McGill University

On the Automated Classification and Progression of Early Breast Cancers

Joseph Szymborski

Faculty of Medicine
Division of Experimental Medicine
McGill University, Montréal
September, 2017

Supervised by:
Prof. Luke M. McCaffrey

A thesis submitted to McGill University in partial fulfilment
of the requirements of the Master's degree.

Abstract

Résumé

Acknowledgements

Contributions

Abbreviations

Contents

1. Introduction

1.1 An Anatomy of Early Breast Cancers

Breast cancer is both a common and lethal diseases, having earned the dubious distinction of being both the most common and second most fatal cancer amongst females in Canada and around the world (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2015). Breast cancer most commonly arises in the epithelium of the mammary gland's many lactiferous ducts, which form a network that delivers to the nipple the milk that is secreted by the lobules of the mammary gland; which is another origin of breast carcinomas. The epithelium of the lactiferous duct is highly organised, with well-defined tissue and cell polarity that is integral to the structure and function of the duct. The tube-like lactiferous duct is a bilayered structure comprised of the outer myoepithelial and inner epithelial monolayers, both surrounding a hollow lumen at the duct's core. This epithelial inner-layer is surrounded by an outer layer of myoepithelial cells which express smooth-muscle actin (SMA) whose muscle-like contractile properties biomechanically deliver milk along the duct in response to hormonal signal (Hamperl, 1970).

1.1.1 Stages of Early Breast Cancer Progression

When diagnosing a suspected early breast cancer, pathologists analyse needle-core biopsies with the aim of identifying and classifying any lesions that may be present. Classification of lesions allow medical professionals to better understand the nature of the particular disease, what treatment is most appropriate, and what statistical outcomes

are associated with the lesion.

The early stages of breast cancer manifest as pre-invasive, hyper-proliferative lesions that exhibit progressive and gradual deterioration of this epithelial organisation. Of these lesions, there are four histologically distinct classes: Usual Hyperplasia (UH), Flat Epithelial Atypia (FEA), Atypical Ductal Hyperplasia (ADH) (or Atypical Lobular Hyperplasia [ALH] when referring to the less common lobular lesion), and Ductal Carcinoma *In Situ* (DCIS).

Ductal or lobular hyperplasias that do not present with abnormal tissue architecture or dysplasia are classified as Usual Hyperplasia (UH), or alternatively Proliferative Disease without Atypia (PDWA). These lesions confer a relative risk of later developping breast cancer as high as 1.9, although this increase in risk is not considered sufficient to warrant any prophylactic measures, including increased follow-up (Mommers *et al.*, 2001). While UH is traditionally believed to progress serially through ADH, DCIS and ultimately IDC due to early Loss of Homozygosity (LOH) analysis, more recent cytokeratin immunophenotype and genetic hybridisation analysis has contested the evolutionary relationship between UH and other proliferative breast lesions (O'Connell *et al.*, 1994; Boecker *et al.*, 2002).

ADH lesions are neoplasias of the lactiferous duct that exhibit subtle dysplasia (as evidenced by nuclear hyperchromaticity), and can form micropapillary or cribiform patterns (Page *et al.*, 1959; Dion *et al.*, 2016). Of the estimated one million instances of benign breast cancer detected in the USA each year, 10% are classified as ADH (Simpson, 2009). While these lesions have been long-known and extensively proven to impart a low relative risk (approximately 4), recent long-term follow-up studies have shown that one in eight individuals will develop more advanced (local or invasive) breast cancers ten years after their diagnosis. This proportion increases to 46% in individuals with more than one atypical foci twenty-five years after diagnosis (Hartmann *et al.*, 2015).

Arising in the terminal duct-lobule unit of the breast, FEA lesions are a purported precursor to early low-grade ductal carcinomas, and in this regard are similar to ADH. Unlike ADH however, FEA lesions are far-more uncommon, never present with complex

architectural patterns (thus the indication "flat"), and are characterised by multi-layered dilated ascini often made-up of columnar cells (Pinder, 2017). While ADH is suspected to arise from FEA lesions due their frequent coincidence, FEA is not independently associated with a long-term increased risk of breast cancer, leaving the matter unclear (Bombonati & Sgroi, 2011; Lerwill, 2008; Acott & Mancino, 2016).

Benign early lesions go on to progress into localised malignant disease, which in the lactiferous duct is termed ductal carcinoma *in situ* (DCIS). DCIS is classified as a Stage 0 cancer and accounts for 20% of all diagnosed breast cancers in the USA in 2003; representing a 500% increase in occurrence over 20 years (Bleicher, 2013; Kerlikowske, 2010).

