UNRAVELING THE TRANSCRIPTOME OF ODONTOGENIC TUMORS

Shijia Hu

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum of Oral Biology in the School of Dentistry.

Chapel Hill
2016

Approved by:

J. Timothy Wright

Kimon Divaris

Eric Everett

Joel Parker

Bernard Weissman

ABSTRACT

Shijia Hu: Unraveling the Transcriptome of Odontogenic Tumors
(Under the direction of J. Timothy Wright)

Odontogenic tumors, represent 31% of the oral tumors in children, are phenotypically diverse neoplasms of tissues that are responsible for tooth formation. Ameloblastoma and keratocystic odontogenic tumor, both believed to be derived from the odontogenic epithelium, constitute more than 50% of odontogenic tumors. Although both are usually slow-growing, they are locally invasive and have a recurrence rate as high as 50-80%. The lack of established adjunctive therapy means that surgical removal of the tumor with extensive margins to ensure complete excision remains the primary treatment of choice, resulting in significant morbidity.

Tremendous advances are being made in the understanding of molecular mechanisms and pathways involved in tumorigenesis, improving diagnostic and therapeutic approaches. Most research on odontogenic tumors focused on candidate-genes with only a handful of studies employing whole genome and transcriptome approaches. Another limitation is the question of what normal tissue is most biologically-relevant for gene expression comparison with odontogenic tumors. Obtaining and characterizing the gene expression profile of odontogenic epithelium at different stages of differentiation will provide a biologically-relevant reference for comparison.

This study aims to bridge the current gap of knowledge in odontogenic tumor biology by characterizing the transcriptome of ameloblastoma and KCOT which up till this point has not been fully explored.

The results showed that ameloblastoma separated into 2 distinct molecular clusters that were associated with 2 types of odontogenic tissue. Importantly, we found that 9/10 of the samples in the pre-secretory cluster were of the follicular type while 6/7 of the samples in the odontoblast cluster were of the plexiform type. Analysis of differential gene expression revealed alteration of common pathways in both clusters including cell cycle regulation, inflammatory and MAPkinase. Similarly, 2 distinct molecular subtypes of KCOT were found with several canonical inflammatory pathways activated in both subtypes of KCOT. Of note, the AKT pathway was activated in one subtype while MAPkinase pathway was activated in the other.

Our results are suggestive of underlying molecular heterogeneity of odontogenic tumors which could indicate different receptiveness to treatment protocols. These findings have implications in the tailored use of chemotherapeutic agents or other treatment modalities.

ACKNOWLEDGEMENTS

The Dunning-Kruger effect (DuKE) is real. It is a cognitive situation whereby an individual is unable to recognize and evaluate their own ineptitude in a task accurately due to a lack of comprehension in the task in the first place (i.e. the stupid are too stupid to realize they're stupid). When I started this program 5 years ago I definitely suffered from DuKE. Fortunately, I had many mentors guide me in this journey towards enlightenment. I could never fully express how thankful I am to this upstanding and noble community (unc) who have helped me in this chapter of my life; but like all PhD students, I can only try.

First and foremost, I need to thank Dr. Tim Wright. He has worn so many hats in my post-graduate career I am no longer sure how to define his impact on me. He was my mentor, my principal investigator, my committee chair, my department chair, attending and awesome dinner host. When I grow up, I want to be like Dr. Wright.

This project would never have been finished without the insightful comments from Dr. Divaris and Dr. Parker. You were always available when I needed help. The patience you shown trying to overcome my DuKEness is something I will carry with me forever. I can never thank you enough but I promise I will pay it forward and become a productive member of this unc.

I think the smartest thing I did for my time here was assembling a great committee. The intellectual exchange that occurred in those meetings had me in awe. Thank you Dr. Everett and Dr. Weissman, your time was greatly appreciated.

In addition, Ms. Cynthia Suggs has to be mentioned. She worked so hard on the experiments and I greatly benefited from it. I would also like to thank my other co-authors Dr. Padilla and Dr. Murrah for your expertise and kind words when the papers got rejected time and again.

To all my comrades-in-arms from the Oral Biology program and the Pediatric Dentistry residency, I would never have survived these years in this foreign land without you guys. You welcomed me with open arms and picked me up when I was down. I came to America alone, but will be leaving family behind.

Last but not least, the unconditional support I received from my family and Bien. Thank you for loving me, I know it can be hard sometimes. I hope I make all of you proud.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AKT1	Protein kinase B Alpha
ASTN1	Astrotactin 1
AMBN	Ameloblastin
AMELX	Amelogenin, X isoform
B3GALTL	Beta 1,3-Galactosyltransferase-Like
BRAF	v-Raf murine sarcoma viral oncogene homolog B
cDNA	Complementary Deoxyribonucleic acid
COPZ2	Coatomer Protein Complex, Subunit Zeta 2
DMP1	Dentin matrix protein
EDTA	Ethylenediaminetetraacetic acid
EFGR	Epidermal growth factor receptor
ENAM	Enamelin
EZH2	Enhancer Of Zeste 2 Polycomb Repressive Complex 2 Subunit
FDR	False discovery rate
FFPE	Formalin-fixed paraffin-embedded
FN1	Fibronectin

GEO	Gene Expression Omnibus
GSEA	Gene set enrichment analysis
GWAS	Genome-wide association study
HIF1A	Hypoxia Inducible Factor 1, Alpha Subunit
IPA	Ingenuity pathway analysis
IRB	Institutional review board
KCOT	Keratocystic odontogenic tumor
K-ras	V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
KRT19	Keratin 19
LCM	Laser capture microdissection
LUM	Lumican
MAPK	Mitogen-activated protein kinases
MDM2	Mouse double minute 2 homolog
MEK	Mitogen-activated protein kinase kinase
mRNA	Messenger Ribonucleic acid
MSigDB	Molecular Signatures Database
mTOR	Mechanistic target of rapamycin

NBCCS	Nevoid basal cell carcinoma syndrome
NSAID	Nonsteroidal anti-inflammatory drug
oAM	Odontoblast-like ameloblastoma
OB	Odontoblast
oKC	Odontoblast-like Keratocystic odontogenic tumor
p53	Tumor protein p53
PA	Pre-secretory ameloblast
pAM	Pre-secretory ameloblast-like ameloblastoma
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
PTCH1	Protein patched homolog 1
RIN	RNA Integrity Number
RNA	Ribonucleic acid
RTPCR	Reverse transcription polymerase chain reaction
SA	Secretory ameloblast
SAM	Significance Analysis of Microarrays
SEZ6L	Seizure 6-like
SHH	Sonic hedgehog

sKC	Secretory ameloblast-like Keratocystic odontogenic tumor
SOX2	SRY (sex determining region Y)-box 2
TCGA	The Cancer Genome Atlas
WHO	World Health Organization
WNT	Wingless-type MMTV integration site family

CHAPTER 1: INTRODUCTION

Odontogenic Tumors in Perspective

Although odontogenic tumors are rare when compared to other head and neck tumors, they represent 31% of the oral tumors in children (Adebayo et al. 2002) and often are clinically challenging to manage. They represent a phenotypically diverse group of tumors that have shown differing prevalence among different populations, suggesting different genetic and/or environmental etiological factors. The 2 most prevalent odontogenic tumors are the ameloblastoma (Avelar et al. 2011) and keratocystic odontogenic tumor (KCOT) (Servato et al. 2012). Both are believed to be derived from the odontogenic epithelium. Together, they constitute more than 50% of all odontogenic tumors (Buchner et al. 2006).

Ameloblastoma is a slow-growing, locally invasive, benign epithelial odontogenic neoplasm. It is thought to arise from remnants of odontogenic epithelium of the tooth-forming enamel organ, also known as the cell rests of the dental lamina (Sehdev et al. 1974). This tumor exhibits epithelial cells resembling pre-ameloblasts on a basement membrane in loosely arranged cells resembling stellate reticulum while the stroma consists of loose connective tissue. Although rare, ameloblastoma has been known to show malignant potential (Bedi et al. 2012) and occasionally metastasize (Luo et al. 2012). Currently, the WHO divides ameloblastoma into 4 subtypes namely, 1) solid, 2) peripheral, 3) desmoplastic, and 4) unicystic. Histopathologically, the solid and peripheral ameloblastoma subtype can be further

divided into a follicular and plexiform type while the much rarer unicystic ameloblastoma can be divided into the intraluminal and luminal types (Barnes L 2005). Current treatment modalities range from conservative enucleation and curettage for the unicystic subtype to radical maxillectomy for follicular subtypes (Singh et al. 2014). A striking feature of ameloblastoma is its penchant for recurrence with high recurrence rates (50-80%) being observed in cases of conservative treatment (Mendenhall et al. 2007). For this reason, and despite recent advances in imaging-assisted surgical margin localization, post-operative histological confirmation is still required. Meanwhile, intra-operative histological margin confirmation is challenging due to the calcified nature of the tumor and surrounding healthy tissue. This forces surgeons to either act conservatively risking the need for a second surgery, or act aggressively thus increasing morbidity (De Silva et al. 2012) and the need for extensive reconstructive surgery. This can be exceptionally devastating for young children due to their continued growth and development. Moreover, histological diagnosis can take up to 4 weeks and in cases of inadequate margins, patients will have to undergo further radiotherapy increasing the morbidity associated with radiation to the head and neck region. Recently, a case report of a recalcitrant stage 4 ameloblastoma carrying a *BRAF* V600E mutation was successfully treated with a combination of *BRAF/MEK* inhibitors (Kaye et al. 2015) raising the possibility of the increased use of chemotherapeutics in the treatment of ameloblastoma.

In 2005, the World Health Organization reclassified the KCOT from a cyst to a tumor to better reflect its neoplastic nature (Barnes L 2005). KCOT is a benign but locally aggressive developmental cystic neoplasm historically thought to arise from the odontogenic epithelium and frequently is associated with the follicle of unerupted teeth (Partridge and Towers 1987).

Histological features includes a thin fibrous wall of connective tissue surrounding an epithelial layer of columnar, pallisading, hyperchromatic basal cells with a cystic lumen filled with keratin (Grasmuck and Nelson 2010). As with the ameloblastoma, it can in rare cases undergo malignant transformation (Tan et al. 2013). A high recurrence rate of up to 30% has been found following conservative treatment such as enucleation and curettage, creating a dilemma in determining the optimal extent of surgical resection to balance the risk of recurrence with the morbidity associated with extensive reconstruction. Although conservative treatment involving enucleation with application of Carnoy's solution has recently been shown to be effective in reducing recurrence (Guler et al. 2012), larger lesions require more aggressive treatment. Of note, mutations in *PTCH1*, a tumor suppressor gene, causes the autosomal disorder Nevoid basal cell carcinoma syndrome (NBCCS) characterized by multiple basal cell carcinomas, KCOT and bifid ribs (Lam et al. 2013). These cases have higher recurrence rates when compared to isolated unifocal cases not associated with NBCCS (Johnson et al. 2012).

Tumorigenesis: What is currently known about Ameloblastoma and KCOT

Despite tremendous interest in these tumors, most published scholarship on this topic consists of case reports with a scant handful of genetic and molecular studies. Furthermore, in the majority of these studies, a bias selection of specific candidate genes or proteins were evaluated with little attention to genome-wide methods such as microarrays to describe the transcriptome of the tumors.

Of the few molecular studies of ameloblastoma, differential expression of K-ras and Mitogen-activated protein kinase (MAPK) pathway were identified in 22 ameloblastomas with 1

tumor showing a K-ras mutation (Kumamoto et al. 2004b). The follow up study showed altered expression of activated members of the MAPK pathway in ameloblastomas suggesting an involvement in oncogenesis and tumor cell differentiation (Kumamoto and Ooya 2007). Recently, involvement of the MAPK pathway in ameloblastoma tumorigenesis was further explored. Specifically, a V600E mutation in *BRAF* was described by 3 different groups (Brown et al. 2014; Kurppa et al. 2014; Sweeney et al. 2014) with Kruppa and colleagues finding *BRAF* mutations in 63% of ameloblastomas. Additionally, transgenic mouse models carrying the *v-Ha-ras* oncogene also have been found to develop ameloblastoma-like tumors spontaneously which is rare in the wild type variant (Dodds et al. 2003). Interestingly, activation of the nuclear mTOR pathway has been found to increase the recurrence of ameloblastomas by a factor of 6.4 fold (Li et al. 2012).

Molecular studies of KCOT have linked *PTCH1* mutations to the aggressiveness and recurrence of KCOT with *PTCH1* mutation associated with a higher recurrence rate; however, patients carrying *PTCH1* mutation also have nevoid basal cell carcinoma syndrome (Kadlub et al. 2013) while mutation of *PTCH1* were found only infrequently (30%) in non-syndromic patients (Pan et al. 2009). Conversely, a recent study employing more sensitive techniques suggests that the proportion of sporadic KCOT with *PTCH1* mutations could be greater than originally believed with up to 80% of the cases affected (Qu et al. 2015). To date, the entire *PTCH1* pathway has yet to be studied in the context of KCOT and other molecules in the pathway could be involved in isolated cases. The Sonic Hedgehog–Patched–Gli Pathway is a highly conserved and essential developmental pathway, with mutations being implicated in a number of birth

defects. If the pathway is aberrantly activated during adult life, it causes uncontrolled cell proliferation that manifests as cancer (Villavicencio et al. 2000).

Interestingly, there are a number of genes and pathways that could be active and important in the tumorigenesis of both ameloblastoma and KCOT. The elevated expression of p53 and MDM2 has been implicated in the development of both ameloblastoma and KCOT (Kumamoto et al. 2004a; Sharifi-Sistani et al. 2011). Recently, there has been interest in the WNT pathway being involved in the pathogenesis of ameloblastoma and KCOT. The WNT pathways are essential for tooth development and altered expression of WNT associated genes had been found in ameloblastoma (Kibe et al. 2013) and KCOT (Hakim et al. 2011). Specifically, WNT5a was found to be highly expressed in a mouse tumor model (Sukarawan et al. 2010) leading to speculation that differential activation of these molecular pathways leads to tumorigenesis.

Currently, there are 5 published microarray studies on human ameloblastoma and each relied upon different normal comparison tissues such as human gingiva (Carinci et al. 2003; Heikinheimo et al. 2015), human tooth buds (Heikinheimo et al. 2002), dentigerous cysts (Lim et al. 2006) and a universal human reference RNA (DeVilliers et al. 2011) to compare gene expression. This renders it impossible to compare the results across the different studies. The 2 microarray studies on KCOTs used a 588 cancer-related human cDNA array against tooth buds (Heikinheimo et al. 2007) and a more extensive whole genome array against human gingival tissue (Heikinheimo et al. 2015). Although these studies provides valuable insight into the molecular basis of ameloblastoma and KCOT development, the lack of uniformity in choosing a reference tissue hampers the understanding of tumorigenesis of these odontogenic tumors.

This heterogeneity is a reflection of the lack of agreement on what constitutes the most biologically-relevant tissue for comparison. Given that odontogenic tumors are thought to be derived from the normal tissues involved in tooth formation, an appropriate comparison would be the normal odontogenic tissue transcriptome.

The human dentome

There is a need to address this lack of a normal tissue/cell-of-origin material for the examination of differential gene expression with the odontogenic tumors. The tissue of comparison will determine which genes and pathways are deemed different in the tumor. Using an inappropriate tissue for comparison can cloud the picture of pathways that are important to tumorigenesis. Ameloblastomas and KCOTs arise from the odontogenic epithelium. Obtaining the gene expression profile of odontogenic epithelium at different stages of development will help identify the tissue and developmental stage that most closely resembles the tumor tissue. This will aid in elucidating aberrant pathways present in odontogenic tumors compared with normal odontogenic tissue.

Tooth formation or odontogenesis is strictly regulated at the molecular level and involves multiple complex processes including development of highly specialized cells that produce unique extracellular matrices and ultimately mineralized tissues. Human enamel and dentin formation involves cells derived from epithelial and mesenchymal tissues and in many ways is unlike any other hard tissue formation in the human body (Figure 1.1).

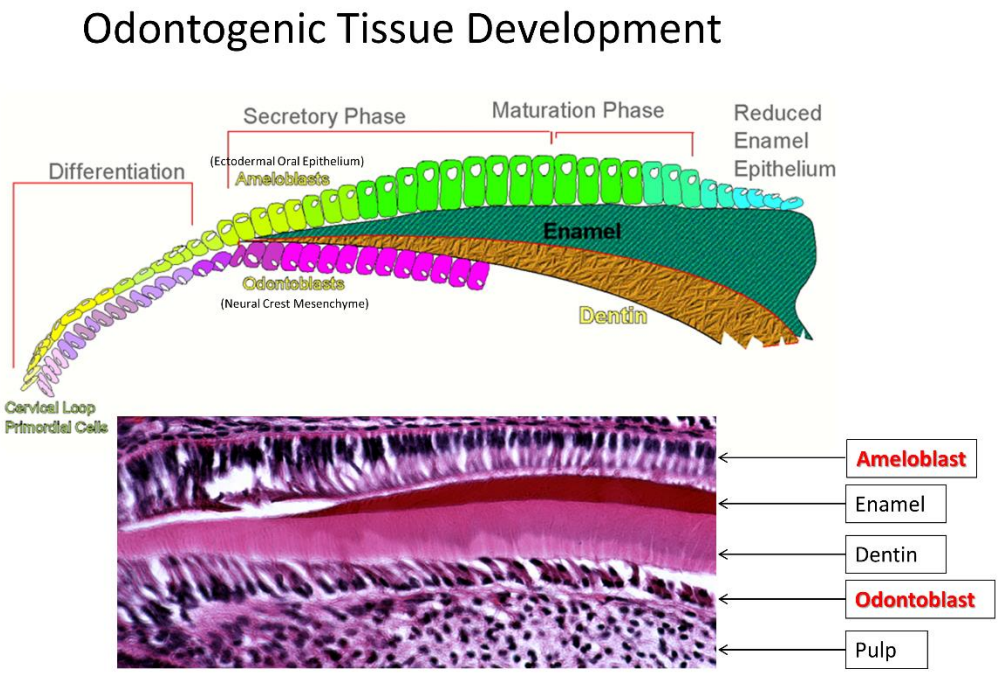


Figure 1.1: Development of ameloblast and odontoblast. Ameloblast is derived from ectodermal oral epithelium and undergo stages of differentiation (pre-secretory), secretory and maturation before undergoing apoptosis with the remaining cells contributing in the development of gingival attachment. Odontoblast is derived from the neural crest mesenchyme and remain functional throughout the tooth's life. Both tissue exist in a pseudo-single layer while odontoblast is sandwiched by hard tissue (dentin) and a cell rich matrix (pulp). With permission from Dr. Tim Wright.

The oral epithelium invaginates into the underlying mesenchymal tissue during the early stages of odontogenesis and gives rise to ameloblasts, the cells that form enamel. The ameloblasts undergo extensive histodifferentiation during their life cycle, going from cuboidal to columnar to squamous morphologies while creating and regulating a unique and changing microenvironment and extracellular matrix (Simmer et al. 2012). During the process of producing a unique extracellular matrix, the ameloblasts moves in a highly organized manner to produce enamel prisms that are oriented directionally into three dimensional species-specific patterns (Bartlett and Smith 2013). Once the full thickness of enamel has formed, the

ameloblasts alter their morphology again to facilitate maturation and further mineralization of the enamel. Upon completion of the enamel maturation process, the ameloblasts then become a protective covering until the tooth crown emerges into the oral cavity. At which point, most of the ameloblasts undergo apoptosis and cease to exist with the remaining cells contributing in the development of gingival attachment. Developing ameloblasts exist as a pseudo-single cell layer making it extremely challenging to the harvest and examine them. It is also known that some less differentiated cells of the dental lamina persists in the oral cavity (e.g. epithelial cell rests of Malassez) after odontogenesis and it is theorized that these cells undergo tumorigenesis and form odontogenic tumors (Juuri et al. 2013; Partridge and Towers 1987; Sehdev et al. 1974).

In contrast, dentin-forming odontoblasts that are derived from the odontogenic mesenchyme continue to lay down matrix and remain functional throughout the tooth's life (Couve et al. 2013). These cells are able to react to stimuli and lay down reparative or reactionary dentin when the tooth experiences environmental insults. During odontogenesis, ameloblasts from the dental epithelium are involved in molecular cross talk with the underlying mesenchymal cells which ultimately form odontoblasts (Jernvall and Thesleff 2000). Although odontoblasts persists throughout the life of the tooth, like ameloblasts, they present as a pseudo-single layer. In addition, the odontoblast cell layer is sandwiched between a highly mineralized layer (the dentin), and a cell rich layer (the pulp) making its study just as, if not more, difficult than ameloblast.

Moreover, the acquisition of the transcriptome of odontogenic tissue (ameloblast and odontoblast), also known as the "dentome", has potential impact in the study of other diseases

associated with ameloblast and odontoblast development and function such as tooth agenesis (Thesleff et al. 2001) and amelogenesis imperfecta (Stephanopoulos et al. 2005).

Investigative approaches

Laser capture microdissection (LCM) is a technique that allows the isolation of specific cells from microscopic regions of tissue samples (Decarlo et al. 2011) (Figure 1.2). This technique can be applied to frozen sections (Hayashi et al. 2010) or archival tissues embedded in paraffin (Salmon et al. 2012). Because LCM does not change or damage the target cell morphology and chemical content, it has been coupled to high-density microarrays (Maxwell et al. 2013) to obtain expression profiles from specific cell populations. LCM has been shown to be able to isolate junctional epithelium of a thickness of 1-4 cells for gene expression studies, a level of precision not possible using traditional histological methods. This allows isolation and pooling of homogenous cell samples of the developing oral epithelium which is frequently only 1 cell thick. The cystic nature of KCOTs and some subtypes of ameloblastomas also mean that the use of LCM will reduce the contamination of surrounding stroma tissue in the samples. Using this approach, we are able to obtain homogenous isolates of specific subpopulations of cells in complex samples (DeVilliers et al. 2011). The characterization of the odontogenic epithelium provides the opportunity to establish a consistent and reliable cell type for the comparison of differential gene expression between normal and tumor cells.

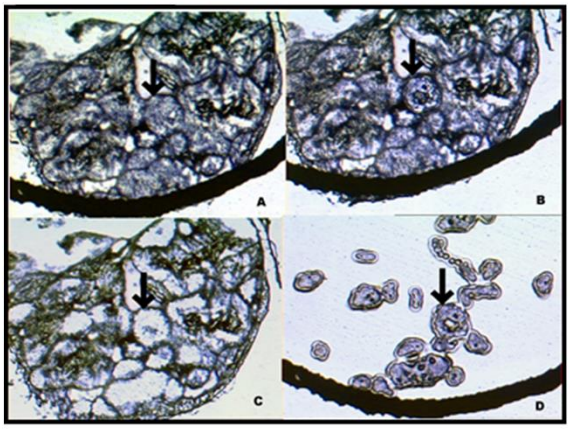


Figure 1.2: Micrograph of Laser microdissection. The arrow shows the cells of interest in all panels. Panel C shows the tissue left behind after LCM while panel D shows the capture of the target cells. With permission from Dr. Tim Wright.

Whole genome microarray allows for entire transcriptome examination of target tissue samples providing a global picture of cellular activity. Traditionally, a 2-color microarray examines mRNA content between control and experimental tissue. Total RNA from the control sample is tagged with Cy3 while the experimental sample is tagged with Cy5. The labelled RNA is then hybridized to a chip containing oligonucleotide probes of interest and scanned for expression levels which is detected as differential luminescent color intensity depending on the level of expression present in the respective samples. When a standardized reference RNA is used between microarrays, it allows for the normalization across the microarray chips and experiments (Devonshire et al. 2010). Using this approach means that it will be possible to

compare data from future studies when the same commercially available standardized reference RNA is used for normalization.

The development and refinement of molecular pathway analysis software aids in the analysis of gene expression data. Using this analytical approach allows us to build on existing knowledge from candidate gene studies, to examine upstream regulators and downstream effectors responses to gene expression changes of these genes of interest, as well as identify driver genes and derive a better understanding of the tumorigenesis of odontogenic tumors (Thomas and Bonchev 2010). Pathway analysis organizes genes into categories based on location, cellular components and reported molecular functions, facilitating assembly of nodes for the analysis of genetic networks and canonical pathways available in the database which will aid in our understanding of the pathogenesis of ameloblastoma and KCOT. Most diseases are complex and multi-factorial and result from the interaction of numerous genetic and environmental factors. The use of pathway analysis has advantages over the candidate gene approaches in this regard. The molecular approach using multi-omic techniques had helped unravel the molecular heterogeneity of breast cancer (Cancer Genome Atlas 2012) and understand the molecular pathogenesis of human hepatocellular carcinoma (Thorgeirsson and Grisham 2002).

Future implications

Characterization of the dentome will aid in the identification of genes and pathways not previously known to play a role in tooth formation. For example, the genes and pathways involved in cellular function such as ion and water regulation and cell movement, to name just a

couple, are likely controlling essential processes involved in normal odontogenesis and are worth closer examination. Unravelling the human dentome will advance our knowledge of tooth formation and the critical events that are regulated by gene expression and that control cell function and development of the tooth. This information can be used for establishing models for tooth development studies and ultimately pave the way for novel treatment for several tooth malformation disorders. For example, the use of a recombinant ectodsyplasin A in dogs can significantly reduce the burden of ectodermal dysplasia, a syndrome in human which includes multiple missing teeth (Casal et al. 2007). Establishing a good tooth development model can be used for developing similar novel therapeutics for the hundreds of hereditary defects affecting teeth.

Comprehensive molecular portraits of many cancers utilizing whole transcriptome approaches (Cancer Genome Atlas 2012) have led to a better understanding of genomic aberrations and other events driving tumor biology. This knowledge is leading to the identification of therapeutically tractable pathways and more effective therapies. The current paucity of knowledge of tumorigenesis in odontogenic tumors means that treatment is still limited to surgical intervention and the need for extensive and devastating reconstructive surgery. Recently, chemotherapeutic agents have been used successfully in dogs for the treatment of ameloblastoma (Kelly et al. 2010). Of note, the SHH pathway inhibitor Vismodegib has been used as an adjunctive therapy in patients with NBCCS to reduce tumor size and reduce the margins needed for surgical resection (Booms et al. 2015). The use of chemotherapeutic agents to reduce tumor size can be very helpful in reducing recurrence and morbidity associated with extensive surgical margins.

Furthermore, we anticipate that the characterization of the transcriptome will elucidate molecular pathways and provide a rationale for the mutational analysis of genes strongly implicated in tumorigenesis. These findings can then be used to identify more specific molecular markers for tumor diagnosis and improve the identification of surgical margins using biological markers rather than traditional histopathology (Otero et al. 2013). As a corollary example, it was shown recently that the use of mass spectrometry to identify molecular markers in brain tumors correlated well with histological diagnosis, offering the potential for surgeons to maximize tumor resection while preserving function (Eberlin et al. 2013). A similar discovery of detectable molecular markers in odontogenic tumors could greatly enhance surgical margin determination and improve current treatment strategies.

Conclusion

In summary, this study aims to bridge the current gap of knowledge in odontogenic tumor biology by characterizing the transcriptome of ameloblastomas and KCOTs which up until now has not been fully explored. A comprehensive molecular profile can serve as a hypothesis-generating resource for the advancement of precision medicine for the diagnosis and treatment of odontogenic tumors. In addition, the genes and pathways found to be associated with ameloblastoma and KCOT will provide a foundation for advancing our understanding of tumorigenesis, the development of new prognostic markers and ultimately novel treatment strategies.

