

UNIVERSIDADE ESTADUAL PAULISTA - UNESP  
Instituto de Biociências  
Pós-Graduação em Ciências Biológicas - Genética

## **BIOMARCADORES DE PROGRESSÃO EM CARCINOMAS ORAIS**

Tese de Doutorado

**Nilva de Karla Cervigne**

Doutorado – Pós Graduação em Ciências Biológicas, Genética

Departamento de Genética – IBB – UNESP/Botucatu

Orientadora: Dra. Patricia Pintor dos Reis

Co-Orientadora: Dra. Suzanne Kamel-Reid

Botucatu - SP  
Novembro/2009

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

# **Biomarcadores de Progressão em Carcinomas Orais**

**Por**

**Nilva de Karla Cervigne**

**Orientadoras:**

**Dra. Patricia Pintor dos Reis**

**Dra. Suzanne Kamel-Reid**

**Tese submetida de acordo com as exigências  
para a certificação do título de Doutor em Filosofia  
junto ao Curso de Pós-Graduação em Ciências Biológicas - Genética  
Universidade Estadual Paulista – UNESP**

**© Por Nilva K. Cervigne, 2009  
Botucatu, SP**

“No porto de antes, apreensivo, eu tentava imaginar as dificuldades e lutas futuras. No de agora, dono do tempo que eu conquistara, simplesmente admirava o que estava ao meu redor e desfrutava do que estava feito. Não era a sensação de uma batalha ganha, de uma luta em que os obstáculos estavam vencidos. Muito mais do que isso; era o prazer interior de ter realizado algo que tanto desejei, de ter feito e visto o que eu fiz e vi”.

(Amyr Klink)

... Dedico esta tese à minha família,

Aos meus pais, Neide e  
Nilson, pelo apoio, encorajamento, amor e  
pelos ensinamentos que formaram os  
alicerces de minha história.

Aos meus irmãos Neiva e Neto, por todo o carinho e apoio  
necessários para a realização desse trabalho.

Ao meu sempre companheiro Carlos, por todo  
apoio, amor, compreensão e pela  
companhia ao longo da trajetória que me  
levou à concretização deste sonho.

Agradecimento Especial...

À **Dra. Patrícia P. Reis**, pela competência com que orientou esta minha tese e o tempo que generosamente me dedicou transmitindo-me os melhores e mais úteis ensinamentos, com paciência, lucidez e confiança. Obrigada por transmitir seus conhecimentos e experiências profissionais e de vida, por acreditar em mim, por ter me guiado para além da teoria e das técnicas... Expresso aqui meus agradecimentos e profundo respeito, que sempre serão mínimos diante do muito que me foi oferecido.

À **Prof Dra. Suzanne Kamel-Reid** pela oportunidade, confiança e competência de sua orientação, o meu sincero agradecimento. Pelo acesso que me facilitou a uma pesquisa mais alargada e enriquecedora e pela sua crítica sempre tão atempada, como construtiva, bem-haja estou-lhe muito, muito grata.

Agradecimentos...

Ao *Galoway Fund* e *Cancer Research Society* pelo auxílio financeiro, essencial para o desenvolvimento deste trabalho.

A *Dra. Natalie Naranjo Galloni*, pelo estímulo e contribuição teórica na fase inicial deste trabalho, bem como coleta dos dados dos pacientes utilizados neste estudo. Ao *Dr. Bayardo Perez-Ordonez*, pela coleta das amostras utilizadas neste projeto.

As *Prof. Dra. Grace Bradley, Dra. Fei-Fei Liu, Dra. Suzan Done*, banca de pesquisadores do *Ontario Cancer Institute*, pelas suas reflexões criativas, críticas e sugestões durante todo o desenvolvimento desse projeto de doutorado, meus sinceros agradecimentos.

A *Dr. Igor Jurisica, Melania Pintilie, Dr. Bekim Sadikovic e Dr. Anath Lionel*, pela ajuda com as análises de biotinformatica e estatística, e contribuição na elaboração dos dados deste trabalho.

Aos companheiros do laboratório e amigos de todas as horas *Jerry Machado, Rashmi Goswami, Mahadeo Sukhai, Mariam Thomas, Yali Xuan, Rikki Bharadwaj, Miranda Tomenson*, meus sinceros e profundos agradecimentos pelo apoio e a grande ajuda, profissional e pessoal, na realização desse projeto. Vocês certamente são parte desta vitória!

A todos os colegas de trabalho do *Ontario Cancer Institute*, que de alguma maneira contribuíram para a concretização desse trabalho, assim como pelas discussões científicas compartilhadas entre um café e outro.

A todos os amigos que encontrei ao longo desses quatro anos em Toronto, que me acolheram e certamente contribuíram para minha adaptação, facilitando o meu desenvolvimento profissional.

Aos funcionários da seção de Pós-graduação, por toda atenção ao longo do meu curso de doutorado e pelos serviços prestados para a finalizacao dessa Tese.

A querida família Furlan, *Carlos, Rosana e Michele*, gostaria de expressar a minha gratidão pelo apoio incondicional em todos os momentos dessa caminhada.

## **Biomarcadores de Progressão em Carcinomas Orais**

Nilva K. Cervigne

Doutorado

Departamento de Genética, IBB, Universidade Estadual Paulista-UNESP, Botucatu, SP

Department of Applied Molecular Oncology, University Health Network, Toronto, ON, Canada.

2009

### **Resumo**

Os carcinomas de células escamosas orais (CCEOs) são tumores malignos da cavidade oral que abrangem aproximadamente 50% de todos os cânceres de cabeça e pescoço, e são a sexta causa de morte por câncer em todo o mundo. Acredita-se que 16-62% destes carcinomas se desenvolvam de lesões potencialmente malignas (OPMLs) conhecidas como leucoplasias orais. Embora se saiba que os CCEOs possam se desenvolver a partir de leucoplasias, a avaliação clínica e histológica existente possui um valor prognóstico limitado para prever quais dessas OPMLs progredirão a carcinoma. Neste contexto, a busca por marcadores genéticos associados ao prognóstico de leucoplasias é importante, uma vez que estes seriam utilizados como ferramentas mais robustas para a predição da transformação maligna das leucoplasias orais. Apesar de alguns estudos mostrarem alterações genéticas envolvidas na progressão do câncer oral, genes capazes de prever a transformação maligna de leucoplasias à carcinoma permanecem desconhecidos. A aplicação de análises genômicas em ampla escala, como as utilizadas neste estudo, possui potencial para superar este problema, uma vez que permite a identificação de genes/vias específicas envolvidos nesta doença. Além disso, a inclusão de amostras sequenciais de leucoplasias que progrediram e CCEOs de um mesmo paciente, permite uma comparação mais acurada de modificações genéticas que ocorrem nestas lesões durante a progressão oral. Para tal, aplicamos a análise de expressão de microRNAs e a análise do número de cópias do DNA para a identificação de alterações genéticas associadas com a progressão do



câncer oral. Este estudo foi o primeiro a identificar uma assinatura de miRNAs (miR-146a, miR-181b, miR-184, miR-21, miR-345, miR-518b, miR-520g, miR-649, miR-196a e miR-206) foi identificada como associada com a progressão ao câncer oral. Adicionalmente, identificamos potenciais marcadores genômicos associados com carcinomas orais, mapeados em 1p, 2p, 5q e 14q, utilizando a análise de aCGH. As alterações genéticas identificadas foram comuns entre as OPMLs progressivas e os CCEOs correspondentes, sugerindo que tais alterações possam desempenhar um importante papel durante a progressão da doença. Os genes identificados possuem o potencial de serem utilizados como biomarcadores para melhorar a avaliação dos riscos de transformação de leucoplasias orais.

## Índice

|   |           |
|---|-----------|
| <b>LISTA DE TABELAS .....</b>   | <b>4</b>  |
| <b>LISTA DE FIGURAS .....</b>   | <b>5</b>  |
| <b>ABREVIACÕES .....</b>  | <b>6</b>  |
| <br>  |           |
| <b>1. Câncer de Cabeça e Pescoço: definição e fatores etiológicos .....</b>                     | <b>10</b> |
| 1.1. Definição .....  | 10        |
| 1.2. Fatores etiológicos do câncer de cabeça e pescoço .....                                    | 10        |
| <b>2. Carcinomas de Células Escamosas da Cavidade Oral .....</b>                                | <b>14</b> |
| 2.1. Definição e Prevalência .....  | 14        |
| 2.2. Prognóstico e Tratamento de Pacientes com CCEO .....                                       | 15        |
| <b>3. Cancerização de Campo e a Progressão de CCEO .....</b>                                    | <b>17</b> |
| <b>4. Lesões Oraís Potencialmente Malignas (OPMLs) .....</b>                                    | <b>18</b> |
| 4.1. Definição e Prevalência Mundial .....  | 19        |
| 4.3. Fatores de Risco da Transformação Maligna .....  | 24        |
| 4.4. Fatores Etiológicos das Leucoplasias Oraís .....   | 25        |
| 4.5. Diagnóstico e Monitoramento das Leucoplasias Oraís .....                                   | 26        |
| <b>5. A Progressão ao Câncer Oral .....</b>   | <b>29</b> |
| 5.1. Anormalidades Genéticas e a Progressão do Câncer Oral .....                                | 32        |
| 5.2. Biomarcadores de Progressão Tumoral .....  | 36        |
| <br>  |           |
| <b>2. ANÁLISE GENÔMICA DO CÂNCER .....</b>  | <b>39</b> |
| <br>  |           |
| <b>1. microRNAs .....</b>   | <b>40</b> |
| 1.1. Biogênese e Função dos microRNAs .....   | 40        |
| 1.2. miRNAs e Câncer .....  | 43        |
| 1.3. Papel dos miRNAs em CCEOs .....  | 53        |
| 1.4. Métodos de Análise do Perfil de Expressão de miRNAs .....                                  | 55        |
| 1.5. Análise de miRNAs com Amostras FFPE .....  | 62        |
| <b>2. Alteração Global do Número de Cópias do DNA (CNAs) .....</b>                              | <b>64</b> |
| 2.1. Variação no Número de Cópias (CNV) e Alterações no Número de Cópias (CNA) do DNA .....     | 64        |
| 2.2. Análise no número de cópias do DNA em Câncer .....   | 65        |
| 2.3. aCGH em Amostras FFPE .....  | 67        |
| 2.4. Amplificação Genômica Global (WGA) para a Análise de aCGH .....                            | 69        |
| 2.5. Comparação de Diferentes Plataformas de a-CGH – BACs vs. Arrays de Oligonucleotídeos ..... | 72        |