While DCIS has a relatively low average standardised mortality ratio (SMR) of 1.8, an estimated 30-50% of cases reoccur as invasive breast cancers (Narod *et al.*, 2015; Page *et al.*, 1982; Betsill *et al.*, 1978). When further stratified by how well the lesion is differentiated, individuals with lesions classified as poorly differentiated (using the European Pathologists Working Group guidelines) have recurrence rates above 60% (Badve *et al.*, 1998). At this early stage of cancer progression, the apical domain of the lumenal epithelium has begun to shrink, resulting in abnormally small lumen (a phenotype referred to as "lumenal collapse" herein). Our understanding of the processes by which transformed mammary duct epithelium undergoes lumenal collapse is still developing, but recent studies have described a mechanism by which lumenal tension is lost as myosin II and RhoA activity is greatly decreased at the lumenal membrane of DCIS lesions (Halaoui et al. 2017, in review).

The lesion becomes an invasive ductal carcinoma (IDC, or ILC in the lobular instances) when epithelial cells breach the surrounding myoepithelial layer of the duct and infiltrate into extra-cellular matrix (ECM). By this stage, cellular polarity is entirely disrupted and the apical membrane domain has completely disappeared.

1.2 Automated Detection and Classification of Early Breast Lesions

1.2.1 Current Practices for the Diagnosis of Breast Cancers

The current standard of care for the diagnosis of breast cancer is the histopathological analysis of tissue biopsies (National Comprehensive Cancer Network, 2017). Sections from biopsy tissue are routinely stained with hematoxilin and eosin (H&E) and immunohistochemical studies performed to detect the presence of the HER2 oncogene and the estrogen and progesterone receptors (Lakhani *et al.*, 2012).

Despite the standardised and continually refined methods and guidelines clinical pathologists rely on to identify breast cancer lesions from histological sections, there is a great deal of inconsistency and uncertainty that is becoming increasingly apparent. While some types of breast cancer (such as high-grade DCIS and LCIS) are more consistently and reliably identified than others, inter-observer agreement between clinical pathologists is mixed. In a retrospective study, agreement between pathologists for ADH, FEA, and low-grade DCIS regions was only moderate (0.44, 0.47, and 0.47 Cohen's κ statistic, respectively) (Gomes *et al.*, 2014).

1.2.2 Conventional Machine-Learning Models for the Diagnosis of Breast Cancers

While computer-aided detection (CADe) is sometimes used to enhance interpretation of imaging from screening mammographies, computational models are not currently used in most clinical settings to assist the diagnosis of breast cancers.

Models using conventional supervised machine-learning (ML) algorithms, such as support-vector machines (SVMs) or random decision forests (RDFs), have been previously described with varying success rates (Anuranjeeta *et al.*, 2017; Gertych *et al.*, 2015). These models depend on automatically extracted features hand-picked by their designers. The features chosen to train such models are directly related to existing pathological guidelines (*e.g.*: nuclear size and spacing) or inferred by them (*e.g.*: texture-based

features such as Gabor filters and Harlick transforms) (Anuranjeeta *et al.*, 2017; Doyle *et al.*, 2008).

1.2.3 Convolutional Neural Network Models for the Diagnosis of Breast Cancers

Convolutional Neural Networks (CNNs or ConvNet) are deep machine learning algorithms that use multiple weighted hidden layers of convolutional filters to make decisions about a given input. Recent developments in general-purpose computing on graphics processing units (GPGPU) have made the use of ConvNets practical, and as a result ConvNets have since been shown to be particularly well suited for the task of classifying and analysing images (Cireşan *et al.*, 2011, 2012). Neural networks have the advantage of dynamically "learning" and optimising features, as opposed to relying on human-selected features. This manner of feature selection allows for the potential discovery of new underlying concepts that fundamentally define a class from others, and also prevents assumptions and misconceptions from biasing features and introducing inaccuracies.

The automated nature of feature selection in neural networks can also pose challenges if the dataset is of low-quality or very small; as differences between classes that are present in the dataset, but do not reflect differences between the classes as a concept, can severely bias a model. This is epitomised by "the tank problem", whereby a neural network trained by the US government to identify concealed tanks were actually identifying dark skies as the images of tanks in the dataset were taken on a cloudy day (Dreyfus & Dreyfus, 1992). Errors such as these are mitigated by assuring datasets are large and representative of the variation of the population.