CHAPTER 2: TOWARDS UNRAVELING THE HUMAN TOOTH TRANSCRIPTOME: THE DENTOME

Shijia Hu[1,2], Joel Parker[3], John Timothy Wright[1]

[1]Pediatric Dentistry, University of North Carolina, Chapel Hill, NC, USA

[2]Oral Biology Curriculum, University of North Carolina, Chapel Hill, NC, USA

[3]Cancer Genetics, University of North Carolina, Chapel Hill, NC, USA

Abstract

The goal of the study was to characterize the transcriptome profiles of human ameloblasts and odontoblasts, evaluate molecular pathways and advance our knowledge of the human “dentome”. Laser capture microdissection was used to isolate odontoblasts and ameloblasts from human tooth buds (15-20week gestational age) from 4 fetuses. RNA was examined using Agilent 41k whole genome arrays at 2 different stages of enamel formation, presecretory and secretory. Probe detection was considered against the array negative control to control for background noise. Differential expression was examined using Significance Analysis of Microarrays (SAM) 4.0 between different cell types and developmental stages with a false discovery rate of 20%. Pathway analysis was conducted using Ingenuity Pathway Analysis software. We found that during primary tooth formation, odontoblasts expressed 14,802 genes, presecretory ameloblasts 15,179 genes and secretory ameloblasts 14,526 genes. Genes known to be active during tooth development for each cell type (eg *COL1A1*, *AMELX*) were shown to be expressed by our approach. Exploring further into the list of differentially expressed genes between the motile odontoblasts and non-motile presecretory ameloblasts we found several genes of interest that could be involved in cell movement (*FN1*, *LUM*, *ASTN1*). Furthermore, our analysis indicated that the Phospholipase C and ERK5 pathways, that are important for cell movement, were activated in the motile odontoblasts. In addition our pathway analysis identified *WNT3A* and *TGFB1* as important upstream contributors. Recent studies implicate these genes in the development of Schimke immuno-osseous dysplasia. The utility of laser capture microdissection can be a valuable tool in the examination of specific tissues or cell populations present in human tooth buds. Advancing our knowledge of the

human dentome and related molecular pathways provides new insights into the complex mechanisms regulating odontogenesis and biomineralization. This knowledge could prove useful in future studies of odontogenic related pathologies.

Introduction

Tooth formation or odontogenesis is strictly regulated at the molecular level and involves multiple complex processes including development of highly specialized cells that produce unique extracellular matrices and ultimately mineralized tissues including the hardest tissue in the body, enamel (Thesleff et al. 2001). Ameloblasts, the cells that form enamel, undergo extensive histodifferentiation during their life cycle going from cuboidal to columnar to squamous morphologies while creating and regulating a unique and changing microenvironment and extracellular matrix (Deutsch et al. 1995; Simmer et al. 2012). During the process of producing a unique extracellular matrix, the ameloblasts move in a highly organized manner to produce enamel prisms that are directionally oriented into three dimensional patterns that are species specific (Bartlett and Smith 2013).

Dentin forming odontoblasts, on the other hand, continue to lay down matrix and remain functional throughout the life of a tooth (Couve et al. 2013). These cells are able to react to stimuli and lay down reparative or reactionary dentin when the tooth experiences environmental insults. During odontogenesis, ameloblasts, derived from the dental epithelium, are involved in molecular cross talk with the underlying mesenchymal cells that ultimately form odontoblasts (Jernvall and Thesleff 2000). Many of the molecular mechanisms involved in tooth formation and the specific genes and interactions that control odontogenesis remain unknown.

The roles of specific genes and pathways involved in tooth development have been queried by numerous investigators using the murine model (D'Souza et al. 1999; Dassule and McMahon 1998; Jarvinen et al. 2006; Kim et al. 2012). Many human studies of odontogenesis have focused on single genes and pathways that are disease driven (Bergendal et al. 2011; Liu and Millar 2010; Rufini et al. 2011). The study of human odontogenesis is challenging due to the issue of obtaining samples at different developmental stages and the difficulty in isolating the different tissue components of the developing tooth bud. Most research has been based on the examination of entire tooth buds (Heikinheimo et al. 2002; Lin et al. 2007) which does not allow interrogation of the disparate tissues present in a developing tooth.

Laser capture microdissection (Decarlo et al. 2011) allows the isolation of specific cells from microscopic regions of tissue samples (Chokechanachaisakul et al. 2012; Sun et al. 2012). Using this technique cells can be harvested from frozen sections or archival tissues embedded in paraffin (Hayashi et al. 2010; Salmon et al. 2012). As laser capture does not change or damage the target cell morphology and chemical content, it can be used for DNA, RNA or protein analyses. Recent development in microarray techniques and reduction in costs has led to novel approaches for the study of tissue and organ development (Heikinheimo et al. 2002; Tranasi et al. 2009). Microarray technology allows examination of the entire genome with very small samples thereby allowing targeted interrogation of gene expression. New bioinformatics approaches and the ability to examine entire pathways rather than individual genes is exciting as small changes in individual gene expression levels may be significant when examined in the context of overall pathway changes (Ganter et al. 2008).

The objective of this study was to characterize the gene expression profiles of human ameloblasts and odontoblasts and to further unravel the transcriptome of human teeth that we call the dentome. The investigation reveals many genes and molecular pathways not previously known to be involved in tooth formation that appear to be important.

Materials and methods

Tissue collection and preparation

Written consent was obtained from mothers that were preparing for elective abortions in this IRB approved protocol through the University of North Carolina-Chapel Hill Office of Human Research Ethics. Four human fetuses were obtained at ages 15-20 weeks gestation, immediately placed on ice and the tooth buds dissected from the jaws, placed in RNAlater and refrigerated at 4C for 1-4 weeks to allow decalcification in Richard Allan Scientifics' decalcifiying solution (ThermoScientific, Grand Island, NY, USA). The tissue was then frozen and stored at -80^{o}C. The tissue was sectioned at -35C at a thickness of 7 microns and lightly stained with haematoxylin and eosin to allow better visualization of the different cell types. Only anterior teeth from both the maxillary and mandibular jaws were used due to similar stage of dental development.

Laser capture of specific tissue

AutoPix^{TM} automated LCM system from Arcturus (Arcturus Engineering, Santa Clara, CA, USA) was used to isolate the human odontoblasts and ameloblasts in different stages of enamel formation, using static image settings. Ameloblasts have different morphological features and organelle content during different stages of enamel formation. In this study ameloblasts were

assigned to specific developmental stages based on the presence or absence of visible enamel matrix at the light microscope level and cell morphology (e.g. presence of Tomes Process). Cells isolated before enamel apposition were designated as being in the pre-secretory stage and if enamel extracellular matrix was visible the ameloblasts were classified as being in the secretory stage. Odontoblasts adjacent to the enamel epithelium secrete the predentin matrix and were harvested predominantly after some dentin matrix secretion. This process allowed standardization of the cell's developmental stage despite slight differences in an individual tooth bud's stage of development.

During the microdissection procedure, a CapSure (ThermoFisher, Grand Island, NY, USA) cap is positioned over the tissue section. At the end of the LCM procedure the transparent thermoplastic film that covers the cap was peeled off with the attached cells of interest and placed into RNA extraction buffer. Images were obtained of the tissue sections before and after LCM, including the captured regions.

Total RNA was isolated from the microdissected cells with the PicoPure RNA Isolation kit (Arcturus Bioscience, Santa Clara, CA, USA). The quality and yield of total RNA were assessed on an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). Samples from each cell type were then sent for analysis to determine the RNA Integrity Number (RIN).

RNA Microarray

Four samples each of odontoblasts, pre-secretory ameloblasts and secretory ameloblasts from 4 fetuses were obtained and RNA extracted. The RNA from each CapSure cap was separately isolated and RNA from each individual tooth bud then pooled to obtain at least

200ng of RNA for each sample. This gave us a total of 4 different samples of each cell type for microarray analysis. Whole genome human oligonucleotide microarrays (41K Agilent) were used to examine gene expression of the different tissue. The arrays contain 44K 60-mer oligonucleotides representing over 41K human genes and transcripts.

Two hundred nanograms of total RNA was converted into labeled cRNA with nucleotides coupled to fluorescent dye Cy3 using the Low RNA Input Linear Amplification Kit (Agilent Technologies, Palo Alto, CA) following the manufacturer's protocol. The Human Universal Reference RNA from Stratagene (Santa Clara, CA, USA) was coupled with Cy5.

Cy3-labeled cRNA (1.65 ng) from each sample was hybridized to Agilent whole genome array 41k formated chips. The hybridized array was then washed, scanned and data was extracted from the scanned image using Feature Extraction version 9.5 (Agilent Technologies, Palo Alto, CA). The microarray data is then submitted to the Gene Expression Omnibus (GEO) microarray database (accession number GSE63289).

RTPCR

RNA from a sample each of pre-secretory ameloblasts, secretory ameloblasts and odontoblasts were compared to the Human Universal Reference RNA from Stratagene by probing for high intensity, medium intensity and low intensity levels of expression based on the microarray data. The probes *ACTIN*, *AMELX*, *COL6A3*, *FAM40B*, *HPRT1*, *IL11* were selected based on intensity levels obtained from the microarrays. RTPCR was performed using QIAGEN RT2 qPCR Primer Assays (Frederick ,MD, USA) in an Eppendorf Mastercycler gradient thermocycler. Experiments were repeated in triplicate. Intraclass correlation coefficient were

calculated between microarray log ratio and RTPCR expression ratio using SPSS 21 (IBM, Armonk, NY, USA).

Data analysis

Unbiased cluster analysis was carried out on the samples using Cluster 3.0 (open source) and the heat maps visualized using Java TreeView-1.1.6r2 (open source).

For the expression data, each microarray was examined and the minimum intensity for expression was set at 95% confidence of the negative controls on that array. The lists were crossed referenced and only genes that were expressed in all 4 samples for each tissue type was determined to be expressed. Differential expression was examined using Significance analysis of microarrays (SAM) 4.0 between the different stages of ameloblasts and different tissues. The false discovery rates of the SAM analyses were set at 20%.

Ingenuity pathway analysis was used to examine the pathways of the differential expression between the samples specifically focusing on the upstream analysis.

Results

We performed laser capture microdissection and microarray of 3 different cell types from tooth buds removed from 4 individual fetuses. Cells from each fetus were analyzed separately. According to the Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA), the microdissection yielded an estimated 10-15 ng total RNA per sample with a 260/280 ratio between 2.12 and 2.14. The samples had an average RIN number of 5.4 ± 0.74 SD.

Genes expressed during tooth formation

Laser capture microdissection allowed us to obtain small discrete areas of cells at relatively specific developmental stages from human tooth bud tissues (Figure 2.1). Using the negative control spots to eliminate background at 95% confidence level, we found that during tooth formation odontoblasts expressed 14,802 genes, pre-secretory ameloblast expresses 15,179 genes and secretory ameloblast expresses 14,526 genes.

Differential gene expression in enamel and dentin formation

To validate the overall gene expression levels observed in the microarrays we performed RTPCR on selected genes in the samples and compared them with the standard reference RNA used in the microarray. The RTPCR analysis showed a intraclass correlation coefficient of 0.838 ($p<0.05$) for single measures, suggesting good correlation for genes (*ACTIN, AMELX, COL6A3, FAM40B, HPRT1, IL11*) that had high, moderate and low levels of expression on the microarray (Figure 2.2).

Using SAM analysis, we looked at differential gene expression between the different tissues at a false discovery rate of 20%. From the SAM plots (Figure 2.3) and unbiased cluster analysis (Figure 2.4), we observed that the greatest gene expression difference was between odontoblasts and pre-secretory ameloblasts.

Odontoblasts had 131 genes expressed at significantly higher levels and 15 genes at lower levels compared with pre-secretory ameloblasts. In addition, 4 genes (*ENAM, ASTN1, AMELX, SEZ6L*) were expressed at lower levels in odontoblasts compared with secretory

ameloblasts; 4 genes (*DMP1, AMBN, COPZ2, B3GALTL*) were expressed at lower levels in pre-secretory compared with secretory ameloblasts.

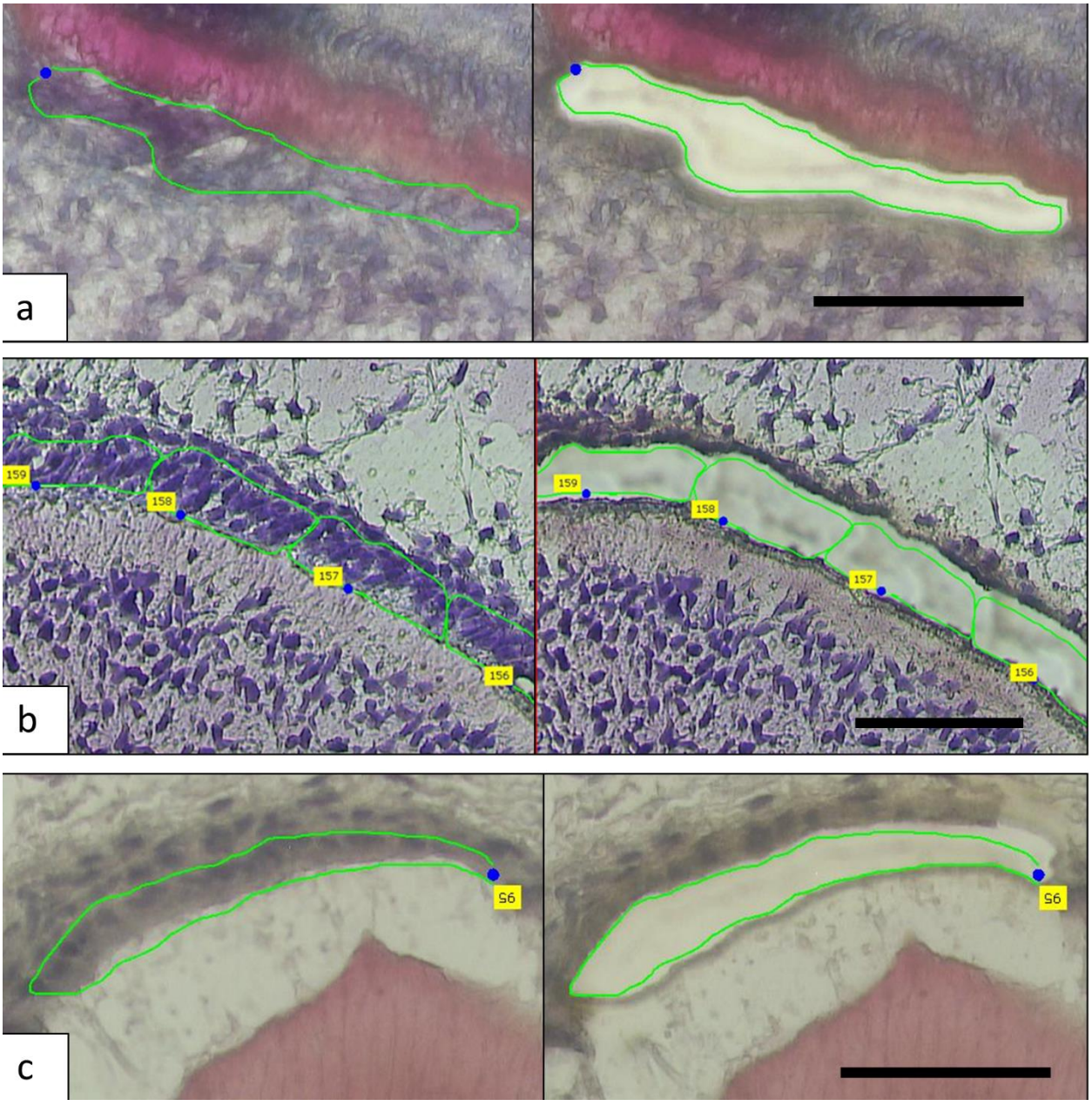
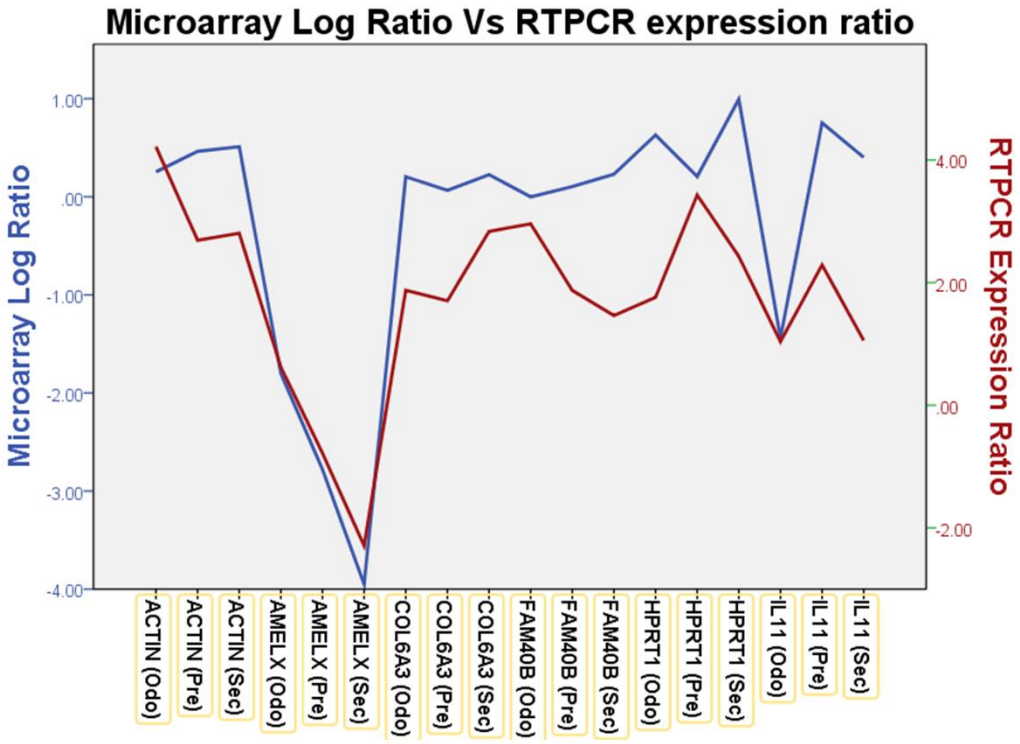


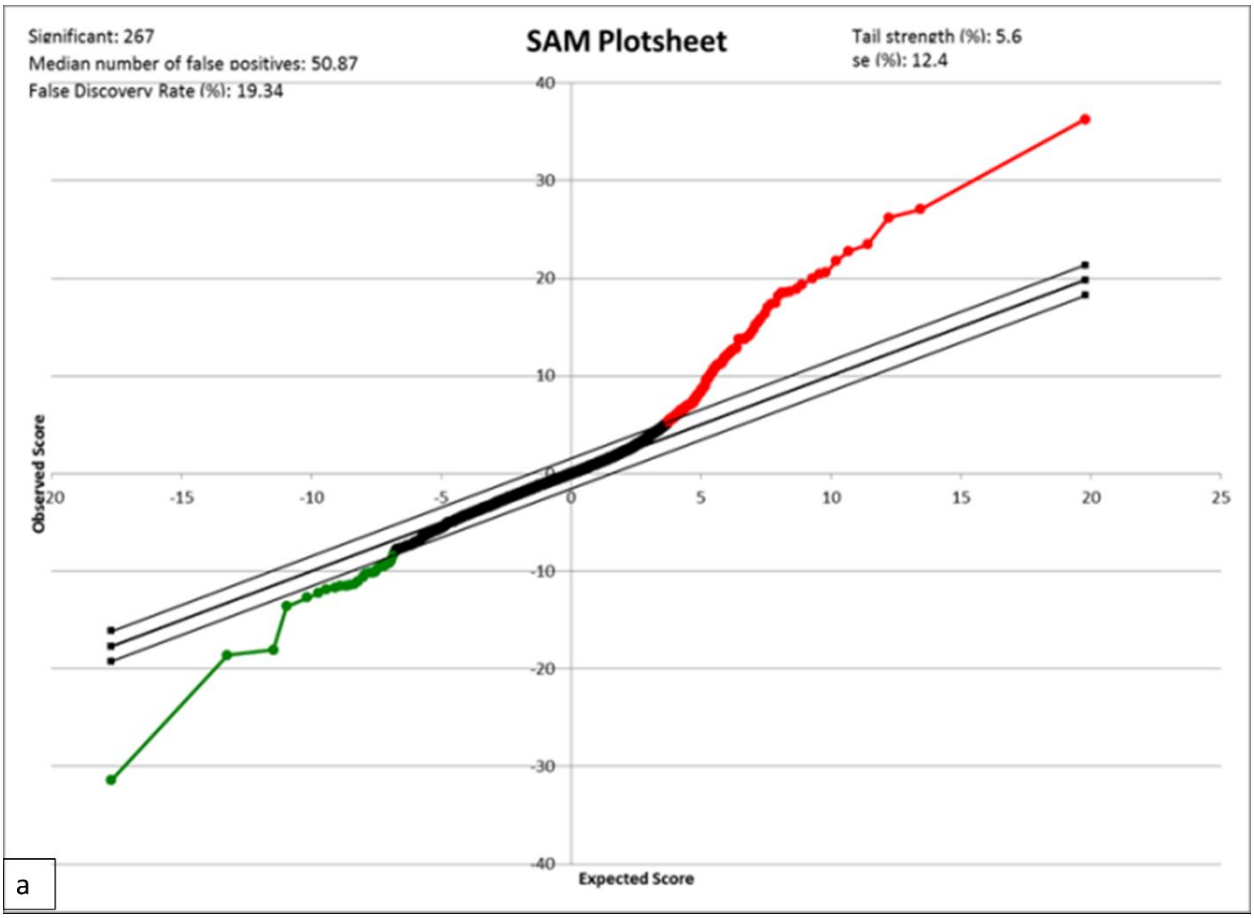
Figure 2.1: The micrographs show the laser capture of a – odontoblasts, b - pre-secretory ameloblasts and c – secretory ameloblasts. The left panel shows pre-capture and area to be captured while the right panel shows post-capture and the removal of target cells. Scale bar: 50 µm.

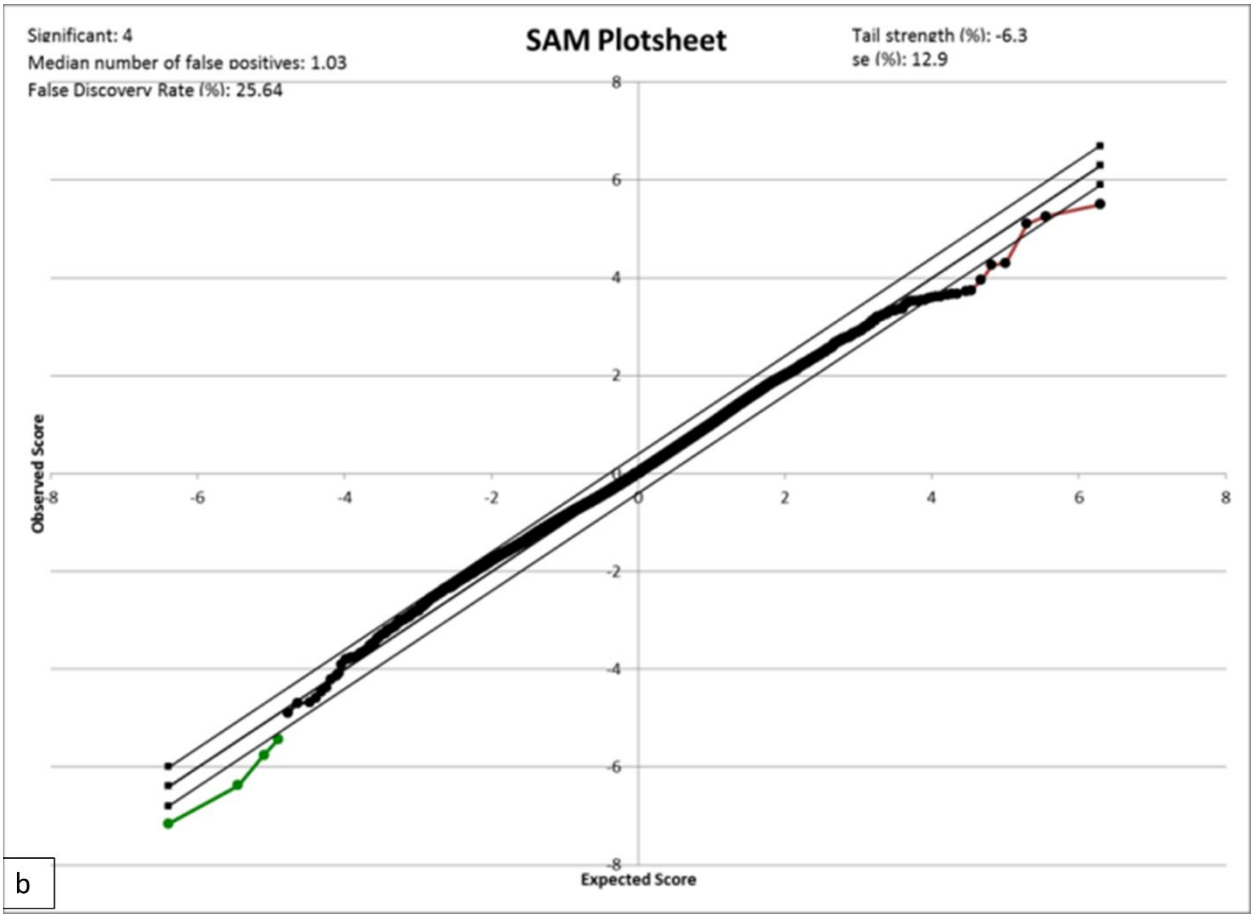


Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.838[a]	.619	.936	11.370	17	17	.000
Average Measures	.912[c]	.765	.967	11.370	17	17	.000

Figure 2.2: The graph shows comparison between the microarray Log ratio and RTPCR Log expression ratio. The microarray Log ratio is calculated using the ΔΔCt Method. Intraclass correlation coefficient analysis showed a significant correlation of 0.838 for single measure.


Significant: 267
Median number of false positives: 50.87
False Discovery Rate (%): 19.34
SAM Plotsheet
Tail strength (%): 5.6
se (%): 12.4
Observed Score
Expected Score
40
30
20
10
0
-10
-20
-30
-40
-20
-15
-10
-5
0
5
10
15
20
25
a


Significant: 4
Median number of false positives: 1.03
False Discovery Rate (%): 25.64
SAM Plotsheet
Tail strength (%): -6.3
se (%): 12.9
Observed Score
Expected Score
8
6
4
2
0
-2
-4
-6
-8
-8
-6
-4
-2
0
2
4
6
8
b

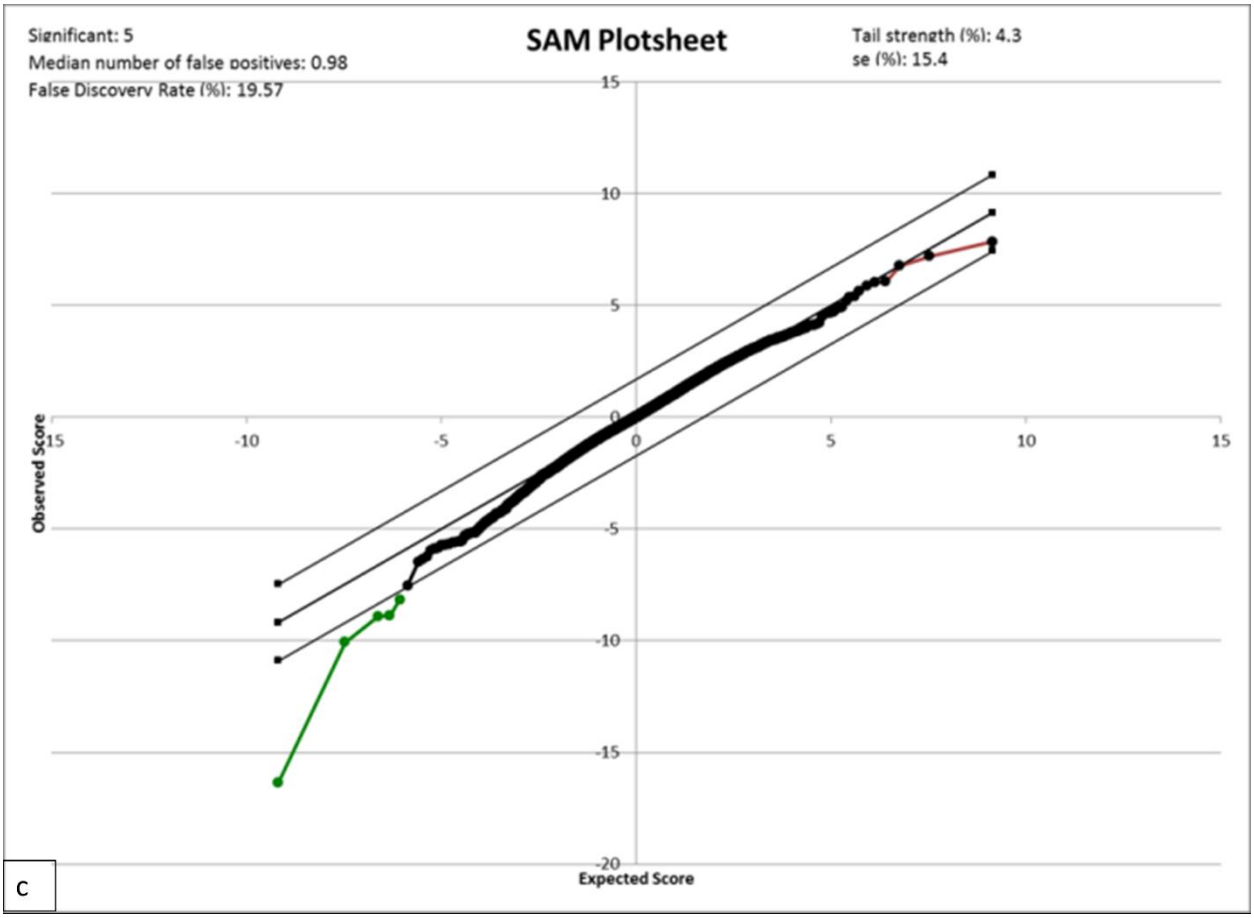


Figure 2.3: SAM plots: a – odontoblasts vs pre-secretory ameloblasts, b – odontoblast vs secretory ameloblasts, c – pre-secretory vs secretory ameloblasts. The red portion shows genes that are expressed at a higher level while the green portion shows genes that are expressed at a lower level with a FDR of 20%.

