,226,155	6p22.2	79852
5	CN Gain	chr6 27,746,874-27,874,436	6p22.1	127563
5	CN Loss	chr7 56,295,982-57,878,090	7p11.2	1582109
5	CN Loss	chr7 154,873,687-156,025,329	7q36.2 - q36.3	1151643
5	CN Loss	chr7 157,281,738-158,423,097	7q36.3	1141360
5	CN Gain	chr7 94,027,737-94,050,306	7q21.3	22570
5	CN Gain	chr7 134,501,890-134,516,257	7q33	14368
5	CN Loss	chr8 462,508-1,565,363	8p23.3	1102856
5	CN Loss	chr8 139,743,394-141,074,233	8q24.23 - q24.3	1330840
5	CN Loss	chr8 142,358,487-143,987,939	8q24.3	1629453
5	CN Gain	chr8 39,958,745-40,048,822	8p11.21	90078
5	CN Loss	chr9 36,778,731-37,014,719	9p13.2	235989
5	CN Loss	chr9 101,062,838-101,172,425	9q22.33	109588
5	CN Loss	chr10 1,177,416-2,211,229	10p15.3	1033814
5	CN Loss	chr10 26,744,455-26,977,870	10p12.1	233416
5	CN Loss	chr10 87,844,547-88,134,907	10q23.1 - q23.2	290361
5	CN Loss	chr10 132,084,219-135,534,747	10q26.3	3450529
5	CN Loss	chr11 116,180,571-116,525,551	11q23.3	344981
5	CN Loss	chr11 126,356,718-126,958,052	11q24.2	601335
5	CN Loss	chr11 133,115,479-133,701,477	11q25	585999
5	CN Gain	chr11 65,049,910-65,275,146	11q13.1	225237
5	CN Loss	chr12 112,891,067-112,948,019	12q24.13	56953
5	CN Loss	chr12 113,936,155-114,267,840	12q24.13	331686
5	CN Loss	chr12 118,839,701-120,103,511	12q24.23	1263811
5	CN Loss	chr12 128,413,609-130,552,157	12q24.32 - q24.33	2138549
5	CN Loss	chr12 130,801,237-131,293,433	12q24.33	492197
5	CN Loss	chr13 111,867,463-113,134,406	13q34	1266944
5	CN Gain	chr13 37,916,644-38,192,355	13q13.3	275712

5	CN Loss	chr14 24,921,826-25,102,911	14q12	181086
5	CN Loss	chr14 96,171,889-96,442,430	14q32.13 - q32.2	270542
5	CN Gain	chr14 68,366,107-68,774,161	14q24.1	408055
5	CN Gain	chr14 68,831,621-68,886,193	14q24.1	54573
5	CN Gain	chr14 77,482,341-77,515,382	14q24.3	33042
5	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
5	CN Loss	chr15 27,890,540-28,332,292	15q12 - q13.1	441753
5	CN Loss	chr15 88,724,961-89,003,474	15q25.3	278514
5	CN Gain	chr15 93,338,187-93,479,204	15q26.1	141018
5	CN Loss	chr16 5,173,205-9,768,469	16p13.3 - p13.2	4595265
5	CN Loss	chr16 23,613,369-24,631,421	16p12.2 - p12.1	1018053
5	CN Loss	chr16 33,828,054-35,219,549	16p11.2 - p11.1	1391496
5	CN Loss	chr16 49,470,113-49,653,214	16q12.1	183102
5	CN Loss	chr16 83,378,232-83,710,981	16q23.3	332750
5	CN Gain	chr16 79,621,593-79,673,887	16q23.2	52295
5	CN Loss	chr17 32,732,800-33,104,907	17q12	372108
5	CN Loss	chr17 35,046,000-35,295,966	17q12	249967
5	CN Loss	chr18 54,709,502-55,207,482	18q21.31	497981
5	CN Loss	chr18 76,061,828-76,627,951	18q23	566124
5	CN Gain	chr18 3,592,059-3,608,238	18p11.31	16180
5	CN Gain	chr18 41,424,543-41,588,101	18q12.3	163559
5	CN Loss	chr19 8,786,085-9,452,179	19p13.2	666095
5	CN Loss	chr19 28,079,145-30,343,191	19q11 - q12	2264047
5	CN Loss	chr19 42,074,179-42,247,949	19q13.2	173771
5	CN Loss	chr19 42,967,112-43,921,371	19q13.2 - q13.31	954260
5	CN Loss	chr19 54,123,228-55,512,121	19q13.42	1388894
5	CN Loss	chr19 57,113,246-57,672,244	19q13.43	558999
5	CN Loss	chr20 23,466,226-24,895,517	20p11.21	1429292
5	CN Gain	chr20 45,892,532-45,992,048	20q13.12	99517
5	CN Loss	chr22 17,299,929-18,163,987	22q11.1 - q11.21	864059
5	CN Loss	chr22 48,856,519-50,140,926	22q13.32 - q13.33	1284408
6	CN Gain	chr7 94,015,482-94,059,477	7q21.3	43996
6	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
6	CN Gain	chr17 25,326,941-25,476,780	17q11.1	149840
7	CN Loss	chr9 11,910,201-12,189,666	9p23	279466
8	CN Loss	chr6 94,728,264-95,063,724	6q16.1	335461
8	CN Gain	chr6 26,146,304-26,226,155	6p22.2	79852
8	CN Gain	chr16 68,612,846-68,860,224	16q22.1	247379
8	CN Loss	chrX 110,219,820-110,495,293	Xq23	275474
9	CN Gain	chr1 21,567,243-21,666,377	1p36.12	99135
9	CN Gain	chr1 68,662,447-68,693,759	1p31.3	31313
9	CN Gain	chr1 145,419,163-145,527,032	1q21.1	107870
9	CN Gain	chr1 219,593,792-219,652,927	1q41	59136
9	CN Gain	chr1 234,602,305-234,929,566	1q42.2 - q42.3	327262
9	CN Gain	chr2 189,537,698-189,964,459	2q32.2	426762
9	CN Gain	chr2 190,396,012-190,468,951	2q32.2	72940