ConvNets have been successfully used to create very accurate models for the classification of breast cancer lesions. Binary models for classifying benign and malignant lesions, as well as multi-class models for distinguishing between multiple subtypes of breast lesions from H&E stained biopsy slides have established with very high ($> 90\%$) accuracy (Wei *et al.*, 2017; Han *et al.*, 2017).

1.3 Epithelial Polarity of the Lactiferous Duct

Loss of cellular organisation and polarity is a common feature across epithelial cancers, but unlike some other cancers like those that present in the colon where cell polarity is lost at late stages of the disease, loss of polarity is a hallmark of early breast cancer (Hinck & Näthke, 2014). The mechanisms by which epithelial cell polarity is lost in carcinomas, however, remains elusive and poorly understood.

1.3.1 The Role of Proteins in Ductal Polarity & Cancer

When discussing the epithelial polarity of the lactiferous duct, one may be referring to asymmetric distribution at either the intercellular or intracellular level.

At the macro, intercellular scale, ductal epithelia is said to exhibit tissue polarity when cells organise into a monolayer forming a single lumen (Bissell *et al.*, 2003).

On the other hand, establishment of cellular apical-basolateral polarity is achieved by the intracellular asymmetric distribution of proteins, phospholipids , and carbohydrates within the inner epithelial monolayer of the mammary duct. Two protein complexes, the Crumbs complex and the Par complex, are particularly important determinants of the apical identity (Horikoshi *et al.*, 2009; Whiteman *et al.*, 2014).

The Crumbs Complex

Localisation of the Crumbs complex to the plasma membrane both contributes to the establishment and maintenance of its apical identity. The Crumbs complex converges upon the apical transmembrane glycoprotein for which it is named, Crumbs3 (*Crb3*), which serves as a scaffold for the complex. Crumbs3 directly binds two proteins via its carboxy-terminal PDZ domain (`ERLI`): protein associated with *Lin-7* one (Pals1) and partitioning-defective protein six (Par6) (Lemmers *et al.*, 2004; Roh *et al.*, 2002).The presence of Pals1 also brings to the Crumbs complex the Pals1 associated tight junction (PATJ) protein, which is essential for proper polarisation and contributes to the establishment of tight-junctions in mammalian cells (Shin *et al.*, 2005).

Crumbs3 has also been known to interact with FERM (4.1 protein, ezrin, radixin and moesin) domain proteins through its PDZ domain. Crumbs3 also interacts with the FERM-domain proteins EHM2 (also known as Lulu2) and YMO1, homologues of *Drosophila melanogastor* protein Yurt, which helps to establish apical-basolateral polarity and maintain the size of the apical membrane by regulating Crumbs3 (Laprise *et al.*, 2006). Crumbs3 has also been shown to recruit EHM2 and p114RhoGEF to maintain the actomyosin belt and promote cell-cell adhesion in a cancer cell-lines, requiring both the C-terminal FERM-binding and PDZ-binding motifs of Crumbs3 (Loie *et al.*, 2015).

The crumbs complex has been also shown to regulate important proliferative programmes such as organ growth and mammary gland contact inhibition through the Salvador/Warts/Hippo (hereafter Hippo) signalling pathway. Crumbs3 regulates the Hippo pathway through interactions with, among other proteins, the FERM domain-containing protein 6 (*FRMD6*), a mammalian homologue of the *D. melanogastor* gene *Ex* (Robinson *et al.*, 2010). Crumbs3 also regulates the Hippo pathway through direct interaction with WW-domain proteins. One such instance is the direct interaction between the Hippo pathway co-effectors yes-associated protein 1 (YAP1), Tafazzin (TAZ), and Crumbs3; this occurs in response to changes in cell density, which require changes to the cells proliferative program (Varelas *et al.*, 2010; Szymaniak *et al.*, 2015). In a similar, cell-density-sensing manner, Crumbs3 interacts directly with Kibra's WW-domain to stabilise it, preventing its degradation and promoting Hippo-pathway-mediated proliferation (Moleirinho *et al.*, 2013; Mao *et al.*, 2017).

Cell density is also coupled with transforming growth factor-β (TGF-β)-induced epithelial-mesenchymal transition (EMT) through the Crumbs3-mediated inhibition of SMAD; effectively reducing downstream activation Snail (Varelas *et al.*, 2010). In addition to being an important to the EMT transcriptional programme, the zinc-finger protein Snail (*SNAI1*) is a potent transcriptional inhibitor of Crumbs3 and to a lesser extent, PATJ and PALS1; resulting in mislocalisation of the Crumbs and Par complexes and disruption of tight-junction and polarity formation (Wang *et al.*, 2013; Whiteman *et al.*, 2014).