|   |            |
|---|------------|
| 2.6. Análises de CNAs em CCECP .....                                    | 77         |
| 2.7. CNAs Aplicados à Progressão do Câncer Oral.....                    | 79         |
| <b>Referências.....</b>   | <b>83</b>  |
| <b>3. HIPÓTESE, OBJETIVOS E MATERIAL E MÉTODOS .....</b>                | <b>100</b> |
| <b>Hipótese.....</b>  | <b>100</b> |
| <b>Objetivos .....</b>  | <b>100</b> |
| <b>Delineamento Experimental.....</b>                                   | <b>101</b> |
| <b>Material e Métodos .....</b>   | <b>102</b> |
| Grupo Amostral .....  | 102        |
| <b>4. MANUSCRITO 1: ANÁLISE DE EXPRESSÃO DE MIRNAS .....</b>            | <b>105</b> |
| <b>Abstract.....</b>  | <b>107</b> |
| <b>Introduction.....</b>  | <b>108</b> |
| <b>Material and Methods .....</b>                                       | <b>109</b> |
| <b>Results .....</b>  | <b>113</b> |
| <b>Discussion.....</b>  | <b>116</b> |
| <b>References .....</b>   | <b>131</b> |
| <b>5. MANUSCRITO 2: ANÁLISE DO NÚMERO DE CÓPIAS DO DNA USING ACGH .</b> | <b>134</b> |
| <b>Abstract.....</b>  | <b>136</b> |
| <b>Introduction.....</b>  | <b>138</b> |
| <b>Material and Methods .....</b>                                       | <b>139</b> |
| <b>Results .....</b>  | <b>143</b> |
| <b>Discussion.....</b>  | <b>146</b> |
| <b>References .....</b>   | <b>162</b> |
| <b>6. CONCLUSÕES E DIREÇÕES FUTURAS .....</b>                           | <b>165</b> |

|   |            |
|---|------------|
| <b>Referências.....</b>   | <b>173</b> |
| <b>7. APÊNDICE.....</b>   | <b>175</b> |
| <b>7.1. Material Suplementar.....</b>   | <b>175</b> |
| <b>7.2. Outros Manuscritos submetidos e/ou em Preparo para publicação .....</b> | <b>190</b> |
| 7.2.1. Submetidos:.....   | 190        |
| 7.2.2. Em preparação: .....   | 194        |

## Lista de Tabelas

| <b>Tabelas</b>                     | <b>Página</b> |
|------------------------------------|---------------|
| <b>Tabela 1.1.....</b>             | <b>21</b>     |
| <b>Tabela 1.2.....</b>             | <b>36</b>     |
| <b>Tabela 2.1.....</b>             | <b>62</b>     |
| <b>Tabela 4.1.....</b>             | <b>124</b>    |
| <b>Tabela 4.2.....</b>             | <b>125</b>    |
| <b>Tabela 4.3.....</b>             | <b>126</b>    |
| <b>Tabela 5.1.....</b>             | <b>153</b>    |
| <b>Tabela 5.2.....</b>             | <b>154</b>    |
| <b>Tabela 5.3.....</b>             | <b>155</b>    |
| <b>Tabela Suplementar 3.1.....</b> | <b>176</b>    |
| <b>Tabela Suplementar 3.2.....</b> | <b>178</b>    |
| <b>Tabela Suplementar 5.1.....</b> | <b>179</b>    |
| <b>Tabela Suplementar 5.2.....</b> | <b>180</b>    |

## Lista de Figuras

| <b>Figuras</b>         | <b>Página</b> |
|------------------------|---------------|
| <b>Figura 1.1.....</b> | <b>24</b>     |
| <b>Figura 1.2.....</b> | <b>28</b>     |
| <b>Figura 1.3.....</b> | <b>32</b>     |
| <b>Figura 2.1.....</b> | <b>43</b>     |
| <b>Figura 2.2.....</b> | <b>49</b>     |
| <b>Figura 2.3.....</b> | <b>58</b>     |
| <b>Figura 3.1.....</b> | <b>105</b>    |
| <b>Figura 4.1.....</b> | <b>127</b>    |
| <b>Figura 4.2.....</b> | <b>128</b>    |
| <b>Figura 4.3.....</b> | <b>129</b>    |
| <b>Figura 4.4.....</b> | <b>130</b>    |
| <b>Figura 4.5.....</b> | <b>131</b>    |
| <b>Figura 5.1.....</b> | <b>158</b>    |
| <b>Figura 5.2.....</b> | <b>159</b>    |
| <b>Figura 5.3.....</b> | <b>160</b>    |
| <b>Figura 5.4.....</b> | <b>162</b>    |

## Abreviações

aCGH = Hibridação Genômica Comparativa em *arrays*

AI = Instabilidade Alélica

BAC = Cromossomo artificial de bactéria

BCL2 = Gene da leucemia de células B/linfoma 2

B-LLC = Leucemia linfocítica crônica de células B

CBS = Segmentação binária circular

CCE = Carcinoma de células escamosas

CCECP = Câncer de células escamosas de cabeça e pescoço

CCEO = Carcinoma de células escamosas oral

CCND1 = Ciclina D1

CCNL1 = Ciclina L1

CGH = Hibridação genômica comparativa

CHCs = Carcinomas hepatocelulares

CIS = Carcinoma *in situ*

CNA = Alteração no número de cópias do DNA

CNV = Variação no número de cópias do DNA

COX-2 = Citocromo c Oxidase subunidade II

CPT = Carcinoma papilar de tireóide

CV = Coeficiente de variação

DIA1 = Diaforase (NADH) citocromo b-5 redutase

DOP-PCR = PCR de oligonucleotídeo degenerado

DSCNAs = Alteração no número de cópias na seqüência do DNA

E2F1 = Fator 1 de transcrição E2F

EBV = Vírus Epstein-Barr

EDD1 = Proteína ubiquitinina ligase E3

EGFR = Receptor do fator de crescimento epidermal

EMS1 = Receptor transmembrano da proteína kinase

ER = Estrogênio

ERBB2 = Homólogo humano do oncogene 2 de v-erb-b2 leucemia eritroblástica viral

FFPE = Fixada em formalina, embebida em parafina

FGFR1 = Receptor do fator de crescimento de fibroblasto 1

FISH = Hibridação *in situ* fluorescente

GST = Gene supressor tumoral

HGD = Displasia de alto grau

HMGA2 = Grupo de alta mobilidade A2

HPV = Vírus do papiloma humano

hrHPV = Vírus do papiloma humano de alto risco

KIT = Homólogo humano do oncogene de sarcoma viral v-kit Hardy-Zuckerman 4

LCM = Linfomas de células do manto

LGD = Displasia de baixo grau

LLC = Leucemia linfocítica crônica

LOH = Perda de heterozigose

LRP12 = Lipoproteína de baixa densidade 12

MARG = Grupo de Microarray

MDM2 = gene *Mdm2* de ligação à proteína p53



miRNAs = microRNAs

MMPs = Metaloproteinases

MSI = Instabilidade de microsatélite

MYC = Homólogo humano do oncogene viral mielocistomatoses v-myc

MYCL1 = Homólogo humano do oncogene de v-myc mielocistomatose viral 1

ncRNA = RNA não codificador de proteína

OPML = Lesão oral potencialmente maligna

P16 = Inibidor da quinase ciclina-dependente 2A

P21 = Inibidor da quinase ciclina-dependente 1A

PCA = Análise de componentes principais

PCR = Reação em cadeia da polimerase

PDCD4 = Morte celular programada 4 (inibidor de transformação neoplásica)

PDGF = Fator de crescimento derivado de plaquetas

PGS = Partek Genome Suite

PIK3CA = Fostfoinosítídeo quinase 3

piRNA = RNA Piwi

PR = Progesterona

PTEN = Homólogo humano da fosfatase e tensina

qPCR = PCR quantitativa em tempo real

QRT-PCR = PCR de transcrição reversa quantitativa em tempo real

RAS = Oncogene do sarcoma viral

rasiRNA = RNA com repetições associadas

RB = Retinoblastoma 1

RISC = Complexo indutor do silenciamento de RNA

SCOMP = Hibridação genômica comparativa de célula única

SEER = Programa de vigilância epidemiológica e de resultados finais

siRNA = RNA de interferência

SKY = Cariótipo espectral

smRNA = RNA modulatório

SNP = Polimorfismo de nucleotídeo único

sRNA = Pequeno RNA

TGF- $\alpha$  = Fator de crescimento transformante alfa

TGF- $\beta$  = Fator de crescimento transformante beta

TLDA = *TaqMan Low Density Arrays*

tncRNA = RNA anão

TP53 (P53) = Proteína tumoral P53

TP73 = Proteína tumoral P73

TPM1 = Tropomiosina 1

TRBP = Proteína conjugal de transferência

VEGF = Fator de crescimento endotelial vascular

WGA = Amplificação do genoma total

WHO = Organização Mundial da Saúde

## 1. Introdução

### *1. Câncer de Cabeça e Pescoço: definição e fatores etiológicos*

#### *1.1. Definição*

O câncer de cabeça e pescoço inclui carcinomas que se originam em diversos sítios anatômicos, tais como a cavidade oral, orofaringe, faringe, laringe e cavidade nasal; estes tumores, apesar de possuírem diferentes fatores etiológicos e prognósticos, compartilham alguns fatores de risco[1]. Aproximadamente 90% desses tumores são carcinomas de células escamosas (CCE), que se originam na superfície epitelial. Os carcinomas de células escamosas da cavidade oral (CCEO) são lesões malignas que compreendem mais de 50% de todos os cânceres de cabeça e pescoço e são a sexta causa de morte por câncer em todo o mundo[2].