9	CN Gain	chr2 238,215,913-238,376,003	2q37.3	160091
9	CN Loss	chr3 10,437,802-10,973,215	3p25.3	535414
9	CN Loss	chr4 177,952,803-178,143,805	4q34.3	191003
9	CN Loss	chr4 188,956,020-191,154,276	4q35.2	2198257
9	CN Gain	chr5 148,776,185-148,977,145	5q32	200961
9	CN Gain	chr6 26,003,552-26,064,704	6p22.2	61153
9	CN Gain	chr6 26,123,254-26,244,409	6p22.2	121156
9	CN Loss	chr7 0-5,719,311	7p22.3 - p22.1	5719312
9	CN Loss	chr7 6,488,745-7,583,246	7p22.1 - p21.3	1094502
9	CN Loss	chr7 7,935,502-11,465,014	7p21.3	3529513
9	CN Loss	chr7 11,777,627-12,386,445	7p21.3	608819
9	CN Loss	chr7 13,284,052-16,483,845	7p21.3 - p21.2	3199794
9	CN Loss	chr7 17,745,173-28,949,939	7p21.1 - p14.3	11204767
9	CN Loss	chr7 29,159,440-36,124,742	7p14.3 - p14.2	6965303
9	CN Loss	chr7 36,437,831-41,988,603	7p14.2 - p14.1	5550773
9	CN Loss	chr7 43,157,773-47,787,860	7p14.1 - p12.3	4630088
9	CN Loss	chr7 47,881,384-55,018,866	7p12.3 - p11.2	7137483
9	CN Loss	chr7 55,192,609-57,878,090	7p11.2	2685482
9	CN Loss	chr7 61,064,518-80,287,194	7q11.1 - q21.11	19222677
9	CN Loss	chr7 82,011,079-90,869,405	7q21.11 - q21.13	8858327
9	CN Loss	chr7 91,254,845-93,810,982	7q21.2 - q21.3	2556138
9	CN Loss	chr7 94,056,185-99,543,334	7q21.3 - q22.1	5487150
9	CN Loss	chr7 100,535,078-129,255,365	7q22.1 - q32.2	28720288
9	CN Loss	chr7 130,860,464-134,239,928	7q32.3 - q33	3379465
9	CN Loss	chr7 134,907,932-142,912,836	7q33 - q34	8004905
9	CN Loss	chr7 143,657,572-159,138,663	7q35 - q36.3	15481092
9	CN Loss	chr8 142,491,110-143,987,939	8q24.3	1496830
9	CN Gain	chr8 17,564,038-17,729,015	8p22	164978
9	CN Loss	chr9 11,793,164-12,189,666	9p23	396503
9	CN Loss	chr10 132,735,673-135,534,747	10q26.3	2799075
9	CN Gain	chr10 33,555,319-33,614,606	10p11.22	59288
9	CN Gain	chr10 95,182,850-95,243,902	10q23.33	61053
9	CN Gain	chr11 64,854,390-65,275,146	11q13.1	420757
9	CN Gain	chr13 48,977,891-48,988,059	13q14.2	10169
9	CN Gain	chr13 110,924,799-111,062,282	13q34	137484
9	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
9	CN Gain	chr15 60,654,900-60,696,050	15q22.2	41151
9	CN Gain	chr15 93,367,501-93,479,204	15q26.1	111704
9	CN Loss	chr16 6,752,964-6,834,966	16p13.3	82003
9	CN Loss	chr16 7,174,796-8,791,398	16p13.3 - p13.2	1616603
9	CN Loss	chr16 23,689,373-24,093,371	16p12.2	403999
9	CN Gain	chr16 79,627,608-79,675,910	16q23.2	48303
9	CN Loss	chr17 32,732,800-33,094,213	17q12	361414
9	CN Loss	chr17 45,749,349-45,896,180	17q21.32	146832
9	CN Gain	chr17 57,784,507-57,996,457	17q23.1	211951
9	CN Gain	chr18 3,446,907-3,458,458	18p11.31	11552

9	CN Gain	chr18 41,394,247-41,525,672	18q12.3	131426
9	CN Gain	chr18 52,898,161-53,090,008	18q21.2	191848
9	CN Loss	chr19 8,984,956-9,066,324	19p13.2	81369
9	CN Loss	chr19 28,079,145-30,766,524	19q11 - q12	2687380
9	CN Loss	chr19 54,123,228-55,512,121	19q13.42	1388894
9	CN Loss	chr22 17,299,929-17,721,349	22q11.1	421421
9	CN Loss	chr22 48,686,018-49,860,392	22q13.32 - q13.33	1174375
9	CN Gain	chr22 36,707,127-36,858,770	22q12.3	151644
9	CN Loss	chrX 21,263,462-22,162,506	Xp22.12 - p22.11	899045
10	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
10	CN Loss	chr5 100,769,075-161,769,011	5q21.1 - q34	60999937
10	CN Loss	chr8 5,993,004-6,093,800	8p23.2	100797
10	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
12	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
12	CN Loss	chr5 0-2,882,758	5p15.33	2882759
12	CN Loss	chr5 3,268,942-5,249,875	5p15.33 - p15.32	1980934
12	CN Loss	chr5 5,409,897-5,505,160	5p15.32	95264
12	CN Loss	chr5 5,672,866-9,623,641	5p15.32 - p15.31	3950776
12	CN Loss	chr5 10,050,032-10,236,091	5p15.2	186060
12	CN Loss	chr5 10,333,382-11,792,960	5p15.2	1459579
12	CN Loss	chr5 12,004,786-13,390,161	5p15.2	1385376
12	CN Loss	chr5 13,457,974-14,150,975	5p15.2	693002
12	CN Loss	chr5 14,231,227-14,924,193	5p15.2	692967
12	CN Loss	chr5 15,527,063-23,373,037	5p15.1 - p14.2	7845975
12	CN Loss	chr5 24,053,709-27,316,842	5p14.2 - p14.1	3263134
12	CN Gain	chr5 2,882,758-3,268,942	5p15.33	386185
12	CN Gain	chr5 5,505,160-5,672,866	5p15.32	167707
12	CN Gain	chr5 9,623,641-10,050,032	5p15.31 - p15.2	426392
12	CN Gain	chr5 10,236,091-10,333,382	5p15.2	97292
12	CN Gain	chr5 13,390,161-13,457,974	5p15.2	67814
12	CN Gain	chr5 14,150,975-14,231,227	5p15.2	80253
12	CN Gain	chr5 14,924,193-15,527,063	5p15.2 - p15.1	602871
12	CN Gain	chr5 23,373,037-24,053,709	5p14.2	680673
12	CN Gain	chr5 27,316,842-29,053,732	5p14.1 - p13.3	1736891
12	CN Gain	chr5 29,538,140-34,728,143	5p13.3 - p13.2	5190004
12	CN Gain	chr5 36,710,291-39,402,423	5p13.2 - p13.1	2692133
12	CN Gain	chr5 40,227,613-40,788,044	5p13.1	560432
12	CN Gain	chr7 94,015,482-94,055,161	7q21.3	39680
12	CN Gain	chr8 2,207,062-2,402,151	8p23.2	195090
12	CN Loss	chr9 20,630,484-22,378,924	9p21.3	1748441
12	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
13	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
13	CN Loss	chr5 0-30,572,608	5p15.33 - p13.3	30572609
13	CN Loss	chr9 20,608,137-21,906,166	9p21.3	1298030
13	Homozygous Copy Loss	chr9 21,906,166-21,993,506	9p21.3	87341