The Par Complex

The Par complex is named after the eponymous family of proteins first discovered in the 1980s as part of screen to identify maternal effect lethal mutations in the model nematode *Caenorhabditis elegans* (Kemphues *et al.*, 1988; Goldstein & Macara, 2007). Of the six, the par proteins most relevant to apical membrane specification are Par3 and Par6; both PDZ-domain scaffolding proteins and core members of the Par complex (Yu *et al.*, 2014; Hung & Kemphues, 1999). Other members of the Par complex include atypical protein kinase C (aPKC) and the cell division control protein 42 homologue (Cdc42) GTPase, binding to Par3 through Par6 which here acts as an adaptor (Joberty *et al.*, 2000). Par3 is also capable of binding aPKC directly; an interaction that is essential to establishing proper cell polarity, normal ductal architecture, and mammary gland morphogenesis (Nagai-Tamai *et al.*, 2002; McCaffrey & Macara, 2009).

The formation of the Par complex is cued by the establishment of cell-cell contacts. This comes as a result of the complex being anchored to the tight- junctions of the cell by Par3, which is tethered through its binding of phosphotidyl inositols and the junctional adhesion molecule (JAM) through the second and first of Par3s three PDZ domains, respectively (Wu *et al.*, 2007; Ebnet *et al.*, 2001).

At their basal levels of expression, the members of the Par complex are at a regulatory equilibrium that is often disrupted in neoplasias, resulting in aberrant signal integration and epithelial disorganisation. Human breast cancers often express dramatically reduced levels of Par3, freeing aPKC to inappropriately activate signalling pathways that lead to increased invasive and metastatic potential; namely the human epidermal growth factor receptor 2 (HER2) and janus kinase/two Signal Transducer and Activator of Transcription (JAK/STAT) pathways (Xue *et al.*, 2013; McCaffrey *et al.*, 2012).

aPKC is but one member of the mammalian Protein Kinase C (PKC) super family, of which there are three additional members: the classical/conventional PKCs (cPKCs), the novel PKCs (nPKCs), and the later-discovered PKC-related kinases (PRKs). These families are distinguished by their dependence/independence on Ca^{2+}, and whether they are activated by diacylglycerols (DAGs). aPKC is both Ca^{2+}-independent and DAG-

insensitive, and in this respect are identical to PRKs. The two families are primarily differentiated by PRKs association with RhoA, which is unique among the PKCs (Mellor & Parker, 1998).

Each PKC sub-family contains multiple isoforms, which each confer their own unique function. The aPKC exists in two isoforms, aPKCλ/ι and aPKCζ.

1.3.2 The Role of Phospholipids in Ductal Polarity & Cancer

2. Materials & Methods

3. Results

3.1 Machine-Learning-Based Regional Early-Lesion Recognition

Whole-image machine-learning-based classifiers of early breast lesions from microscopy images of breast biopsies with various accuracy rates have been previously described. While these models are effective at detecting which stage is most represented within a lesion, this poorly reflects the heterogenous nature of breast cancers. Furthermore, whole-image classifiers offer little information when presented with so-called "borderline" lesions, that exhibit characteristics of many early lesions in near-equal measure.

A model capable of classifying regions of a given image of a lesion would offer an added dimension of information that takes into account the heterogenous nature of cancers, while also providing reproducible, quantifiable insight into currently enigmatic "'borderline'" lesions.

3.1.1 Per-Pixel Gabor-Filter-Based Classification kMKNN Model of Early Breast Lesions

In a first attempt at developing such a classifier, a kMKNN model was trained on per-pixel Gabor-filter features of confocal imagery of mammary glands immunohistochemically stained for E-cadherin and Par-6. This model will be referred to as the Gabor-kMKNN model hereafter.

As the characteristic "cobblestone" epithelial phenotype is altered distinctly in both ADH and DCIS, it was hypothesised that a model trained on texture-based features

would be sufficient in classifying those lesions. For this reason, the Gabor-filter was chosen. Analagous to the manner in which simple cortical cells perceive patterns and texture, the Gabor filter and has long been used to programmatically discern textures from one another (Fogel & Sagi 1989, Marĉelja 1980).