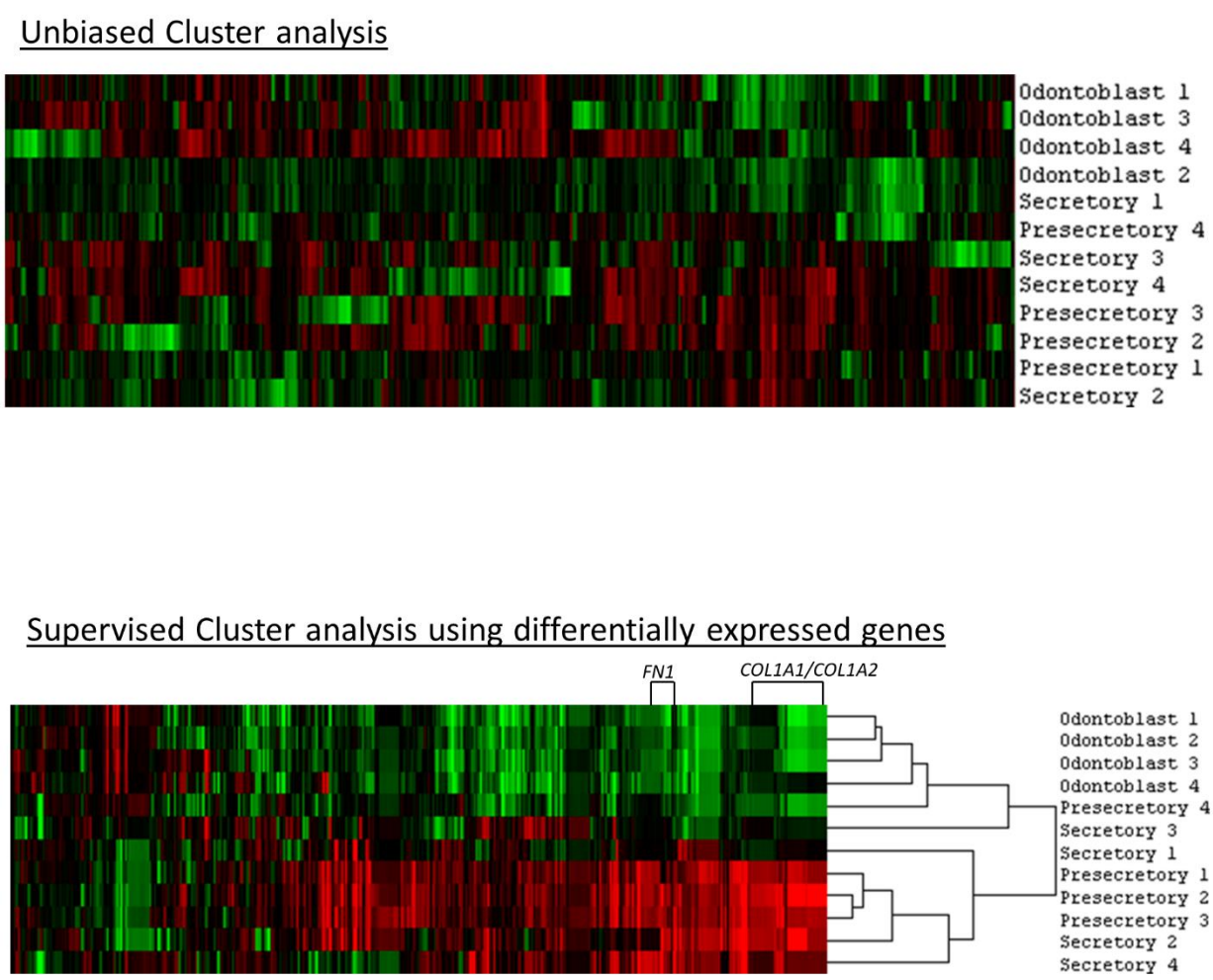


Figure 2.4: Heat map from cluster analysis. The map shows clustering of genes associated with collagen and extra cellular matrix formation.

Pathways important to enamel and dentin formation

The Ingenuity Pathway Analysis program uses an extensive database of canonical pathways to analyze the differentially expressed genes, the pathways they are involved in and examine the activated and/or inhibited pathways. Pathway analysis shows that the 2 networks with the greatest difference between odontoblasts and pre-secretory ameloblasts are mainly collagen and NF-κB driven. There were minimal differences for odontoblasts compared to secretory ameloblasts and pre-secretory compared to secretory ameloblasts.

Some of the canonical pathways that are different between odontoblasts and pre-secretory ameloblasts include - Intrinsic Prothrombin Activation Pathway, Hepatic Fibrosis / Hepatic Stellate Cell Activation, Atherosclerosis Signaling, Phospholipase C signaling, ERK5 signaling and Sphingosine-1-phosphate signaling. (Table 2.1)

Upstream Analysis

An upstream analysis examines current levels of gene expression detected by the microarray analysis and predicts which upstream regulators were most likely to be involved. Our analysis indicated that numerous upstream regulators were predicted to be different between odontoblasts and pre-secretory ameloblasts. Of note we found that *WNT3A*, *TGFB1*, *IGF2BP1*, *SHH*, *GLI1* and *FGF2* were predicted to be significantly more active in odontoblast while Alpha catenin (Figure 2.5) was inhibited in odontoblasts when compared with pre-secretory ameloblasts.

Table 2.1: Top 20 differentially expressed canonical pathways between odontoblasts and pre-secretory ameloblasts.

Ingenuity Canonical Pathways	-log(p-value)	Ratio	Molecules
Intrinsic Prothrombin Activation Pathway	3.84E00	1.14E-01	COL1A2,COL1A1,COL11A2,COL3A1
Hepatic Fibrosis / Hepatic Stellate Cell Activation	3.44E00	4.79E-02	COL1A2,COL1A1,FN1,PDGFRA,TGFB2,PDGFC,COL3A1
Atherosclerosis Signaling	3.12E00	4.41E-02	COL1A2,COL1A1,CLU,COL11A2,PDGFC,COL3A1
Phospholipase C Signaling	2.89E00	3.08E-02	RAP1B,GNAS,RALB,RHOU,RPS6KA3,ATF4,MEF2C,PRKCB
Dendritic Cell Maturation	2.48E00	2.9E-02	B2M,COL1A2,COL1A1,ATF4,COL11A2,COL3A1
ERK5 Signaling	2.46E00	6.15E-02	SGK1,RPS6KA3,ATF4,MEF2C
Corticotropin Releasing Hormone Signaling	2.44E00	3.68E-02	RAP1B,GNAS,ATF4,MEF2C,PRKCB
Sphingosine-1-phosphate Signaling	2.42E00	4.17E-02	GNAS,PDGFRA,RHOU,SMPD3,PDGFC
Guanosine Nucleotides Degradation III	2.21E00	9.09E-02	NT5E,PNP
Urate Biosynthesis/Inosine 5'-phosphate Degradation	2.14E00	9.09E-02	NT5E,PNP
Adenosine Nucleotides Degradation II	2.14E00	7.69E-02	NT5E,PNP
ILK Signaling	2.07E00	3.12E-02	FN1,TMSB10/TMSB4X,RHOU,VIM,ATF4,PDGFC
Melanocyte Development and Pigmentation Signaling	2.01E00	4.4E-02	GNAS,RPS6KA3,ATF4,KIT
Xanthine and Xanthosine Salvage	2E00	1.11E-01	PNP
Adenine and Adenosine Salvage I	2E00	1.11E-01	PNP
Purine Nucleotides Degradation II (Aerobic)	1.96E00	5.71E-02	NT5E,PNP
Glioma Signaling	1.83E00	3.57E-02	IGF2,PDGFRA,PDGFC,PRKCB
Guanine and Guanosine Salvage I	1.7E00	1.11E-01	PNP
Sulfate Activation for Sulfonation	1.7E00	1.25E-01	PAPSS2

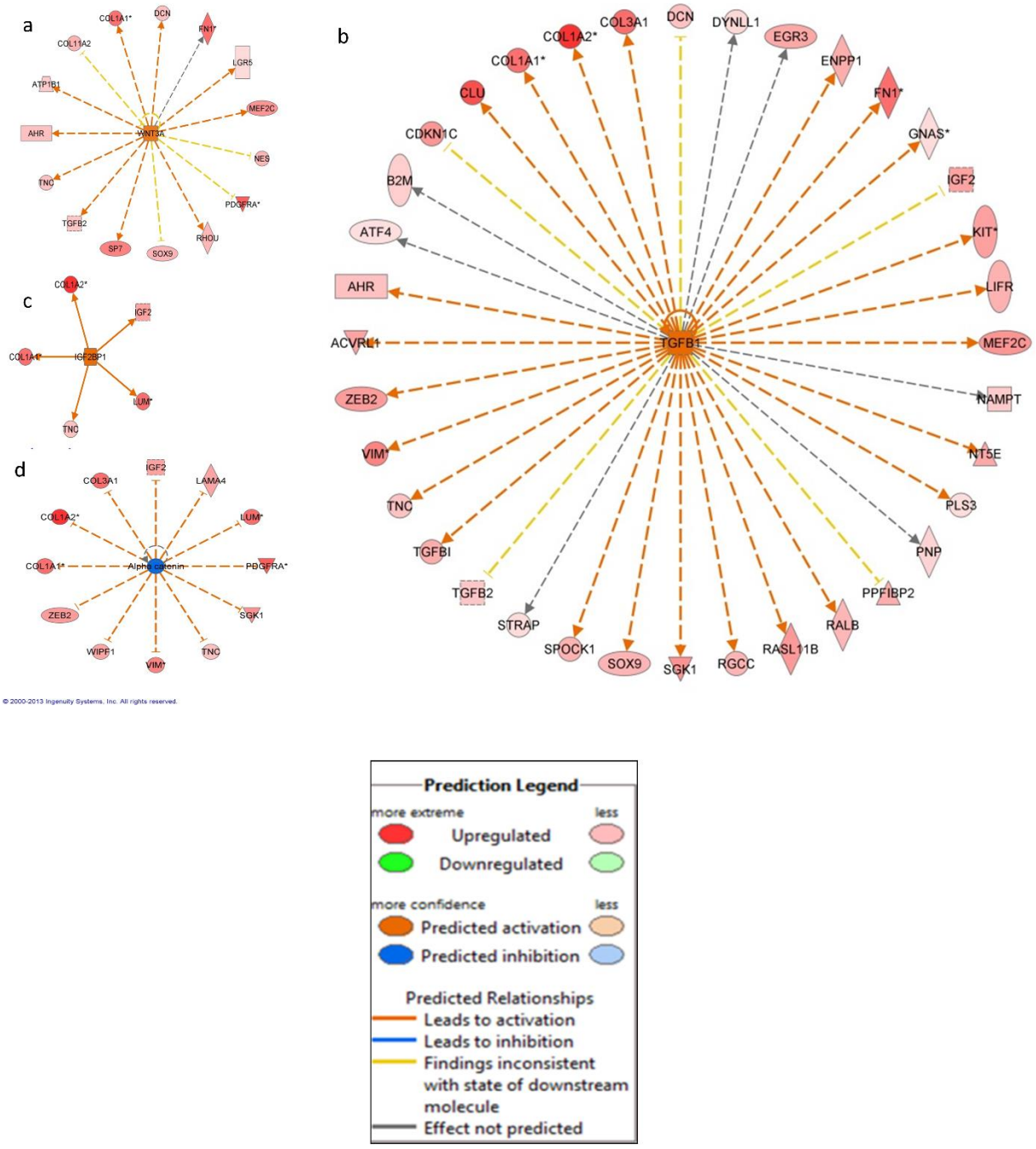


Figure 2.5: Upstream analysis: a – molecule network map for WNT3A, b – molecule network map for TGFB1, c – molecule network map for IGF2BP1, d – molecule network map for Alpha catenin. The predicted molecule is in the center with linkage to downstream targets based on differential expression from SAM analysis. The figure is generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, www.qiagen.com/ingenuity).

Discussion

Examination of murine tooth development using many different approaches suggested that more than 300 genes are involved in the tooth formation (Nieminen et al. 1998). More recently, the use of microarrays showed that 4362 genes are differentially expressed (Landin et al. 2012), suggesting that in fact large numbers of genes are involved. We showed that a much larger portion of the human genome is involved in the development of the human tooth. Many of the 14,802 genes and 15,179 genes expressed by odontoblasts and ameloblasts respectively are involved in basic cellular functions rather than tooth development specific functions. Other investigators found that various types of tissue express between 10,000 to 14,000 genes with an upper limit of 17,000 genes (Jongeneel et al. 2003). The present study clearly shows that the molecular control of odontogenesis is extremely complex and involves many more genes and molecular pathways than previously known.

Our observation that secretory ameloblast expression profile was more similar to odontoblast expression as compared with pre-secretory ameloblast was unexpected. This finding is likely due to secretory ameloblasts and odontoblasts sharing many similar characteristics as both are differentiated cells, are secreting an extracellular matrix, are controlling the microenvironment of that matrix, and are motile and moving away from the secreted matrix. This is in contrast to the pre-secretory ameloblasts that are in the process of differentiation, are not motile, and not secreting an extracellular matrix. As ameloblasts differentiate and mature, their expression profile becomes more similar to that of odontoblasts with both cells functioning to produce a mineralized extracellular matrix that involves similar yet distinct processes (Figures 2.3 and 2.4).

As expected, the SAM analysis showed significantly higher expression of known dentin products such as collagen type 1, collagen type 3 in odontoblasts when compared to pre-secretory ameloblasts. Our findings showed that the *ENAM* and *AMELX* genes which code for the enamel specific extracellular matrix proteins enamelin and amelogenin respectively were expressed at higher levels in enamel forming secretory ameloblasts. These differences in gene expression are reflective of the unique extracellular matrices produced by these different cell types and ultimately the compositionally and structurally different mineralized tissue they create.

Interestingly, fibronectin (*FN1*) and lumican (*LUM*), showed greater levels of expression in odontoblasts compared with pre-secretory ameloblasts. The proteins derived from these genes are thought to be involved in the organization of extracellular matrix (Kadler et al. 2008; Matheson et al. 2005) that is secreted by the odontoblasts during dentinogenesis. In addition, lumican is also shown to affect actin organization in the cell cytoskeleton (Radwanska et al. 2008) which has implications in the control of cell movement and shape. It is noteworthy that once the pre-secretory ameloblasts (non-motile cells) differentiate further to secretory ameloblasts (motile cells) *FN1* and *LUM* are not differently expressed compared with the odontoblasts (motile cells). Furthermore, *ASTN1*, which codes for the neuronal protein astrotactin, is expressed at a higher level in secretory ameloblasts compared to odontoblasts. Astrotactin is a major player in the migration and movement of neurons (Wilson et al. 2010) and could be an important component in the control of ameloblast movement, potentially reflecting the ectodermal origin of ameloblasts and conservation of this protein function during enamel formation. There is virtually nothing known about the complex regulation of motility in

these different odontogenic cell types but it is known that motility is critical in the normal formation and structure of the mineralized dentin and enamel. The other differentially expressed genes have been implicated in various other pathologies such as Peters' plus syndrome with cleft lip/palate (B3GALTL) (Schoner et al. 2013), tumor formation (COPZ2) (Shtutman et al. 2011) and lung cancer development (SEZ6L) (Gorlov et al. 2007) but have not yet been associated with tooth development.

Not surprisingly, the top canonical pathways are driven by clusters of collagen genes (Intrinsic Prothrombin Activation Pathway, Hepatic Fibrosis / Hepatic Stellate Cell Activation, Atherosclerosis Signaling). Of note, the Phospholipase C signaling pathway is an important signal transduction pathway that is implicated in cancer development (Bunney and Katan 2010) and regulates chemotaxis through the modulation of Phosphoinositide 3-kinase activity to provide a direction-sensing machinery (Kolsch et al. 2008). In addition, the activation of ERK5 signaling pathway is associated with disruption to cell actin cytoskeleton. This change in cell actin dynamics can lead to increased cell motility and decreased cell adhesion (Barros and Marshall 2005). Furthermore, Sphingosine-1-phosphate is thought to be a more potent chemoattractant of dental pulp stem cells than transforming growth factor β1 (TGF-β1), fibroblast growth factors (FGFs), and epidermal growth factors (EGFs) (Howard et al. 2010). Dental pulp stem cells are postulated to form reparative dentin in the event of tooth injury by migrating and differentiating into odontoblasts (Qvist 1975). The signaling pathways identified in this study could be contributing to the motility of odontoblasts compared to non-motile pre-secretory ameloblasts.

Examination of upstream regulators showed that *SHH*, *GLI1*, *FGF2* and *TGFB1* are likely involved in dentinogenesis. This finding was expected as these genes have been well characterized to be involved in tooth formation and the disturbances of these genes are known to lead to tooth malformation and agenesis (Cobourne et al. 2001; Lee et al. 2011; Pispa et al. 1999). Upstream analysis also indicated that both *WNT3A* and *TGFB1* are activated in odontoblasts (Figure 2.5a-b). Recent investigation of Schimke immuno-osseous dysplasia, a syndrome associated with mutations of the *SMARCAL1* gene shows there are cellular disturbance in the expression of *WNT3A* and *TGFB1* as a result of abnormal *SMARCAL1*. Patients with this rare condition exhibit microdontia, hypodontia and severe molar root hypoplasia (Morimoto et al. 2012). These findings suggest that these pathways could be important in root formation and tooth morphogenesis.

The present study also found that *IGF2BP1* was identified as an active upstream regulator in odontoblasts (Figure 2.5c) and a human GWAS showed that it is one of the loci associated with tooth agenesis (Pillas et al. 2010). *IGF2BP1* has also been implicated in the upregulation of betaTrCP1 leading to the activation of beta-catenin/Tcf signaling (Noubissi et al. 2006) that is important in beta-actin mRNA translation and cell migration. This pathway provides a possible mechanism for the movement of odontoblasts as it deposits newly formed dentin.

Conversely, Alpha catenin, that was predicted to be expressed at a lower level in odontoblasts when compared to pre-secretory ameloblasts (Figure 2.5d), is known to be important in ameloblast development (Sorkin et al. 2000). The cadherin–catenin complex, with Wnt/beta-catenin signaling, also has critical roles in regulating cell motility/adhesion (Van den

Bossche et al. 2012). The nature and complexity of movement has been shown to be quite different between odontoblasts and ameloblasts. A recent paper described the importance of MMP20 and the cadherin complexes during ameloblast maturation to allow movement of the epithelial cells by the switching of cadherin types (Guan and Bartlett 2013).

One of the shortcomings of this study is that we were only able to obtain primary tooth buds for examination. It is likely that there are differences between the dentome of primary and permanent teeth; however, permanent tooth buds were not available in our study sample. Future studies of permanent tooth buds using similar protocols will shed light on the differences between the dentome of primary and permanent teeth. Later stages of tooth formation were not evaluated as the later stages of tooth formation (e.g. maturation) were not available. In addition, although the use of the LCM aids in providing a relatively discrete sample, there is still a possibility of contamination from adjacent cell layers such as the stratum intermedium collected with the ameloblasts. Furthermore, we did not analyze cementum or root forming cells, leaving the dentome incomplete at the moment.

In summary, our results support the utility of laser capture microdissection as a valuable tool that allows interrogation of different tissues and cell types present in human teeth during different stages of development. The use of laser capture and RNA microarrays shows that the early developing human tooth transcriptome involves more genes than anticipated and diverse molecular pathways that are differentially activated in the tooth forming cells. We identified genes and pathways not previously known to play a role in tooth formation. For example, the genes and pathways involved in cell movement may be essential processes in normal odontogenesis that are worth closer examination. Unravelling the human dentome will advance

our knowledge of tooth formation and the critical events that are regulated by gene expression and that control cell function and development of the tooth.

CHAPTER 3: AMELOBLASTOMA PHENOTYPES REFLECTED IN DISTINCT TRANSCRIPTOME PROFILES

Shijia Hu[1,2], Joel Parker[3], Kimon Divaris[1,4], Ricardo Padilla[5], Valerie Murrah[5], John Timothy Wright[1]

[1]Pediatric Dentistry, School of Dentistry University of North Carolina-Chapel Hill, Chapel Hill, NC, USA

[2]Faculty of Dentistry, National University of Singapore, Singapore

[3]Cancer Genetics, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA

[4]Epidemiology, Gillings School of Global Public Health, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA

[5]Diagnostic Sciences, School of Dentistry, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA

Abstract

Ameloblastoma is a slow growing, locally invasive benign epithelial odontogenic neoplasm derived from odontogenic epithelium. Although relatively rare, it makes up 40-50% of odontogenic tumors. A high recurrence rate of 50-80% with conservative treatment in some sub-types warrants radical surgical resections resulting in high morbidity. Ameloblastoma presents with diverse phenotypes which have yet to be characterized at the molecular level. The objective of the study was to characterize the transcriptome of ameloblastoma and identify relevant genes and molecular pathways using normal odontogenic tissue for comparison. Seventeen patients with ameloblastoma formed the study sample. Laser capture microdissection was used to obtain discrete samples of neoplastic epithelial tissue from excised tumors which were examined using the Agilent 44k whole genome microarray and the NanoString nCounter system. Odontogenic tissue transcriptome (human "dentome") was used as reference for cluster, fold change, pathway and gene set enrichment analyses. The ameloblastomas separated into 2 distinct molecular clusters that were associated with 2 types of odontogenic tissue, namely pre-secretory ameloblast and odontoblast. Within the pre-secretory cluster, 9/10 of ameloblastoma samples were of the follicular type while 6/7 of the samples in the odontoblast cluster were of the plexiform type (p<0.05). Analysis of differential gene expression revealed alteration of common pathways in both clusters including cell cycle regulation, inflammatory and MAP kinase. The pre-secretory ameloblast cluster exhibited higher activation of inflammatory pathways while the odontoblast cluster showed greater disturbances in transcription regulators. Known cancer-driving genes such as *TP53* and members of the MAP kinase pathways were predicted to be altered in upstream analyses. Our

results are suggestive of underlying inter-tumor molecular heterogeneity of ameloblastoma, where different clinical phenotypes show different molecular signatures and involve different pathways. These findings highlight the molecular heterogeneity of ameloblastoma sub-types and may have implications in the tailored use of other treatment modalities.

Introduction

Ameloblastoma is a slow-growing, locally invasive, benign epithelial odontogenic neoplasm. It is thought to be arise from *SOX2*-expressing dental lamina epithelium (Juuri et al. 2013), remnants of the tooth-forming enamel organ (Sehdev et al. 1974). The tumor exhibits epithelial cells resembling pre-ameloblasts on a basement membrane in loosely arranged cells resembling stellate reticulum while the stroma consists of loose connective tissue. Although odontogenic tumors are relatively rare, they constitute 3.8% of head and neck pathology, of which 40-50% are ameloblastoma (Avelar et al. 2011; Siriwardena et al. 2012). Occasionally, ameloblastomas show malignant features or transform into malignancy (Uzawa et al. 2015) and in rare cases metastasize (Luo et al. 2012). Current treatment modalities range from conservative enucleation to radical excision and vary according to tumor subtypes and location (Reichart et al. 1995; Singh et al. 2014). High recurrence rates (50-80%) have been observed in cases of conservative treatment (Mendenhall et al. 2007). Consequently, and despite recent advances in imaging-assisted surgical margin localization, post-operative histological confirmation is still required. This forces surgeons to either act conservatively, risking the need for a second surgery, or act aggressively thus increasing morbidity (De Silva et al. 2012) and the need for extensive reconstructive surgery.

There are few genomics and transcriptomics studies of ameloblastoma, with most investigations focusing on candidate-genes. Moreover, different comparison tissues were used in the handful of microarray studies; including gingival tissue (Carinci et al. 2003; Heikinheimo et al. 2015b), whole tooth buds (Heikinheimo et al. 2002), dentigerous cysts (Lim et al. 2006) and a universal human reference RNA (DeVilliers et al. 2011). Furthermore, whole tumor samples were used in these studies which includes large portions of stromal tissue. In spite of the heterogeneity in comparison tissue, there have been advances in understanding tumorigenesis of ameloblastoma.

A recent study examining the whole transcriptome of ameloblastoma suggested the existence of distinct molecular subtypes (Heikinheimo et al. 2015b). It is envisaged that better understanding of the molecular basis of ameloblastoma can aid the identification of diagnostic and prognostic markers and may lead to the development of novel, personalized treatment protocols (Gomes et al. 2014). To address this knowledge gap, we embarked on this study aiming to characterize the transcriptome of neoplastic ameloblastoma tissue and identify relevant molecular pathways and genes, using a whole genome microarray.

Materials and methods

Tumor collection and preparation

Written informed consent was obtained from 17 subjects diagnosed with ameloblastoma and slated to undergo surgical treatment at the University of North Carolina (UNC) Department of Oral and Maxillofacial Surgery in this IRB-approved study. Between 2005 and 2008, 2 fresh frozen samples were obtained during surgical resection and 15 formalin-fixed

paraffin-embedded (FFPE) samples were retrieved from the archives of the Department of Oral and Maxillofacial Pathology Laboratory, UNC School of Dentistry. All samples were evaluated by a board-certified oral and maxillofacial pathologist and at least one other author and diagnoses were classified based on the 2005 WHO Histologic Classification of Odontogenic Tumors. Additional demographic data including gender, age, race, and tumor recurrence were recorded and examined for potential associations.

Dissected tumors were placed in RNA*later* and Richard Allan Scientifics' decalcifying solution (water, hydrochloric acid, EDTA, tetrasodium tartrate and potassium tartrate) at 4^{o}C for 1-4 weeks before 7µm sections were obtained under RNAse-free conditions (DeVilliers et al. 2011). The sections were then used for laser capture microdissection (LCM).

Laser capture microdissection

The ability of LCM to isolate one cell-thick discrete tissue populations (Hayashi et al. 2010) facilitates the targeting and pooling of neoplastic epithelial portions of ameloblastoma. The AutoPix^{TM} automated LCM system (Arcturus Engineering, Santa Clara, CA, USA) was used to isolate tumor cells (basal epithelial cells adjacent to the basement membrane). Images of the tissue sections including the captured regions were obtained before and after LCM.

RNA extraction and microarray

Laser-captured cells from each tumor were pooled and total RNA was isolated with the PicoPure RNA Isolation kit (Arcturus Bioscience, Santa Clara, CA, USA). The Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA) was used to assess the yield and quality of total

RNA. Amplification was completed on all samples using TargetAmp^{TM} 2-Round Aminoallyl-aRNA Amplification Kit (Epicentre Biotechnologies, Madison, WI, USA). RIN values ranged from 2.3 to 4.9 with a mean of 3.0 ± 0.85.