Em 2002, surgiram aproximadamente 650,000 novos casos e 350,000 pacientes morreram do câncer de células escamosas de cabeça e pescoço (CCECP)[3]. Somente nos EUA, a estimativa foi de 45,660 novos casos e 11,210 mortes em 2007[4]. No Canadá, as taxas atingiram 10.200 novos casos e 2.102 mortes por cânceres de cabeça e pescoço no ano de 2007.

#### *1.2. Fatores etiológicos do câncer de cabeça e pescoço*

Os principais fatores etiológicos associados com o desenvolvimento de CCECP são o o consumo excessivo de tabaco e álcool. Porém, infecções virais e hábitos alimentares não saudáveis também podem estar associados com o surgimento de CCECP.

### *1.2.1. Consumo de tabaco e álcool*

Em 1957, o consumo de cigarro foi identificado como um fator de risco independente para o desenvolvimento de câncer oral e de orofaringe[5]. Posteriormente, o uso de produtos do fumo (isto é, cigarros, cachimbos e/ou charutos ou fumo de mascar), juntamente com o uso de álcool, foi confirmado como o principal fator de risco para o desenvolvimento de CCECP[6-9].

Em suporte a essa associação entre o consumo excessivo de cigarro e CCECP, existe evidência que associa o consumo de tabaco com um risco aumentado do desenvolvimento de displasia oral[10]. Além disso, estudos demonstraram que cessando o consumo de cigarro (por pelo menos 15 anos) resultava em uma redução no risco de desenvolver lesões orais pré-malignas e malignas; este risco reduzido foi semelhante àquele determinado em indivíduos que nunca fumaram[10]. Em estudo semelhante, o risco de desenvolver o câncer oral foi reduzido em 30% para indivíduos que pararam de fumar por 1 a 9 anos, e em 50% para aqueles que deixaram de fumar durante mais de 9 anos[9].

Apesar do consumo de cigarro ser, por si só, o fator de risco mais importante estabelecido para o desenvolvimento de CCECP, também existem relatos de que o consumo de álcool contribua para o aumento do risco de CCECP, especialmente em indivíduos com forte hábito tabagista. Vários estudos mostraram que o risco de desenvolver CCECP aumenta de acordo com a quantidade de tabaco e álcool consumida, sendo que os efeitos deste consumo agem de maneira sinérgica[11-14]. Além disso, os CCECP ocorrem com maior frequência em homens, geralmente com mais de 60 anos de idade[15]. Estudos na literatura têm demonstrado a associação entre o consumo de tabaco/álcool e a idade dos pacientes com CCECP. Em um estudo incluindo 1.648 pacientes com CCECP, apenas 40 pacientes (2.5%), com idade média de 60 anos (27-90 yrs), não possuíam história prévia de uso de álcool e/ou fumo[16]. Achados semelhantes

foram relatados por Koch e colaboradores[17], nos quais pacientes com CCECP foram estratificados em 46 não fumantes, 233 fumantes, e 29 ex-fumantes. Neste estudo, os não-fumantes possuíam um intervalo bastante variado entre as idades e um número maior de pacientes jovens (17-93), enquanto os grupos de ex-fumantes e ainda fumantes compunham-se de indivíduos mais velhos, 33-84 e 46-85 anos de idade, respectivamente. Apesar dos achados similares dos dois estudos, deve-se considerar cuidadosamente certas diferenças que eles trazem, tais como as definições usadas para fumantes/não-alcoólicas. Por exemplo, no estudo por Weiseman e colaboradores[16], os indivíduos não-fumantes foram definidos como aqueles que nunca tinham consumido fumo durante a sua vida, enquanto no estudo realizado por Koch e colaboradores[17], os não-fumantes foram definidos como indivíduos que não fumavam regularmente.

Embora a maioria dos casos de CCEOs seja atribuída ao abuso de tabaco e álcool, um estudo mostrou que 25% (29/116) dos pacientes, com menos de 45 anos, não possuíam estes fatores de risco[18]. Acredita-se que vias moleculares distintas de tumorigênese podem existir entre indivíduos que apresentam os fatores de risco para o desenvolvimento de CCEOs comparado a indivíduos que não estão expostos a carcinógenos do tabaco e álcool[16-17]. Estudos recentes realizados em nosso laboratório levaram à identificação de regiões genômicas distintas de ganhos e perdas de DNA em CCEOs de pacientes jovens (<40 anos) comparados a pacientes mais velhos (>60 anos) (Machado J e colaboradores, em preparação). Estas análises são importantes para determinar se os tumores orais de pacientes jovens possuem alterações genômicas que possam elucidar os mecanismos genéticos de tumorigênese neste grupo de pacientes que não apresenta fatores de risco cumulativo para o desenvolvimento do câncer oral.

### 1.2.2. Infecções virais

Os vírus são considerados fatores de risco ao desenvolvimento do câncer, uma vez que eles podem alterar o controle de eventos moleculares, tais como a regulação do ciclo celular. Os vírus mais comuns implicados na transformação dos CCECP são o vírus do papiloma humano (HPV) e os vírus da herpes[19-20]. Entre os vírus da herpes, o vírus Epstein-Barr (EBV) é considerado um fator etiológico dos carcinomas da nasofaringe[21-23]. Entretanto, o HPV é o fator etiológico viral mais estudado em CCECP[24].

Os HPV são vírus de DNA circulares de 8 KBs, que infectam especificamente as células basais da mucosa epitelial[25]. A família dos vírus HPV é composta por mais de 100 genótipos, sendo que os tipos virais HPV6, 11, 16, e 18 infectam células epiteliais da mucosa da cavidade oral, orofaringe, trato anogenital e cérvix uterino[25]. De acordo com sua capacidade de transformar células epiteliais, os vírus HPV podem ser classificados em HPV de alto risco e HPV de baixo risco; os tipos de alto risco se associam frequentemente com o desenvolvimento de lesões malignas[26]. Os HPV de alto risco podem codificar pelo menos 3 proteínas (E5, E6, e E7), que possuem potencial oncogênico devido à sua capacidade de afetar as funções dos genes supressores de tumor *P21*, *P53*, e *RB*, levando à disfunção de processos celulares como a apoptose, o reparo a danos no DNA e o controle do ciclo celular[27-29].

Aproximadamente 20-25% de CCECP contêm o vírus HPV com propriedades oncogênicas[30-31]. Os subtipos de HPV de alto risco mais comumente detectados em CCECP são o HPV 16 e 18[32-50].

Entre os CCECP, o HPV é considerado um fator de risco independente para o desenvolvimento do câncer de orofaringe[43, 51-54]. Recentemente, um estudo realizado em nosso laboratório, utilizando uma técnica robusta e sensível (*Roche linear array*), mostrou um

índice baixo (2/53, 4%) de positividade para HPV em CCEOs, e um maior número de casos (16/22, 73%) positivos para o HPV em carcinomas da orofaringe (Machado J e colaboradores, em preparação). Estes dados concordam com a literatura atual de HPV em CCECP[55], que sugerem que o HPV esteja principalmente associado a carcinomas da orofaringe, e que não deve desempenhar um papel relevante na tumorigênese dos CCEOs.

## *2. Carcinomas de Células Escamosas da Cavidade Oral*

### *2.1. Definição e Prevalência*

Os carcinomas de células escamosas da cavidade oral (CCEO) representam mais de 50% de todos os cânceres de cabeça e pescoço e são a sexta causa de morte por câncer em todo o mundo[56]. Globalmente, as regiões de alta incidência de CCEO incluem a Melanésia (uma subregião da Oceania, ao nordeste da Austrália), região centro-sul da Ásia, Sudoeste Europeu e África do Sul[3]. Apesar das taxas de incidência e mortalidade por cancer oral serem menores nos EUA, quando comparadas com a média mundial, aproximadamente 22.900 novos casos e 5.390 mortes foram relatados em 2008[57]. No Canadá, 3.200 novos casos foram diagnosticados e 1.100 mortes foram relatadas por câncer oral no ano de 2008[57].

A América do Sul possui uma das taxas mais altas de câncer de boca (18,7 por 100.000 habitantes)[58]. No Brasil, a análise da distribuição proporcional das causas de morte mostra que, desde a década de 30, as neoplasias, juntamente com as doenças do aparelho circulatório, vêm sendo responsáveis por um número cada vez maior de óbitos, indicando uma mudança no perfil da mortalidade, semelhante àquela observada nos países desenvolvidos[3]. O Instituto Nacional do Câncer classifica os carcinomas orais como os de maior incidência na região sudeste do Brasil: em 2008/2009, a estimativa de foi de aproximadamente 8.010 novos casos de câncer

entre homens e mulheres ([www.inca.gov.br](http://www.inca.gov.br)); para um total de aproximadamente 14.160 novos casos em todo o Brasil.