14	CN Loss	chr4 52,700,771-191,154,276	4q12 - q35.2	1.38E+08
14	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
15	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
16	CN Loss	chr1 0-3,038,656	1p36.33 - p36.32	3038657
16	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
16	CN Loss	chr2 149,280,179-243,199,373	2q23.1 - q37.3	93919195
16	CN Gain	chr2 144,086,159-144,217,694	2q22.2 - q22.3	131536
16	CN Loss	chr5 62,009,001-68,740,653	5q12.1 - q13.2	6731653
16	CN Loss	chr5 70,744,658-180,915,260	5q13.2 - q35.3	1.1E+08
16	CN Gain	chr5 49,560,859-55,070,314	5q11.1 - q11.2	5509456
16	CN Gain	chr5 56,298,730-59,315,214	5q11.2 - q12.1	3016485
16	CN Gain	chr5 61,198,344-62,009,001	5q12.1	810658
16	CN Loss	chr7 38,525,684-47,748,989	7p14.1 - p12.3	9223306
16	CN Loss	chr7 48,028,854-49,906,334	7p12.3 - p12.2	1877481
16	CN Loss	chr7 50,487,083-50,853,145	7p12.2 - p12.1	366063
16	CN Loss	chr7 51,801,876-52,754,655	7p12.1	952780
16	CN Loss	chr7 61,064,518-94,010,901	7q11.1 - q21.3	32946384
16	CN Loss	chr7 94,054,371-159,138,663	7q21.3 - q36.3	65084293
16	CN Gain	chr7 0-10,298,426	7p22.3 - p21.3	10298427
16	CN Gain	chr7 52,916,459-55,451,949	7p12.1 - p11.2	2535491
16	CN Loss	chr10 130,904,734-131,860,192	10q26.3	955459
16	CN Gain	chr11 19,112,441-22,974,066	11p15.1 - p14.3	3861626
16	CN Gain	chr11 26,185,156-29,969,403	11p14.2 - p14.1	3784248
16	CN Gain	chr11 36,097,891-36,262,802	11p13	164912
16	CN Loss	chr13 19,043,558-115,169,878	13q11 - q34	96126321
16	CN Loss	chr14 20,443,778-107,349,540	14q11.2 - q32.33	86905763
16	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
16	CN Loss	chr16 0-32,482,955	16p13.3 - p11.2	32482956
16	CN Loss	chr16 33,828,054-35,219,549	16p11.2 - p11.1	1391496
16	CN Loss	chr16 46,505,821-75,617,935	16q11.2 - q23.1	29112115
16	CN Loss	chr16 76,350,929-79,504,460	16q23.1 - q23.2	3153532
16	CN Loss	chr16 80,084,226-90,354,753	16q23.2 - q24.3	10270528
16	CN Loss	chr17 21,430,743-22,193,626	17p11.2	762884
16	CN Loss	chr17 25,326,941-27,998,913	17q11.1 - q11.2	2671973
16	CN Gain	chr17 5,296,112-9,174,658	17p13.2 - p13.1	3878547
16	CN Loss	chr18 66,094,637-78,077,248	18q22.1 - q23	11982612
16	CN Gain	chr18 40,066,661-40,830,761	18q12.3	764101
16	CN Gain	chr18 41,452,789-45,214,777	18q12.3 - q21.1	3761989
16	CN Gain	chr18 45,372,096-47,778,681	18q21.1	2406586
16	CN Gain	chr18 49,427,917-58,873,990	18q21.2 - q21.32	9446074
16	CN Gain	chr18 62,165,709-64,788,445	18q22.1	2622737
16	CN Gain	chr18 65,234,990-66,094,637	18q22.1	859648
16	CN Gain	chr20 0-981,273	20p13	981274
16	CN Gain	chr20 2,097,839-2,591,738	20p13	493900
16	CN Gain	chr20 3,746,679-5,068,313	20p13	1321635
16	CN Gain	chr20 7,216,928-8,011,273	20p12.3	794346

16	CN Gain	chr20 11,872,003-17,363,946	20p12.2 - p12.1	5491944
16	CN Gain	chr20 19,213,523-20,397,189	20p11.23	1183667
16	CN Loss	chr22 40,039,177-51,304,566	22q13.1 - q13.33	11265390
17	CN Gain	chr1 0-2,314,475	1p36.33 - p36.32	2314476
17	CN Loss	chr2 0-1,420,263	2p25.3	1420264
17	CN Loss	chr2 2,697,484-5,652,163	2p25.3 - p25.2	2954680
17	CN Loss	chr2 29,429,367-29,727,380	2p23.2	298014
17	CN Loss	chr2 70,733,189-70,945,031	2p13.3	211843
17	CN Loss	chr2 120,129,954-120,451,007	2q14.2	321054
17	CN Gain	chr2 197,010,274-197,129,976	2q32.3	119703
17	CN Loss	chr3 12,747,735-14,864,713	3p25.2 - p25.1	2116979
17	CN Loss	chr3 54,238,667-54,612,156	3p21.1 - p14.3	373490
17	CN Loss	chr3 116,328,487-117,124,459	3q13.31	795973
17	CN Loss	chr3 125,013,911-126,285,171	3q21.2 - q21.3	1271261
17	CN Gain	chr3 9,322,373-10,004,672	3p25.3	682300
17	CN Gain	chr3 10,246,327-12,747,735	3p25.3 - p25.2	2501409
17	CN Gain	chr3 14,864,713-15,393,244	3p25.1	528532
17	CN Loss	chr4 0-12,236,950	4p16.3 - p15.33	12236951
17	CN Loss	chr4 36,418,529-37,417,493	4p14	998965
17	CN Loss	chr4 181,900,098-182,825,656	4q34.3	925559
17	CN Gain	chr6 27,746,874-27,889,619	6p22.1	142746
17	CN Gain	chr7 93,918,681-94,054,371	7q21.3	135691
17	CN Gain	chr7 134,462,262-134,627,046	7q33	164785
17	CN Loss	chr10 0-3,783,382	10p15.3 - p15.2	3783383
17	CN Loss	chr10 3,924,266-14,009,283	10p15.1 - p13	10085018
17	CN Loss	chr10 14,095,291-39,075,616	10p13 - p11.1	24980326
17	CN Loss	chr10 87,931,325-88,366,633	10q23.2	435309
17	CN Loss	chr10 117,706,974-118,558,199	10q25.3	851226
17	CN Loss	chr10 124,327,316-124,875,101	10q26.13	547786
17	CN Loss	chr10 130,164,780-135,534,747	10q26.2 - q26.3	5369968
17	CN Loss	chr11 4,167,279-4,731,401	11p15.4	564123
17	CN Loss	chr11 15,295,502-19,135,407	11p15.2 - p15.1	3839906
17	CN Loss	chr11 29,366,931-30,252,668	11p14.1	885738
17	CN Loss	chr11 83,697,258-85,228,552	11q14.1	1531295
17	CN Loss	chr11 112,360,547-112,952,421	11q23.1 - q23.2	591875
17	CN Loss	chr11 114,485,868-115,052,710	11q23.2 - q23.3	566843
17	CN Loss	chr11 117,221,881-117,680,274	11q23.3	458394
17	CN Loss	chr11 120,349,250-120,741,124	11q23.3	391875
17	CN Loss	chr11 126,366,633-128,233,120	11q24.2 - q24.3	1866488
17	CN Loss	chr11 130,724,393-135,006,516	11q24.3 - q25	4282124
17	CN Loss	chr12 109,173,054-115,175,763	12q24.11 - q24.21	6002710
17	CN Loss	chr12 117,383,517-118,006,963	12q24.22	623447
17	CN Loss	chr12 126,531,315-130,884,425	12q24.32 - q24.33	4353111
17	CN Gain	chr13 24,215,307-24,286,444	13q12.12	71138
17	CN Loss	chr14 44,655,207-45,450,045	14q21.2	794839
17	CN Loss	chr14 57,135,339-57,362,794	14q22.3	227456