RNA was then analyzed using whole genome human oligonucleotide microarrays (41K Agilent) containing 44 thousand 60-mer oligonucleotides representing over 41 thousand human genes and transcripts. For this step, 200ng of RNA was converted into labeled cRNA with nucleotides coupled to fluorescent dye Cy3 using the Low RNA Input Linear Amplification Kit (Agilent Technologies, Palo Alto, CA) according to the manufacturer's protocol. The Human Universal Reference RNA from Stratagene (Santa Clara, CA, USA) was coupled with Cy5. Cy3-labeled cRNA (1.65 ng) from each sample and the Cy5-labeled universal reference was hybridized to the Agilent whole genome array 41k formatted chips. Data were extracted using Feature Extraction version 9.5 (Agilent Technologies, Palo Alto, CA). Background subtraction and Loess normalization were performed using default setting of the Agilent extractor. The use of the universal RNA facilitated the use of the dentome as a comparison. It acts as a technical intra/inter normalizing control, decreasing variability by measuring signal output ratio of experimental to reference RNA rather than relying on absolute signal intensity two-color hybridization experiments (Novoradovskaya et al. 2004). The dentome consists of odontogenic tissue (microdissected samples of human odontoblasts, pre-secretory ameloblasts and secretory ameloblasts) expression data from previous work that employed the universal reference as a normalizing control (Hu et al. 2015). The data set included 4 samples of each type of odontogenic tissue. It was collected from 12 anterior tooth buds (incisor and canine) from 4 different fetuses with each fetus contributing 3 tooth buds. Each type of odontogenic

tissue was collected from a single tooth bud with each of the 3 buds providing a single type of odontogenic tissue or developmental stage. (Gene Expression Omnibus microarray database accession number GSE63289)

The ameloblastoma expression data were submitted to the Gene Expression Omnibus microarray database (accession number GSE68531).

Microarray data analysis

A multiclass analysis was conducted between the 3 types of odontogenic tissue and the 60 genes differentially expressed at a false discovery rate (FDR) < 20% were designated as the odontogenic tissue-defining genes. The most appropriate comparison tissue was decided to be the normal tissue with the most similar profile to ameloblastoma, such that identified differences would be tumor specific. Cluster analysis was conducted using Cluster 3.0 between the 3 normal and tumor samples and visualized using Java TreeView-1.1.6r2.

Differential gene expression between tumors and comparison tissue were examined using Significance analysis of microarrays (SAM) 4.0. Ingenuity pathway analysis was used to identify differentially expressed pathways. In addition, upstream analysis from the ingenuity pathway analysis software was conducted. Gene set enrichment analysis (GSEA) (Subramanian et al. 2005) was conducted using GSEA v2.1.0 from the Broad institute (Cambridge, MA, USA) and the “all curated gene sets v4.0” available via the Molecular Signatures Database (MSigDB) (Liberzon et al. 2011). Additionally, the ameloblastoma transcriptome was compared with the 13 cancer molecular subtypes from The Cancer Genome Atlas (TCGA) project (Hoadley et al. 2014) to investigate possible correlation with other known cancer types.

Microarray gene expression validation using NanoString

A variety of approaches have been used to validate microarray data in the literature and NanoString was selected for the study. A random subset of 3 ameloblastoma and 2 control odontogenic tissue samples was used to validate the microarray gene expression data. NanoString nCounter (Seattle, WA, USA) high throughput gene expression analysis (Geiss et al. 2008) was performed using the Human Cancer Reference codeset (http://www.nanostring.com/products/gene_expression_panels). Each reaction contained 50 ng of total sample RNA plus reporter and capture probes. Digital counts were extracted, normalized and analyzed using nSolver v2.5 software. Differential expression between ameloblastoma and pre-secretory ameloblast from nanoString was compared with that obtained from the microarray.

Results

LCM facilitated the isolation of basal epithelial (neoplastic) cells from the tumor samples (Figure 3.1) without contamination from surrounding stroma cells. RNA was extracted from the LCM samples, with the 260/280 ratio for the 17 samples between 1.7-2.1 and a yield of between 88ng to 928ng. Quality control analysis conducted on the microarray chips indicated expected values for positive and negative controls, as well as uniformly high detected genes in both the red and green channels. Overall, no outliers were detected in either the normal tissue or tumor arrays indicating consistency in hybridization between samples.

NanoString validation

Fold changes obtained with the nCounter system were correlated with those obtained from the microarray for the same samples (Table 3.1). The 2 sets of expression data showed a good Pearson correlation ($r = 0.61$) in the scatter-plot (Figure 3.2).

Determination of comparison tissue

Unsupervised hierarchical cluster analysis was conducted for the tumor and normal tissue samples which showed the presence of 2 distinct clusters of ameloblastoma and a separate cluster of normal odontogenic tissue (Figure 3.3). Supervised cluster analysis using the 60 odontogenic tissue defining genes showed that the 2 clusters of ameloblastoma associated most closely with pre-secretory ameloblast (PA) and odontoblast (OB) (Figure 3.4A, 4B). As such, these 2 clusters were designated as the pre-secretory ameloblast-like ameloblastoma (pAM) and odontoblast-like ameloblastoma (oAM), respectively. Out of 10 samples in the pAM cluster, 9 were of the follicular type while 6/7 of the samples in the oAM cluster were of the plexiform type. (Figure 3.4C). A Chi-square analysis showed that the molecular clusters were significantly associated with a histological subtype ($p<0.05$). A single comparison tissue could not be designated as the 2 clusters associated most closely with different odontogenic tissue; instead, a multiclass approach was employed.

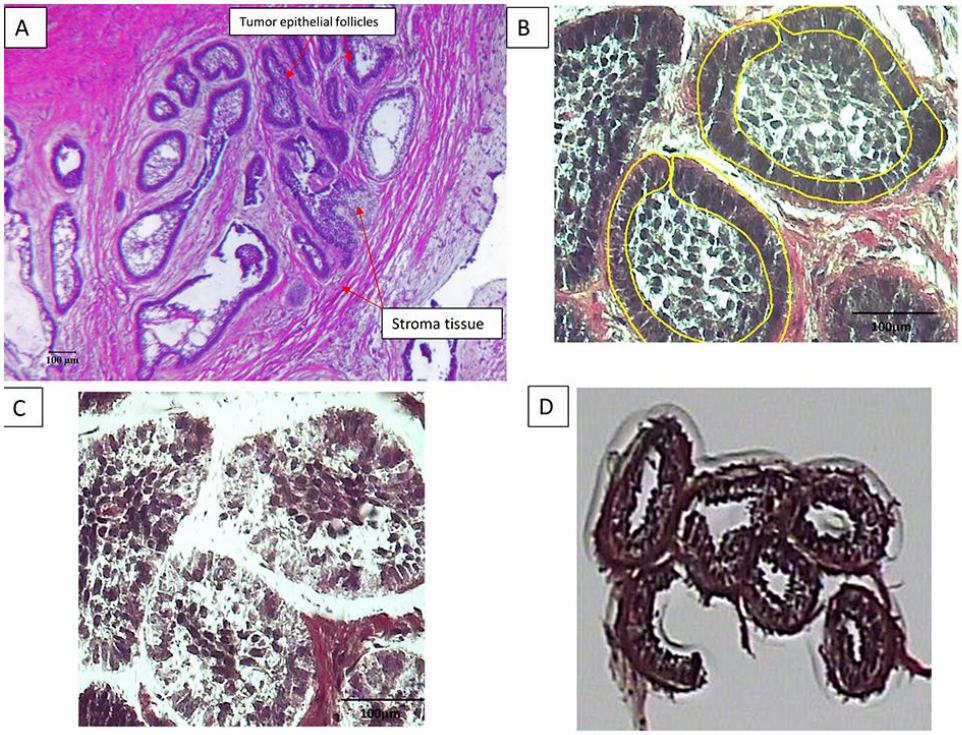


Figure 3.1: The micrographs show the laser capture of the epithelial portion of an ameloblastoma sample. A – light micrograph of follicular ameloblastoma at 4X showing tumor epithelial follicles that are single-cell thick with surrounding stroma, B – laser capture outline of epithelial cells, C – remnants of the stroma tissue after LCM and D – captured cells on capsure cap. Scale bar: 100 µm.

Table 3.1

Top 20 upregulated genes in ameloblastoma compared to pre-secretory ameloblast

Gene Name	NanoString Fold Change	Microarray Fold change
IL1A	8.39	1.45
IL8	7.16	1.43
AREG	6.96	-1.74
IL1B	6.74	-1.19
MMP1	6.38	1.41
SERPINE1	6.23	-2.41
CASP10	6.19	-1.29
CSF3	6.03	1.02
MMP3	5.91	1.27
NTRK1	5.67	1.79
CYP1A1	5.43	1.41
GATA1	5.43	1.26
IL4	5.43	1.06
MPL	5.43	1.15
PLG	5.43	1.44
TERT	5.43	-1.02
WEE1	5.43	1.20
RET	5.40	1.33
IL6	5.36	1.18
EGF	4.84	1.13

Top 20 downregulated genes in ameloblastoma compared to pre-secretory ameloblast

Gene Name	NanoString Fold Change	Microarray Fold change
SPP1	-5.83	-1.19
TYMS	-4.03	-1.18
KDR	-3.67	-1.18
DLC1	-3.63	-1.87
KRAS	-3.59	-2.39
MYC	-3.53	-2.29
FYN	-3.29	-1.02
MSH6	-3.17	-1.16
TOP2A	-3.05	-1.53
CCNA2	-2.91	-1.12
COL1A1	-2.79	1.30
JUNB	-2.73	-1.05
IGF1	-2.68	-1.76
PDGFRA	-2.56	-2.13
EPS8	-2.36	-2.36
CCND2	-2.36	-1.76

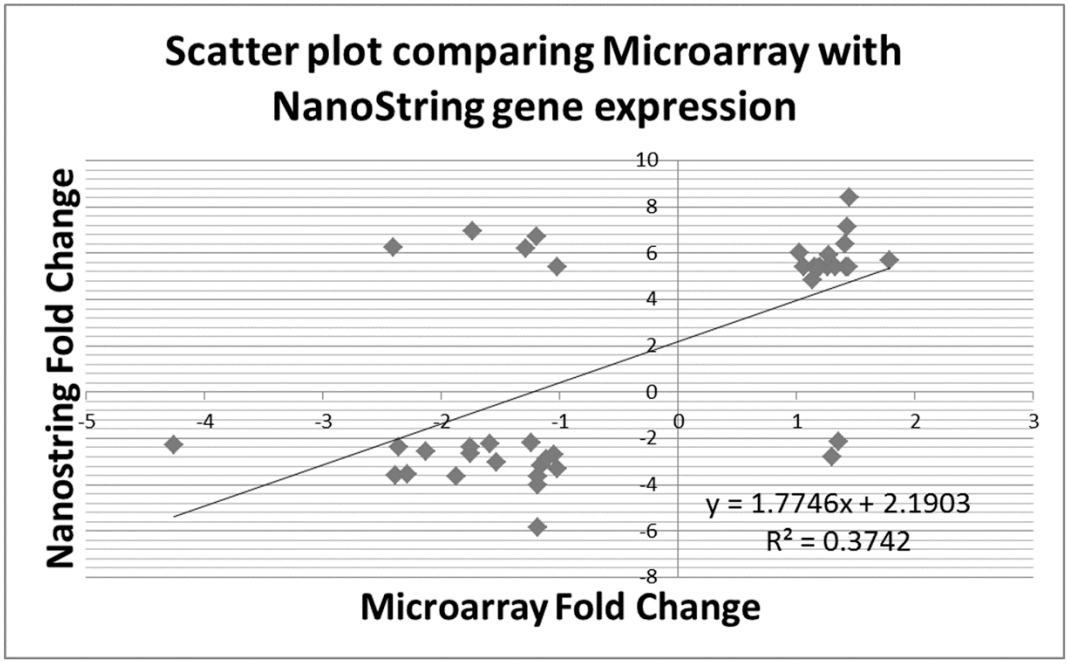


Figure 3.2: Scatterplot of Microarray (x-axis) differential expression versus NanoString (y-axis) differential expression. There is Pearson's correlation of 0.61 between the 2 platforms.

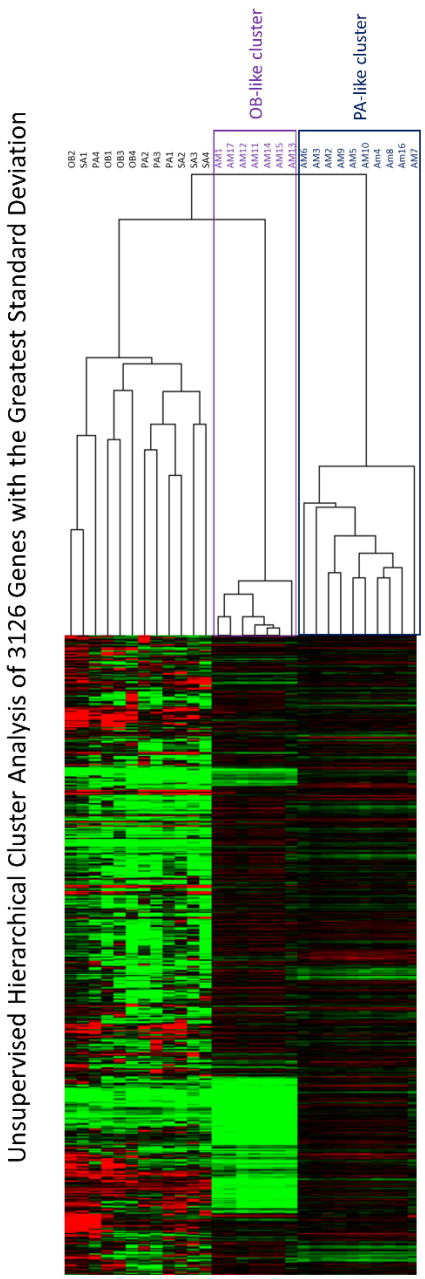
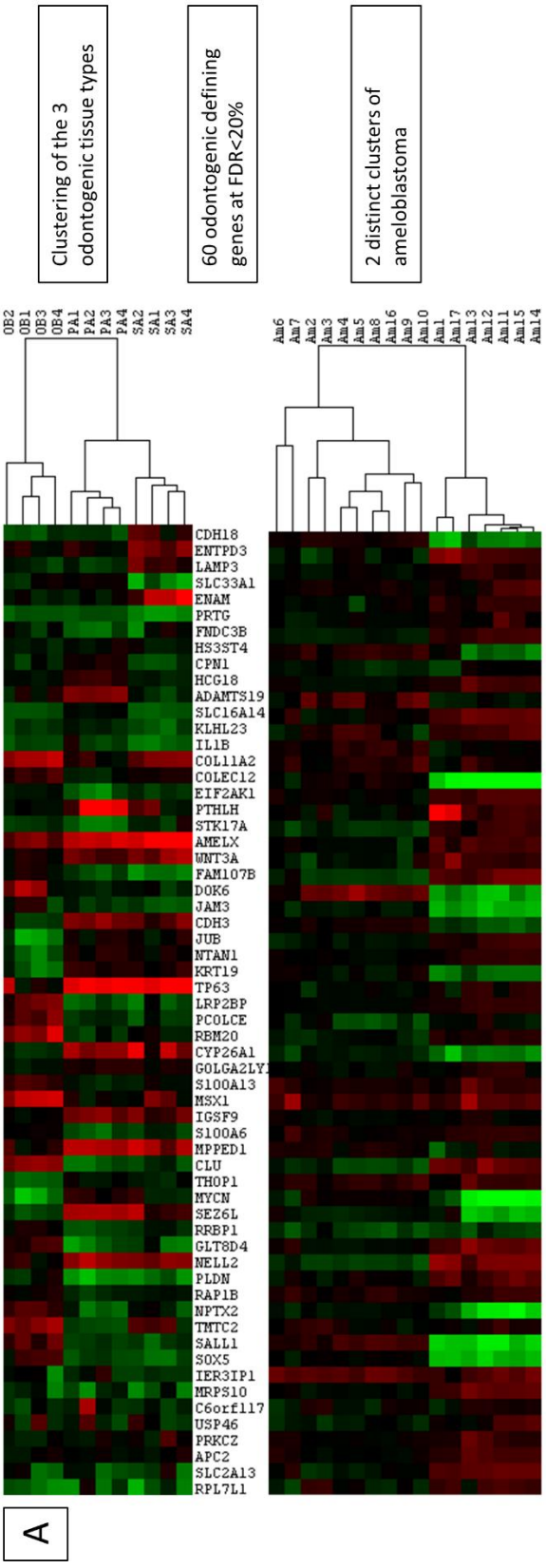


Figure 3.3: Unsupervised hierarchical cluster analysis of reference tissue with tumor samples. 3126 genes with the greatest standard deviation between the samples were used for the cluster analysis.


Clustering of the 3 odontogenic tissue types
60 odontogenic defining genes at FDR<20%
2 distinct clusters of ameloblastoma
OB2
OB1
OB3
OB4
PA1
PA2
PA3
PA4
SA2
SA1
SA3
SA4
Am6
Am7
Am2
Am3
Am4
Am5
Am8
Am16
Am9
Am10
Am1
Am17
Am13
Am12
Am11
Am15
Am14
CDH18
ENTPD3
LAMP3
SLC33A1
ENAM
PRTG
FNDC3B
HS3ST4
CPN1
HCG18
ADAMTS19
SLC16A14
KLHL23
IL1B
COL11A2
COLEC12
EIF2AK1
PTHLH
STK17A
AMELX
WNT3A
FAM107B
DOK6
JAM3
CDH3
JUB
NTAN1
KRT19
TP63
LRP2BP
PCOLCE
RBM20
CYP26A1
GOLGA2LY1
S100A13
MSX1
IGSF9
S100A6
MPPED1
CLU
THOP1
MYCN
SEZ6L
RRBP1
GLT8D4
NELL2
PLDN
RAP1B
NPTX2
TMTC2
SALL1
SOX5
IER3IP1
MRPS10
C6orf117
USP46
PRKCZ
APC2
SLC2A13
RPL7L1
A

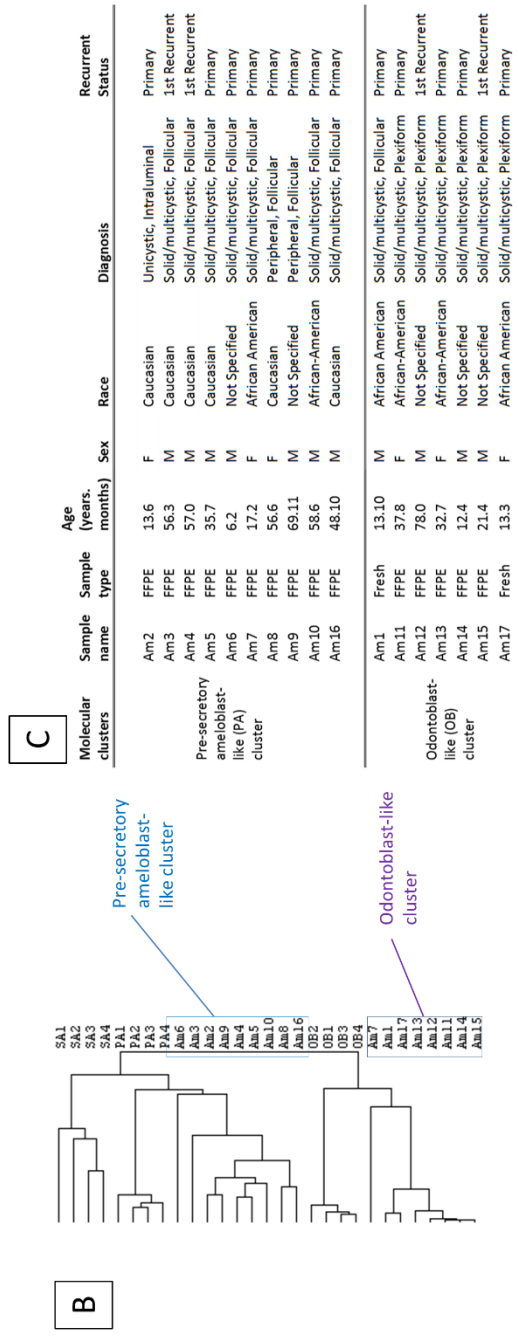


Figure 3.4: Cluster analysis used to determine reference tissue. A – Heat map of the 3 different odontogenic tissue (OB – Odontoblast, PA – Pre-secretory ameloblast, SA – Secretory ameloblast) and the 2 distinct clusters of Ameloblastoma (AM) clustered using the 60 odontogenic epithelium-defining genes. B – Array tree showing grouping of the 2 clusters with odontoblast and pre-secretory ameloblast. C – Demographics and tumor phenotype.

Multiclass analysis

Multiclass analysis was conducted between the 2 tumor clusters and 2 associated normal tissue (pAM, oAM, PA, OB) and differentially expressed genes were carried forward in a cluster analysis (Figure 3.5A). To characterize the transcriptome of ameloblastoma and the 2 molecular sub-clusters, the gene expression data were analyzed in 3 groups. The common tumor cluster describes differential gene expression common in both tumor clusters compared to normal tissue and comprises 2592 genes that were expressed at a higher and lower level in the 2 tumor clusters (pAM, oAM) compared to the 2 normal tissue clusters (OB, PA). The pAM cluster describes differential gene expression unique to that cluster and consists of 1287 genes expressed at higher and lower levels compared to the other 3 groups. The oAM cluster describes differential gene expression unique to oAM tumors and consists of 1516 genes expressed at a higher and lower levels compared to the other 3 groups. The genes with fold changes at FDR < 1% were used for pathway analysis (Data uploaded at http://genomewide.net/public/transcriptome/ameloblastoma/Supplemental_Table_2.xlsx).


A
Heat map of gene expression derived from multiclass analysis
OB1
OB2
OB3
OB4
PA1
PA2
PA3
PA4
oAM1
oAM17
oAM12
oAM11
oAM14
oAM15
oAM13
pAM7
pAM6
pAM3
pAM2
pAM9
pAM5
pAM10
pAM4
pAM8
pAM16
Genes differentially expressed in both clusters of ameloblastoma compared to normal tissue
Genes differentially expressed in Pre-secretory cluster
Genes differentially expressed in Odontoblast cluster

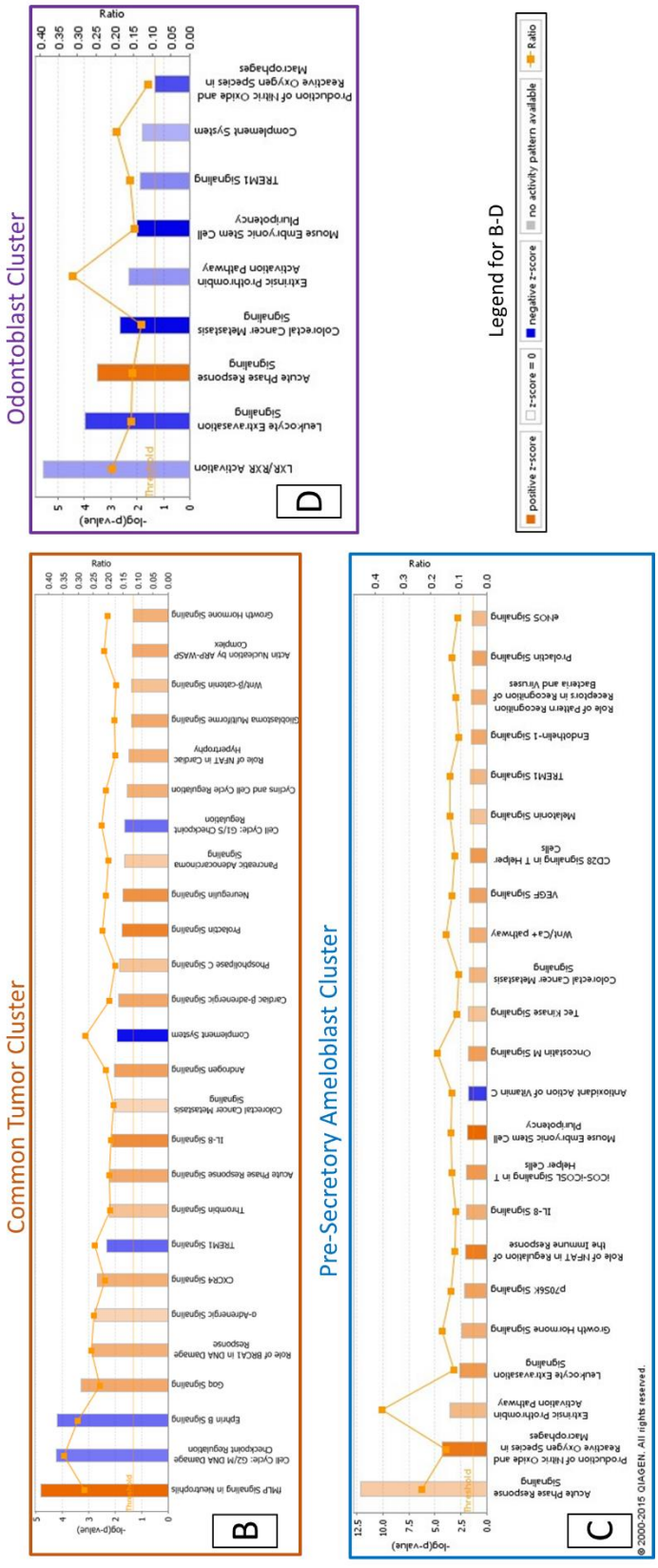

Odontoblast Cluster
Common Tumor Cluster
Pre-Secretory Ameloblast Cluster
Legend for B-D
positive z-score
z-score = 0
negative z-score
no activity pattern available
Ratio
-log(p-value)
Threshold
B
C
D
LXR/RXR Activation
Leukocyte Extravasation Signaling
Acute Phase Response Signaling
Colorectal Cancer Metastasis Signaling
Extrinsic Prothrombin Activation Pathway
Mouse Embryonic Stem Cell Pluripotency
TREM1 Signaling
Complement System
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages
HMGB1 Signaling in Neutrophils
Cell Cycle: G2/M DNA Damage Checkpoint Regulation
Ephrin B Signaling
GaQ Signaling
Role of BRCA1 in DNA Damage Response
α-Adrenergic Signaling
CXCR4 Signaling
TREM1 Signaling
Thrombin Signaling
Acute Phase Response Signaling
IL-8 Signaling
Colorectal Cancer Metastasis Signaling
Androgen Signaling
Complement System
Cardiac β-adrenergic Signaling
Phospholipase C Signaling
Prolactin Signaling
Neuregulin Signaling
Pancreatic Adenocarcinoma Signaling
Cell Cycle: G1/S Checkpoint Regulation
Cyclins and Cell Cycle Regulation
Role of NFAT in Cardiac Hypertrophy
Glioblastoma Multiforme Signaling
Wnt/β-catenin Signaling
Actin Nucleation by ARP-WASP Complex
Growth Hormone Signaling
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages
Leukocyte Extravasation Signaling
Extrinsic Prothrombin Activation Pathway
Growth Hormone Signaling
p70S6K Signaling
Role of NFAT in Regulation of the Immune Response
IL-8 Signaling
ICOS-ICOSL Signaling in T Helper Cells
Mouse Embryonic Stem Cell Pluripotency
Antioxidant Action of Vitamin C
Oncostatin M Signaling
Tec Kinase Signaling
Colorectal Cancer Metastasis Signaling
Wnt/Ca+ pathway
VEGF Signaling
CD28 Signaling in T Helper Cells
Melatonin Signaling
TREM1 Signaling
Endothelin-1 Signaling
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses
Prolactin Signaling
eNOS Signaling
© 2000-2015 QIAGEN. All rights reserved.

Figure 3.5: Multiclass and pathway analysis of the different tumor clusters. A – Heat map of the genes with a FDR < 1% that are differentially expressed in the 4 clusters (OB – odontoblast, PA – pre-Secretory Ameloblast, oAM – odontoblast-like ameloblastoma, pAM – pre-secretory Ameloblast-like ameloblastoma) from a SAM multiclass analysis. The cluster tree on the right showed the 2 distinct clusters of ameloblastoma. Groups of genes with similar expression (identified by colored bars at the bottom of the heat map) were used for pathway analysis for the different clusters of ameloblastoma which were shown in B-D. B – Canonical pathways that are differentially expressed for the Common tumor cluster in IPA. C – Canonical pathways that are differentially expressed for the Pre-secretory ameloblast cluster in IPA. D – Canonical pathways that are differentially expressed for the odontoblast cluster in IPA. The figures (B-D) are generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, www.qiagen.com/ingenuity).

Pathway and gene set enrichment analysis

Ingenuity pathway analysis was used to examine the activated and inhibited canonical pathways for each tumor cluster (Table 3.2).

The common tumor cluster had 21 activated (*z*-score>1) and 5 inhibited (*z*-score <-1) pathways (Figure 3.5B) at *p*-value<0.05. Genes associated with notable biological processes that were differentially expressed in all the ameloblastoma tumors included prevention of damage to cell cycle regulation, cancer pathways, inflammatory pathways and Map kinase related pathways. In addition, GSEA conducted between the common tumor cluster and normal tissues showed that 1860 out of the 2381 genes sets in the "all curated gene sets v4.0" were up-regulated in the common tumor cluster. Nineteen upregulated gene sets were significantly enriched at the nominal *p*-value<0.05. (Table 3.3).