De acordo com dados descritos pelo “*National Cancer Institute's Surveillance, Epidemiology and End Results Program (SEER)*”, EUA, idade média dos pacientes diagnosticados com CCEO foi de 63 anos, e a idade média para a morte por este câncer foi de 68 anos, entre 1997 e 2002[59]. No mundo, a incidência de CCEOs em homens é mais alta do que em mulheres; porém, devido a recente quedas na incidência destes carcinomas no sexo masculino, a disparidade entre os gêneros têm diminuído progressivamente, principalmente por mudanças no estilo de vida das mulhes nas últimas décadas. A incidência de CCEO em homens, ajustada para idade, diminuiu de 21,2 para 15,9 por 100.000 (média: 18,8 por 100.000) entre 1975 e 2002, enquanto para mulheres, essa incidência diminuiu menos, de 7,1 para 6,5 por 100.000 (média: 7,2 por 100.000)[59].

Diferenças na incidência de câncer oral também são observadas entre etnias diferentes, por exemplo, os homens afro-americanos têm uma incidência mais alta de câncer do que os caucasianos (razões ajustadas para idade [1992-2002]: 20,7 por 100.000), e as taxas de mortalidade em afro-americanos são aproximadamente duas vezes mais altas (razões ajustadas para idade [1992-2002]: 8,2 por 100.000)[59-60].

## *2.2. Prognóstico e Tratamento de Pacientes com CCEO*

Os CCEOs são caracterizados por possuírem pobre prognóstico e baixa taxa de sobrevida[61]. Embora os pacientes com CCEO tenham se beneficiado dos últimos avanços em terapia radioativa, quimioterapia e técnicas cirúrgicas, a taxa de sobrevida é de 5 anos para aproximadamente 50-60% dos pacientes com CCEO (SEER database 2009), sendo ainda



considerada uma doença de curso fatal[62-63]. Isso acontece especialmente pela ocorrência do diagnóstico tardio e pela recorrência do tumor primário. Estes fatores reforçam a importância de estudos direcionados ao conhecimento dos eventos genéticos associados a progressão dos CCEO. A determinação desses eventos genéticos, bem como o enfoque na busca de novos métodos de detecção precoce, prognóstico e terapêutica do câncer oral, seriam fundamentais para a melhoria da sobrevida dos pacientes com CCEO.

Apesar de estudos demonstrarem que a transformação maligna resulta do acúmulo de danos em genes e/ou vias genéticas, os eventos genéticos envolvidos na tumorigênese oral ainda não estão totalmente esclarecidos[64]. Neste aspecto, é muito importante o esforço contínuo na busca de biomarcadores em tumores orais que irão contribuir enormemente para a melhor caracterização e, conseqüentemente, para a melhoria na resposta ao tratamento desses carcinomas.

Os principais fatores prognósticos dos CCEOs são a presença de metástase em linfonodos, invasão vascular ou linfática, margens cirúrgicas positivas (exemplo: presença de células displásicas) e extensão extracapsular de células tumorais dos linfonodos envolvidos[65]. A progressão metastática de células tumorais é um processo complexo de múltiplas etapas[66]. Cada etapa parece envolver interações moleculares entre as células tumorais e o microambiente circundante, que tem sido cada vez mais estudado com o uso da tecnologia genômica e proteômica[67]. Estudos moleculares utilizando microarranjos de DNA (DNA *microarray*) sugerem proteínas putativas associadas a um prognóstico pobre dos CCEOs. Entre as proteínas identificadas estão várias queratinas, proteases da superfície celular, marcadores mesenquimais celulares, moléculas de adesão celular, quemoquinas, e fatores implicados na modulação da matriz extracelular e transição epitelial-mesenquimal[68-72].

O tratamento dos CCEOs inclui cirurgia, radioterapia e quimioterapia. A remoção cirúrgica completa do tumor é o fator prognóstico mais importante em CCEO[73]. Os pacientes com estadiamentos avançados (doença nos estágios III ou IV) normalmente recebem uma combinação de cirurgia, radiação pós-operatória e/ou quimioterapia. Os pacientes que apresentam doença nos estágios I ou II geralmente são tratados com cirurgia e/ou radioterapia, e possuem um melhor prognóstico[74]. Contudo, tais pacientes ainda possuem um alto risco de recorrência do tumor primário, e por isso devem ser monitorados. Embora ainda em etapas iniciais, alguns estudos têm examinado a utilidade da terapia molecular para o tratamento de cânceres de cabeça e pescoço; os alvos testados nestes estudos incluem *VEGF*, *PDGF*, *FGF*, *TGF- $\alpha$*  and *- $\beta$*  [75-77]. O uso de drogas contra estes alvos genéticos, com propriedades antiangiogênicas ou antimetastáticas, pode levar à uma melhora na sobrevida dos pacientes com CCEO.

### *3. Cancerização de Campo e a Progressão de CCEO*

O fenômeno conhecido como cancerização de campo foi primeiramente descrito por Slaughter e colaboradores[78], em um esforço para entender como as alterações histológicas em tecido epitelial de epitélios adjacente ao tumor estariam relacionadas com o desenvolvimento do câncer oral em áreas multifocais pré-malignas, e como a persistência do tecido anormal após a cirurgia pode levar ao desenvolvimento de uma segunda malignidade. Quanto a alterações moleculares, o fenômeno de cancerização de campo possui implicações clínicas importantes[79], uma vez que um campo da lesão é considerado pré-neoplásico e pode conter alterações genéticas que precedem as alterações fenotípicas. De fato, a progressão molecular de CCECP está correlacionada com o grau de displasia das lesões pré-malignas[80].

Em relação à progressão de CCECP, o fenômeno de cancerização de campo postula que, num primeiro evento, "uma mancha" pode se desenvolver no tecido normal. Esta "mancha" pré-neoplásica está geralmente associada com a perda de heterozigose (LOH) em 17p e um clone de células TP53-transformadas. O evento seguinte é a extensão "da mancha" a um campo de cancerização, que consiste de uma lesão epitelial na qual as células possuem alterações genéticas responsáveis pelo desenvolvimento do câncer. Entre os eventos genéticos associados com a progressão tumoral de CCECP (por exemplo, LOH em 3p, 9p, 8p, 18q), a amplificação em 11q13 foi um dos principais eventos associado com estágios tardios da progressão[81]. Desse modo, a presença de um campo com células geneticamente alteradas parece ser um fator de risco contínuo para carcinogênese[79]. Com respeito a marcadores moleculares, os achados preliminares em lesões orais potencialmente malignas mostraram que a presença de LOH em 3p e 9p dentro destes campos geneticamente alterados estava associada com um risco aumentado para a transformação maligna[82-84].

O conceito de múltiplas etapas da carcinogênese é atualmente bem estabelecido, e é definido pela acumulação de alterações genéticas que levam à progressão de uma célula normal a uma tumoral[85]. Uma vez que até 62% dos CCEOs podem resultar de lesões pré-malignas[86], a avaliação das alterações genéticas nessas lesões potencialmente malignas (OPMLs) e CCEOs correspondentes, pode ser capaz de identificar as importantes vias genéticas envolvidas na progressão da doença. **A identificação de novos marcadores com poder de predição da transformação maligna é útil em nível diagnóstico e de tratamento, resultando na melhora da sobrevida dos pacientes com CCEO.**

#### *4. Lesões Oraís Potencialmente Malignas (OPMLs)*

#### *4.1. Definição e Prevalência Mundial*

De acordo com dados recentes[87], lesões pré-malignas na cavidade oral são definidas como ‘lesões orais potencialmente malignas’ (OPML). Evidências sugerem que alterações genéticas são responsáveis pela progressão de OPMLs a carcinoma invasivo[79, 88]. As lesões orais mais comuns com propensão a transformação maligna são as leucoplasias e as eritropias[89]. Outras lesões, tais como linha alba, leucoedema e queratose de atrito são comuns na cavidade oral, mas raramente sofrem transformação maligna.

Embora a leucoplasia e a eritropiasia são consideradas lesões potencialmente malignas[90], o uso de características clínicas para classificar estas lesões é difícil, uma vez que elas variam em aparência. O diagnóstico histopatológico é geralmente mais indicativo de uma modificação pré-maligna do que alterações clinicamente evidentes. Neste contexto, durante os anos, tem ocorrido várias definições para leucoplasia oral, como apresentado na tabela 1.1.

Leucoplasia oral é atualmente definida como "uma mancha ou placa de aspecto esbranquiçado na mucosa oral, que possui risco duvidoso, uma vez excluídas (outras) doenças conhecidas ou desordens que não possuem nenhum risco aumentado de transformação maligna[89].