17	CN Loss	chr14 95,761,768-96,518,919	14q32.13 - q32.2	757152
17	CN Loss	chr14 96,861,107-100,164,568	14q32.2	3303462
17	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
17	CN Loss	chr15 22,796,596-27,936,377	15q11.2 - q12	5139782
17	CN Loss	chr15 53,422,076-54,430,253	15q21.3	1008178
17	CN Loss	chr15 59,604,216-60,565,271	15q22.2	961056
17	CN Loss	chr15 69,256,599-70,370,501	15q23	1113903
17	CN Loss	chr15 88,163,141-89,003,474	15q25.3	840334
17	CN Loss	chr15 93,944,364-94,309,537	15q26.1 - q26.2	365174
17	CN Gain	chr15 59,577,752-59,604,216	15q22.2	26465
17	CN Loss	chr16 5,533,187-10,200,494	16p13.3 - p13.2	4667308
17	CN Loss	chr16 12,721,412-13,234,505	16p13.12	513094
17	CN Loss	chr16 16,220,446-24,441,534	16p13.11 - p12.1	8221089
17	CN Loss	chr16 25,870,610-26,944,503	16p12.1	1073894
17	CN Loss	chr16 27,795,877-28,714,254	16p12.1 - p11.2	918378
17	CN Loss	chr16 33,828,054-35,219,549	16p11.2 - p11.1	1391496
17	CN Loss	chr16 68,340,169-68,834,307	16q22.1	494139
17	CN Loss	chr16 73,161,378-73,747,376	16q22.3	585999
17	CN Loss	chr16 79,992,404-80,478,566	16q23.2	486163
17	CN Loss	chr16 82,119,482-83,726,871	16q23.3	1607390
17	CN Loss	chr16 85,944,631-86,589,416	16q24.1	644786
17	CN Loss	chr17 5,423,721-6,488,051	17p13.2	1064331
17	CN Loss	chr17 12,050,896-13,028,504	17p12	977609
17	CN Loss	chr17 25,936,898-26,412,879	17q11.2	475982
17	CN Loss	chr17 31,817,719-33,557,464	17q12	1739746
17	CN Gain	chr17 49,070,647-49,343,480	17q21.33	272834
17	CN Loss	chr18 34,449,292-45,360,061	18q12.2 - q21.1	10910770
17	CN Loss	chr18 47,335,576-48,327,909	18q21.1 - q21.2	992334
17	CN Loss	chr18 74,865,252-76,622,610	18q23	1757359
17	CN Loss	chr19 7,698,752-9,225,812	19p13.2	1527061
17	CN Loss	chr19 12,276,628-13,688,274	19p13.2	1411647
17	CN Loss	chr19 14,422,621-16,074,712	19p13.12	1652092
17	CN Loss	chr19 28,079,145-32,698,583	19q11 - q13.11	4619439
17	CN Loss	chr19 52,574,218-52,809,216	19q13.41	234999
17	CN Loss	chr19 54,072,769-56,799,114	19q13.42 - q13.43	2726346
17	CN Loss	chr21 14,414,872-15,862,900	21q11.2	1448029
17	CN Loss	chr21 20,241,040-23,408,822	21q21.1	3167783
17	CN Loss	chr21 31,262,362-32,351,832	21q21.3 - q22.11	1089471
17	CN Loss	chr21 39,130,486-39,570,730	21q22.13	440245
17	CN Loss	chr21 40,951,501-44,171,368	21q22.2 - q22.3	3219868
18	CN Gain	chr1 19,698,557-19,811,828	1p36.13	113272
18	CN Gain	chr1 174,996,846-175,058,961	1q25.1	62116
18	CN Gain	chr1 201,596,821-205,478,501	1q32.1	3881681
18	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
18	CN Gain	chr2 38,319,218-38,451,906	2p22.2	132689
18	CN Gain	chr2 189,537,698-190,180,902	2q32.2	643205

18	CN Gain	chr2 197,006,464-197,043,354	2q32.3	36891
18	CN Gain	chr6 132,293,826-132,553,270	6q23.2	259445
18	CN Gain	chr7 93,918,681-94,056,840	7q21.3	138160
18	CN Gain	chr8 38,046,708-38,874,853	8p11.23 - p11.22	828146
18	CN Gain	chr8 69,958,098-70,136,952	8q13.2	178855
18	CN Gain	chr8 102,569,632-103,077,130	8q22.3	507499
18	CN Gain	chr9 129,657,681-129,990,717	9q33.3	333037
18	Homozygous Copy Loss	chr9 21,917,917-22,021,617	9p21.3	103701
18	CN Gain	chr11 44,869,678-44,907,690	11p11.2	38013
18	CN Loss	chr16 23,901,688-24,179,253	16p12.2	277566
18	CN Loss	chr17 32,737,405-32,979,825	17q12	242421
18	CN Gain	chr18 6,740,271-6,901,288	18p11.31	161018
18	CN Gain	chr21 39,612,056-39,686,483	21q22.13	74428
19	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
19	CN Loss	chr3 0-29,288,635	3p26.3 - p24.1	29288636
19	CN Loss	chr3 29,487,919-37,004,234	3p24.1 - p22.2	7516316
19	CN Loss	chr3 37,083,307-81,295,869	3p22.2 - p12.2	44212563
19	CN Loss	chr3 174,254,000-174,894,768	3q26.31	640769
19	CN Loss	chr8 0-1,503,749	8p23.3	1503750
19	CN Loss	chr8 1,697,226-3,897,423	8p23.3 - p23.2	2200198
19	CN Loss	chr8 4,610,151-6,911,399	8p23.2 - p23.1	2301249
19	CN Loss	chr9 21,033,382-39,184,065	9p21.3 - p13.1	18150684
19	CN Loss	chr11 0-38,971,287	11p15.5 - p12	38971288
19	CN Loss	chr11 44,598,088-49,056,481	11p11.2 - p11.12	4458394
19	CN Loss	chr11 89,150,559-103,748,871	11q14.3 - q22.3	14598313
19	CN Loss	chr11 105,717,448-108,782,298	11q22.3	3064851
19	CN Loss	chr11 109,281,240-113,122,295	11q22.3 - q23.2	3841056
19	CN Loss	chr11 113,164,240-114,176,087	11q23.2	1011848
19	CN Loss	chr11 117,243,797-118,755,045	11q23.3	1511249
19	CN Loss	chr11 126,569,415-131,635,707	11q24.2 - q25	5066293
19	CN Loss	chr11 134,345,105-135,006,516	11q25	661412
19	CN Gain	chr11 43,701,726-44,222,946	11p11.2	521221
19	CN Gain	chr11 44,435,842-44,598,088	11p11.2	162247
19	CN Gain	chr11 116,763,882-117,016,922	11q23.3	253041
19	CN Loss	chr12 40,129,002-65,296,554	12q12 - q14.3	25167553
19	CN Loss	chr12 65,855,703-70,269,536	12q14.3 - q15	4413834
19	CN Loss	chr12 73,797,248-77,636,690	12q21.1 - q21.2	3839443
19	CN Loss	chr12 78,600,338-84,055,732	12q21.2 - q21.31	5455395
19	CN Loss	chr12 85,862,313-91,426,392	12q21.31 - q21.33	5564080
19	CN Loss	chr12 92,181,161-95,649,238	12q21.33 - q22	3468078
19	CN Loss	chr12 96,260,508-96,451,805	12q23.1	191298
19	CN Loss	chr12 97,070,286-97,469,345	12q23.1	399060
19	CN Loss	chr12 98,194,254-133,851,895	12q23.1 - q24.33	35657642
19	CN Loss	chr17 0-22,193,626	17p13.3 - p11.2	22193627
19	CN Loss	chr17 25,326,941-41,984,726	17q11.1 - q21.31	16657786