The pre-secretory ameloblast tumor cluster had 22 activated and 1 inhibited pathway below the critical *p*-value threshold (Figure 3.5C). Pathway analysis showed activation in the known cancer pathways, several inflammatory pathways and EGFR pathways.

The odontoblast tumor cluster had 1 activated and 8 inhibited pathways (Figure 3.5D). Several inflammatory pathways were found to be inhibited in this cluster.

Upstream analysis

Upstream regulators that were predicted to be activated or inhibited are listed in Table 3.4. Most of the predicted upstream regulators in the common tumor cluster were transcription

regulators, kinases and cytokines. Specifically, several Map kinase members and inflammatory cytokines were predicted to be activated.

Correlation with The Cancer Genome Atlas

The 2 molecular subtypes of ameloblastoma were compared with the transcriptome of the cancer subtypes in TCGA (Figure 3.6). The analysis did not show any significant correlation of ameloblastoma with any of the 13 subtypes of cancers that are well studied and has established treatment protocols. As ameloblastoma does not seem to correlate molecularly with the cancer subtypes, more investigation into ameloblastoma tumorigenesis is needed for the development of effective treatment.

Table 3.2 **Canonical pathways differentially expressed in ameloblastoma clusters**

Canonical pathways different between the ameloblastoma common tumor cluster and normal cells

Biological process	Ingenuity Canonical Pathways	-log(p-value*)	Ratio	z-score
Inflammatory (Immune) response/cytokine signaling	fMLP Signaling in Neutrophils	4.81	0.28	3.674
	CXCR4 Signaling	2.68	0.21	1.961
	TREM1 Signaling	2.34	0.25	-1.213
	Thrombin Signaling	2.21	0.19	1.512
	Acute Phase Response Signaling	2.20	0.20	2.117
	IL-8 Signaling	2.12	0.19	2.785
	Complement System	1.94	0.28	-1.89
Cell cycle regulation	Cell Cycle: G2/M DNA Damage Checkpoint Regulation	4.23	0.35	-1.069
	Role of BRCA1 in DNA Damage Response	2.88	0.26	1.941
	Cell Cycle: G1/S Checkpoint Regulation	1.63	0.22	-1.155
	Cyclins and Cell Cycle Regulation	1.53	0.21	1.941
Cancer	Colorectal Cancer Metastasis Signaling	2.04	0.18	1
	Pancreatic Adenocarcinoma Signaling	1.63	0.20	1.291
	Glioblastoma Multiforme Signaling	1.40	0.18	2.132
	Wnt/β-catenin Signaling	1.37	0.18	1.528
	Actin Nucleation by ARP-WASP Complex	1.35	0.21	2.111
Map kinase related	Gαq Signaling	3.31	0.23	2.117
	α-Adrenergic Signaling	2.77	0.25	1.069
	Phospholipase C Signaling	1.83	0.18	1.461
	Prolactin Signaling	1.74	0.22	3.051
Receptor tyrosine kinase (RTK)	Ephrin B Signaling	4.21	0.30	-1.155
	Neuregulin Signaling	1.69	0.21	2.668
Nuclear receptor signaling	Androgen Signaling	2.01	0.21	2.333
Cellular growth and proliferation	Growth Hormone Signaling	1.32	0.20	1.941
Others	Cardiac β-adrenergic Signaling	1.87	0.20	2.138
	Role of NFAT in Cardiac Hypertrophy	1.47	0.18	1.961

Canonical pathways different for the ameloblastoma presecretory cluster

Biological process	Ingenuity Canonical Pathways	-log(p-value)	Ratio	z-score
Inflammatory (Immune) response/cytokine signaling	Acute Phase Response Signaling	12.20	0.23	1.225
	Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	4.26	0.15	3.138
	Extrinsic Prothrombin Activation Pathway	3.55	0.38	1.633
	Leukocyte Extravasation Signaling	2.63	0.12	2.324
	Role of NFAT in Regulation of the Immune Response	2.08	0.11	3
	IL-8 Signaling	1.95	0.11	1.886
	iCOS-iCOSL Signaling in T Helper Cells	1.91	0.13	2.333
	Antioxidant Action of Vitamin C	1.80	0.13	-1.897
	CD28 Signaling in T Helper Cells	1.60	0.12	2.333
	TREM1 Signaling	1.55	0.13	1.667
	Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	1.44	0.11	1.897
	eNOS Signaling	1.35	0.10	1.604
Cellular growth and proliferation	Growth Hormone Signaling	2.44	0.16	1.897
	p70S6K Signaling	2.15	0.13	2.138
	Mouse Embryonic Stem Cell Pluripotency	1.84	0.13	3.464
	Oncostatin M Signaling	1.75	0.18	2.236
	VEGF Signaling	1.63	0.12	2.121
Protein tyrosine kinase (PTK)	Tec Kinase Signaling	1.73	0.11	1.387
Cancer	Colorectal Cancer Metastasis Signaling	1.71	0.10	1.706
	Wnt/Ca+ pathway	1.68	0.15	1.89
Others	Melatonin Signaling	1.59	0.13	1.414
	Endothelin-1 Signaling	1.45	0.10	2
	Prolactin Signaling	1.41	0.12	2.121

Canonical pathways different for the ameloblastoma odontoblast cluster

Biological process	Ingenuity Canonical Pathways	-log(p-value)	Ratio	z-score
Inflammatory (Immune) response/cytokine signaling	Leukocyte Extravasation Signaling	3.99	0.16	-2.524
	Acute Phase Response Signaling	3.51	0.16	1.897
	Extrinsic Prothrombin Activation Pathway	2.30	0.31	-1.342
	TREM1 Signaling	1.89	0.16	-1.508
	Complement System	1.80	0.19	-1
	Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	1.33	0.11	-2.236
Nuclear receptor signaling	LXR/RXR Activation	5.57	0.21	-1.225
Cancer	Colorectal Cancer Metastasis Signaling	2.67	0.13	-3.157
Cellular growth and proliferation	Mouse Embryonic Stem Cell Pluripotency	2.01	0.15	-3.207

* p-value < 0.05

Table 3.3

Upregulated gene sets of the ameloblastoma common tumor cluster compared to normal cells

Gene Set	Number of genes	Enrichment Score	Normalized enrichment score	Nominal p-value
GAUSSMANN_MLL_AF4_FUSION_TARGETS_F_DN	12	0.70	1.57	> 0.01
LEE_LIVER_CANCER_E2F1_DN	14	0.58	1.68	> 0.01
BROWNE_HCMV_INFECTION_10HR_UP	24	0.45	1.55	> 0.01
PACHER_TARGETS_OF_IGF1_AND_IGF2_UP	14	0.54	1.52	0.01
TONKS_TARGETS_OF_RUNX1_RUNX1T1_FUSION_SUSTAINDED_IN_ERYTHROCYTE_UP	14	0.61	1.68	0.01
HOUSTIS_ROS	15	0.63	1.47	0.01
NAKAYAMA_SOFT_TISSUE_TUMORS_PCA1_UP	22	0.52	1.72	0.02
SINGH_NFE2L2_TARGETS	10	0.77	1.43	0.02
GERHOLD_ADIPOGENESIS_UP	19	0.49	1.62	0.02
SENESE_HDAC1_AND_HDAC2_TARGETS_UP	76	0.42	1.31	0.03
DACOSTA_UV_RESPONSE_VIA_ERCC3_COMMON_UP	20	0.56	1.40	0.03
KEGG_METABOLISM_OF_XENOBIOTICS_BY_CYTOCHROME_P450	12	0.56	1.59	0.03
LEE_LIVER_CANCER_CIPROFIBRATE_DN	12	0.47	1.52	0.03
KRIGE_AMINO_ACID_DEPRIVATION	13	0.65	1.39	0.03
MANTOVANI_VIRAL_GPCR_SIGNALING_UP	28	0.37	1.42	0.03
SU_LIVER	14	0.54	1.63	0.04
ADDYA_ERYTHROID_DIFFERENTIATION_BY_HEMIN	21	0.54	1.42	0.04
KERLEY_RESPONSE_TO_CISPLATIN_UP	11	0.51	1.57	0.04
SMID_BREAST_CANCER_LUMINAL_A_DN	12	0.62	1.45	0.05

Table 3.4 **Upstream analysis using Ingenuity Pathway Analysis**

Activated and inhibited molecules in Common Tumor Cluster

Molecule Type	Upstream Regulator	Predicted Activation State	Activation z-score	p-value of overlap
transcription regulator	SOX11	Inhibited	-2.16	0.47
	NUPR1	Inhibited	-2.46	0.00
	NEUROG1	Inhibited	-2.18	0.00
	KDM5B	Inhibited	-3.66	0.00
	TP53	Inhibited	-2.37	0.00
	HIF1A	Activated	2.90	0.00
	JUN	Activated	2.55	0.00
	FOXM1	Activated	2.99	0.00
	EZH2	Activated	2.06	0.10
	YAP1	Activated	2.21	0.01
	MYB	Activated	2.39	0.00
kinase	TRIB3	Inhibited	-2.00	0.10
	CDKN1A	Inhibited	-2.60	0.00
	AKT1	Activated	2.65	0.01
	EGFR	Activated	2.20	0.04
	MAPK9	Activated	2.31	0.02
	PIM1	Activated	2.42	0.05
	MAP3K14	Activated	2.21	0.13
	EIF2AK2	Activated	2.43	0.01
	ERBB2	Activated	2.46	0.00
cytokine	CSF2	Activated	2.43	0.01
	TNF	Activated	2.27	0.00
	IL1A	Activated	2.12	0.00
	IL6	Activated	3.33	0.01
	OSM	Activated	2.41	0.00
enzyme	STUB1	Inhibited	-2.24	0.03
	HRAS	Activated	2.16	0.00
ligand-dependent nuclear receptor	PGR	Activated	3.22	0.00
	ESRRA	Activated	2.20	0.07
transporter	SLC29A1	Activated	2.00	0.14
	S100A6	Activated	2.12	0.01
complex	IgG	Inhibited	-2.04	0.00
	NFkB (complex)	Activated	2.35	0.00
	Cg	Activated	2.69	0.00
group	STAT5a/b	Activated	2.00	0.55
	Mek	Activated	2.36	0.05
	Growth hormone	Activated	2.55	0.06
transmembrane receptor	TREM1	Activated	2.82	0.00
growth factor	NRG1	Activated	2.17	0.06
other	RBM5	Inhibited	-2.34	0.07
	UXT	Inhibited	-2.44	0.01
	AHI1	Activated	2.00	0.24
	RABL6	Activated	3.30	0.00
	HSPB2	Activated	2.22	0.00

Activated and inhibited molecules in Presecretory Cluster

Molecule Type	Upstream Regulator	Predicted Activation State	Activation z-score	p-value of overlap
enzyme	MGEA5	Inhibited	-2.04	0.00
ligand-dependent nuclear receptor	PPARG	Activated	2.01	0.00
transcription regulator	NUPR1	Activated	2.11	0.00
growth factor	WISP2	Activated	2.24	0.06

Activated and inhibited molecules in Odontoblast Cluster

Molecule Type	Upstream Regulator	Predicted Activation State	Activation z-score	p-value of overlap
transcription regulator	SOX11	Inhibited	-2.98	0.05
	MDM2	Inhibited	-2.24	0.04
	CEBPA	Inhibited	-2.14	0.00
	NUPR1	Inhibited	-2.47	0.36
	TP53	Inhibited	-2.19	0.01
enzyme	TGM2	Inhibited	-3.59	0.00
	HMOX1	Activated	2.00	0.27
ligand-dependent nuclear receptor	AR	Inhibited	-2.24	0.25
	NR3C1	Inhibited	-2.56	0.17
transporter	S100A6	Activated	2.45	0.01
group	estrogen receptor	Inhibited	-2.28	0.00

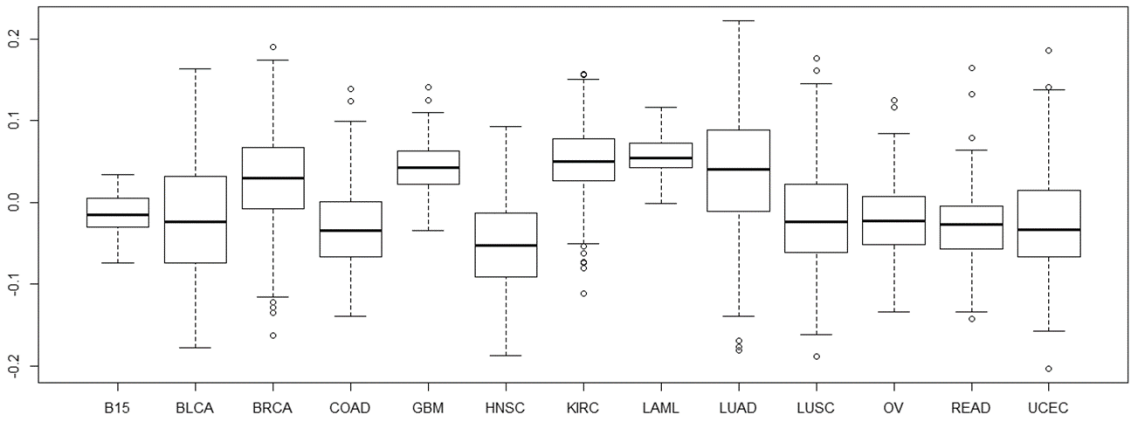


Figure 3.6: Correlation analysis with The Cancer Genome Atlas subtypes. This suggests no significant correlation between ameloblastoma and the 13 cancer subtypes.

Discussion

The major finding of this study was the molecular heterogeneity of ameloblastoma that was strongly associated with its histological subtypes. Gene expression profiles of follicular and plexiform subtypes were more closely related to gene expression profiles of different normal odontogenic tissues and the follicular subtype showed activation of different molecular pathways compared with the plexiform subtype. This new knowledge can serve as a rich hypothesis-generating resource for the study of molecular and phenotypic characteristics of ameloblastoma.

Similar to the present study, Heikinheimo and colleagues found that ameloblastoma gene expression is heterogeneous, and identified 2 distinct tumor clusters with gene expression profile that were most similar to gene expression in the cap/bell stage of tooth development (Heikinheimo et al. 2015b). Using supervised cluster analysis we found that more than half of ameloblastoma samples were most similar in gene expression to pre-secretory ameloblast, similar to those observed in the early cap/bell stage as described by Heikinheimo. It was surprising that using the supervised cluster analysis the remaining ameloblastoma samples associated with mesenchymal derived odontoblasts rather than with the epithelial derived ameloblasts. This finding appears to be driven by differences in inflammatory pathways and was associated with a different histological appearance.

The examination of pathways common to both tumors show that inflammation appears to play an important role in ameloblastoma tumorigenesis and proliferation. Canonical pathway analysis showed that several immune/inflammatory pathways are activated in addition to

several predicted activated upstream cytokines. The association between dysregulated inflammation and cancer progression has been studied extensively (Coussens and Werb 2002). The greater number of pathways activated in the pre-secretory cluster suggests that this process may be more important in the follicular subtype.

Additionally, both tumor clusters revealed that damage to cell cycle regulation pathways play important roles. A key regulator in the cell cycle damage prevention pathways is *TP53* which is also predicted to be inhibited in our upstream analysis. *TP53* is a major tumor suppression gene (Rivlin et al. 2011) and the loss of a tumor suppressor gene activity in ameloblastoma may be important in the tumorigenesis process.

Several canonical pathways involving the MAPK pathways and upstream members were found to be activated in the common tumor cluster. The MAPK pathways have long been considered tumor driver pathways in the pathogenesis of various cancers and also is thought to be important in ameloblastoma tumorigenesis (Brown et al. 2014; Sweeney et al. 2014). Recently, Kurppa and colleagues found *BRAF* gene mutations, specifically V600E mutation, in 63% of ameloblastomas (Kurppa et al. 2014). *BRAF* is in the RAS pathway and MAPK cascade. The V600E mutation in *BRAF* is a promising oncogene target for the anti-neoplastic drug dabrafenib; which was used in conjunction with a MEK inhibitor in a patient with stage 4 ameloblastoma with good results (Kaye et al. 2015). The use of chemotherapeutic agents to reduce tumor size can be very helpful in cases of ameloblastoma requiring extensive surgical margins and major post-surgical reconstruction (Heikinheimo et al. 2015a).

In GSEA, the gene sets with the highest activation scores, were cancer related including "SMID_BREAST_CANCER_LUMINAL_A_DN". Cancer related pathways were also found to be differentially expressed in our pathway analysis. Moreover, breast cancer specific "Role of BRCA1 in DNA Damage Response" pathway was found to be differentially expressed, with the 2 different analyses highlighting similar molecular pathways.

One of the short-comings of this study is that most of the samples were FFPE. Formalin fixing can cause the degradation of RNA (Ravo et al. 2008) and affect the accuracy of microarrays. However, recent studies supported the use of such samples for gene expression analysis (Abdueva et al. 2010) and NanoString has been shown to produce consistent results independent of the sample type (fresh frozen versus FFPE) (Malkov et al. 2009). Genes in our microarray data that had the greatest fold changes showed good correlation with the nanoString expression. In addition, there were no outliers in the cluster analysis among the ameloblastoma cluster analysis indicating consistent results between fresh frozen and FFPE samples.

Strengths of the study included the use of LCM and universal RNA as a means of normalizing between arrays. Ameloblastoma presents with neoplastic epithelial tissue surrounded by stromal tissue making isolation very difficult. As a result, most investigations of ameloblastoma used samples that contain diverse cell populations such as the surrounding stroma which can obscure driver pathways from the actual neoplastic epithelial cells. The use of the universal RNA facilitated the use of the dentome as a comparison and also allows the use of the microarray data by other investigators using universal RNA for normalization.

In conclusion, our study isolated ameloblastoma epithelial and normal odontogenic cells using LCM to identify gene expression profiles and molecular pathways that are potentially important in the tumorigenesis of ameloblastoma. Ameloblastoma showed 2 distinct molecular profiles that were associated with different histological subtypes suggesting they could be receptive to different chemotherapeutic protocols. These results provide a wealth of information that can be used in future experimental and mechanistic studies, involving animal models and new pharmacogenomic approaches.

CHAPTER 4: TRANSCRIPTOME VARIABILITY IN KERATOCYSTIC ODONTOGENIC TUMOR SUGGESTS DISTINCT MOLECULAR SUBTYPES

Shijia Hu[1,2], Kimon Divaris[1,3], Joel Parker[4], Ricardo Padilla[5], Valerie Murrah[5], John Timothy Wright[1]

[1]Pediatric Dentistry, University of North Carolina, Chapel Hill, NC, USA

[2]Faculty of Dentistry, National University of Singapore, Singapore

[3]Epidemiology, Gillings School of Global Public Health, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA

[4]Cancer Genetics, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA

[5]Diagnostic Sciences, School of Dentistry, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA

Abstract

Keratocystic Odontogenic Tumor (KCOT) is a locally aggressive developmental cystic neoplasm thought to arise from the odontogenic epithelium. A high recurrence rate of up to 30% has been found following conservative treatment. Aggressive tumor resection can lead to the need for extensive reconstructive surgery, resulting in significant morbidity and impacting quality of life. Most research has focused on candidate-genes with a handful of studies employing whole transcriptome approaches. There is also the question of which reference tissue is most biologically-relevant. This study characterizes the transcriptome of KCOT using whole genome microarray and compare it with gene expression of different odontogenic tissues ("dentome"). Laser capture microdissection was used to isolate the neoplastic epithelial tissue in 20 cases. KCOT gene expression was compared with the "dentome" and relevant pathways were examined. Cluster analysis revealed 2 distinct molecular subtypes of KCOT. Several inflammatory pathways were activated in both subtypes. The AKT pathway was activated in one subtype while MAP kinase pathway was activated in the other. Additionally, PTCH1 expression was downregulated in both clusters suggesting involvement in KCOT tumorigenesis. In conclusion, this study provides new insights into the transcriptome of KCOT and highlights pathways that could be of diagnostic and prognostic value.

Introduction

In 2005, the World Health Organization reclassified the Keratocystic Odontogenic Tumor (KCOT) from a cyst to a tumor to better reflect its neoplastic nature (Barnes L 2005). KCOT is a benign but locally aggressive developmental cystic neoplasm historically thought to arise from the odontogenic epithelium (Partridge and Towers 1987) and frequently is associated with the follicle of unerupted teeth. During the last decade there has been a resurgence of interest in and efforts to understand tumorigenesis of KCOTs with the ultimate goal of developing better diagnostic and treatment approaches.

Despite advancements in antineoplastic therapies, surgical intervention remains the treatment of choice for KCOTs. A high recurrence rate of up to 30% has been found following conservative treatment such as enucleation and curettage, creating a challenge in determining the optimal extent of surgical resection. Nevoid basal cell carcinoma syndrome (NBCCS) is a disorder that presents with multifocal KCOT; these cases have even higher recurrence rates when compared to isolated unifocal cases not associated with NBCCS (Johnson et al. 2013). Aggressive tumor resection can lead to the need for extensive reconstructive surgery and rehabilitation for patients with KCOTs, causing significant morbidity and negatively impact their quality of life.

Little is known about the molecular profile of KCOTs; however, recent reports have shed some light on molecular pathways driving the tumorigenesis of KCOT. For example, the PTCH1 pathway has been the primary focus of candidate-gene studies due to its association with the NBCCS (Guo et al. 2013; Kadlub et al. 2013; Qu et al. 2015). Other potential aberrant pathways

including SHH (Ren et al. 2012), WNT signaling (Hakim et al. 2011), p53 (Alur et al. 2014) and matrix metalloproteinases (Amm et al. 2014) are of interest. These candidate-gene studies aid in the advancement of precision medicine (Collins and Varmus 2015) for improved diagnostics for predicting prognosis and managing KCOT. However, only 2 published studies to date have used next-generation approaches including genomics and transcriptomics methods to characterize the tumorigenesis of KCOT. In a 2007 report, Heikinheimo et al used a 588 cancer-related human cDNA array (Heikinheimo et al. 2007) for the study of KCOT and more recently they employed a more extensive whole genome array (Heikinheimo et al. 2015). These studies provided valuable initial insights into the molecular basis of KCOT development.

Unraveling the molecular basis of KCOT can help identify altered pathways involved in its tumorigenesis, as well as discover potentially informative cell markers for its diagnosis and treatment. Thus, the objective of this study was to characterize the transcriptome of KCOTs using a whole genome microarray and compare it with the expression profile of a biologically-relevant odontogenic tissue (referred to as “dentome”) (Hu et al. 2015).

Materials and methods

Patient recruitment and sample collection

The study was conducted in accordance with approved human subject research guidelines and was approved by the local institutional review board and the ethics committee of the University of North Carolina-Chapel Hill. Patients diagnosed with isolated KCOT and scheduled for tumor removal at the Department of Oral and Maxillofacial Surgery at the University of North Carolina-Chapel Hill between 2005 and 2008 formed the study sample. All

participants (N=20) provided written informed consent. Diagnosis of KCOT was confirmed by a board-certified oral pathologist and another investigator using the 2005 WHO Histologic Classification of Odontogenic Tumors. Patients presenting with NBCCS were excluded from this study. Six fresh frozen samples were obtained during the surgical procedures and 14 formalin-fixed paraffin-embedded (FFPE) samples were obtained from the Department of Oral and Maxillofacial Pathology at UNC-Chapel Hill after their histological diagnosis was confirmed. Potential associations of gene expression with participants' demographic characteristics including gender, age, race, and tumor recurrence were examined using bivariate methods (Fisher's exact test) and a conventional $p<0.05$ statistical significance criterion.

Sample preparation

All 20 samples were decalcified for 1-4 weeks using a solution containing water, hydrochloric acid, EDTA, tetrasodium tartrate and potassium sodium tartrate from Richard Allan Scientifics (ThermoScientific, Grand Island, NY, USA). The samples were then sectioned at -35^{o}C at a thickness of 7 microns and lightly stained with hematoxylin and eosin. These slides were then used for the microdissection of target tumor epithelial cells.

Laser capture microdissection (LCM) (Sun et al. 2012) is a technique that allows the isolation of tissue that is 1 cell thick. KCOT is a cystic tumor containing loose surrounding stroma and a keratotic layer that arises from a neoplastic basal layer with numerous satellite cysts (Figure 4.1A and 4.1B). LCM allowed isolation of discrete areas of the basal neoplastic layer (Fig. 4.1C and 4.1D) for the examination of a purer cell population. The technique has been successfully used in microarray studies for both fresh (Hayashi et al. 2010) and formalin

fixed samples (Salmon et al. 2012). The AutoPix™ automated system (Arcturus Engineering, Santa Clara, CA, USA) was used for LCM. The captured cells were pooled for each sample and placed in RNA extraction buffer. RNA was isolated from the tumor cells using the PicoPure RNA Isolation kit (Arcturus Bioscience, Santa Clara, CA, USA). The quality and yield of total RNA were assessed using the Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA).

Microarray analysis

Agilent whole genome human oligonucleotide microarrays which contain 44 thousand 60-mer oligonucleotides representing over 41 thousand probes and transcripts was used for whole transcriptome analysis. The Human Universal Reference RNA from Stratagene (Santa Clara, CA, USA) served as a control to standardize hybridization levels between microarrays and was coupled to Cy5. Two hundred nanograms of total RNA from the KCOT samples were converted into labeled cRNA with nucleotides coupled to fluorescent dye Cy3 using the Low RNA Input Linear Amplification Kit (Agilent Technologies, Palo Alto, CA). The Cy3-labeled cRNA (1.65 ng) from each sample and Cy5 coupled universal reference RNA was hybridized to whole genome array formatted chips. The hybridized array was washed, scanned and data extracted from the scanned image using Feature Extraction version 9.5 (Agilent Technologies, Palo Alto, CA).

The microarray data were submitted to the Gene Expression Omnibus (GEO) microarray database (accession number GSE68532).

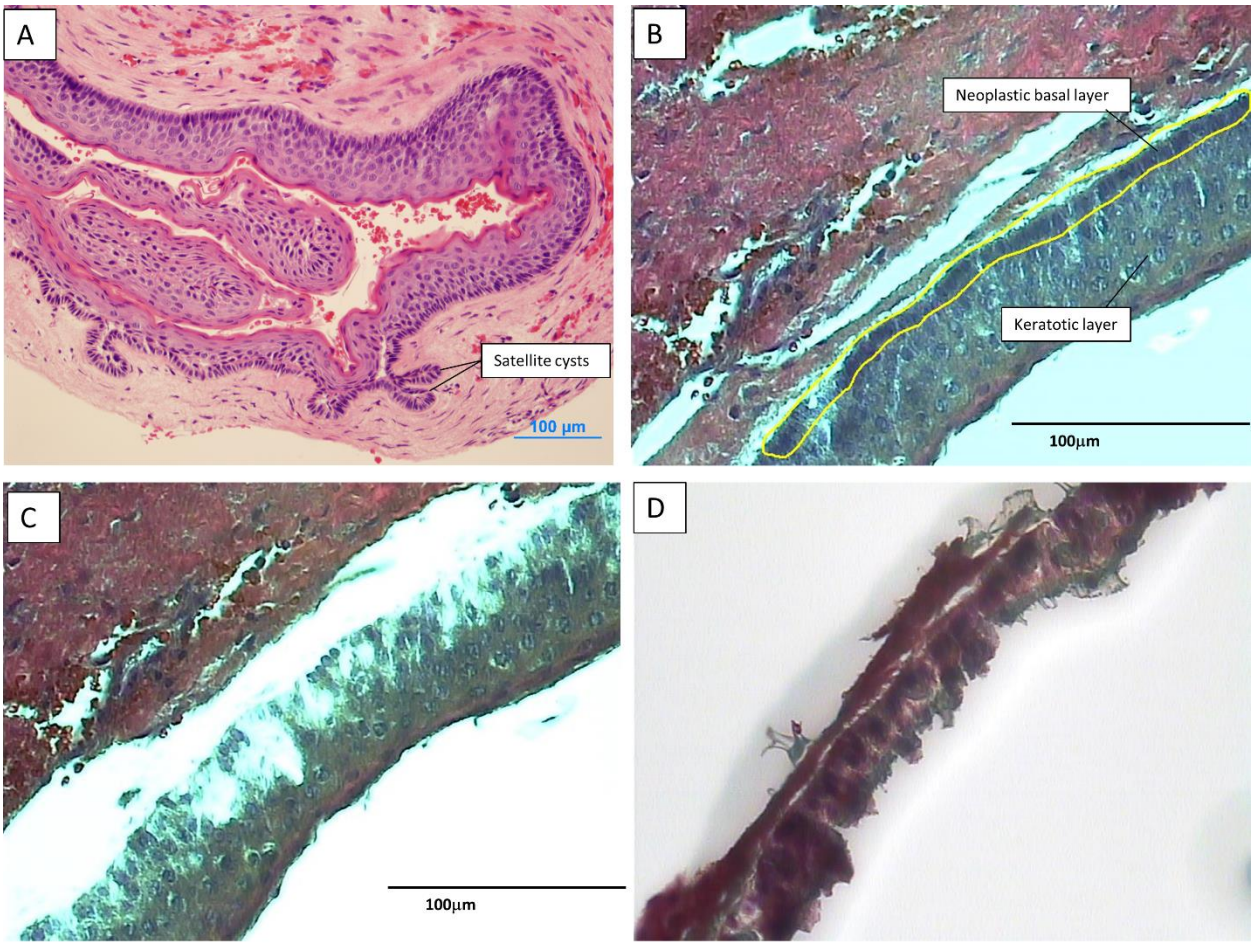


Figure 4.1: The micrographs show the laser capture of the epithelial portion of a KCOT sample. A – light micrograph of KCOT at 4X showing formation of satellite cysts, B – laser capture outline of epithelial cells showing target basal epithelial cells, C – remnants of the stroma tissue after LCM and D – captured cells on Capsure cap. Scale bar: 100 µm.