**Tabela 1.1.** Definições de leucoplasias orais que foram propostas nas décadas passadas, e definição atualmente utilizada (modificado de [89]).

| <b>Grupo de Pesquisa</b>  | <b>Definição</b>  | <b>Referência</b> |
|---|---|-------------------|
| WHO (1978)  | <p><b>“Uma mancha placa branca na mucosa oral que não pode ser clinicamente ou patologicamente caracterizada como qualquer outra doença”</b></p> <p><i>“A white patch or plaque that cannot be characterized clinically or pathologically as any other disease”</i></p>   | [91]              |
| I Conferência Internacional em Leucoplasias Orais. Malmö, Sweden (1984) | <p><b>“Uma mancha ou placa branca na mucosa oral que não pode ser clinicamente ou patologicamente caracterizada como qualquer outra doença, e que não esteja associada com nenhum agente etiológico físico ou químico, exceto o uso do fumo”</b></p> <p><i>“A white patch or plaque that cannot be characterized clinically or pathologically as any other disease and is not associated with any physical or chemical causative agent except use of tobacco”</i></p> | [92]              |
| Simposio Internacional. Uppsala, Sweden (1994)                          | <p><b>“Uma lesão predominantemente branca na mucosa oral que não pode ser caracterizada como nenhuma outra doença definível”</b></p> <p><i>“A predominantly white lesion of the oral mucosa that cannot be characterized as any other definable disease”</i></p>  | [90]              |
| WHO (1997)  | <p><b>“Uma lesão predominantemente branca na mucosa oral que não pode ser caracterizada como nenhuma outra doença definível”</b></p> <p><i>“A predominantly white lesion of the oral mucosa that cannot be characterized as any other definable lesion”</i></p>   | [93]              |
| WHO (2005)  | <p><b>“Não definida – nenhuma distinção é feita de outras lesões de placas brancas”</b></p> <p><i>“Not defined – no distinction is made from other white patch lesions”</i></p>   | [94]              |
| Warnakulasuriya e colaboradores (2007)                                  | <p><b>“Leucoplasia deve ser utilizada para definir feridas brancas de risco duvidoso, uma vez excluída (outras) doenças conhecidas que não trazem risco aumentado para o câncer”</b></p> <p><i>“Leukoplakia should be used to define white plaques of questionable risk having excluded (other) known diseases or disorders that carry no increased risk for cancer”</i></p>  | [89]              |

A prevalência mundial relatada para leucoplasias orais tem sido contrastante na literatura; relatos anteriores citam uma prevalência de aproximadamente 2%[95], enquanto um estudo mais recente (baseado apenas na transformação maligna das leucoplasias orais) modificou esta prevalência para aproximadamente 0.5%[96].

As lesões leucoplásicas têm uma taxa de transformação de até 18%[97-99], embora a frequência da transformação maligna não seja a mesma para todas as formas de leucoplasias[100-101]. Há relatos na literatura de que uma proporção (16-62%) dos carcinomas orais podem se desenvolver a partir de OPMLs, tais como a leucoplasia[86, 102].

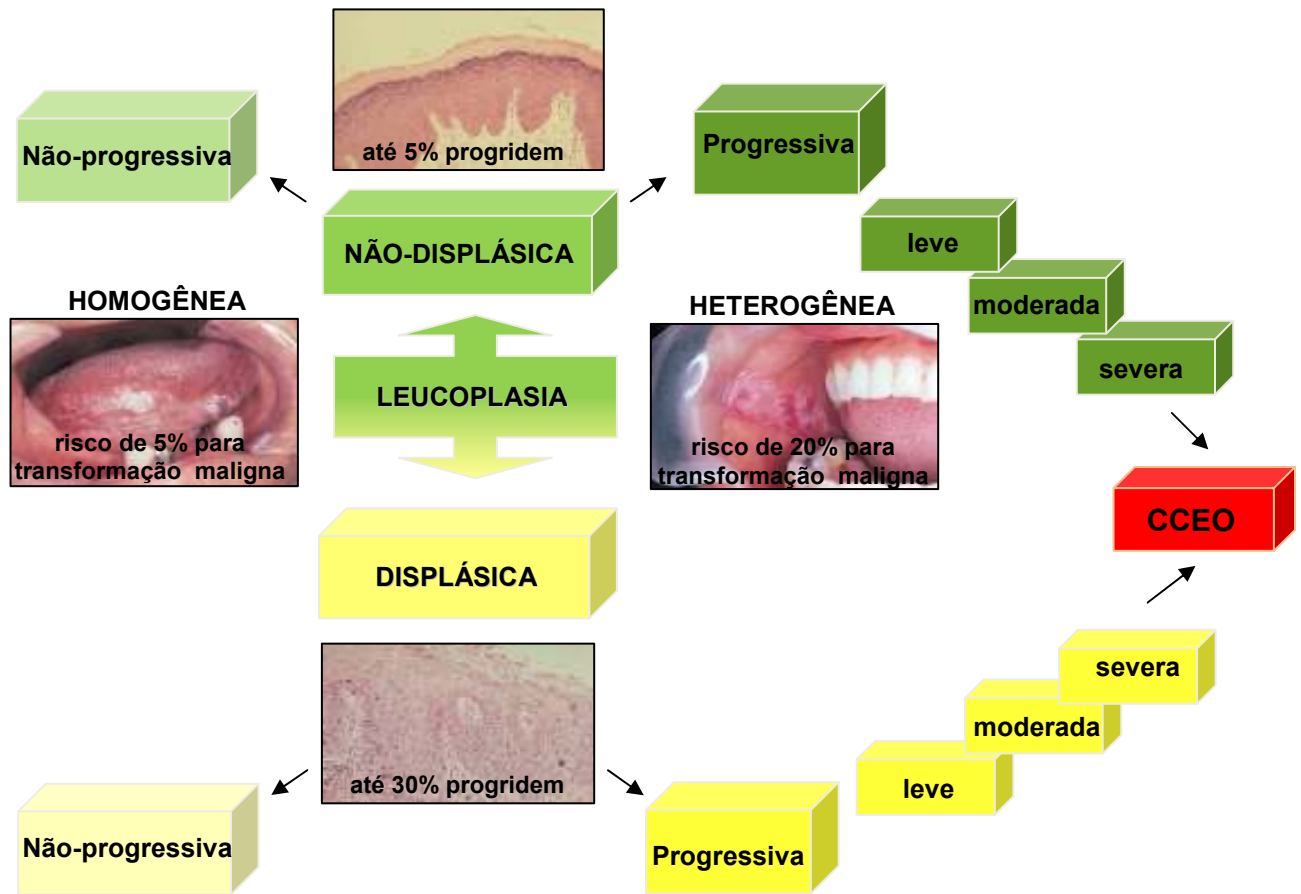
As eritroplasias consistem em OPMLs com alto potencial de transformação maligna. Estas lesões apresentam coloração avermelhada e não podem ser diagnosticadas, clínica ou histopatologicamente, como outras condições, como por exemplo, aquelas originadas a partir de causas traumáticas, vasculares ou inflamatórias[103-104]. Por serem muito menos comuns do que as leucoplasias, poucos estudos são inteiramente dedicados às eritroplasias[104-105]. Apesar disso, as taxas de transformação de eritroplasias são relatadas como as mais altas entre todas as lesões orais pré-malignas[103]. Alguns estudos sugerem ainda que as eritroplasias podem se desenvolver ‘de novo’ ou a partir de leucoplasias orais por várias etapas intermediárias das lesões de aspecto esbranquiçado/avermelhado[103]. Porém, mais estudos são necessários para esclarecer a origem das eritroplasias.

#### *4.2. Apresentação Clínica e Classificação Histológica de Leucoplasias Orais*

A literatura em CCEO relata várias tentativas para classificar corretamente as leucoplasias segundo o seu potencial de transformação maligna[90]. Atualmente, elas são

classificadas baseadas na avaliação clínica e histopatológica (Figura 1.1). Clinicamente, existem dois tipos de leucoplasias: homogênea e não-homogênea[106]. As leucoplasias homogêneas são lesões predominantemente branca de aparência chata, fina, uniforme e com superfície lisa, franzida ou enrugada de textura constante[90]. As leucoplasias não-homogêneas são predominantemente brancas ou de coloração branco-avermelhada com aparência irregularmente chata ou nodulosa[107], e possuem um risco mais alto de transformação maligna[98-99, 108]. De fato, a presença da vermelhidão ou textura verrucosa pode indicar presença de malignidade[109]. Quanto aos sintomas associados com estas lesões, o tipo homogêneo é normalmente assintomático, enquanto que o tipo não-homogêneo se associa muitas vezes com dor localizada e/ou desconforto[109].

Histologicamente, lesões leucoplásicas são classificadas como não-displásicas ou displásicas (leve, moderada ou severa)[108], e a presença de displasia epitelial parece estar associada com um risco aumentado de transformação de até 31%[110]. Embora as características clínicas e histológicas das leucoplasias sejam úteis para definir o diagnóstico clínico, tais características possuem valor limitado para prever quais leucoplasias progredirão à carcinoma.



**Figura 1.1.** Classificação clínica e histológica de leucoplasias orais proposto pelo WHO (*World Health Organization*). Clinicamente estas lesões podem se classificadas como homogêneas ou não-homogêneas; e histologicamente como displásicas ou não-displásicas (ambos os grupos com possibilidades de progredir a lesão leve, moderada, severa, ou carcinoma *in situ* para CCEO invasivo). Os critérios para tal classificação consideram a presença de atipia celular, e as características arquitetônicas baseadas na espessura das camadas displásicas comparadas com a espessura total do epitélio normal.



#### 4.3. Fatores de Risco da Transformação Maligna

Acredita-se que os CCEOs se desenvolvam através de estágios seqüenciais que evoluem de lesões premalignas, tais como displasia leve, moderada e grave, para carcinomas invasivos.

Os critérios histológicos são razoavelmente eficazes na determinação do risco de transformação maligna das lesões pré-invasivas de alto grau displásico (displasia severa e carcinoma *in situ*); a maioria destas lesões, quando não tratada, progride à carcinoma invasivo[111]. Entretanto, lesões sem displasia ou com displasia mínima representam a maioria das OPMLs, e somente uma pequena porcentagem destas lesões progride ao câncer.

Alguns estudos têm examinado o risco de câncer em diferentes graus de displasias; Napier SS e colaboradores[100] relataram um valor preditivo positivo de desenvolvimento de CCEO em pacientes cuja biópsia inicial mostrou displasia. Em outro estudo, o risco de progressão ao câncer, em indivíduos com displasia moderada ou severa, se mostrou duas vezes maior (OR=2.30; 95%CI = 0.8-3.4) do que aquele observado para displasias de grau leve ou hiperplasias[86]. A presença de leucoplasia com displasia moderada ou severa foi associada com um risco significativamente maior ( $p < 0.01$ ) de desenvolver CCEO do que leucoplasias com um baixo grau de displasia[98]. No geral, o valor preditivo de transformação maligna de displasias depende da prevalência da leucoplasia em uma dada população. A maioria dos estudos sugere que o estado histológico da displasia, embora limitado, seria um marcador útil para o manejo e o monitoramento destas desordens potencialmente malignas.