19	CN Loss	chr18 0-15,322,428	18p11.32 - p11.21	15322429
19	CN Loss	chr18 18,535,946-51,759,108	18q11.1 - q21.2	33223163
19	CN Loss	chr18 58,315,573-78,077,248	18q21.32 - q23	19761676
19	CN Loss	chr20 62,600,145-63,025,520	20q13.33	425376
19	CN Gain	chr20 47,030,432-50,242,155	20q13.13 - q13.2	3211724
19	CN Gain	chr20 60,072,489-62,600,145	20q13.33	2527657
19	Homozygous Copy Loss	chr20 14,799,707-14,836,662	20p12.1	36956
19	CN Loss	chr21 24,593,236-25,385,284	21q21.2	792049
19	CN Loss	chr22 16,267,567-32,170,230	22q11.1 - q12.2	15902664
19	CN Loss	chr22 34,613,912-43,747,390	22q12.3 - q13.2	9133479
19	CN Loss	chr22 45,196,461-48,220,095	22q13.31	3023635
19	CN Loss	chr22 48,415,323-51,304,566	22q13.32 - q13.33	2889244
20	CN Loss	chr2 91,813,414-92,196,400	2p11.1	382987
20	CN Loss	chr15 20,161,372-20,192,951	15q11.1	31580
20	CN Gain	chr19 53,320,036-53,363,140	19q13.41	43105

Supplemental Table 2 541 somatic mutation assays interrogated on the OncoScan™ platform

1	ABL1_pE255K_c763G_A
2	ABL1_pE355G_c1064A_G
3	ABL1_pF311L_c931T_C
4	ABL1_pF359V_c1075T_G
5	ABL1_pG250E_c749G_A
6	ABL1_pH396R_c1187A_G
7	ABL1_pM351T_c1052T_C
8	ABL1_pT315I_c944C_T
9	ABL1_pY253H_c757T_C
10	ALK_pD1091N_c3271G_A
11	ALK_pF1174V_c3520T_G
12	ALK_pF1245C_c3734T_G
13	ALK_pF1245V_c3733T_G
14	ALK_pR1275Q_c3824G_A
15	ALK_pT1211T_c3633C_A
16	APC_p_c835_minus_8A_G
17	APC_pE1286X_c3856G_T
18	APC_pE1309fs4_c3921_3925delAAAAG_allele1
19	APC_pE1309fs4_c3927_3931delAAAGA_allele1
20	APC_pE1309fs6_c3923_3924insA_allele1
21	APC_pE1309X_c3925G_T
22	APC_pE1322X_c3964G_T
23	APC_pE1345X_c4033G_T
24	APC_pE1353X_c4057G_T
25	APC_pE1379X_c4135G_T
26	APC_pE1397X_c4189G_T
27	APC_pE1408X_c4222G_T
28	APC_pE1451X_c4351G_T
29	APC_pE1464X_c4390G_T
30	APC_pE1577X_c4729G_T
31	APC_pE853X_c2557G_T
32	APC_pG1499X_c4495G_T
33	APC_pQ1291X_c3871C_T
34	APC_pQ1294X_c3880C_T
35	APC_pQ1328X_c3982C_T
36	APC_pQ1338X_c4012C_T
37	APC_pQ1367X_c4099C_T
38	APC_pQ1378X_c4132C_T
39	APC_pQ1406X_c4216C_T
40	APC_pQ1429X_c4285C_T
41	APC_pQ1447X_c4339C_T
42	APC_pQ1469X_c4405C_T

43	APC_pQ789X_c2365C_T
44	APC_pR1114X_c3340C_T
45	APC_pR1450X_c4348C_T
46	APC_pR213X_c637C_T
47	APC_pR232X_c694C_T
48	APC_pR283X_c847C_T
49	APC_pR302X_c904C_T
50	APC_pR332X_c994C_T
51	APC_pR564X_c1690C_T
52	APC_pR876X_c2626C_T
53	APC_pS1281X_c3842C_A
54	APC_pS1341R_c4023T_G
55	APC_pS1465fs3_c4393_4394delAG_allele1
56	APC_pT1537K_c4610C_A
57	APC_pT1556fs3_c4660_4661insA_allele1
58	APC_pT1556fs3_c4666_4667insA_allele1
59	AR_pA749A_c2247C_T
60	ATM_pA1309T_c3925G_A
61	ATM_pF858L_c2572T_C
62	ATM_pG2867E_c8600G_A
63	ATM_pI1681V_c5041A_G
64	ATM_pP604S_c1810C_T
65	ATM_pQ2442P_c7325A_C
66	ATM_pR2443Q_c7328G_A
67	ATM_pR3008C_c9022C_T
68	ATM_pR3008H_c9023G_A
69	ATM_pR3047X_c9139C_T
70	ATM_pR337C_c1009C_T
71	ATM_pR337S_c1009C_A
72	ATM_pS707P_c2119T_C
73	ATM_pT2666A_c7996A_G
74	ATM_pV410A_c1229T_C
75	BRAF_pD594G_c1781A_G
76	BRAF_pF583F_c1749T_C
77	BRAF_pG464E_c1391G_A
78	BRAF_pG466E_c1397G_A
79	BRAF_pG469E_c1406G_A
80	BRAF_pG469R_c1405G_A
81	BRAF_pI326T_c977T_C
82	BRAF_pL618W_c1853T_G
83	BRAF_pV600E_c1799T_A_allele1