Gabor features were extracted in a similar fashion as Melendez *et al.* 2008 and is described by *Figure 1.* Namely, for each pixel in an image, six windows of increasing size (3×3, 5×5, 9×9, 17×17, 33×33, 65×65) centered on the pixel are defined. Each window is then filtered through four Gabor kernels with quarter-turn orientations (*i.e.*: $\theta = \left\{\frac{1}{2}\pi, \pi, \frac{3}{2}\pi, 2\pi\right\}$). Each Gabor kernel also has a sinusoidal wavelength of 0.25 pixels ($\lambda = 0.25$), which has been previously described as providing good discrimination in general-purpose texture classification (Manjunath & Ma 1996).

The mean and the standard deviation of the resulting Gabor energies are then added to a vector for the relevant pixel. This results in a total of 48 features per pixel (6 windows $\times$ 4 orientations $\times$ [1 mean + 1 standard deviation] = 48 feature per pixel).

To reduce computational complexity, images are resized to a resolution of 256×256 pixels. Computation time of training can be further reduced by extracting the features of a smaller randomly sampling of pixels from the total population.

Following the k-means for k-nearest neighbor (kMkNN) model described by Wang 2011, the resultant 48-dimensional matrix is clustered using the k-means algorithm. The number of clusters is determined using the heuristic:

$$k_c = \left\lceil 2\sqrt{(n)} \right\rceil$$

Where k_c is the number of clusters to compute, and n is the number of elements to cluster. If features were computed for all 65,536 pixels in a 256×256 image, a total of 512 clusters would be computed $\left(k_c = \left\lceil 2\sqrt{(65,536)} \right\rceil = 512\right)$.

The process of classifying a pixel (p) begins by calculating the 48 Gabor energy means and standard deviations of p as described above. The nearest cluster (C') is determined by choosing the cluster with the minimum Euclidean distances between its centroid and the feature vector of p. p is then classified using the standard K-nearest neighbor algorithm

with feature vectors contained in C' and a K-value of 3.

While the Gabor-kMKNN model was efficient at identifying distinct textures in images composed of multiple MeasTex

4. Discussion

Bibliography

Acott, A. A., & Mancino, A. T. 2016. Flat epithelial atypia on core needle biopsy, must we surgically excise? *The American Journal of Surgery*, **212**(6), 1211–1213.

Anuranjeeta, K.K.Shukla, Tiwari, Anoop, & Sharma, Shiru. 2017. Classification of Histopathological images of Breast Cancerous and Non Cancerous Cells Based on Morphological features. *Biomedical and Pharmacology Journal*, **10**(1), 353–366.

Badve, Sunil, A'hern, Roger P, Ward, Ann M, Millis, Rosemary R, Pinder, Sarah E, Ellis, Ian O, Gusterson, Barry A, & Sloane, John P. 1998. Prediction of local recurrence of ductal carcinoma in situ of the breast using five histological classifications: A comparative study with long follow-up. *Human Pathology*, **29**(9), 915–923.

Betsill, William L., Rosen, Paul Peter, Lieberman, Philip H., & Robbins, Guy F. 1978. Intraductal Carcinoma: Long-Term Follow-up After Treatment by Biopsy Alone. *JAMA*, **239**(18), 1863–1867.

Bissell, Mina J, Rizki, Aylin, & Mian, I Saira. 2003. Tissue architecture: the ultimate regulator of breast epithelial function. *Current Opinion in Cell Biology*, **15**(6), 753–762.

Bleicher, Richard J. 2013. Ductal carcinoma in situ. *Surgical Clinics of North America*, **93**(2), 393–410.

Boecker, Werner, Moll, Roland, Dervan, Peter, Buerger, Horst, Poremba, Christopher, Diallo, Raihanatou Ina, Herbst, Hermann, Schmidt, Ansgar, Lerch, Markus M, & Buchwalow, Igor B. 2002. Usual ductal hyperplasia of the breast is a committed stem (progenitor) cell lesion distinct from atypical ductal hyperplasia and ductal carcinoma in situ. *The Journal of Pathology*, **198**(4), 458–467.

Bombonati, Alessandro, & Sgroi, Dennis C. 2011. The molecular pathology of breast cancer progression. *The Journal of Pathology*, **223**(2), 308–318.

Canadian Cancer Society's Advisory Committee on Cancer Statistics. 2015. *Canadian Cancer Statistics 2015*.