Determination of reference tissue and bioinformatics analysis

Human odontogenic tissue whole transcriptome data, also known as the “dentome” was used to help identify the most appropriate (based on being the most similar) odontogenic reference tissue to be used for comparison with the tumor samples. As part of earlier work conducted by our group, multiple embryonic teeth were used to obtain 4 samples each of human odontoblasts, pre-secretory ameloblasts and secretory ameloblasts using microdissection to isolate discrete samples of the different cell types and development stages from which gene expression data were obtained. (Gene Expression Omnibus microarray database accession number GSE63289) (Hu et al. 2015). Multiclass significance analysis of microarrays 4.0 (SAM) was conducted between the 3 types of normal odontogenic tissue and 60 genes were found to be expressed differentially at a FDR of < 20%. A cluster analysis, using these 60 genes that differentiated the normal odontogenic tissue, was conducted using Cluster 3.0 between the KCOT samples and the normal odontogenic tissues. Results of this analysis between normal odontogenic tissues and KCOT were then visualized using Java TreeView-1.1.6r2 to establish the normal tissue with a gene expression profile most similar to KCOT.

Gene set enrichment analysis (GSEA) (Subramanian et al. 2005) was conducted with GSEA v2.1.0 from the Broad institute (Cambridge, MA, USA) using the Molecular Signatures Database (MSigDB) curated gene set(Liberzon et al. 2011) “all curated gene sets v4.0” and 1000 permutations. Differential expression between the tumor and comparison normal tissue was calculated with SAM. Genes of interest were carried forward and interrogated using the Ingenuity pathway analysis (IPA), including canonical pathway and upstream analyses.

Upstream analysis is a model that predicts activation or inhibition of upstream regulators based on the expression levels of downstream molecules. The model has the advantage of detecting genes with possible gain-of-function mutation that did not show an increase level of expression.

Microarray data validation via NanoString analysis

The NanoString nCounter system (Seattle, WA, USA) was used to validate the microarray gene expression data. The nCounter system is based on a direct multiplexed measurement of gene expression and offers high levels of precision and sensitivity (Malkov et al. 2009). The nCounter Human Cancer Reference codeset (http://www.nanostring.com/products/gene_expression_panels) which profiles 230 cancer-related human genes and 6 internal reference genes was used. In this analysis, the microarray data of a subset of 3 KCOT and 2 normal secretory ameloblast samples were examined. Hybridization reactions were performed using 50 ng of total RNA with reporter and capture probes, in accordance to the manufacturer's instructions. The NanoString nCounter digital count readings were extracted, normalized and analyzed using nSolver v2.5 to obtain fold-change values between KCOT and normal tissue. These data were then compared to the corresponding KCOT and normal tissue microarray fold change values. A scatterplot was constructed using the 20 most upregulated and downregulated genes in the nanoString analysis between KCOT and secretory ameloblast with differential expression from the microarray data (Table 4.1). The 2 sets of data showed good correlation (r-0.65) between the microarray and nanoString data for genes showing the highest fold change differences (Figure 4.2).

Table 4.1: Nano string validation of microarray

Top 20 upregulated genes in KCOT compared to secretory ameloblast		
Gene name	Nanostring fold change	Sam fold change
WT1	7.68	1.11
IL1B	7.01	-1.71
NTRK1	6.03	1.25
IL1A	5.80	1.54
CASP10	5.65	-1.22
AKT1	5.23	1.25
CYP1A1	5.23	1.37
FLT3	5.23	1.25
GATA1	5.23	1.03
TERT	5.23	1.21
CSF3	5.19	1.03
AREG	4.88	-2.00
TNF	4.59	1.55
PTGS2	4.55	-1.12
FGR	4.53	1.51
THPO	4.53	1.26
ATM	4.53	1.03
ABCB1	4.28	1.28
RET	4.14	1.26
EGF	4.10	1.19
Top 20 downregulated genes in KCOT compared to secretory ameloblast		
Gene name	Nanostring fold change	sam fold change
SPP1	-10.26	-2.49
MMP2	-6.17	-2.44
PDGFRA	-6.12	-1.77
DLC1	-4.65	-1.95
FRZB	-4.33	-1.32
GAS1	-4.25	-2.04
IGFBP6	-4.25	1.66
CD34	-3.96	-1.37
ABL1	-3.57	-1.87
KDR	-3.37	-1.14
FGF2	-2.84	-1.51
LMO2	-2.70	1.21
ETS1	-2.60	-1.50
TEK	-2.58	-2.23
FYN	-2.29	-1.16
CCND2	-2.23	-1.26
GNAS	-2.13	-1.15
TGFBR3	-2.08	-1.03
FGFR1	-2.07	-1.44
CAV1	-2.01	-4.38
APC	-1.73	-1.44

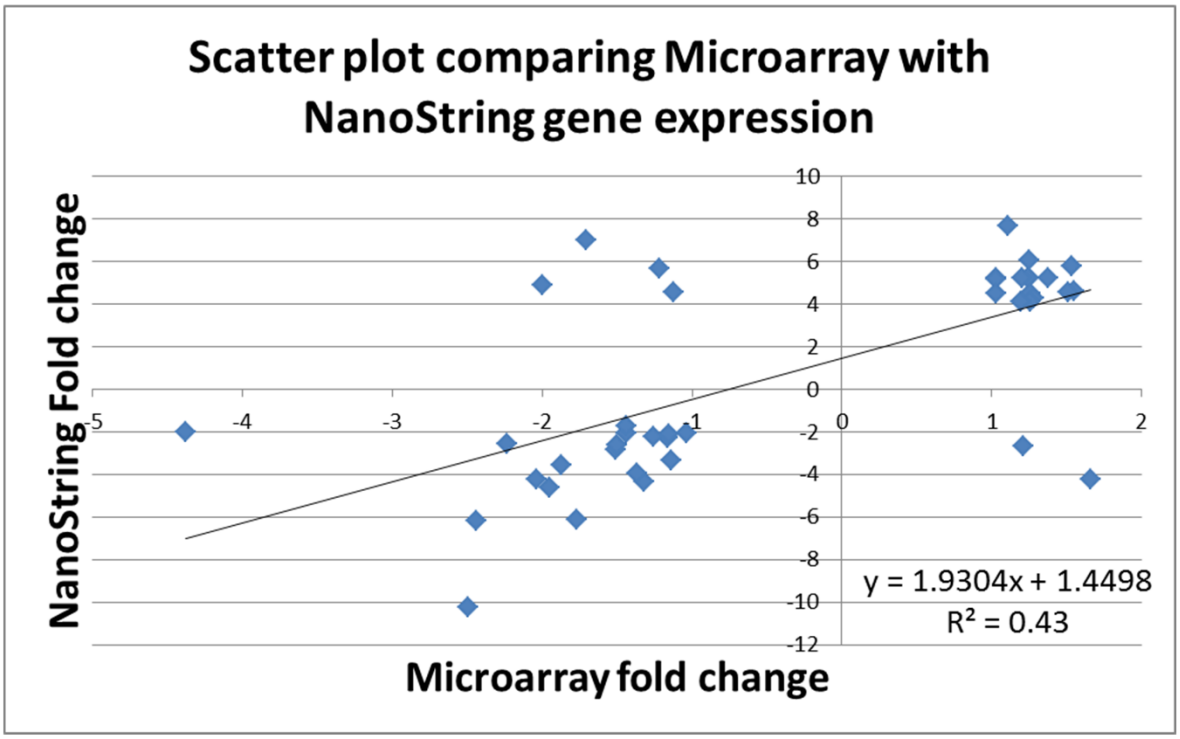


Figure 4.2: Scatterplot of Microarray differential expression versus NanoString differential expression. Microarray fold change is on the x-axis while NanoString fold change is on the y-axis.

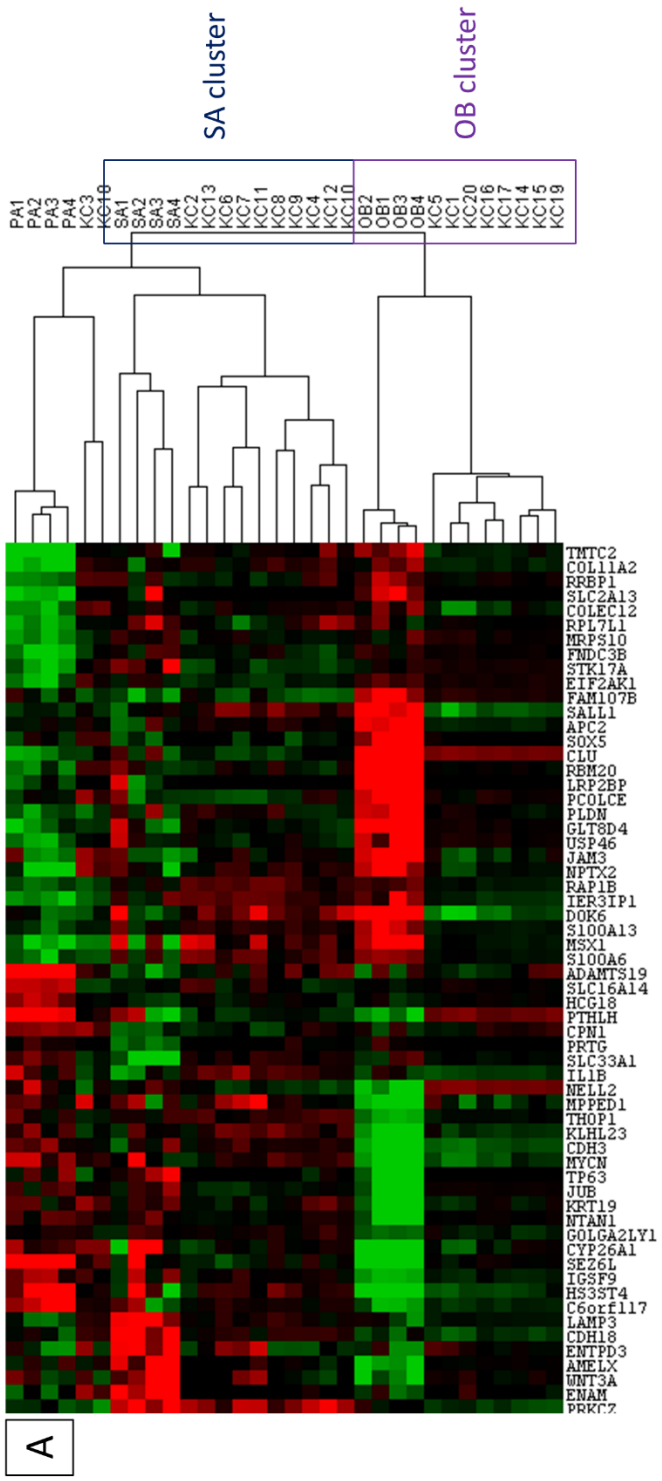
Results

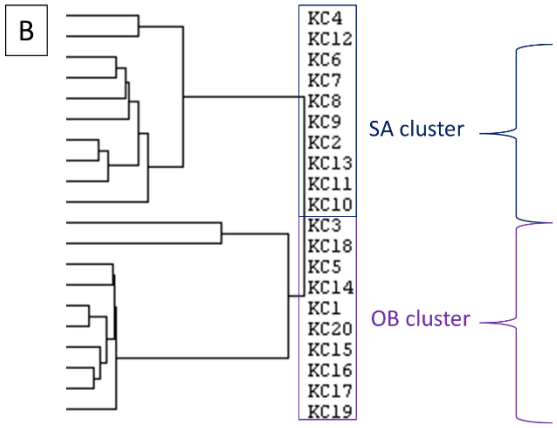
RNA extracted from the LCM samples had a 260/280 ratio of between 2.17 and 2.20 and a yield of between 914ng to 1225ng per sample, demonstrating reasonable quality and yield.

Reference tissue and multiclass analysis

When cluster analysis was conducted with the 60 odontogenic tissue defining genes, the KCOT samples appeared to cluster into 2 distinct molecular subtypes (Figure 4.3A). One of the subtypes clustered with secretory ameloblast (SA) while the other clustered with odontoblast (OB), as such the clusters were designated as secretory ameloblast-like KCOT (sKC) and odontoblast-like KCOT (oKC) respectively. In order to confirm the finding of 2 distinct subtypes, an unsupervised cluster analysis of all the genes were performed with the 20 KCOT samples. The tumor samples separated into 2 distinct subtypes as shown in the cluster tree (Figure 4.3B). The Fisher exact test did not show significant association of the clusters (Figure 4.3C) with race, recurrence, gender, tumor type or age.

The discovery of 2 distinct molecular clusters of KCOT did not allow the designation of a single normal odontogenic tissue for gene expression comparison, as such, a multiclass analysis approach was employed. SAM multiclass analysis of genes differentially expressed at a FDR < 1% was conducted between the 2 KCOT clusters (sKC, oKC) and 2 associated normal tissues (SA, OB) (Figure 4.4A) (Data uploaded at http://genomewide.net/public/transcriptome/kcot/Supplemental_Table_2.xlsx). The gene expression data were analyzed in 3 clusters. The “common tumor cluster” consists of 3166 genes which were differentially expressed in the 2 tumor clusters (sKC, oKC) compared to the 2

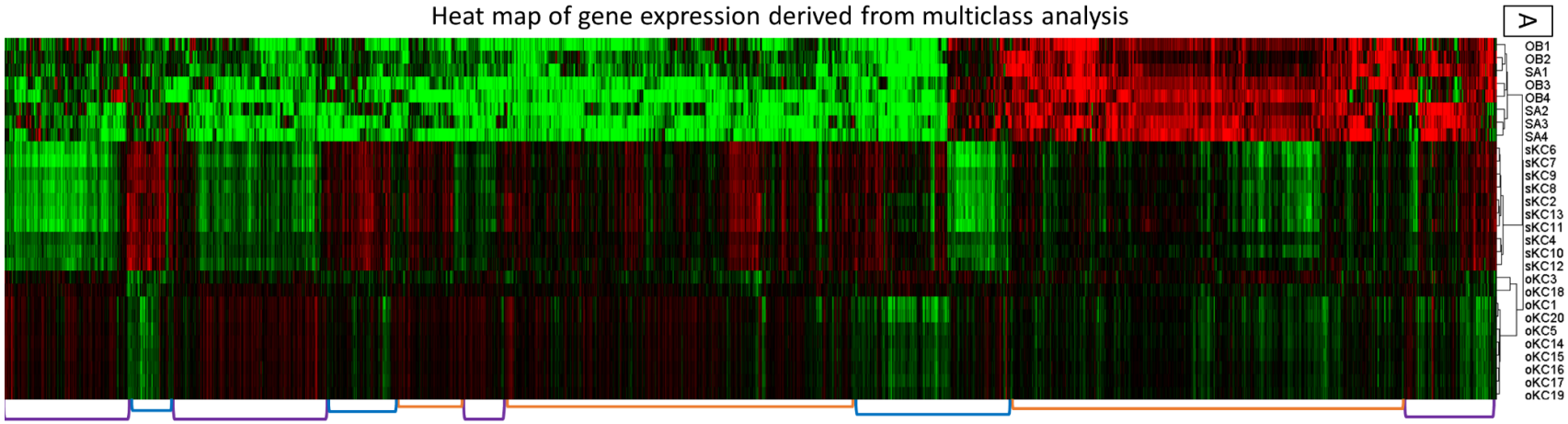

SA cluster
OB cluster
PA1
PA2
PA3
PA4
KC3
KC18
SA1
SA2
SA3
SA4
KC2
KC13
KC6
KC7
KC11
KC8
KC9
KC4
KC12
KC10
OB2
OB1
OB3
OB4
KC5
KC1
KC20
KC16
KC17
KC14
KC15
KC19
TMTC2
COL11A2
RRBP1
SLC2A13
COLEC12
RPL7L1
MRPS10
FNDC3B
STK17A
EIF2AK1
FAM107B
SALL1
APC2
SOX5
CLU
RBM20
LRP2BP
PCOLCE
PLDN
GLT8D4
USP46
JAM3
NPTX2
RAP1B
IER3IP1
DOK6
S100A13
MSX1
S100A6
ADAMTS19
SLC16A14
HCG18
PTHLH
CPN1
PRTG
SLC33A1
IL1B
NELL2
MPPED1
THOP1
KLHL23
CDH3
MYCN
TP63
JUB
KRT19
NTAN1
GOLGA2LY1
CYP26A1
SEZ6L
IGSF9
HS3ST4
C6orf117
LAMP3
CDH18
ENTPD3
AMELX
WNT3A
ENAM
PRKCZ
A



C Demographics

Sample Name	Sample type	Age (years.months)	Sex	Race	Recurrent tumor
KC4	Fresh	15.5	M	Caucasian	No
KC12	FFPE	49.11	F	Caucasian	Yes
KC6	Fresh	16.0	F	African American	Yes
KC7	FFPE	17.2	M	African-American	No
KC8	FFPE	43.5	F	African-American	No
KC9	FFPE	50.5	M	-	No
KC2	Fresh	18.4	F	African American	No
KC13	FFPE	27.3	M	Caucasian	No
KC11	FFPE	13.0	F	African-American	No
KC10	FFPE	18.8	F	Caucasian	No
KC3	Fresh	16.5	M	Caucasian	No
KC18	FFPE	87.7	M	-	No
KC5	Fresh	16.10	F	African American	Yes
KC14	FFPE	67.8	F	-	No
KC1	Fresh	17.4	F	African American	No
KC20	FFPE	16.5	F	African-American	No
KC15	FFPE	52.9	M	Caucasian	No
KC16	FFPE	46.0	M	-	Yes
KC17	FFPE	15.5	M	Caucasian	No
KC19	FFPE	47.6	F	African-American	No

Figure 4.3: Cluster analysis to determine reference tissue. A – Heat map of the 3 different odontogenic tissues (OB – Odontoblast, PA – Pre-secretory ameloblast, SA – Secretory ameloblast) and KCOT (KC) clustered using the 60 odontogenic tissue defining genes. B – Array tree showing 2 clusters of the 2 KCOT molecular subtypes (odontoblast and secretory ameloblast clusters). C – Demographics and tumor recurrence



Legend for Figure A

- Genes differentially expressed in both clusters of KCOT compared to normal tissue
- Genes differentially expressed in Secretory cluster
- Genes differentially expressed in Odontoblast cluster

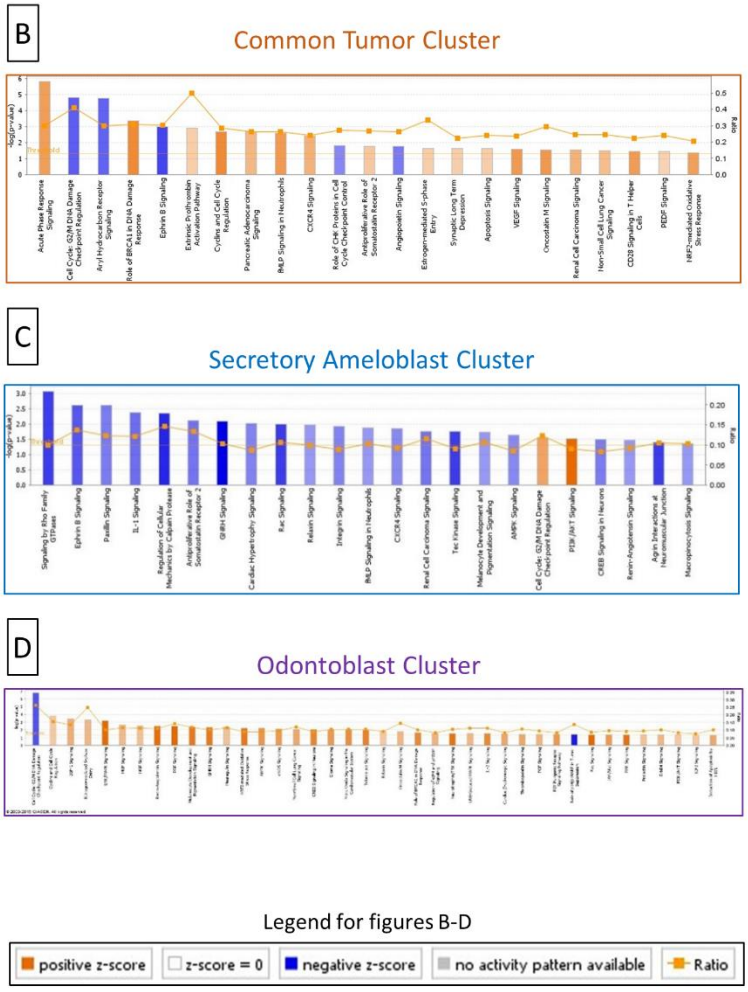


Figure 4.4: Multiclass analysis and pathway analysis of the different tumor clusters. A – Heat map of the genes with a FDR < 1% that are differentially expressed in the 4 clusters (OB – odontoblast, SA – Secretory Ameloblast, oKC – odontoblast-like KCOT, sKC – secretory ameloblast-like KCOT) from a SAM multiclass analysis. The color bars on the right of the heat map shows the groups of genes used to define each cluster's differential gene expression. B – Canonical pathways that are differentially expressed for the common tumor cluster in IPA. C – Canonical pathways that are differentially expressed for the secretory ameloblast cluster in IPA. D – Canonical pathways that are differentially expressed for the odontoblast cluster in IPA. The figures (B-D) are generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA, QIAGEN Redwood City, www.qiagen.com/ingenuity).

normal tissue clusters (OB, SA). The “secretory ameloblast cluster” consists of the 985 differentially expressed genes in the sKC cluster compared to the other 3 groups. The “odontoblast cluster” consists of 902 differentially expressed genes in the oKC cluster compared to the other 3 groups. The fold changes were then used for pathway and GSEA analyses.

Canonical pathways and gene set enrichment analysis

Ingenuity pathway analysis was used to examine the activated and inhibited canonical pathways for each tumor cluster. Canonical pathways with an absolute z-score > 1 and *p*-value < 0.05 are presented in Table 4.2 showing the pathways exhibiting the greatest activation/inhibition that were significantly different.

The common tumor cluster had 18 activated and 5 inhibited pathways (Figure 4.4B). There were 7 pathways involved in cell cycle regulation. Notably, 6 Inflammatory (Immune) response/cytokine signaling pathways were found to be differentially expressed in addition to 5 Cellular growth, proliferation and apoptosis pathways and 3 Cancer pathways. GSEA conducted between the common tumor cluster and normal tissues showed that 14 gene sets were significantly enriched (nominal p-value < 0.05) and 1 gene set was inhibited (nominal p-value < 0.05) (Table 4.3). Specifically, the Sonic hedgehog signaling (SHH/PTCH1) pathway was not found to be significantly differentially expressed between normal and tumor tissue. Closer examination of the SHH/PTCH1 pathway showed downregulation of *PTCH1* and its downstream target *GLI* without changes to *SHH* (Figure 4.5).