84	BRAF_pV600K_c1798_1799GT_AA_allele1
85	BRAF_pV600R_c1798_1799GT_AG_allele1
86	BRCA1_p_c134_plus_1G_T
87	BRCA1_p_c5278_minus_1G_T
88	BRCA1_pE1725X_c5173G_T
89	BRCA1_pG1077W_c3229G_T
90	BRCA1_pG778C_c2332G_T
91	BRCA1_pH448H_c1344C_T
92	BRCA1_pL30F_c90G_T
93	BRCA1_pR1751X_c5251C_T
94	BRCA1_pS1009X_c3026C_A
95	BRCA1_pS1140G_c3418A_G
96	BRCA1_pW372X_c1116G_A
97	BRCA2_pD3095E_c9285C_A
98	BRCA2_pE1593X_c4777G_T
99	BRCA2_pE187K_c559G_A
100	BRCA2_pG1338G_c4014C_T
101	BRCA2_pH2415N_c7243C_A
102	BRCA2_pI1017S_c3050T_G
103	BRCA2_pI3103M_c9309A_G
104	BRCA2_pP920S_c2758C_T
105	BRCA2_pQ2934X_c8800C_T
106	BRCA2_pR18H_c53G_A
107	BRCA2_pR2678S_c8034G_T
108	BRCA2_pR2787H_c8360G_A
109	BRCA2_pR2842C_c8524C_T
110	BRCA2_pR3128X_c9382C_T
111	BRCA2_pS1682S_c5046T_C
112	BRCA2_pT630I_c1889C_T
113	BRCA2_pV1988I_c5962G_A
114	CBL_p_c1228_minus_2A_G
115	CBL_pR420Q_c1259G_A
116	CBL_pV391I_c1171G_A
117	CBL_pY371H_c1111T_C
118	CDC73_pW43X_c128G_A
119	CDH1_p_c1009_minus_1G_A
120	CDH1_pI374I_c1122C_T
121	CDKN2A_p_c1_minus_25C_T
122	CDKN2A_p_c150_plus_2T_C
123	CDKN2A_p_c151_minus_1G_A
124	CDKN2A_p_c457_plus_1G_T
125	CDKN2A_p_c457_plus_2T_C
126	CDKN2A_pA30V_c89C_T
127	CDKN2A_pD108N_c322G_A

128	CDKN2A_pE120K_c358G_A
129	CDKN2A_pE26X_c76G_T
130	CDKN2A_pE61X_c181G_T
131	CDKN2A_pE69X_c205G_T
132	CDKN2A_pE88E_c264G_A
133	CDKN2A_pG35E_c104G_A
134	CDKN2A_pH83Y_c247C_T
135	CDKN2A_pP48L_c143C_T
136	CDKN2A_pR131H_c392G_A
137	CDKN2A_pR58X_c172C_T
138	CDKN2A_pR80X_c238C_T
139	CDKN2A_pS43I_c128G_T
140	CDKN2A_pW110X_c329G_A
141	CDKN2A_pW110X_c330G_A
142	CDKN2A_pW15X_c44G_A
143	CDKN2A_pY44X_c132C_A
144	CSF1R_pL301S_c902T_C
145	CSF1R_pY969C_c2906A_G
146	CSF1R_pY969H_c2905T_C
147	CSF1R_pY969X_c2907T_G
148	CTNNB1_pA13T_c37G_A
149	CTNNB1_pA5_A80del_c14_241del228_allele1
150	CTNNB1_pD32N_c94G_A
151	CTNNB1_pG34E_c101G_A
152	CTNNB1_pS33Y_c98C_A
153	CTNNB1_pS45P_c133T_C
154	CTNNB1_pT41A_c121A_G
155	CTNNB1_pT41I_c122C_T
156	EGFR_pA289V_c866C_T
157	EGFR_pD761Y_c2281G_T
158	EGFR_pE709A_c2126A_C
159	EGFR_pE709K_c2125G_A
160	EGFR_pE746_A750del_c2235_2249del15_allele1
161	EGFR_pE746_A750del_c2236_2250del15_allele1
162	EGFR_pE746_S752_V_c2237_2255_T_allele1
163	EGFR_pG598V_c1793G_T
164	EGFR_pG719C_c2155G_T
165	EGFR_pG719S_c2155G_A
166	EGFR_pG863D_c2588G_A
167	EGFR_pL747_A750_P_c2239_2248TTAAGAGAAG_C_allele1

168	EGFR_pL747_P753_S_c2240_2257del18_allele1
169	EGFR_pL747_T751del_c2240_2254del15_allele1
170	EGFR_pL833V_c2497T_G
171	EGFR_pL858M_c2572C_A
172	EGFR_pL858R_c2573T_G
173	EGFR_pL861Q_c2582T_A_allele1
174	EGFR_pL861R_c2582T_G
175	EGFR_pR108K_c323G_A
176	EGFR_pR677H_c2030G_A
177	EGFR_pS768I_c2303G_T
178	EGFR_pT263P_c787A_C
179	EGFR_pT790M_c2369C_T
180	EGFR_pV774M_c2320G_A
181	EGFR_pV819V_c2457G_A
182	ERBB2_pG776S_c2326G_A
183	ERBB2_pL755S_c2264T_C
184	ERBB2_pV777L_c2329G_T
185	FBXW7_pR224X_c670C_T
186	FBXW7_pR278X_c832C_T
187	FBXW7_pR393X_c1177C_T
188	FBXW7_pR465C_c1393C_T
189	FBXW7_pR465H_c1394G_A
190	FBXW7_pR479Q_c1436G_A
191	FBXW7_pR505C_c1513C_T
192	FBXW7_pS582L_c1745C_T
193	FGFR1_pP252T_c754C_A
194	FGFR1_pS125L_c374C_T
195	FGFR2_pC382R_c1144T_C
196	FGFR2_pE475K_c1423G_A
197	FGFR2_pK310R_c929A_G
198	FGFR2_pN549K_c1647T_G
199	FGFR2_pR203C_c607C_T
200	FGFR2_pY375C_c1124A_G
201	FGFR3_pA369A_c1107G_T
202	FGFR3_pA391E_c1172C_A
203	FGFR3_pG370C_c1108G_T
204	FGFR3_pG697C_c2089G_T
205	FGFR3_pK650M_c1949A_T_allele1
206	FGFR3_pK650Q_c1948A_C
207	FGFR3_pR248C_c742C_T
208	FGFR3_pS249C_c746C_G_allele1
209	FGFR3_pS371C_c1111A_T_allele1
210	FGFR3_pY373C_c1118A_G
211	FLT3_pD835E_c2505T_G
212	FLT3_pD835N_c2503G_A
213	FLT3_pD835V_c2504A_T_allele1
214	FLT3_pL561L_c1683A_G
215	GATA1_pQ17X_c49C_T
216	GATA1_pV74I_c220G_A
217	HNF1A_pW206C_c618G_T
218	HNF1A_pW206L_c617G_T
219	HRAS_pG12D_c35G_A
220	HRAS_pG12S_c34G_A
221	HRAS_pG13D_c38G_A
222	HRAS_pG13S_c37G_A
223	HRAS_pQ61H_c183G_T
224	HRAS_pQ61K_c181C_A
225	HRAS_pQ61P_c182A_C
226	IKBKB_pA360S_c1078G_T
227	INPP4A_pE940D_c2820A_C
228	IRAK1_pS690G_c2068A_G
229	JAK2_pK191Q_c571A_C
230	JAK2_pR683G_c2047A_G
231	JAK2_pV617F_c1849G_T
232	JAK2_pY570Y_c1710C_T
233	KIT_pD52N_c154G_A
234	KIT_pD816Y_c2446G_T
235	KIT_pE839K_c2515G_A
236	KIT_pF584S_c1751T_C
237	KIT_pI798I_c2394C_T
238	KIT_pK642E_c1924A_G
239	KIT_pL576P_c1727T_C
240	KIT_pN822K_c2466T_G
241	KIT_pP585P_c1755C_T
242	KIT_pT670I_c2009C_T
243	KIT_pV559A_c1676T_C
244	KIT_pV560D_c1679T_A_allele1
245	KIT_pV654A_c1961T_C
246	KIT_pV825A_c2474T_C
247	KIT_pW557_K558del_c1669_1674delTGGAAG_allele1
248	KIT_pW557R_c1669T_C
249	KIT_pY503_F504insAY_c1509_1510insGCCTAT_allele1
250	KIT_pY823D_c2467T_G