Considerando que os critérios histológicos são insuficientes para predizer quais destas lesões apresentam risco aumentado de transformação maligna, o entendimento das alterações genéticas moleculares em OPMLs possui grande potencial para a caracterização de biomarcadores de progressão.

#### 4.4. Fatores Etiológicos das Leucoplasias Orais

O uso de fumo e a infecção por *Candida albicans* são os principais fatores etiológicos relatados na literatura para o desenvolvimento de leucoplasias, e ambos têm sido relacionados com o prognóstico dos pacientes[108].

Além de serem lesões comuns em fumantes, as leucoplasias também ocorrem em pacientes com a história prévia de câncer de cabeça e pescoço[112]. Em um estudo prospectivo abrangendo várias áreas geográficas, as taxas anuais de incidência de leucoplasia, ajustadas pela idade, variaram de 1,1 a 2,4 entre homens e de 0,2 a 1,3 entre mulheres por 1.000 indivíduos (0.2 a 4.9%)[113]. Quanto à idade, lesões leucoplásicas geralmente ocorrem depois dos 30 anos, com uma incidência máxima até os 50 anos de idade[112, 114].

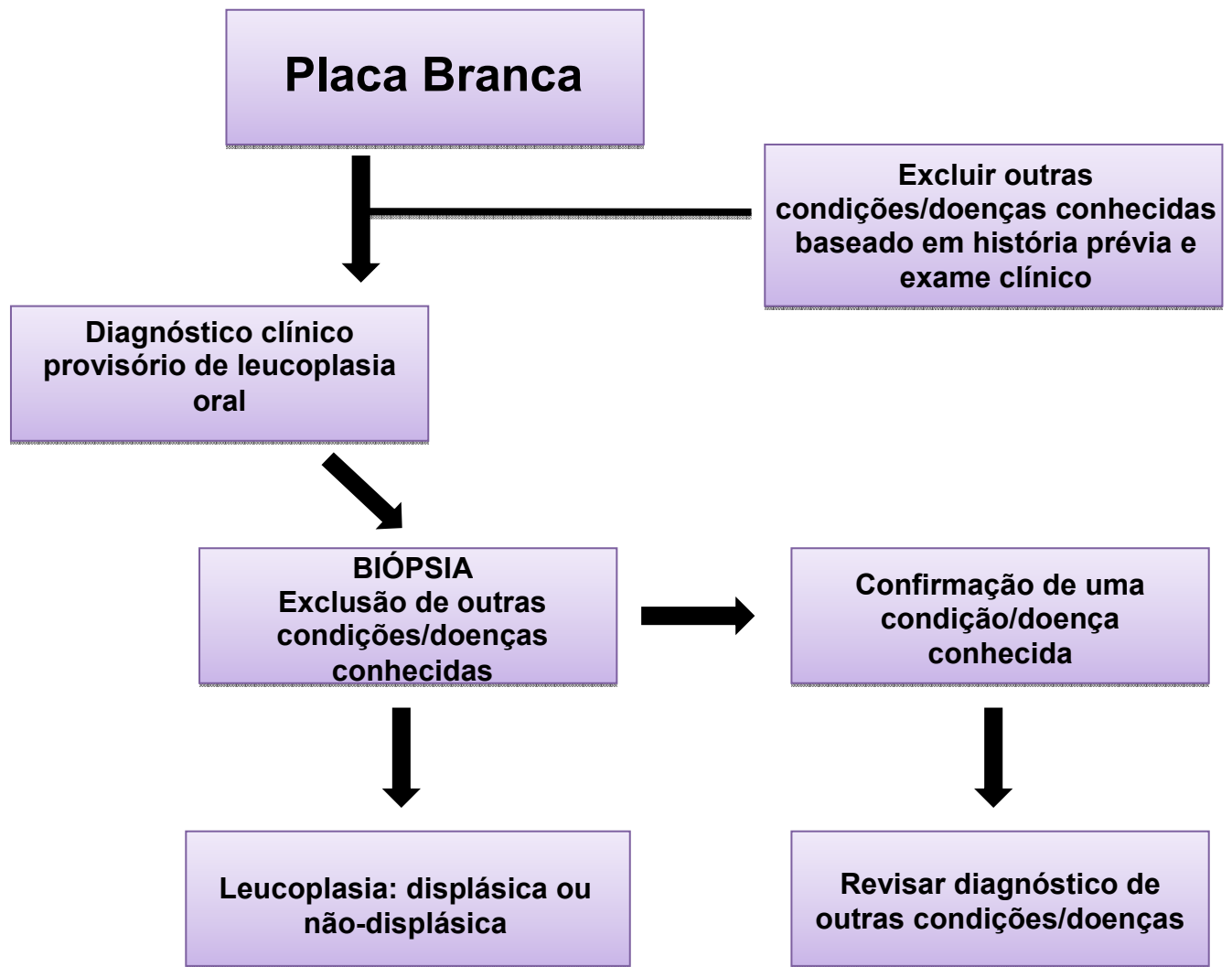
Tanto o soalho da boca como a língua são os dois principais sítios de alto risco para o desenvolvimento de leucoplasias[108]. Estudos mostraram uma associação entre o tabagismo e o desenvolvimento de leucoplasias no soalho da boca, sendo que as leucoplasias na borda da língua são mais comuns entre não fumantes[109, 115].

A associação entre infecção por *Candida albicans* e o risco do desenvolvimento de CCEO foi originalmente relatada em leucoplasias não-homogêneas[116-117], e em leucoplasias displásicas[118-119]. Estudos com modelo animais mostraram que *Candida albicans* possui um efeito promotor da carcinogênese, contribuindo para o desenvolvimento de carcinomas na mucosa oral de ratos[117]. Resultados deste estudo suportam a hipótese de que *C. albicans* pode estar associada com a transformação neoplástica.

#### *4.5. Diagnóstico e Monitoramento das Leucoplasias Orais*

Os critérios apropriados para o diagnóstico e o monitoramento das leucoplasias orais têm sido alvo de debate[109, 120]. Em um esforço para se obter um consenso sobre o monitoramento destas lesões, vários sistemas de classificação foram propostos. Dois dos sistemas atualmente utilizados incluem a aparência clínica e o diagnóstico histopatológico como determinantes do estadiamento das leucoplasias. Outros sistemas de classificação, propostos por Schepman e van der Waal[121], também incluem o sítio anatômico da lesão; que considera soallho de boca e língua como os sítios mais comuns de ocorrência das leucoplasias.

Outro sistema de estadiamento de leucoplasias orais, combinando achados clínicos e histopatológicos, foi proposto por van der Waal e Axell[122]. Tal sistema objetivou promover uma avaliação uniforme do diagnóstico de leucoplasia, incluindo eventos de transformação maligna. Primeiramente, este sistema considera um diagnóstico provisório de leucoplasia quando uma lesão predominantemente branca não pode ser claramente diagnosticada, no exame clínico, como outra doença ou condição patológica da mucosa oral. Posteriormente, um diagnóstico definitivo é dado quando qualquer fator etiológico é excluído, e a histopatologia ainda não confirmou nenhuma outra desordem específica. O processo do reconhecimento de uma leucoplasia oral está ilustrado na Figura 1.2. Nos casos onde outra doença não é confirmada, a lesão é caracterizada como displásica ou não-displásica, após subsequente biópsia.



**Figure 1.2.** Esquema ilustrando as etapas utilizadas no diagnóstico de leucoplasias orais (modificado de[89]).

O uso de outras técnicas para o diagnóstico de leucoplasia oral, tais como a citologia esfoliativa bucal, é ainda controverso[123-124]. A citologia esfoliativa oral pode ter valor como ferramenta para *screening*, mas o exame histopatológico ainda permanece como uso padrão de referência no diagnóstico de leucoplasias. O mesmo se aplica ao uso do azul de toluidina[125-126], que pode ser útil para identificar desordens potencialmente malignas com risco de progressão a CCEO[127]. Tecnologias mais sofisticadas, tais como a espectroscopia ótica[128], a visualização de fluorescência direta[129-130] e a avaliação visual da luminescência da cavidade oral após reações quimiluminescentes[131], estão se tornando instrumentos diagnósticos não-invasivos promissores para a identificação de malignidade. Na cavidade oral, a visualização de autofluorescência tem se mostrado promissora na detecção de malignidade que não são evidentes no exame clínico (resultados do nosso grupo: Barker e colaboradores, em preparo para publicação).

Quanto ao monitoramento e o tratamento de leucoplasias displásicas, a remoção cirúrgica é geralmente o procedimento adotado[112]. Pequenas leucoplasias displásicas podem ser facilmente removidas sob anestesia local, e submetidas à avaliação histopatológica. A excisão à raio laser também pode ser realizada com pouca reação da mucosa local[132]. Outras opções de tratamento incluem a destruição por eletrodissecção e crioterapia com nitrogênio líquido[112].

O monitoramento de leucoplasias orais pode ser afetado por vários fatores, tais como a presença e duração dos sintomas, dúvidas quanto a presença de fatores etiológicos relacionados, e história prévia de câncer oral[121]. Dessa maneira, a avaliação acurada destas lesões é uma ferramenta importante para reduzir a sua ocorrência e prevenir a progressão destas lesões a carcinomas.