Table 4.2 – IPA canonical pathway analysis between the different KCOT molecular clusters and reference odontogenic tissue

Canonical pathways differentially enriched (p<0.05) between KCOT "common tumor cluster" and normal cells

Biological process	Ingenuity Canonical Pathways (z-score)			
Cell cycle regulation	Cell Cycle: G2/M DNA Damage Checkpoint Regulation	(-1.70)	Role of CHK Proteins in Cell Cycle Checkpoint Control	(-1.27)
	Aryl Hydrocarbon Receptor Signaling	(-1.63)	Antiproliferative Role of Somatostatin Receptor 2	(1.27)
	Role of BRCA1 in DNA Damage Response	(2.67)	Estrogen-mediated S-phase Entry	(1.13)
	Cyclins and Cell Cycle Regulation	(2.67)		
Inflammatory (Immune) response/cytokine signaling	Acute Phase Response Signaling	(2.20)	CXCR4 Signaling	(1.77)
	Extrinsic Prothrombin Activation Pathway	(1.13)	CD28 Signaling in T Helper Cells	(2.24)
	fMLP Signaling in Neutrophils	(2.29)	NRF2-mediated Oxidative Stress Response	(2.18)
Cellular growth, proliferation and apoptosis	Angiopoietin Signaling	(-1.39)	Oncostatin M Signaling	(2.33)
	Apoptosis Signaling	(1.09)	PEDF Signaling	(1.00)
	VEGF Signaling	(2.32)		
Cancer	Renal Cell Carcinoma Signaling	(1.73)	Pancreatic Adenocarcinoma Signaling	(1.61)
	Non-Small Cell Lung Cancer Signaling	(1.51)		
Tyrosine kinase signaling	Ephrin B Signaling	(-1.94)		
Others	Synaptic Long Term Depression	(1.10)		

Canonical pathways differentially enriched (p<0.05) between KCOT "secretory ameloblast cluster" and normal cells

Biological process	Ingenuity Canonical Pathways (z-score)			
Cellular growth, proliferation and apoptosis	Paxillin Signaling	(-1.27)	Melanocyte Development and Pigmentation Signaling	(-1.00)
	Regulation of Cellular Mechanics by Calpain Protease	(-2.12)	PI3K/AKT Signaling	(1.67)
	Relaxin Signaling	(-1.00)	CREB Signaling in Neurons	(-1.27)
Inflammatory (Immune) response/cytokine signaling	IL-1 Signaling	(-1.41)	CXCR4 Signaling	(-1.16)
	fMLP Signaling in Neutrophils	(-1.00)	Macropinocytosis Signaling	(-1.00)
Cell cycle regulation	Antiproliferative Role of Somatostatin Receptor 2	(-1.34)	Integrin Signaling	(-1.21)
	Cell Cycle: G2/M DNA Damage Checkpoint Regulation	(1.00)		
G-protein signaling	Signaling by Rho Family GTPases	(-1.89)	Rac Signaling	(-1.90)
Tyrosine kinase signaling	Ephrin B Signaling	(-1.63)	Tec Kinase Signaling	(-1.90)
Cancer	Renal Cell Carcinoma Signaling	(-1.34)		
Others	GNRH Signaling	(-2.50)	Renin-Angiotensin Signaling	(-1.00)
	Cardiac Hypertrophy Signaling	(-1.21)	Agrin Interactions at Neuromuscular Junction	(-1.89)

Canonical pathways differentially enriched (p<0.05) between KCOT "odontoblast cluster" and normal cells

Biological process	Ingenuity Canonical Pathways (z-score)			
Cellular growth, proliferation and apoptosis	IGF-1 Signaling	(1.51)	Oncostatin M Signaling	(1.34)
	HGF Signaling	(1.27)	Thrombopoietin Signaling	(1.63)
	NGF Signaling	(1.16)	FGF Signaling	(1.41)
	Melanocyte Development and Pigmentation Signaling	(2.53)	JAK/Stat Signaling	(1.89)
	AMPK Signaling	(1.94)	ErbB4 Signaling	(1.34)
	CREB Signaling in Neurons	(2.31)	PI3K/AKT Signaling	(1.00)
	Relaxin Signaling	(1.00)	EIF2 Signaling	(1.00)
	Regulation of eIF4 and p70S6K Signaling	(1.41)	Induction of Apoptosis by HIV1	(1.63)
Cell cycle regulation	Cell Cycle: G2/M DNA Damage Checkpoint Regulation	(-1.51)	Estrogen-mediated S-phase Entry	(1.34)
	Cyclins and Cell Cycle Regulation	(1.00)	Role of BRCA1 in DNA Damage Response	(2.24)
Cancer	Non-Small Cell Lung Cancer Signaling	(1.63)	Telomerase Signaling	(2.33)
	Glioma Signaling	(1.89)	Role of p14/p19ARF in Tumor Suppression	(-2.00)
Inflammatory (Immune) response/cytokine signaling	NRF2-mediated Oxidative Stress Response	(2.53)	IL-2 Signaling	(2.24)
	eNOS Signaling	(1.89)		
MAP kinase related	ERK/MAPK Signaling	(2.83)	UVB-Induced MAPK Signaling	(1.34)
	EGF Signaling	(2.83)		
Tyrosine kinase signaling	Neuregulin Signaling	(1.13)	Neurotrophin/TRK Signaling	(2.65)
G-protein signaling	Rac Signaling	(2.83)	PAK Signaling	(2.65)
Others	Renin-Angiotensin Signaling	(2.71)	Cardiac β-adrenergic Signaling	(1.67)
	GNRH Signaling	(2.50)	P2Y Purigenic Receptor Signaling Pathway	(1.90)
	Nitric Oxide Signaling in the Cardiovascular System	(2.12)	Prolactin Signaling	(1.13)

Table 4.3 – Results of Gene set enrichment analysis (GSEA) conducted between KCOT and reference cells using "all curated gene sets v4.0"

Gene sets enriched in KCOT "common tumor cluster" compared to normal cells				
Gene Set	Number of genes	Enrichment Score	Normalized enrichment score	Nominal p-value
SEKI_INFLAMMATORY_RESPONSE_LPS_UP	27	0.57	1.62	0.01
REACTOME_PLATELET_AGGREGATION_PLUG_FORMATION	10	0.66	1.66	0.01
JI_METASTASIS_REPRESSED_BY_STK11	10	0.70	1.42	0.02
HAHTOLA_SEZARY_SYNDROM_UP	38	0.51	1.39	0.02
SMID_BREAST_CANCER_RELAPSE_IN_BRAIN_UP	18	0.50	1.57	0.02
TONKS_TARGETS_OF_RUNX1_RUNX1T1_FUSION_GRANULOCYTE_UP	19	0.41	1.45	0.02
SMID_BREAST_CANCER_LUMINAL_A_DN	11	0.67	1.42	0.03
MOLENAAR_TARGETS_OF_CCND1_AND_CDK4_DN	34	0.68	1.40	0.03
BENPORATH_ES_CORE_NINE_CORRELATED	28	0.55	1.30	0.03
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_6HR_UP	32	0.42	1.39	0.04
ZHOU_TNF_SIGNALING_30MIN	11	0.57	1.46	0.04
SARTIPY_NORMAL_AT_INSULIN_RESISTANCE_UP	10	0.71	1.33	0.04
WAMUNYOKOLI_OVARIAN_CANCER_GRADES_1_2_UP	46	0.49	1.34	0.04
PID_AURORA_A_PATHWAY	13	0.64	1.55	0.05
FLOTHO_PEDIATRIC_ALL_THERAPY_RESPONSE_UP	10	-0.61	-1.56	0.03

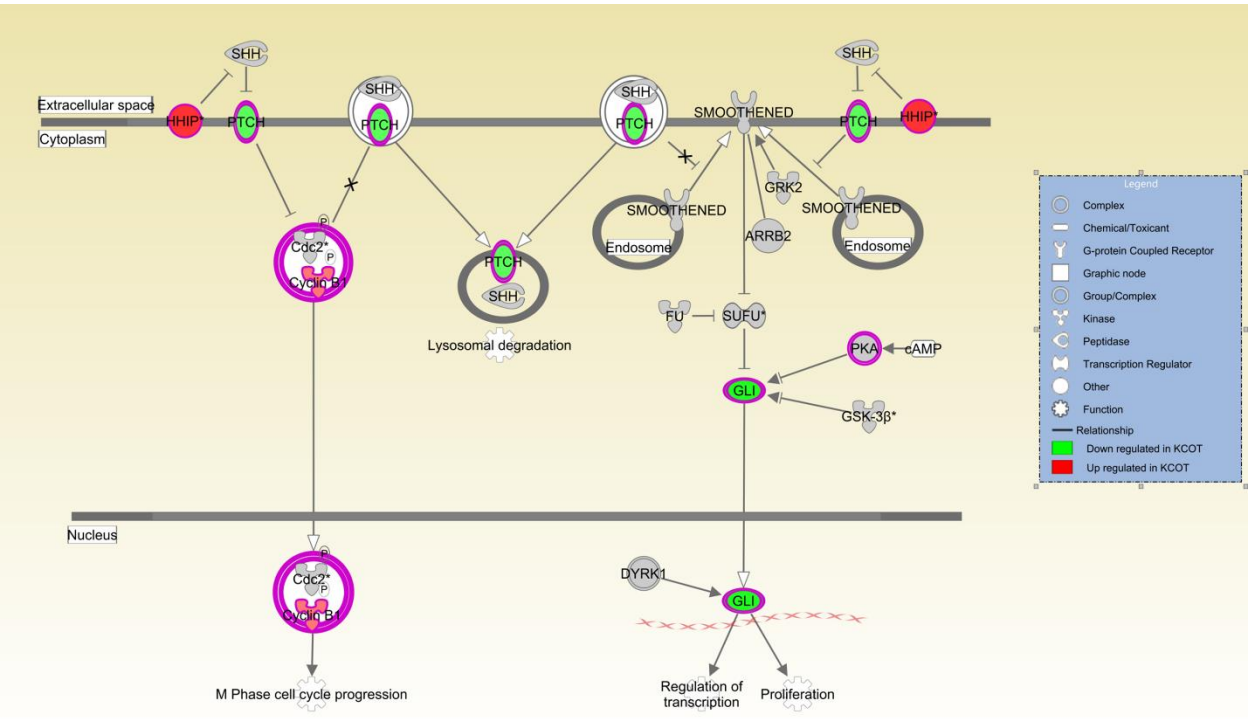


Figure 4.5: Sonic hedgehog signaling pathway. The upregulated and downregulated molecules of the pathway is shown according to the legend.

The sKC cluster had 2 activated and 20 inhibited pathways (Figure 4.4C). All Inflammatory (Immune) response/cytokine signaling pathways and most of the Cellular growth, proliferation and apoptosis pathways were found to be inhibited except for the PI3K/AKT Signaling pathway.

The oKC cluster had 38 activated and 2 inhibited pathways (Figure 4.4D). Most of these pathways involved Cellular growth, proliferation and apoptosis. Activation of MAP kinase-related pathways were also noted.

Upstream analysis

Table 4.4 presents predicted differentially activated genes with *p*-values < 0.05 along with the corresponding z-scores. In the common tumor cluster, 21 molecules were predicted to be inhibited and 57 to be activated when compared to normal tissue. These upstream molecules were mainly transcription regulators (21 molecules), cytokines (10 molecules) and kinases (8 molecules). Two molecules were predicted to be activated in the secretory ameloblast tumor sKC cluster, namely the complex *Cg* and cytokine *CSF2*. In addition, 3 transcription regulators and 2 kinases were predicted to have differential activity in the odontoblast cluster.

Table 4.4- Results of IPA Upstream analysis indicating significantly (p<0.05) differentially expressed molecules in KCOT clusters

Differentially expressed molecules in KCOT "common tumor cluster"

Molecule Type	Upstream Regulator (z-score)			
Transcription regulator	NUPR1	(-5.32)	NFE2L2	(2.59)
	MXI1	(-2.20)	FOXO1	(3.50)
	E2F6	(-2.65)	TP63	(2.61)
	KDM5B	(-4.57)	EZH2	(2.17)
	TP53	(-3.48)	MITF	(4.33)
	FOXL2	(2.65)	RUVBL1	(3.32)
	HIF1A	(2.93)	YAP1	(2.43)
	SP1	(2.17)	MYB	(2.59)
	JUN	(2.55)	RELA	(3.32)
	PPRC1	(3.06)	SMARCA4	(2.03)
	FOXM1	(3.33)		
Cytokine	IL4	(2.30)	IL1A	(3.83)
	CSF2	(3.86)	CCL5	(2.69)
	CD40LG	(2.02)	IL6	(3.81)
	IL17A	(2.24)	OSM	(2.28)
	TNF	(3.54)	IL1B	(3.32)
Kinase	STK17A	(-2.22)	EGFR	(3.10)
	PLK1	(-2.20)	MAP3K14	(2.43)
	TRIB3	(-2.65)	BRD4	(3.30)
	CDKN1A	(-3.28)	CCNK	(2.72)
Ligand-dependent nuclear receptor	PPARA	(2.20)	RARA	(3.45)
	PGR	(2.96)	ESR1	(3.71)
Enzyme	STUB1	(-2.63)	HRAS	(2.62)
	TRAF2	(2.97)	TERT	(2.07)
Complex	CD3	(-2.03)	NFkB (complex)	(3.34)
	IgG	(-2.23)	Cg	(4.54)
	I kappa b kinase	(-2.22)		
Group	STAT5a/b	(3.00)	ERK	(2.59)
	Jnk	(2.01)	Gm-csf	(2.16)
	E2f	(2.53)		
Growth factor	GDF2	(-2.00)	EGF	(2.81)
	FGF2	(2.20)		
Transporter	AZGP1	(-2.20)	SYVN1	(3.80)
Transmembrane receptor	TREM1	(3.72)	TNFRSF1A	(2.60)
Mature microrna	miR-34a-5p (and other miRNAs w/seed GGCAGUG)	(-2.46)	miR-122-5p (miRNAs w/seed GGAGUGU)	(-2.50)
Translation regulator	EIF4G1	(2.83)		
Peptidase	F7	(2.76)		
G-protein coupled receptor	HCAR2	(-2.12)		
Other	CBX7	(-2.24)	CD24	(4.00)
	PPP2R5C	(-2.33)	SELPLG	(2.31)
	UXT	(-2.88)	IGFBP2	(2.12)
	RABL6	(4.70)	RTKN	(2.00)
	HSPB2	(2.00)		

Differentially expressed molecules in KCOT "secretory ameloblast cluster"

Molecule Type	Upstream Regulator (z-score)	
Complex	Cg	(2.41)
Cytokine	CSF2	(2.36)

Differentially expressed molecules in KCOT "odontoblast cluster"

Molecule Type	Upstream Regulator (z-score)			
Transcription regulator	NUPR1	(-3.74)	TP53	(-3.06)
	SMARCE1	(-2.00)		
Kinase	CDKN1A	(-2.42)	MAPK9	(2.43)

Discussion

A major finding of this study was the discovery of 2 distinct clusters of KCOT that exhibit similar phenotype despite differences in molecular pathway activity. In addition, *PTCH1* and *GLI* expression were found to be downregulated in both clusters suggesting involvement in the tumorigenesis of non-NBCCS associated KCOT. Moreover, the study provides a comprehensive characterization of KCOT transcriptome that can serve as a hypothesis-generating resource for the advancement of precision medicine for the diagnosis and treatment of KCOT.

The finding of 2 distinct molecular clusters was in contrast to the findings of Heikinheimo et al. who described a more homogeneous profile of KCOT (Heikinheimo et al. 2015). However, comparing the findings of the 2 studies is difficult due to the use of gingiva tissue as the comparison tissue in their study. In the present study we interrogated the human "Dentome" to determine what tissue was most similar in gene expression profile to KCOT. It remains unclear what would be the optimal comparative oral tissue to use to help delineate differential gene expression in KCOT or other odontogenic tumors. Our use of a universal RNA allows for the normalization between arrays within the study and also can be used across studies which employ the universal RNA as an internal standardization that allows better comparison from one array to another.

The present study did find that most KCOT showed gene expression profiles that more closely resemble secretory ameloblast which is a differentiated cell type of the dental lamina from which KCOT is thought to arise. However, it is surprising that a smaller subset of tumors showed a gene expression profile more closely associated with odontoblast that is derived from

a mesenchymal cell lineage. Differences in the 2 molecular subtypes includes activation of PI3K/AKT Signaling in sKC and activation of the MAP kinase pathway in oKC. Earlier studies have found activation of the AKT pathway in KCOT (Chaisuparat et al. 2013) which have been extensively targeted in cancers for therapy (Hennessy et al. 2005) and may provide a novel treatment modality for this particular subtype of KCOT. Conversely, the oKC subtype had activation of 3 MAP kinase associated pathways. The MAP kinase pathway has been implicated in KCOT development (Gonzalez Moles et al. 2008). In addition, the oKC subtype showed activation of numerous Cellular growth, proliferation and apoptosis pathways that were not seen in the sKC subtype including the JAK/STAT pathway. Recently, a trial targeted this pathway in solid tumors with good results (Houghton et al. 2014) which could be explored as an adjunctive therapy before surgery to reduce the size of the tumors and thus extent of resection and reconstruction needed to reduce the high recurrence rate of KCOT.

The canonical pathway that showed the biggest difference in the “common tumor cluster” was the activation of the Acute Phase Response Signaling pathway. Additionally, 5 other inflammatory pathways were activated in the common tumor cluster compared to the normal tissue. Other investigators (Cottom et al. 2012; Kadlub et al. 2013) described the presence of inflammation around KCOT and a possible link between the presence of inflammation and the aggressiveness of the tumor. Furthermore, upstream analysis predicted the activation of several pro-inflammatory markers. GSEA also showed several inflammatory gene sets that were enriched in KCOT.

In addition, several cell cycle mechanism pathways and GSEA gene sets were found to be disturbed, corresponding to the uncontrolled proliferation found in the neoplasm. Upstream analysis also predicted inhibition of *CDKN1A* which has been implicated heavily in DNA damage response (Cazzalini et al. 2010) and control of cell cycles to prevent proliferation of neoplastic cells. Other notable upstream molecules predicted to be upregulated include members of the MAP kinase pathways such as *JUN*, *MAP3K14*, *ERK* and *JNK*, all of which have been heavily implicated with the development of cancers (Dhillon et al. 2007).

The SHH/PTCH1 pathway which is known to cause NBCCS (OMIM # 10940), was specifically interrogated. Individuals with NBCCS have *PTCH1* mutations and presents clinically with multiple and recurrent KCOT as well as other neoplasms (e.g. basal cell carcinoma). Historically, mutations in the *PTCH1* gene were found in more than 85% of syndromic KCOT but less than 30% of sporadic cases. More recent studies suggest that the proportion of sporadic KCOT with *PTCH1* mutations could be greater than originally believed with up to 80% of the cases affected (Qu et al. 2015). Although our study did not investigate the prevalence of *PTCH1* mutation, we found that the expression level of *PTCH1* in KCOT was decreased. The *PTCH1* gene functions as an important tumor suppressor (Kadlub et al. 2013; Pan et al. 2009) which when suppressed, can lead to development of tumors. More importantly, the SHH pathway inhibitor Vismodegib has been used as an adjunctive therapy in patients with NBCCS to reduce tumor size and reduce the margins needed for surgical resection (Booms et al. 2015). Currently, chemotherapeutic adjuncts are seldom used in sporadic cases; although they may be helpful in cases with decreased *PTCH1* expression in limiting tumor size, surgical margins and recurrence.

Although a large portion of the KCOT samples were FFPE which can result in RNA degradation (Ravo et al. 2008) and affect the performance of microarray analyses, previous studies supported the use of such samples (Abdueva et al. 2010). Furthermore, NanoString analysis has been used for expression analysis in various sample types including fresh-frozen, FFPE and even whole cell lysates to produce excellent results (Malkov et al. 2009). Despite the issue of RNA degradation in FFPE samples, there was good correlation between the microarray and nanoString expression data for genes with the greatest fold changes. Another short-coming is the limited sample size of 20 tumors which is due to KCOT being a relatively rare tumor. However, it is comparable to other similar studies with sample sizes ranging from 10-12 (Heikinheimo et al. 2007; Heikinheimo et al. 2015).

In conclusion, this study provides new insights into the transcriptome of KCOT using a method that provided purer sample populations of neoplastic epithelial cells. By comparing the KCOT transcriptome to the available “dentome”, we found familiar pathways that have been implicated in the formation of KCOT and other cancers. This is in addition to novel pathways that may serve as markers for diagnosis and prognosis. Possible targets for novel therapy were also identified in this study and should be investigated further to development precision medicine for the treatment of KCOT. Future studies should take into account the 2 distinct molecular subtypes when developing specific treatment modalities in order to maximize effectiveness.

CHAPTER 5: FINAL THOUGHTS AND FUTURE DIRECTIONS

The work presented in this dissertation contributes to our understanding of the molecular mechanisms involved in the tumorigenesis of odontogenic tumors and to advance the development of the field of precision medicine for the treatment of odontogenic tumors. In this dissertation, we presented the transcriptome of different odontogenic tissues at different stages of development (the dentome) and used this new information to compare gene expression profiles between normal and pathological odontogenic tissues. This represented a shift in the selection of normal comparison tissue for the examination of odontogenic tumor gene expression. By making this data freely available, we anticipate that future investigations of odontogenic tumors will adopt the concept of comparing tumor transcriptome to normal odontogenic tissue which will make the results more comparable between studies.

In this study, we made several important discoveries about gene expression differences between different odontogenic tissues. In the dentome portion of this study, we found that the early developing human tooth transcriptome involved more genes than anticipated. In addition, many diverse molecular pathways, not previously thought to be involved in the tooth formation process, were differentially activated in the tooth-forming cells. One important example is that we found genes and pathways involved in cell movement that may be essential for processes in normal odontogenesis and are worth closer examination. For example, future molecular studies could examine the role of Astrotactin, a major player in the migration and movement of

neurons (Adams et al. 2002). We identified high levels of Astrotactin mRNA suggesting the protein may be involved in the motility of maturing ameloblast. Understanding these processes is important in the advancement of technologies related to tissue engineering in the production of laboratory-grown replacement teeth for patients suffering from conditions with multiple congenitally missing teeth such as ectodermal dysplasia.

Also of note is the number of dentome defining genes that we found in our multiclass analysis between the odontogenic tissues. We found that there were only 60 differentially expressed genes out of upwards of 14,000 genes expressed by odontogenic tissue during normal development. Multiple differentially expressed genes in this group of 60 genes were identified that could be tissue defining genes. Investigators may be able to utilize this information for the development of viable odontogenic cell lines and models that are currently lacking. Development of ameloblast and odontoblast cell lines that replicate normal cell function would be of great importance to the field. The possible candidate genes include *SOX5*, which is highly expressed in odontoblasts but not ameloblasts, a transcriptional regulator involved in the regulation of embryonic development (Kawasaki et al. 2015) and may determine the differentiation of mesenchyme cells into odontoblasts. On the other hand, *KRT19* was found to be highly expressed in ameloblasts at both stages of development but not odontoblasts. *KRT19* has been implicated in the development of other tissue such as chondrocytes (Rodrigues-Pinto et al. 2016) but may prove to be an important marker in the development of odontogenic epithelium.

In the second paper, we found and reported that ameloblastoma presented with 2 distinct molecular profiles that were associated with 2 histological subtypes. Inter/intra tumor

heterogeneity has been described in many cancer types, including breast cancer, melanoma and prostate cancer with genomic instability, epigenetic changes and the tumor microenvironment implicated as factors causing gene expression heterogeneity (Burrell et al. 2013). This introduces significant challenges in designing effective treatment strategies ultimately resulting in treatment resistance after initial tumor shrinkage. The follicular ameloblastoma subtype showed activation of different molecular pathways compared with the plexiform subtype, suggesting that they could be receptive to different chemotherapeutic protocols. Recent advancements in understanding the tumorigenesis of ameloblastoma have opened up the possibility of novel chemotherapeutic treatment for recalcitrant disease, such as the successful treatment of recurrent stage 4 ameloblastoma with *BRAF* and *MEK* inhibitors (Kaye et al. 2015).

There are several promising targets for therapy from the canonical pathway and upstream analysis, including transcription factors *TP53* and *EZH2*. The guardian gene *TP53* is one of the most important genes in tumor development, acting as a powerful tumor suppressor. It is important in cell cycle regulation, initiating DNA repair during cell division and signaling the cell to undergo apoptosis when DNA damage cannot be corrected. Recently, there have been attempts to target *TP53* in cancer treatment including gene therapy to reintroduce wild type *TP53*, targeting cells that have lost *TP53* with an adenovirus to destroy them and the use of small molecules to upregulate the expression of *TP53* (Wang and Sun 2010). In our study, the frequent inhibition of *TP53* in ameloblastoma suggests that the patients would probably benefit from *TP53* targeted treatment. In addition, we found that *EZH2* was predicted to be activated in ameloblastomas. *EZH2* is responsible for the methylation levels of histone3K27, an

epigenetic change that determines the expression levels of many genes. Overexpression of *EZH2* can lead to suppression of important tumor suppressor genes and have been implicated in several human cancers (Yoo and Hennighausen 2012). Investigators have made rapid progress in the development of anti-*EZH2* therapy with some showing great promise in the use of a protein inhibitor in the treatment of lymphoma (McCabe et al. 2012). Therefore, targeting *EZH2* may be a viable therapy in the treatment of human ameloblastomas.

Another promising target group of pathways is the Inflammatory/immune. The link between inflammation dysregulation and cancer is not a new one, with several anti-inflammatory agents such as NSAIDS, COX-2 inhibitors and corticosteroids being studied as possible therapeutic agents (Rayburn et al. 2009). We found increased expression of inflammatory and immune gene products in laser captured neoplastic cells which suggests that these tumors produce inflammatory cytokines normally provided by immune cells. This implicates inflammatory pathways as viable targets in the development of adjunct therapies in the treatment of ameloblastomas. However, it is also important to note that the pre-secretory cluster (follicular ameloblastoma) showed much more activated inflammatory pathways compared to the odontoblast cluster (plexiform ameloblastoma). Future research should examine if the differences at the molecular level affects the response of the different subtypes of ameloblastoma to chemotherapeutic agents. This has the potential of improving treatment approaches to reduce morbidity and recurrence rates.

In the third paper, we found activation of familiar pathways, such as *PTCH1*, that have been implicated in the formation of multifocal KCOT in NBCCS. However, our data suggest that disturbances to the SHH/PTCH1 pathway are also important in isolated cases too. This finding

affirms existing knowledge of the suspected molecular drivers in the tumorigenesis of KCOT and strengthens the case for employing currently available protocols targeting these pathways. Such as the use of Vismodegib, a SHH pathway inhibitor currently used in the treatment of NBCCS (Booms et al. 2015), in the treatment of isolated cases of KCOT. Alternatively, downstream targets of *PTCH1* can be targeted as well. Other investigators targeted the smoothen gene in KCOT cell lines with cyclopamine that killed the tumor cells (Ren et al. 2012). We can hypothesize from our results, that the *GLI* genes downstream of *PTCH1* and smoothen may be viable targets. As with ameloblastoma, inflammatory dysregulation appears to be an important component in the development of KCOT. Targeting the inflammatory pathways in KCOT could prove to be a viable chemotherapeutic adjunct in the treatment of KCOT.

Several transcription regulators implicated in cancer development were also predicted to be disturbed in our study including the activation of *EZH2*, *HIF1A* and inhibition of *TP53*. The implications of *EZH2* and *TP53* disturbance were discussed earlier in the ameloblastoma section. Activation of *HIF1A* is found in numerous human cancers and promotes tumor growth via promotion of angiogenesis and regulation of cellular hypoxia (Talks et al. 2000). *HIF1A* is an important factor in solid tumor growth and has been targeted successfully in other human cancers. There are currently many approaches to inhibit *HIF1A* (Onnis et al. 2009) and should be examined for their effectiveness on KCOT.

In addition, we discovered 2 distinct clusters of KCOT gene expression profiles that exhibit similar phenotypes despite differences in molecular pathway activity. We found that the AKT pathway is activated in the secretory ameloblast cluster while the MAPKinase pathways is activated in the odontoblast cluster. This could explain the variable expression found by earlier

studies when these pathways were examined in isolation. Activated *AKT* was found in only 40% (Chaisuparat et al. 2013) of KCOT in an immunohistological study with activated *ERK*, the effector of the MAPKinase pathway, found in 60% (Ribeiro et al. 2012) of KCOT in another study. These findings suggest that future treatment of KCOT should examine the possibility of the 2 distinct molecular subtypes responding differently to treatment protocols and the use of additional diagnostic tools to clinically identify the different clusters. Another treatment strategy is the use of combination therapy targeting multiple pathways to combat the issue of intra/inter tumor heterogeneity. AKT inhibitors (Alexander 2011) can be combined with inhibitors of the members of the MAPKinase cascade (Sebolt-Leopold 2000) RAF, MEK and ERK, as an adjunct for the treatment of KCOT to target the different clonal populations, maximizing effectiveness and reducing treatment resistance.

In the near future, the rapidly advancing field of gene expression study will be able to employ techniques with higher resolution and be better able to amalgamate information from fields such as genomics and proteomics to form a clearer picture of the tumorigenesis of odontogenic tumors. Future studies will also benefit from high throughput methods such as deep sequencing that allow for the examination of alternatively spliced transcripts, post-transcriptional modifications and gene mutations. In addition to mRNA, other important populations of RNA such as miRNA, tRNA, and rRNA which are intimately involved in the final expression of protein products are likely informative and should be examined. Using available bioinformatics tools, we can superimpose information from genomic and proteomic studies with gene expression to provide an atlas of odontogenic tumors akin to The Cancer Genome Atlas currently being complied for the major cancer types.

In this age of rapid information sharing, we can overcome another shortcoming of the work presented here, which is the small sample size of tumors. Odontogenic tumors are relatively rare. Therefore, providing a standardized method of gene expression study with a comparable internal control, such as a universal human RNA or using digital methods such as RNA-seq, will allow the pooling of more samples and rare subtypes and better examination of the involved molecular pathways.

In conclusion, we provided a detailed characterization of ameloblastoma and KCOT transcriptome that can serve as a hypothesis-generating resource for the advancement of precision medicine for the diagnosis and treatment of these odontogenic tumors.

REFERENCES

Abdueva D, Wing M, Schaub B, Triche T, Davicioni E. 2010. Quantitative expression profiling in formalin-fixed paraffin-embedded samples by affymetrix microarrays. J Mol Diagn. 12(4):409-417.

Adams NC, Tomoda T, Cooper M, Dietz G, Hatten ME. 2002. Mice that lack astrotactin have slowed neuronal migration. Development. 129(4):965-972.

Adebayo ET, Ajike SO, Adekeye EO. 2002. Odontogenic tumours in children and adolescents: A study of 78 nigerian cases. J Craniomaxillofac Surg. 30(5):267-272.

Alexander W. 2011. Inhibiting the akt pathway in cancer treatment: Three leading candidates. P & T : a peer-reviewed journal for formulary management. 36(4):225-227.

Alur J, Narayan TV, Mohanty L, Shenoy S, Jamadar S, Shetty S. 2014. Ki-67 and p53 expression in solitary sporadic, syndrome associated and recurrent keratocystic odontogenic tumor. J Oral Maxillofac Pathol. 18(Suppl 1):S21-25.

Amm HM, Casimir MD, Clark DB, Sohn P, MacDougall M. 2014. Matrix metalloproteinase expression in keratocystic odontogenic tumors and primary cells. Connect Tissue Res. 55 Suppl 1:97-101.

Avelar RL, Primo BT, Pinheiro-Nogueira CB, Studart-Soares EC, de Oliveira RB, Romulo de Medeiros J, Hernandez PA. 2011. Worldwide incidence of odontogenic tumors. J Craniofac Surg. 22(6):2118-2123.

Barnes L EJ, Reichart P, Sidransky D, eds. 2005. World health organization classification of tumours. Pathology and genetics of head and neck tumours. Lyon: IARC Press.

Barros JC, Marshall CJ. 2005. Activation of either erk1/2 or erk5 map kinase pathways can lead to disruption of the actin cytoskeleton. J Cell Sci. 118(Pt 8):1663-1671.

Bartlett JD, Smith CE. 2013. Modulation of cell-cell junctional complexes by matrix metalloproteinases. J Dent Res. 92(1):10-17.