251	KRAS_pA146T_c436G_A
252	KRAS_pA59T_c175G_A
253	KRAS_pG12D_c35G_A
254	KRAS_pG12S_c34G_A
255	KRAS_pG13D_c38G_A
256	KRAS_pG13S_c37G_A
257	KRAS_pQ61H_c183A_C
258	KRAS_pQ61K_c181C_A
259	KRAS_pQ61P_c182A_C
260	MAP2K4_pR154W_c460C_T
261	MAP2K4_pS184L_c551C_T
262	MAP2K4_pS280X_c839C_A
263	MEN1_p_c654_plus_3A_G
264	MEN1_pR98X_c292C_T
265	MEN1_pW471X_c1413G_A
266	MET_p982_1028del47_c3082_plus_1G_T
267	MET_pM1268T_c3803T_C
268	MET_pR988C_c2962C_T
269	MET_pT1010I_c3029C_T
270	MET_pY1248C_c3743A_G
271	MET_pY1253D_c3757T_G
272	MLH1_p_c1732_minus_1G_A
273	MLH1_p_c790_plus_1G_A
274	MLH1_pC233R_c697T_C
275	MLH1_pS556fs14_c1731G_A
276	MPL_pW515L_c1544G_T
277	MSH2_p_c1276_plus_1G_A
278	MSH2_pR680X_c2038C_T
279	MSH2_pR711X_c2131C_T
280	MSH6_pP1082P_c3246G_T
281	MSH6_pP1087fs3_c3261delC_allele1
282	MSH6_pP1087fs5_c3261_3262insC_allele1
283	NF1_pK1444E_c4330A_G
284	NF1_pR1276X_c3826C_T
285	NF1_pR1362X_c4084C_T
286	NF1_pR1769X_c5305C_T
287	NF1_pR1968X_c5902C_T
288	NF1_pR304X_c910C_T
289	NF1_pR461X_c1381C_T
290	NF1_pR816X_c2446C_T
291	NF2_p_c1340_plus_1G_A
292	NF2_p_c240_plus_2T_C
293	NF2_p_c516_plus_1G_A
294	NF2_p_c675_plus_1G_T

295	NF2_p_c810_plus_2T_C
296	NF2_pQ212X_c634C_T
297	NF2_pQ337X_c1009C_T
298	NF2_pQ362X_c1084C_T
299	NF2_pQ400X_c1198C_T
300	NF2_pQ410X_c1228C_T
301	NF2_pQ456X_c1366C_T
302	NF2_pR196X_c586C_T
303	NF2_pR198X_c592C_T
304	NF2_pR262X_c784C_T
305	NF2_pR341X_c1021C_T
306	NF2_pR466X_c1396C_T
307	NF2_pR57X_c169C_T
308	NF2_pV219M_c655G_A
309	NFKB1_p_c40_minus_1G_A
310	NOTCH1_pL1575P_c4724T_C
311	NOTCH1_pL1586P_c4757T_C
312	NOTCH1_pL1594P_c4781T_C
313	NOTCH1_pL1601P_c4802T_C
314	NOTCH1_pL1679P_c5036T_C
315	NOTCH1_pQ2460X_c7378C_T
316	NPM1_pW288fs12_c863_864insCATG_allele1
317	NPM1_pW288fs12_c863_864insCCTG_allele1
318	NPM1_pW288fs12_c863_864insTCTG_allele1
319	NRAS_pA18T_c52G_A
320	NRAS_pG12D_c35G_A
321	NRAS_pG12S_c34G_A
322	NRAS_pG13D_c38G_A
323	NRAS_pQ61H_c183A_C
324	NRAS_pQ61K_c181C_A
325	NRAS_pQ61P_c182A_C
326	PAK7_pT397K_c1190C_A
327	PDGFRA_pD1071N_c3211G_A
328	PDGFRA_pD842V_c2525A_T_allele1
329	PDGFRA_pD842Y_c2524G_T
330	PDGFRA_pD846Y_c2536G_T
331	PDGFRA_pF808L_c2422T_C
332	PDGFRA_pN659K_c1977C_A
333	PDGFRA_pN870S_c2609A_G
334	PDGFRA_pT674I_c2021C_T
335	PDGFRA_pV824V_c2472C_T
336	PIK3CA_pC420R_c1258T_C
337	PIK3CA_pC901F_c2702G_T

338	PIK3CA_pE453K_c1357G_A
339	PIK3CA_pE542K_c1624G_A
340	PIK3CA_pE545A_c1634A_C
341	PIK3CA_pE545K_c1633G_A
342	PIK3CA_pG1049S_c3145G_A
343	PIK3CA_pG118D_c353G_A
344	PIK3CA_pH1047R_c3140A_G
345	PIK3CA_pH1047Y_c3139C_T
346	PIK3CA_pH701P_c2102A_C
347	PIK3CA_pM1043I_c3129G_T
348	PIK3CA_pM1043V_c3127A_G
349	PIK3CA_pQ546K_c1636C_A
350	PIK3CA_pR108H_c323G_A
351	PIK3CA_pR38H_c113G_A
352	PIK3CA_pR88Q_c263G_A
353	PIK3CA_pT1025A_c3073A_G
354	PIK3CA_pT1025T_c3075C_T
355	PIK3CA_pY1021C_c3062A_G
356	PIK3R1_pG376R_c1126G_A
357	PTCH1_pM561R_c1682T_G
358	PTCH1_pQ365X_c1093C_T
359	PTCH1_pQ417X_c1249C_T
360	PTCH1_pW1018X_c3054G_A
361	PTEN_p_c1026_plus_1G_T
362	PTEN_p_c1027_minus_2A_G
363	PTEN_p_c165_minus_2A_C
364	PTEN_p_c209_plus_5G_A
365	PTEN_p_c253_plus_1G_A
366	PTEN_p_c253_plus_1G_T
367	PTEN_pA126T_c376G_A
368	PTEN_pA151T_c451G_A
369	PTEN_pD107Y_c319G_T
370	PTEN_pE235X_c703G_T
371	PTEN_pE299X_c895G_T
372	PTEN_pE7X_c19G_T
373	PTEN_pG129R_c385G_A
374	PTEN_pG165R_c493G_A
375	PTEN_pG251C_c751G_T
376	PTEN_pH61R_c182A_G
377	PTEN_pH93Q_c279T_G
378	PTEN_pH93Y_c277C_T
379	PTEN_pI101T_c302T_C
380	PTEN_pK267fs9_c800delA_ allele1
381	PTEN_pL23F_c69A_C
382	PTEN_pL42R_c125T_G
383	PTEN_pN323fs2_c968_969 insA_allele1
384	PTEN_pN323fs21_c968delA_ allele1
385	PTEN_pP246L_c737C_T
386	PTEN_pP38S_c112C_T
387	PTEN_pP95L_c284C_T
388	PTEN_pQ110X_c328C_T
389	PTEN_pQ171X_c511C_T
390	PTEN_pQ17X_c49C_T
391	PTEN_pQ214X_c640C_T
392	PTEN_pQ219X_c655C_T
393	PTEN_pQ245X_c733C_T
394	PTEN_pQ298X_c892C_T
395	PTEN_pR130Q_c389G_A
396	PTEN_pR130X_c388C_T
397	PTEN_pR15I_c44G_T
398	PTEN_pR173C_c517C_T
399	PTEN_pR173H_c518G_A
400	PTEN_pR233X_c697C_T
401	PTEN_pR335X_c1003C_T
402	PTEN_pW274X_c822G_A
403	PTEN_pY155C_c464A_G
404	PTEN_pY174D_c520T_G
405	PTEN_pY68H_c202T_C
406	PTPN11_pA72T_c214G_A
407	PTPN11_pD61Y_c181G_T
408	PTPN11_pE76K_c226G_A
409	PTPN11_pG503V_c1508G_T
410	PTPN11_pG60V_c179G_T
411	PTPN11_pS502P_c1504T_C
412	RB1_p_c1215_plus_1G_A
413	RB1_p_c1499_minus_1G_T
414	RB1_p_c1961_minus_1G_A
415	RB1_p_c2107_minus_2A_G
416	RB1_p_c2326_minus_2A_C
417	RB1_p_c380_plus_1G_A
418	RB1_pE137X_c409G_T
419	RB1_pE440X_c1318G_T
420	RB1_pE54X_c160G_T
421	RB1_pE748X_c2242G_T
422	RB1_pQ217X_c649C_T
423	RB1_pQ395X_c1183C_T
424	RB1_pQ685X_c2053C_T
425	RB1_pQ702X_c2104C_T
426	RB1_pR251X_c751C_T