Cireşan, Dan, Meier, Ueli, & Schmidhuber, Juergen. 2012. Multi-column Deep Neural Networks for Image Classification. *arXiv:1202.2745 [cs]*, Feb. arXiv: 1202.2745.

Cireşan, Dan C., Meier, Ueli, Masci, Jonathan, Gambardella, Luca M., & Schmidhuber, Jürgen. 2011. Flexible, High Performance Convolutional Neural Networks for Image Classification. *Pages 1237–1242 of: Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two*. IJCAI'11. AAAI Press.

Dion, L., Racin, A., Brousse, S., Beltjens, F., Cauchois, A., Levêque, J., Coutant, C., & Lavoue, V. 2016. Atypical epithelial hyperplasia of the breast: state of the art. *Expert Review of Anticancer Therapy*, **16**(9), 943–953.

Doyle, Scott, Agner, Shannon, Madabhushi, Anant, Feldman, Michael, & Tomaszewski, John. 2008. Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. *Pages 496–499 of: Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*. IEEE.

Dreyfus, Hubert L., & Dreyfus, Stuart E. 1992. What artificial experts can and cannot do. *AI & SOCIETY*, **6**(1), 18–26.

Ebnet, Klaus, Suzuki, Atsushi, Horikoshi, Yosuke, Hirose, Tomonori, Brickwedde, Maria-Katharina Meyer zu, Ohno, Shigeo, & Vestweber, Dietmar. 2001. The cell polarity protein ASIP/PAR-3 directly associates with junctional adhesion molecule (JAM). *The EMBO Journal*, **20**(14), 3738–3748.

Gertych, Arkadiusz, Ing, Nathan, Ma, Zhaoxuan, Fuchs, Thomas J., Salman, Sadri, Mohanty, Sambit, Bhele, Sanica, Velásquez-Vacca, Adriana, Amin, Mahul B., & Knudsen, Beatrice S. 2015. Machine learning approaches to analyze histological images of tissues from radical prostatectomies. *Computerized Medical Imaging and Graphics: The Official Journal of the Computerized Medical Imaging Society*, **46 Pt 2**(Dec), 197–208.

Goldstein, Bob, & Macara, Ian G. 2007. The PAR Proteins: Fundamental Players in Animal Cell Polarization. *Developmental Cell*, **13**(5), 609–622.

Gomes, Douglas S., Porto, Simone S., Balabram, Débora, & Gobbi, Helenice. 2014. Inter-observer variability between general pathologists and a specialist in breast pathology in the diagnosis of lobular neoplasia, columnar cell lesions, atypical ductal hyperplasia and ductal carcinoma in situ of the breast. *Diagnostic Pathology*, **9**(1), 121.

Hamperl, H. 1970. *The Myothelia (Myoepithelial Cells)*. Springer, Berlin, Heidelberg. Page 161–220.

Han, Zhongyi, Wei, Benzheng, Zheng, Yuanjie, Yin, Yilong, Li, Kejian, & Li, Shuo. 2017. Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model. *Scientific Reports*, **7**(1), 4172.

Hartmann, Lynn C., Degnim, Amy C., Santen, Richard J., Dupont, William D., & Ghosh, Karthik. 2015. Atypical Hyperplasia of the Breast — Risk Assessment and Management Options. *The New England Journal of Medicine*, **372**(1), 78–89.

Hinck, Lindsay, & Näthke, Inke. 2014. Changes in cell and tissue organization in cancer of the breast and colon. *Current opinion in cell biology*, **0**(Feb), 87–95.

Horikoshi, Yosuke, Suzuki, Atsushi, Yamanaka, Tomoyuki, Sasaki, Kazunori, Mizuno, Keiko, Sawada, Hajime, Yonemura, Shigenobu, & Ohno, Shigeo. 2009. Interaction between PAR-3 and the aPKC–PAR-6 complex is indispensable for apical domain development of epithelial cells. *Journal of Cell Science*, **122**(10), 1595–1606.

Hung, T. J., & Kemphues, K. J. 1999. PAR-6 is a conserved PDZ domain-containing protein that colocalizes with PAR-3 in Caenorhabditis elegans embryos. *Development*, **126**(1), 127–135.

Joberty, Gérard, Petersen, Clark, Gao, Lin, & Macara, Ian G. 2000. The cell-polarity protein Par6 links Par3 and atypical protein kinase C to Cdc42. *Nature Cell Biology*, **2**(8), 531–539.