Apesar dos métodos de diagnóstico e prognóstico disponíveis, uma porcentagem das leucoplasias orais progride a CCEO, e quando isto ocorre a sobrevida dos pacientes por mais de cinco anos é baixa[63, 133]. Portanto, a acurácia do prognóstico é importante em pacientes com leucoplasia oral, desde que o conhecimento dos fatores de risco e a intervenção precoce apropriada, podem evitar a progressão para uma doença invasiva[134-135].

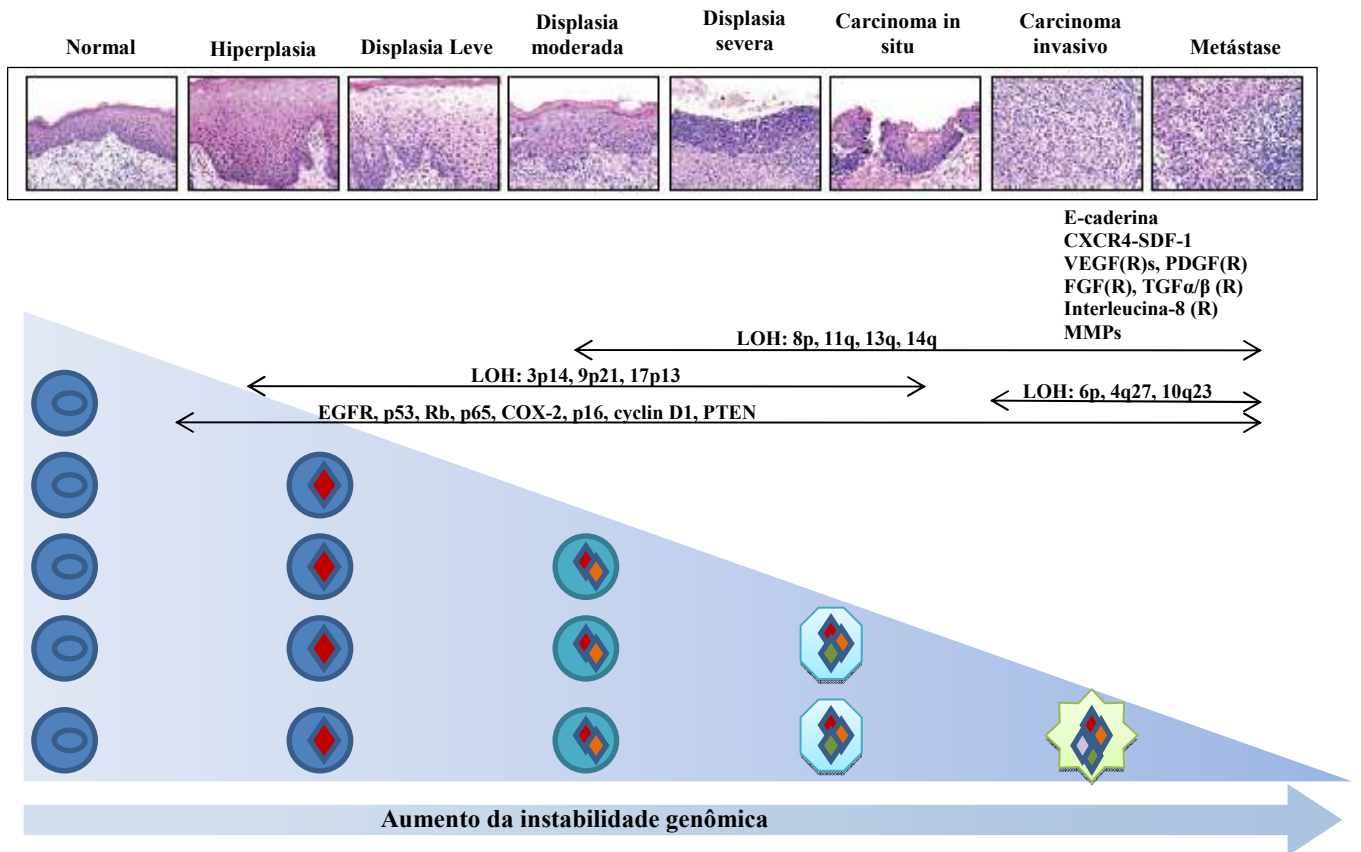
### *5. A Progressão ao Câncer Oral*

O câncer é considerado ser resultado do acúmulo de alterações genéticas que levam à expansão clonal de uma população de células transformadas. Acredita-se que a progressão gradual da tumorigênese está associada a três princípios básicos: (I) o câncer pode resultar da mutação de genes supressores tumorais (TSGs), proto-oncogenes e genes de reparo do DNA[136]; (II) há uma ordem de eventos genéticos que leva à carcinogênese; e (III) a variação na ordem dos eventos genéticos é possível, embora o acúmulo de mutações genéticas invariavelmente resulte no desenvolvimento tumoral[137].

Na cavidade oral, acredita-se que a transformação maligna ocorra devido a múltiplas alterações graduais de vias celulares e moleculares ocorrentes nas células da mucosa epitelial. Estas alterações geralmente estão associadas a características histológicas seqüenciais desde o epitélio normal a hiperplasia, displasia leve, displasia moderada, displasia severa e carcinoma invasivo[138-141] (Figura 1.3). Evidências indicam que as alterações genéticas e epigenéticas se acumulam durante o desenvolvimento tumoral maligno, e, na mucosa oral, isto é representado por alterações histológicas tais como a presença de displasia[142]. Estudos anteriores relataram alterações moleculares associadas com a progressão de lesões pré-malignas a carcinomas orais

invasivos[88, 143-144] (Figura 1.3). Contudo, o estudo da progressão genética utilizando amostras progressivas seqüenciais ainda não foi demonstrada.

Califano e colaboradores[145] propuseram um modelo de progressão genética para o câncer de cabeça e pescoço, baseado na análise de instabilidade de microsátélites (MSI) em amostras retrospectivas de arquivo. Esta análise identificou a freqüência de alterações genéticas em displasias e carcinomas de pacientes diferentes. Segundo tal modelo, uma célula progenitora adquire uma alteração genética que é transmitida a suas células filhas. Subseqüentemente ocorre a expansão dessas células, devido ao acúmulo de alterações genéticas adicionais, que transformam as células de um estado pré-neoplásico para um estado maligno. Este estudo sugeriu que certos eventos genéticos (9p21 LOH, LOH de 3p, e 17p13 LOH) tendem a ocorrer em estágios iniciais da progressão de tumores orais, enquanto outras alterações genéticas (LOH em 13q11, 8) ocorreriam preferencialmente em estágios tardios da progressão.



**Figura 1.3.** O painel superior mostra o diagnóstico histológico das amostras coradas com HE (hematoxilina, Eosina) em paralelo com as alterações genéticas presentes em lesões pré-malignas que progrediram a carcinomas. No painel inferior, uma representação esquemática ilustra como um grupo de células normais adquire a primeira alteração genética, seguida de alterações adicionais que levam a uma expansão clonal, e a subsequente transformação dessas células de um estado pré-maligno para um estado maligno, durante a progressão tumoral.



A progressão genética do câncer também inclui a desregulação transcricional, como consequência do acúmulo de alterações genéticas. O modelo de progressão transcricional, semelhante ao de progressão genética, propõe que a maioria das alterações ocorra antes da malignidade e que a desregulação transcricional ocorra durante a progressão de um estado normal para pré-maligno, e então para um estado maligno. Ha e colaboradores[146] mostraram que o maior número de alterações transcricionais (mRNA) ocorrem durante a transição da mucosa normal à pré-maligna, quando comparados com os eventos de alteração na expressão gênica que ocorrem na transição de tecidos pré-malignos a tumores malignos. Esses resultados sugerem que muitas das alterações genéticas que se acumulam durante a progressão do câncer oral ocorrem preferencialmente nos estágios iniciais da doença. Neste aspecto, os tecidos pré-malignos já possuiriam muitas das alterações transcricionais, mutações e perdas de material genético identificadas em CCEO.

Já que a avaliação clínica e histológica das leucoplasias orais possui um valor prognóstico limitado[147-148], a caracterização de eventos genéticos dessas lesões pré-malignas pode ajudar na identificar de lesões que possuam um risco aumentado de transformação maligna[149]. Esses estudos levariam à intervenção precoce e, conseqüentemente, à melhoria da sobrevida dos pacientes com CCEO.

### *5.1. Anormalidades Genéticas e a Progressão do Câncer Oral*

As aberrações genéticas associadas com a progressão do câncer oral estão apresentadas na Tabela 1.2. A maioria dos estudos de investigação de anormalidades genéticas associadas com progressão de carcinomas orais levaram à identificação de instabilidade alélica (AI), alterações do conteúdo de DNA e de perda de heterozigose (LOH). A ocorrência de AI em mais de duas

regiões genômicas foi significativamente ( $P=0.004$ ) associada com um risco aumentado de progressão[150], assim como alterações no conteúdo de DNA em lesões não-diplóides mostrou risco de progressão ao câncer[151]. Além disso, estudos de LOH em várias regiões genômicas (3p, 8p, 9p e 11q) sugeriram um risco aumentado de progressão ao câncer[84, 152].

Mao e colaboradores[82] investigaram a associação de LOH em 3p e 9p com a progressão do câncer oral. Eles mostraram que a presença de LOH em 9p21 e/ou 3p14 em OPMLs está associada com uma maior probabilidade de progressão da doença; 7 de 19 casos (37%) que possuíam LOH progrediram a CCEO, em comparação com somente 1 de 18 casos (6%) que progrediram e não apresentavam LOH nestas regiões genômicas.