Bedi RS, Chugh A, Pasricha N. 2012. Ameloblastic carcinoma of maxilla. Natl J Maxillofac Surg. 3(1):70-74.

Bergendal B, Klar J, Stecksen-Blicks C, Norderyd J, Dahl N. 2011. Isolated oligodontia associated with mutations in edaradd, axin2, msx1, and pax9 genes. Am J Med Genet A. 155A(7):1616-1622.

Booms P, Harth M, Sader R, Ghanaati S. 2015. Vismodegib hedgehog-signaling inhibition and treatment of basal cell carcinomas as well as keratocystic odontogenic tumors in gorlin syndrome. Ann Maxillofac Surg. 5(1):14-19.

Brown NA, Rolland D, McHugh JB, Weigelin HC, Zhao L, Lim MS, Elenitoba-Johnson KS, Betz BL. 2014. Activating fgfr2-ras-braf mutations in ameloblastoma. Clin Cancer Res. 20(21):5517-5526.

Buchner A, Merrell PW, Carpenter WM. 2006. Relative frequency of central odontogenic tumors: A study of 1,088 cases from northern california and comparison to studies from other parts of the world. J Oral Maxillofac Surg. 64(9):1343-1352.

Bunney TD, Katan M. 2010. Phosphoinositide signalling in cancer: Beyond pi3k and pten. Nat Rev Cancer. 10(5):342-352

Burrell RA, McGranahan N, Bartek J, Swanton C. 2013. The causes and consequences of genetic heterogeneity in cancer evolution. Nature. 501(7467):338-345.

Cancer Genome Atlas N. 2012. Comprehensive molecular portraits of human breast tumours. Nature. 490(7418):61-70.

Carinci F, Francioso F, Piattelli A, Rubini C, Fioroni M, Evangelisti R, Arcelli D, Tosi L, Pezzetti F, Carinci P et al. 2003. Genetic expression profiling of six odontogenic tumors. J Dent Res. 82(7):551-557.

Casal ML, Lewis JR, Mauldin EA, Tardivel A, Ingold K, Favre M, Paradies F, Demotz S, Gaide O, Schneider P. 2007. Significant correction of disease after postnatal administration of recombinant ectodysplasin a in canine x-linked ectodermal dysplasia. Am J Hum Genet. 81(5):1050-1056.

Cazzalini O, Scovassi AI, Savio M, Stivala LA, Prosperi E. 2010. Multiple roles of the cell cycle inhibitor p21(cdkn1a) in the DNA damage response. Mutat Res. 704(1-3):12-20.

Chaisuparat R, Yodsanga S, Montaner S, Jham BC. 2013. Activation of the akt/mtor pathway in dentigerous cysts, odontogenic keratocysts, and ameloblastomas. Oral Surg Oral Med Oral Pathol Oral Radiol. 116(3):336-342.

Chokechanachaisakul U, Kaneko T, Yamanaka Y, Okiji T, Suda H. 2012. A novel whole tooth-in-jaw-bone culture of rat molars: Morphological, immunohistochemical, and laser capture microdissection analysis. Microsc Res Tech. 75(10):1341-1347.

Cobourne MT, Hardcastle Z, Sharpe PT. 2001. Sonic hedgehog regulates epithelial proliferation and cell survival in the developing tooth germ. J Dent Res. 80(11):1974-1979.

Collins FS, Varmus H. 2015. A new initiative on precision medicine. N Engl J Med. 372(9):793-795.

Cottom HE, Bshena FI, Speight PM, Craig GT, Jones AV. 2012. Histopathological features that predict the recurrence of odontogenic keratocysts. J Oral Pathol Med. 41(5):408-414.

Coussens LM, Werb Z. 2002. Inflammation and cancer. Nature. 420(6917):860-867.

Couve E, Osorio R, Schmachtenberg O. 2013. The amazing odontoblast: Activity, autophagy, and aging. J Dent Res. 92(9):765-772.

D'Souza RN, Aberg T, Gaikwad J, Cavender A, Owen M, Karsenty G, Thesleff I. 1999. Cbfa1 is required for epithelial-mesenchymal interactions regulating tooth development in mice. Development. 126(13):2911-2920.

Dassule HR, McMahon AP. 1998. Analysis of epithelial-mesenchymal interactions in the initial morphogenesis of the mammalian tooth. Dev Biol. 202(2):215-227.

De Silva I, Rozen WM, Ramakrishnan A, Mirkazemi M, Baillieu C, Ptasznik R, Leong J. 2012. Achieving adequate margins in ameloblastoma resection: The role for intra-operative specimen imaging. Clinical report and systematic review. PLoS One. 7(10):e47897.

Decarlo K, Emley A, Dadzie OE, Mahalingam M. 2011. Laser capture microdissection: Methods and applications. Methods Mol Biol. 755:1-15.

Deutsch D, Catalano-Sherman J, Dafni L, David S, Palmon A. 1995. Enamel matrix proteins and ameloblast biology. Connect Tissue Res. 32(1-4):97-107.

DeVilliers P, Suggs C, Simmons D, Murrah V, Wright JT. 2011. Microgenomics of ameloblastoma. J Dent Res. 90(4):463-469.

Devonshire AS, Elaswarapu R, Foy CA. 2010. Evaluation of external rna controls for the standardisation of gene expression biomarker measurements. BMC Genomics. 11:662.

Dhillon AS, Hagan S, Rath O, Kolch W. 2007. Map kinase signalling pathways in cancer. Oncogene. 26(22):3279-3290.

Dodds AP, Cannon RE, Suggs CA, Wright JT. 2003. Mrna expression and phenotype of odontogenic tumours in the v-ha-ras transgenic mouse. Arch Oral Biol. 48(12):843-850.

Eberlin LS, Norton I, Orringer D, Dunn IF, Liu X, Ide JL, Jarmusch AK, Ligon KL, Jolesz FA, Golby AJ et al. 2013. Ambient mass spectrometry for the intraoperative molecular diagnosis of human brain tumors. Proc Natl Acad Sci U S A. 110(5):1611-1616.

Ganter B, Zidek N, Hewitt PR, Muller D, Vladimirova A. 2008. Pathway analysis tools and toxicogenomics reference databases for risk assessment. Pharmacogenomics. 9(1):35-54.

Geiss GK, Bumgarner RE, Birditt B, Dahl T, Dowidar N, Dunaway DL, Fell HP, Ferree S, George RD, Grogan T et al. 2008. Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol. 26(3):317-325.

Grasmuck EA, Nelson BL. 2010. Keratocystic odontogenic tumor. Head Neck Pathol. 4(1):94-96.

Gomes CC, Diniz MG, Gomez RS. 2014. Progress towards personalized medicine for ameloblastoma. J Pathol. 232(5):488-491.

Gonzalez Moles MA, Mosqueda-Taylor A, Esteban F, Gil-Montoya JA, Diaz-Franco MA, Delgado M, Munoz M. 2008. Cell proliferation associated with actions of the substance p/nk-1 receptor complex in keratocystic odontogenic tumours. Oral Oncol. 44(12):1127-1133.

Gorlov IP, Meyer P, Liloglou T, Myles J, Boettger MB, Cassidy A, Girard L, Minna JD, Fischer R, Duffy S et al. 2007. Seizure 6-like (sez6l) gene and risk for lung cancer. Cancer Res. 67(17):8406-8411.

Guan X, Bartlett JD. 2013. Mmp20 modulates cadherin expression in ameloblasts as enamel develops. J Dent Res. 92(12):1123-1128.

Guler N, Sencift K, Demirkol O. 2012. Conservative management of keratocystic odontogenic tumors of jaws. ScientificWorldJournal. 2012:680397.

Guo YY, Zhang JY, Li XF, Luo HY, Chen F, Li TJ. 2013. Ptch1 gene mutations in keratocystic odontogenic tumors: A study of 43 chinese patients and a systematic review. PLoS One. 8(10):e77305.

Hakim SG, Kosmehl H, Sieg P, Trenkle T, Jacobsen HC, Attila Benedek G, Ribbat J, Driemel O. 2011. Altered expression of cell-cell adhesion molecules beta-catenin/e-cadherin and related wnt-signaling pathway in sporadic and syndromal keratocystic odontogenic tumors. Clin Oral Investig. 15(3):321-328.

Hayashi Y, Matsunaga T, Yamamoto G, Nishii K, Usui M, Yamamoto M, Tachikawa T. 2010. Comprehensive analysis of gene expression in the junctional epithelium by laser microdissection and microarray analysis. J Periodontal Res. 45(5):618-625.

Heikinheimo K, Jee KJ, Morgan PR, Nagy B, Knuutila S, Leivo I. 2007. Genetic changes in sporadic keratocystic odontogenic tumors (odontogenic keratocysts). J Dent Res. 86(6):544-549.

Heikinheimo K, Jee KJ, Niini T, Aalto Y, Happonen RP, Leivo I, Knuutila S. 2002. Gene expression profiling of ameloblastoma and human tooth germ by means of a cdna microarray. J Dent Res. 81(8):525-530.

Heikinheimo K, Kurppa KJ, Elenius K. 2015a. Novel targets for the treatment of ameloblastoma. J Dent Res. 94(2):237-240.

Heikinheimo K, Kurppa KJ, Laiho A, Peltonen S, Berdal A, Bouattour A, Ruhin B, Caton J, Thesleff I, Leivo I et al. 2015b. Early dental epithelial transcription factors distinguish ameloblastoma from keratocystic odontogenic tumor. J Dent Res. 94(1):101-111.

Hennessy BT, Smith DL, Ram PT, Lu Y, Mills GB. 2005. Exploiting the pi3k/akt pathway for cancer drug discovery. Nature reviews Drug discovery. 4(12):988-1004.

Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, Leiserson MD, Niu B, McLellan MD, Uzunangelov V et al. 2014. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell. 158(4):929-944.

Howard C, Murray PE, Namerow KN. 2010. Dental pulp stem cell migration. J Endod. 36(12):1963-1966.

Houghton PJ, Kurmasheva RT, Lyalin D, Maris JM, Kolb EA, Gorlick R, Reynolds CP, Kang MH, Keir ST, Wu J et al. 2014. Initial solid tumor testing (stage 1) of azd1480, an inhibitor of janus kinases 1 and 2 by the pediatric preclinical testing program. Pediatr Blood Cancer. 61(11):1972-1979

Hu S, Parker J, Wright JT. 2015. Towards unraveling the human tooth transcriptome: The dentome. PLoS One. 10(4):e0124801.

Jarvinen E, Salazar-Ciudad I, Birchmeier W, Taketo MM, Jernvall J, Thesleff I. 2006. Continuous tooth generation in mouse is induced by activated epithelial wnt/beta-catenin signaling. Proc Natl Acad Sci U S A. 103(49):18627-18632.

Jernvall J, Thesleff I. 2000. Reiterative signaling and patterning during mammalian tooth morphogenesis. Mech Dev. 92(1):19-29.

Johnson NR, Batstone MD, Savage NW. 2012. Management and recurrence of keratocystic odontogenic tumor: A systematic review. Oral Surg Oral Med Oral Pathol Oral Radiol.

Jongeneel CV, Iseli C, Stevenson BJ, Riggins GJ, Lal A, Mackay A, Harris RA, O'Hare MJ, Neville AM, Simpson AJ et al. 2003. Comprehensive sampling of gene expression in human cell lines with massively parallel signature sequencing. Proc Natl Acad Sci U S A. 100(8):4702-4705.

Juuri E, Isaksson S, Jussila M, Heikinheimo K, Thesleff I. 2013. Expression of the stem cell marker, sox2, in ameloblastoma and dental epithelium. Eur J Oral Sci. 121(6):509-516.

Kadler KE, Hill A, Canty-Laird EG. 2008. Collagen fibrillogenesis: Fibronectin, integrins, and minor collagens as organizers and nucleators. Curr Opin Cell Biol. 20(5):495-501.

Kadlub N, Coudert A, Gatibelza ME, El Houmami N, Soufir N, Ruhin-Poncet B, L'Hermine A C, Berdal A, Vazquez MP, Descroix V et al. 2013. Ptch1 mutation and local aggressiveness of odontogenic keratocystic tumors in children: Is there a relationship? Hum Pathol. 44(6):1071-1078.

Kawasaki K, Kawasaki M, Watanabe M, Idrus E, Nagai T, Oommen S, Maeda T, Hagiwara N, Que J, Sharpe PT et al. 2015. Expression of sox genes in tooth development. Int J Dev Biol. 59(10-12):471-478.

Kaye FJ, Ivey AM, Drane WE, Mendenhall WM, Allan RW. 2015. Clinical and radiographic response with combined braf-targeted therapy in stage 4 ameloblastoma. J Natl Cancer Inst. 107(1):378.

Kelly JM, Belding BA, Schaefer AK. 2010. Acanthomatous ameloblastoma in dogs treated with intralesional bleomycin. Vet Comp Oncol. 8(2):81-86.

Kibe T, Fuchigami T, Kishida M, Iijima M, Ishihata K, Hijioka H, Miyawaki A, Semba I, Nakamura N, Kiyono T et al. 2013. A novel ameloblastoma cell line (am-3) secretes mmp-9 in response to wnt-3a and induces osteoclastogenesis. Oral Surg Oral Med Oral Pathol Oral Radiol. 115(6):780-788.

Kim KM, Lim J, Choi YA, Kim JY, Shin HI, Park EK. 2012. Gene expression profiling of oral epithelium during tooth development. Arch Oral Biol. 57(8):1100-1107.

Kolsch V, Charest PG, Firtel RA. 2008. The regulation of cell motility and chemotaxis by phospholipid signaling. J Cell Sci. 121(Pt 5):551-559.

Kumamoto H, Izutsu T, Ohki K, Takahashi N, Ooya K. 2004a. P53 gene status and expression of p53, mdm2, and p14 proteins in ameloblastomas. J Oral Pathol Med. 33(5):292-299.

Kumamoto H, Ooya K. 2007. Immunohistochemical detection of phosphorylated jnk, p38 mapk, and erk5 in ameloblastic tumors. J Oral Pathol Med. 36(9):543-549.

Kumamoto H, Takahashi N, Ooya K. 2004b. K-ras gene status and expression of ras/mitogen-activated protein kinase (mapk) signaling molecules in ameloblastomas. J Oral Pathol Med. 33(6):360-367.

Kurppa KJ, Caton J, Morgan PR, Ristimaki A, Ruhin B, Kellokoski J, Elenius K, Heikinheimo K. 2014. High frequency of braf v600e mutations in ameloblastoma. J Pathol. 232(5):492-498.

Landin MA, Shabestari M, Babaie E, Reseland JE, Osmundsen H. 2012. Gene expression profiling during murine tooth development. Front Genet. 3:139.

Lam C, Ou JC, Billingsley EM. 2013. "Ptch"-ing it together: A basal cell nevus syndrome review. Dermatol Surg.

Lee DS, Yoon WJ, Cho ES, Kim HJ, Gronostajski RM, Cho MI, Park JC. 2011. Crosstalk between nuclear factor i-c and transforming growth factor-beta1 signaling regulates odontoblast differentiation and homeostasis. PLoS One. 6(12):e29160.

Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. 2011. Molecular signatures database (msigdb) 3.0. Bioinformatics. 27(12):1739-1740.

Li N, Zhong M, Song M. 2012. Expression of phosphorylated mtor and its regulatory protein is related to biological behaviors of ameloblastoma. Int J Clin Exp Pathol. 5(7):660-667.

Lim J, Ahn H, Min S, Hong SD, Lee JI, Hong SP. 2006. Oligonucleotide microarray analysis of ameloblastoma compared with dentigerous cyst. J Oral Pathol Med. 35(5):278-285.

Lin D, Huang Y, He F, Gu S, Zhang G, Chen Y, Zhang Y. 2007. Expression survey of genes critical for tooth development in the human embryonic tooth germ. Dev Dyn. 236(5):1307-1312.

Liu F, Millar SE. 2010. Wnt/beta-catenin signaling in oral tissue development and disease. J Dent Res. 89(4):318-330.

Luo DY, Feng CJ, Guo JB. 2012. Pulmonary metastases from an ameloblastoma: Case report and review of the literature. J Craniomaxillofac Surg. 40(8):e470-474.

Malkov VA, Serikawa KA, Balantac N, Watters J, Geiss G, Mashadi-Hossein A, Fare T. 2009. Multiplexed measurements of gene signatures in different analytes using the nanostring ncounter assay system. BMC Res Notes. 2:80.

Matheson S, Larjava H, Hakkinen L. 2005. Distinctive localization and function for lumican, fibromodulin and decorin to regulate collagen fibril organization in periodontal tissues. J Periodontal Res. 40(4):312-324.

Maxwell GL, Allard J, Gadisetti CV, Litzi T, Casablanca Y, Chandran U, Darcy KM, Levine DA, Berchuck A, Hamilton CA et al. 2013. Transcript expression in endometrial cancers from black and white patients. Gynecol Oncol. 130(1):169-173.

McCabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS, Liu Y, Graves AP, Della Pietra A, 3rd, Diaz E et al. 2012. Ezh2 inhibition as a therapeutic strategy for lymphoma with ezh2-activating mutations. Nature. 492(7427):108-112.

Mendenhall WM, Werning JW, Fernandes R, Malyapa RS, Mendenhall NP. 2007. Ameloblastoma. Am J Clin Oncol. 30(6):645-648.

Morimoto M, Kerouredan O, Gendronneau M, Shuen C, Baradaran-Heravi A, Asakura Y, Basiratnia M, Bogdanovic R, Bonneau D, Buck A et al. 2012. Dental abnormalities in schimke immuno-osseous dysplasia. J Dent Res. 91(7 Suppl):29S-37S.

Nieminen P, Pekkanen M, Aberg T, Thesleff I. 1998. A graphical www-database on gene expression in tooth. Eur J Oral Sci. 106 Suppl 1:7-11.

Noubissi FK, Elcheva I, Bhatia N, Shakoori A, Ougolkov A, Liu J, Minamoto T, Ross J, Fuchs SY, Spiegelman VS. 2006. Crd-bp mediates stabilization of betatrcp1 and c-myc mrna in response to beta-catenin signalling. Nature. 441(7095):898-901.

Novoradovskaya N, Whitfield ML, Basehore LS, Novoradovsky A, Pesich R, Usary J, Karaca M, Wong WK, Aprelikova O, Fero M et al. 2004. Universal reference rna as a standard for microarray experiments. BMC Genomics. 5(1):20.

Onnis B, Rapisarda A, Melillo G. 2009. Development of hif-1 inhibitors for cancer therapy. J Cell Mol Med. 13(9A):2780-2786.

Otero D, Lourenco SQ, Ruiz-Avila I, Bravo M, Sousa T, de Faria PA, Gonzalez-Moles MA. 2013. Expression of proliferative markers in ameloblastomas and malignant odontogenic tumors. Oral Dis. 19(4):360-365.

Pan S, Xu LL, Sun LS, Li TJ. 2009. Identification of known and novel ptch mutations in both syndromic and non-syndromic keratocystic odontogenic tumors. Int J Oral Sci. 1(1):34-38.

Partridge M, Towers JF. 1987. The primordial cyst (odontogenic keratocyst): Its tumour-like characteristics and behaviour. Br J Oral Maxillofac Surg. 25(4):271-279

Pillas D, Hoggart CJ, Evans DM, O'Reilly PF, Sipila K, Lahdesmaki R, Millwood IY, Kaakinen M, Netuveli G, Blane D et al. 2010. Genome-wide association study reveals multiple loci associated with primary tooth development during infancy. PLoS Genet. 6(2):e1000856.

Pispa J, Jung HS, Jernvall J, Kettunen P, Mustonen T, Tabata MJ, Kere J, Thesleff I. 1999. Cusp patterning defect in tabby mouse teeth and its partial rescue by fgf. Dev Biol. 216(2):521-534.

Qvist V. 1975. Pulp reactions in human teeth to tooth colored filling materials. Scand J Dent Res. 83(2):54-66.

Qu J, Yu F, Hong Y, Guo Y, Sun L, Li X, Zhang J, Zhang H, Shi R, Chen F et al. 2015. Underestimated ptch1 mutation rate in sporadic keratocystic odontogenic tumors. Oral Oncol. 51(1):40-45.

Radwanska A, Baczynska D, Nowak D, Brezillon S, Popow A, Maquart FX, Wegrowski Y, Malicka-Blaszkiewicz M. 2008. Lumican affects actin cytoskeletal organization in human melanoma a375 cells. Life Sci. 83(19-20):651-660.

Ravo M, Mutarelli M, Ferraro L, Grober OM, Paris O, Tarallo R, Vigilante A, Cimino D, De Bortoli M, Nola E et al. 2008. Quantitative expression profiling of highly degraded rna from formalin-fixed, paraffin-embedded breast tumor biopsies by oligonucleotide microarrays. Lab Invest. 88(4):430-440.

Rayburn ER, Ezell SJ, Zhang R. 2009. Anti-inflammatory agents for cancer therapy. Mol Cell Pharmacol. 1(1):29-43.

Reichart PA, Philipsen HP, Sonner S. 1995. Ameloblastoma: Biological profile of 3677 cases. Eur J Cancer B Oral Oncol. 31B(2):86-99.

Ren C, Amm HM, DeVilliers P, Wu Y, Deatherage JR, Liu Z, MacDougall M. 2012. Targeting the sonic hedgehog pathway in keratocystic odontogenic tumor. J Biol Chem. 287(32):27117-27125.

Ribeiro AL, Nobre RM, Alves-Junior SM, Kataoka MS, Barroso RF, Jaeger RG, Pinheiro JJ. 2012. Matrix metalloproteinases, tissue inhibitors of metalloproteinases, and growth factors regulate the aggressiveness and proliferative activity of keratocystic odontogenic tumors. Oral Surg Oral Med Oral Pathol Oral Radiol. 114(4):487-496.

Rivlin N, Brosh R, Oren M, Rotter V. 2011. Mutations in the p53 tumor suppressor gene: Important milestones at the various steps of tumorigenesis. Genes Cancer. 2(4):466-474.

Rodrigues-Pinto R, Berry A, Piper-Hanley K, Hanley N, Richardson SM, Hoyland JA. 2016. Spatiotemporal analysis of putative notochordal cell markers reveals cd24 and keratins 8, 18 and 19 as notochord-specific markers during early human intervertebral disc development. J Orthop Res.

Rufini A, Barlattani A, Docimo R, Velletri T, Niklison-Chirou MV, Agostini M, Melino G. 2011. P63 in tooth development. Biochem Pharmacol. 82(10):1256-1261.

Salmon CR, Silverio KG, Giorgetti AP, Sallum EA, Casati MZ, Nociti FH, Jr. 2012. Gene expression analysis in microdissected samples from decalcified tissues. Diagn Mol Pathol. 21(2):120-126.

Schoner K, Kohlhase J, Muller AM, Schramm T, Plassmann M, Schmitz R, Neesen J, Wieacker P, Rehder H. 2013. Hydrocephalus, agenesis of the corpus callosum, and cleft lip/palate represent frequent associations in fetuses with peters' plus syndrome and b3galtl mutations. Fetal pps phenotypes, expanded by dandy walker cyst and encephalocele. Prenat Diagn. 33(1):75-80.

Sebolt-Leopold JS. 2000. Development of anticancer drugs targeting the map kinase pathway. Oncogene. 19(56):6594-6599.

Sehdev MK, Huvos AG, Strong EW, Gerold FP, Willis GW. 1974. Proceedings: Ameloblastoma of maxilla and mandible. Cancer. 33(2):324-333.

Servato JP, de Souza PE, Horta MC, Ribeiro DC, de Aguiar MC, de Faria PR, Cardoso SV, Loyola AM. 2012. Odontogenic tumours in children and adolescents: A collaborative study of 431 cases. Int J Oral Maxillofac Surg. 41(6):768-773.

Sharifi-Sistani N, Zartab H, Babakoohi S, Saghravanian N, Jamshidi S, Esmaili H, Mohtasham N, Zamanzadeh M, Abbaszadeh-Bidokhty H. 2011. Immunohistochemical comparison of the expression of p53 and mdm2 proteins in ameloblastomas and keratocystic odontogenic tumors. J Craniofac Surg. 22(5):1652-1656.

Shtutman M, Baig M, Levina E, Hurteau G, Lim CU, Broude E, Nikiforov M, Harkins TT, Carmack CS, Ding Y et al. 2011. Tumor-specific silencing of copz2 gene encoding coatomer protein complex subunit zeta 2 renders tumor cells dependent on its paralogous gene copz1. Proc Natl Acad Sci U S A. 108(30):12449-12454.

Simmer JP, Richardson AS, Hu YY, Smith CE, Ching-Chun Hu J. 2012. A post-classical theory of enamel biomineralization... And why we need one. Int J Oral Sci. 4(3):129-134.

Singh M, Shah A, Bhattacharya A, Raman R, Ranganatha N, Prakash P. 2014. Treatment algorithm for ameloblastoma. Case Rep Dent. 2014:121032.

Sorkin BC, Wang MY, Dobeck JM, Albergo KL, Skobe Z. 2000. The cadherin-catenin complex is expressed alternately with the adenomatous polyposis coli protein during rat incisor amelogenesis. J Histochem Cytochem. 48(3):397-406.

Stephanopoulos G, Garefalaki ME, Lyroudia K. 2005. Genes and related proteins involved in amelogenesis imperfecta. J Dent Res. 84(12):1117-1126.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 102(43):15545-15550.

Sukarawan W, Simmons D, Suggs C, Long K, Wright JT. 2010. Wnt5a expression in ameloblastoma and its roles in regulating enamel epithelium tumorigenic behaviors. Am J Pathol. 176(1):461-471.

Sun JX, Horst OV, Bumgarner R, Lakely B, Somerman MJ, Zhang H. 2012. Laser capture microdissection enables cellular and molecular studies of tooth root development. Int J Oral Sci. 4(1):7-13.

Sweeney RT, McClary AC, Myers BR, Biscocho J, Neahring L, Kwei KA, Qu K, Gong X, Ng T, Jones CD et al. 2014. Identification of recurrent smo and braf mutations in ameloblastomas. Nat Genet. 46(7):722-725.

Talks KL, Turley H, Gatter KC, Maxwell PH, Pugh CW, Ratcliffe PJ, Harris AL. 2000. The expression and distribution of the hypoxia-inducible factors hif-1alpha and hif-2alpha in normal human tissues, cancers, and tumor-associated macrophages. Am J Pathol. 157(2):411-421.

Tan B, Yan TS, Shermin L, Teck KC, Yoke PC, Goh C, Balakrishnan A. 2013. Malignant transformation of keratocystic odontogenic tumor: Two case reports. Am J Otolaryngol. 34(4):357-361.

Thesleff I, Keranen S, Jernvall J. 2001. Enamel knots as signaling centers linking tooth morphogenesis and odontoblast differentiation. Adv Dent Res. 15:14-18.

Thomas S, Bonchev D. 2010. A survey of current software for network analysis in molecular biology. Hum Genomics. 4(5):353-360.

Thorgeirsson SS, Grisham JW. 2002. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet. 31(4):339-346.

Tranasi M, Sberna MT, Zizzari V, D'Apolito G, Mastrangelo F, Salini L, Stuppia L, Tete S. 2009. Microarray evaluation of age-related changes in human dental pulp. J Endod. 35(9):1211-1217.

Uzawa N, Suzuki M, Miura C, Tomomatsu N, Izumo T, Harada K. 2015. Primary ameloblastic carcinoma of the maxilla: A case report and literature review. Oncol Lett. 9(1):459-467.

Van den Bossche J, Malissen B, Mantovani A, De Baetselier P, Van Ginderachter JA. 2012. Regulation and function of the e-cadherin/catenin complex in cells of the monocyte-macrophage lineage and dcs. Blood. 119(7):1623-1633.

Villavicencio EH, Walterhouse DO, Iannaccone PM. 2000. The sonic hedgehog-patched-gli pathway in human development and disease. Am J Hum Genet. 67(5):1047-1054.

Wang Z, Sun Y. 2010. Targeting p53 for novel anticancer therapy. Transl Oncol. 3(1):1-12.

Wilson PM, Fryer RH, Fang Y, Hatten ME. 2010. Astn2, a novel member of the astrotactin gene family, regulates the trafficking of astn1 during glial-guided neuronal migration. J Neurosci. 30(25):8529-8540.

Yoo KH, Hennighausen L. 2012. Ezh2 methyltransferase and h3k27 methylation in breast cancer. Int J Biol Sci. 8(1):59-65.