Kemphues, Kenneth J., Priess, James R., Morton, Diane G., & Cheng, Niansheng. 1988. Identification of genes required for cytoplasmic localization in early C. elegans embryos. *Cell*, **52**(3), 311–320.

Kerlikowske, Karla. 2010. Epidemiology of Ductal Carcinoma In Situ. *Journal of the National Cancer Institute. Monographs*, **2010**(41), 139–141.

Lakhani, Sunil R., Ellis, Ian O., Schnitt, Stuart J., Tan, Puay Hoon, & van de Vijver, Marc J. 2012. *WHO classification of tumours of the breast*. 4 edn. International Agency for Research on Cancer. Chap. 1.

Laprise, Patrick, Beronja, Slobodan, Silva-Gagliardi, Nancy F., Pellikka, Milena, Jensen, Abbie M., McGlade, C. Jane, & Tepass, Ulrich. 2006. The FERM Protein Yurt Is a Negative Regulatory Component of the Crumbs Complex that Controls Epithelial Polarity and Apical Membrane Size. *Developmental Cell*, **11**(3), 363–374.

Lemmers, Céline, Michel, Didier, Lane-Guermonprez, Lydie, Delgrossi, Marie-Hélène, Médina, Emmanuelle, Arsanto, Jean-Pierre, & Bivic, André Le. 2004. CRB3 Binds Directly to Par6 and Regulates the Morphogenesis of the Tight Junctions in Mammalian Epithelial Cells. *Molecular Biology of the Cell*, **15**(3), 1324–1333.

Lerwill, MF. 2008. Flat epithelial atypia of the breast. *Archives of Pathology & Laboratory Medicine*, **132**(4), 615–621.

Loie, Elise, Charrier, Lucie E., Sollier, Kévin, Masson, Jean-Yves, & Laprise, Patrick. 2015. CRB3A Controls the Morphology and Cohesion of Cancer Cells through Ehm2/p114RhoGEF-Dependent Signaling. *Molecular and Cellular Biology*, **35**(19), 3423–3435.

Mao, Xiaona, Li, Pingping, Wang, Yaochun, Liang, Zheyong, Liu, Jie, Li, Juan, Jiang, Yina, Bao, Gang, Li, Lei, Zhu, Bofeng, & et al. 2017. CRB3 regulates contact inhibition by activating the Hippo pathway in mammary epithelial cells. *Cell Death & Disease*, **8**(1), e2546.

McCaffrey, Luke Martin, & Macara, Ian G. 2009. The Par3/aPKC interaction is essential for end bud remodeling and progenitor differentiation during mammary gland morphogenesis. *Genes & Development*, **23**(12), 1450–1460.

McCaffrey, Luke Martin, Montalbano, JoAnne, Mihai, Constantina, & Macara, Ian G. 2012. Loss of the Par3 Polarity Protein Promotes Breast Tumorigenesis and Metastasis. *Cancer Cell*, **22**(5), 601–614.

Mellor, Harry, & Parker, Peter J. 1998. The extended protein kinase C superfamily. *Biochemical Journal*, **332**(2), 281–292.

Moleirinho, S., Chang, N., Sims, A. H., Tilston-Lünel, A. M., Angus, L., Steele, A., Boswell, V., Barnett, S. C., Ormandy, C., Faratian, D., & et al. 2013. KIBRA exhibits MST-independent functional regulation of the Hippo signaling pathway in mammals. *Oncogene*, **32**(14), 1821–1830.

Mommers, Ellen C.M., Page, David L., Dupont, William D., Schuyler, Peggy, Leonhart, Angelique M., Baak, Jan P.A., Meijer, Chris J.L.M., & Van Diest, Paul J. 2001. Prognostic value of morphometry in patients with normal breast tissue or usual ductal hyperplasia of the breast. *International Journal of Cancer*, **95**(5), 282–285.

Nagai-Tamai, Yoko, Mizuno, Keiko, Hirose, Tomonori, Suzuki, Atsushi, & Ohno, Shigeo. 2002. Regulated protein–protein interaction between aPKC and PAR-3 plays an essential role in the polarization of epithelial cells. *Genes to Cells*, **7**(11), 1161–1171.

Narod, Steven A., Iqbal, Javaid, Giannakeas, Vasily, Sopik, Victoria, & Sun, Ping. 2015. Breast Cancer Mortality After a Diagnosis of Ductal Carcinoma In Situ. *JAMA Oncology*, **1**(7), 888–896.

National Comprehensive Cancer Network. 2017 (June). *Breast Cancer Screening and Diagnosis*. Tech. rept. 1.2017.