427	RB1_pR320X_c958C_T
428	RB1_pR358X_c1072C_T
429	RB1_pR445X_c1333C_T
430	RB1_pR455X_c1363C_T
431	RB1_pR552X_c1654C_T
432	RB1_pR556X_c1666C_T
433	RB1_pR579X_c1735C_T
434	RB1_pR787X_c2359C_T
435	RB1_pS82X_c245C_A
436	RET_pA664D_c1991C_A
437	RET_pC634R_c1900T_C
438	RET_pC634Y_c1901G_A
439	RET_pM918T_c2753T_C
440	RUNX1_pD198G_c593A_G
441	RUNX1_pD198N_c592G_A
442	RUNX1_pI114I_c342C_A
443	RUNX1_pR107C_c319C_T
444	RUNX1_pR166X_c496C_T
445	RUNX1_pR204Q_c611G_A
446	SMAD4_pD537Y_c1609G_T
447	SMAD4_pE330A_c989A_C
448	SMAD4_pG168X_c502G_T
449	SMAD4_pG358X_c1072G_T
450	SMAD4_pK507Q_c1519A_C
451	SMAD4_pQ245X_c733C_T
452	SMAD4_pR361H_c1082G_A
453	SMAD4_pR445X_c1333C_T
454	SMAD4_pR497H_c1490G_A
455	SMAD4_pY353C_c1058A_G
456	SMARCB1_p_c1119_minus_41G_A
457	SMARCB1_pE216X_c646G_T
458	SMARCB1_pR158X_c472C_T
459	SMARCB1_pR201X_c601C_T
460	SMARCB1_pR377H_c1130G_A
461	SMARCB1_pR40X_c118C_T
462	SMARCB1_pS299S_c897G_A
463	SMARCB1_pY47X_c141C_A
464	SMO_pW535L_c1604G_T
465	SRC_pQ531X_c1591C_T
466	STK11_pD194Y_c580G_T
467	STK11_pP281L_c842C_T
468	STK11_pQ170X_c508C_T
469	STK11_pQ37X_c109C_T
470	TGFBR2_pR497X_c1489C_T
471	TP53_p_c376_minus_1G_A
472	TP53_p_c672_plus_1G_A
473	TP53_p_c782_plus_1G_T
474	TP53_p_c920_minus_1G_A
475	TP53_pA159V_c476C_T
476	TP53_pC124R_c370T_C
477	TP53_pC135F_c404G_T
478	TP53_pC176F_c527G_T
479	TP53_pE285K_c853G_A
480	TP53_pE298X_c892G_T
481	TP53_pE336X_c1006G_T
482	TP53_pF113C_c338T_G
483	TP53_pG245S_c733G_A
484	TP53_pG266E_c797G_A
485	TP53_pH179Q_c537T_G
486	TP53_pH179R_c536A_G
487	TP53_pH193R_c578A_G
488	TP53_pK132Q_c394A_C
489	TP53_pQ331X_c991C_T
490	TP53_pR158H_c473G_A
491	TP53_pR175H_c524G_A
492	TP53_pR196X_c586C_T
493	TP53_pR213X_c637C_T
494	TP53_pR248Q_c743G_A
495	TP53_pR248W_c742C_T
496	TP53_pR249S_c747G_T
497	TP53_pR273C_c817C_T
498	TP53_pR273H_c818G_A
499	TP53_pR306X_c916C_T
500	TP53_pR342X_c1024C_T
501	TP53_pV157F_c469G_T
502	TP53_pY163C_c488A_G
503	TP53_pY220C_c659A_G
504	TP53_pY236C_c707A_G
505	TSHR_pA623S_c1867G_T
506	TSHR_pA623V_c1868C_T
507	TSHR_pI630L_c1888A_C
508	TSHR_pM453T_c1358T_C
509	TSHR_pT632I_c1895C_T
510	VHL_p_c463_plus_1G_T
511	VHL_p_c463_plus_2T_C
512	VHL_p_c464_minus_1G_A
513	VHL_pE160K_c478G_A
514	VHL_pE189K_c565G_A
515	VHL_pG114C_c340G_T
516	VHL_pH115Y_c343C_T
517	VHL_pI151S_c452T_G

518	VHL_pL118P_c353T_C
519	VHL_pL153P_c458T_C
520	VHL_pL184P_c551T_C
521	VHL_pL85P_c254T_C
522	VHL_pP61P_c183C_T
523	VHL_pP81S_c241C_T
524	VHL_pQ132X_c394C_T
525	VHL_pQ96X_c286C_T
526	VHL_pR161X_c481C_T
527	VHL_pR167W_c499C_T
528	VHL_pS183X_c548C_A
529	VHL_pS65L_c194C_T
530	VHL_pS68X_c203C_A
531	VHL_pW117R_c349T_C
533	VHL_pW88X_c263G_A

534	WT1_pF154S_c461T_C
535	WT1_pG379V_c1136G_T
536	WT1_pR301X_c901C_T
537	WT1_pR362X_c1084C_T
538	WT1_pR390X_c1168C_T
539	WT1_pR394W_c1180C_T
540	WT1_pS313X_c938C_A
541	WT1_pS46X_c137C_A