O'Connell, Peter, Pekkel, Vladimir, Fuqua, Suzanne, Osborne, C. Kent, & Allred, D. Craig. 1994. Molecular genetic studies of early breast cancer evolution. *Breast Cancer Research and Treatment*, **32**(1), 5–12.

Page, David L, Dupont, William D, Rogers, LW, & Rados, MS. 1959. Atypical hyperplastic lesions of the female breast. *A long-term follow-up study. Cancer.*

Page, David L., Dupont, William D., Rogers, Lowell W., & Landenberger, Marie. 1982. Intraductal carcinoma of the breast: Follow-up after biopsy only. *Cancer*, **49**(4), 751–758.

Pinder, Sarah E. 2017. *Proliferative Breast Lesions.* Springer International Publishing. Page 33–41.

Robinson, Brian S., Huang, Juang, Hong, Yang, & Moberg, Kenneth H. 2010. Crumbs Regulates Salvador/Warts/Hippo Signaling in Drosophila via the FERM-Domain Protein Expanded. *Current Biology*, **20**(7), 582–590.

Roh, Michael H., Makarova, Olga, Liu, Chia-Jen, Shin, KunYoo, Lee, Seonok, Laurinec, Stephanie, Goyal, Meera, Wiggins, Roger, & Margolis, Ben. 2002. The Maguk protein, Pals1, functions as an adapter, linking mammalian homologues of Crumbs and Discs Lost. *The Journal of Cell Biology*, **157**(1), 161–172.

Shin, Kunyoo, Straight, Sam, & Margolis, Ben. 2005. PATJ regulates tight junction formation and polarity in mammalian epithelial cells. *The Journal of Cell Biology*, **168**(5), 705–711.

Simpson, Jean F. 2009. Update on atypical epithelial hyperplasia and ductal carcinoma in situ. *Pathology*, **41**(1), 36–39.

Szymaniak, Aleksander D., Mahoney, John E., Cardoso, Wellington V., & Varelas, Xaralabos. 2015. Crumbs3-Mediated Polarity Directs Airway Epithelial Cell Fate through the Hippo Pathway Effector Yap. *Developmental Cell*, **34**(3), 283–296.

Varelas, Xaralabos, Samavarchi-Tehrani, Payman, Narimatsu, Masahiro, Weiss, Alexander, Cockburn, Katie, Larsen, Brett G., Rossant, Janet, & Wrana, Jeffrey L. 2010. The Crumbs Complex Couples Cell Density Sensing to Hippo-Dependent Control of the TGF-Beta-SMAD Pathway. *Developmental Cell*, **19**(6), 831–844.

Wang, Yifan, Shi, Jian, Chai, Kequn, Ying, Xuhua, & Zhou, Binhua P. 2013. The Role of Snail in EMT and Tumorigenesis. *Current cancer drug targets*, **13**(9), 963–972.

Wei, Benzheng, Han, Zhongyi, He, Xueying, & Yin, Yilong. 2017. *Deep learning model based breast cancer histopathological image classification.* IEEE. Page 348–353.

Whiteman, Eileen L., Fan, Shuling, Harder, Jennifer L., Walton, Katherine D., Liu, Chia-Jen, Soofi, Abdul, Fogg, Vanessa C., Hershenson, Marc B., Dressler, Gregory R., Deutsch, Gail H., & et al. 2014. Crumbs3 Is Essential for Proper Epithelial Development and Viability. *Molecular and Cellular Biology*, **34**(1), 43–56.

Wu, Hao, Feng, Wei, Chen, Jia, Chan, Ling-Nga, Huang, Siyi, & Zhang, Mingjie. 2007. PDZ Domains of Par-3 as Potential Phosphoinositide Signaling Integrators. *Molecular Cell*, **28**(5), 886–898.

Xue, Bin, Krishnamurthy, Kannan, Allred, D. Craig, & Muthuswamy, Senthil K. 2013. Loss of Par3 promotes breast cancer metastasis by compromising cell–cell cohesion. *Nature Cell Biology*, **15**(2), 189–200.

Yu, Cao Guo, Tonikian, Raffi, Felsensteiner, Corinna, Jhingree, Jacquelyn R., Desveaux, Darrell, Sidhu, Sachdev S., & Harris, Tony J. C. 2014. Peptide Binding Properties of the Three PDZ Domains of Bazooka (Drosophila Par-3). *PLOS ONE*, **9**(1), e86412.