Com o objetivo de investigar LOH em vários cromossomos durante os estágios iniciais de carcinogênese, outro estudo examinou um grupo 116 biópsias de pacientes com OPMLs (hiperplasia ou displasia leve/moderada) para LOH em 7 braços cromossômicos (3p, 4q, 8p, 9p, 11q, 13q e 17p), e associou estas alterações com um risco aumentado de progressão[153]. Nenhum dos pacientes possuía história de câncer antes do surgimento das OPMLs; em 29 pacientes, essas lesões progrediram para CCEO. Quando comparadas com lesões que não progrediram ao câncer, as lesões progressivas mostraram não só um número significativamente mais alto de LOH, mas também padrões de LOH característicos. Um achado interessante indicou que a maioria (97%) das lesões progressivas possuía LOH em 3p e/ou 9p, sugerindo que a perda nestes braços cromossômicos seria um pré-requisito para a progressão.

Outro estudo[83] também identificou LOH em 3p ou 9p em OPMLs que progrediram a CCEO. Este estudo incluiu biópsias de 78 pacientes com hiperplasia ou displasia, e sem história prévia de câncer oral. A progressão a carcinoma foi observada em 50% destes pacientes, normalmente em regiões dentro ou adjacente às OPMLs originalmente identificadas. Perdas em

9p ou 3p foram observadas em 94% das lesões que progrediram ao câncer, com maior frequência comparada a lesões que não progrediram. Contudo, a presença de LOH nas regiões 3p e/ou 9p é insuficiente para a transformação maligna. Alterações genéticas adicionais, inclusive LOH em outros braços cromossômicos, seriam provavelmente necessárias e aumentariam significativamente o risco de câncer oral.

Em uma revisão sistemática dos estudos de biomarcadores em displasias da cavidade oral[154] a expressão alterada da proteína TP53 nas camadas basais e suprabasais foi associada com a progressão ao câncer. Esta mesma revisão incluiu estudos que sugeriram que a desregulação na expressão de TP73[155], metaloproteinases (MMPs)[156] e survivina[157] estavam associadas com um risco aumentado de progressão de displasias orais a carcinoma.

**Tabela 1.2.** Marcadores associados com a transformação maligna de displasias para carcinomas orais.

| <b>Marcadores</b>                    | <b>Referência</b>              |
|--------------------------------------|--------------------------------|
| <b>P53</b>                           | Rich AM et al (1999)[158]      |
|                                      | Cruz IB et al (1998)[159]      |
|                                      | Shahnavaz SA et al (2000)[160] |
|                                      | Regezi JA et al (1995)[161]    |
|                                      | Hogmo A et al (1998)[162]      |
| <b>LOH 3p±9p</b>                     | Rosin MP et al (2000)[84]      |
| <b>3p</b>                            | Rosin MP et al (2000)[84]      |
| <b>9p</b>                            | Rosin MP et al (2000)[84]      |
|                                      | Zhou X et al (2005)[152]       |
| <b>8p</b>                            | Rosin MP et al (2000)[84]      |
|                                      | Zhou X et al (2005)[163]       |
| <b>11q</b>                           | Rosin MP et al (2000)[84]      |
|                                      | Zhou X et al (2005)[163]       |
| <b>AI&gt;2</b>                       | Partridge M et al (1998)[150]  |
| <b>Survivina (mRNA)</b>              | Lo Muzio L et al (2003)[157]   |
| <b>MMP1 (mRNA)</b>                   | Jordan RCK et al (2004)[156]   |
| <b>MMP2 (mRNA)</b>                   | Jordan RCK et al (2004)[156]   |
| <b>MMP9 (mRNA)</b>                   | Jordan RCK et al (2004)[156]   |
| <b>P73</b>                           | Chen YK et al (2004)[155]      |
| <b>Alterações do conteúdo de DNA</b> | Hogmo A et al (1998)[162]      |

Em resumo, as alterações genéticas associadas com as etapas iniciais da progressão de CCEOs incluem LOH em 3p14, 9p21, 17p13, 8p, 11q, 13q, 14q, 6p, 4q27, e 10q23; expressão aumentada ou diminuída de oncogenes e genes supressores tumorais, tais como *EGFR*, *TP53*, *Rb*, *p65*, *COX-2*, *p16*, *CCND1*, e *PTEN* [88, 144].

Estas alterações foram associadas com etapas distintas da progressão a carcinoma oral [82, 164-165]. Particularmente, alterações no gene supressor tumoral *TP53* representam um evento importante em estágios iniciais da progressão, ao passo que mutações no gene *p16*, um inibidor da quinase ciclina-dependente (proteína reguladora do ciclo celular), são consideradas alterações tardias durante a progressão tumoral [164, 166-167]. Aproximadamente 50% dos CCEO contém mutações no gene *TP53* [168], e aproximadamente 1/3 destes tumores também contém mutações no gene *CCND1* [166, 169].

Embora muitos estudos tenham se concentrado na definição de alterações genéticas associadas com a progressão dos carcinomas orais, ainda existe a necessidade de uma melhor caracterização de tais eventos genéticos em lesões pré-malignas. Novos estudos devem se focar na identificação de alterações genéticas correlacionadas com o comportamento clínico mais agressivo dessas lesões [149]. De fato, o significado prognóstico destas alterações genéticas identificadas em displasias permanece desconhecido. Dessa maneira, o delineamento de alterações genéticas em amostras sequenciais progressivas poderia justificar estratégias de tratamento mais agressivas, quando combinadas com o diagnóstico histopatológico.

## 5.2. Biomarcadores de Progressão Tumoral

No campo de pesquisa de biomarcadores em leucoplasias orais, o desafio é a capacidade de identificar as lesões que possuem risco de progressão a carcinoma, conseqüentemente

melhorando o diagnóstico precoce e sobrevida destes pacientes. A aplicação de técnicas moleculares no estudo da biologia dos tumores tem contribuído para a melhoria da compreensão dos processos celulares e da desregulação de funções celulares normais associadas com a iniciação e progressão do câncer oral[167]. Os avanços na tecnologia de *arrays* tem permitido a análise, em ampla escala, de expressão de RNA ou do número de cópias de DNA. Todas estas técnicas podem ser utilizadas para a descoberta de múltiplos genes/vias gênicas implicados na progressão dos CCEOs. O uso de tais alterações genéticas como biomarcadores para a detecção de doenças também oferece a possibilidade de melhores modalidades de tratamento. Tais marcadores genéticos, uma vez estabelecidos, poderiam ser utilizados na clínica, para prever quais leucoplasias possuem um risco aumentado de progressão.

Apesar de todos estes estudos, os genes implicados na progressão de leucoplasias à carcinoma permanecem desconhecidos, tornando-se difícil o seu uso como marcadores genéticos para prever qual dessas lesões se transformarão. Essa incapacidade de prever a progressão pode ser, em parte, devido ao fato de que a maior parte dos estudos se concentram na identificação de alterações a nível cromossômico. Embora esses estudos sejam sugestivos, outros estudos em larga escala ainda são necessários para confirmar os resultados até então obtidos e aplicá-los à prática clínica. Um dos enfoques potenciais de estudos futuros é a análise genômica global, que permite a identificação de ampliações e deleções de genes específicos, bem como de alterações na expressão gênica. Tais alterações, quando associadas com a progressão, podem ser úteis como biomarcadores de predição de transformação maligna. Além disso, alterações genômicas identificadas podem levar à caracterização de vias implicadas na progressão de CCEO, relevantes ao melhor entendimento da biologia destes tumores. Neste projeto de pesquisa objetivamos delinear alterações genéticas globais que ocorrem em amostras de leucoplasias

progressivas sequenciais e CCEOs dos mesmos pacientes. Detalhes das análises utilizadas neste estudo para identificar tais marcadores genéticos estão descritos no Capítulo 2. Os resultados destas análises estão descritos nos Capítulos 4-5 (manuscritos).

## 2. Análise Genômica do Câncer

Apesar de estudos anteriores que buscaram identificar genes/vias envolvidos na progressão de câncer oral, os genes implicados na transformação maligna de leucoplasia à carcinoma invasivo permanecem desconhecidos, o que dificulta prever quais dessas lesões apresentam risco aumentado de transformação à carcinoma. Essa incapacidade de prever a progressão maligna pode ser, em parte, devido à ausência de seguimento dos pacientes com leucoplasias apresentando baixo grau displásico, além da falta de metodologias de análise genômica robustas, compatíveis com quantidades limitadas de DNA extraído a partir de tecidos parafinizados (FFPE) e microdissecados. Além disso, a maior parte dos estudos têm se concentrado na identificação de modificações cromossômicas ou regiões da perda de heterozigose, o que dificulta a identificação de genes específicos como marcadores de progressão em câncer oral.

A aplicação de técnicas de análise genômica global possui o potencial para superar este problema, uma vez que permite a identificação de genes específicos implicados na doença. Além disso, a inclusão de amostras sequenciais de leucoplasias e CCEOs, de mesmo sítio de um mesmo paciente, permite uma comparação mais precisa das alterações genéticas que ocorrem nessas lesões pré-malignas. Este conjunto de amostras sequenciais é importante na identificação de biomarcadores eficazes no uso clínico; tais biomarcadores, quando associados com a avaliação clínica e histopatológica, possuem o potencial de prever quais leucoplasias têm um risco aumentado de transformação maligna.

O presente estudo foi delineado enfocando a análise genômica global da expressão de *microRNAs* e do número de cópias do DNA, as quais tem sido amplamente utilizadas para a identificação de alterações genéticas associadas com o desenvolvimento e progressão do câncer.



Estas análises foram utilizadas para identificar alterações genéticas comuns entre lesões pré-malignas e CCEOs correspondentes; tais alterações têm aplicação potencial para predição da transformação maligna de leucoplasia orais.

## *1. microRNAs*

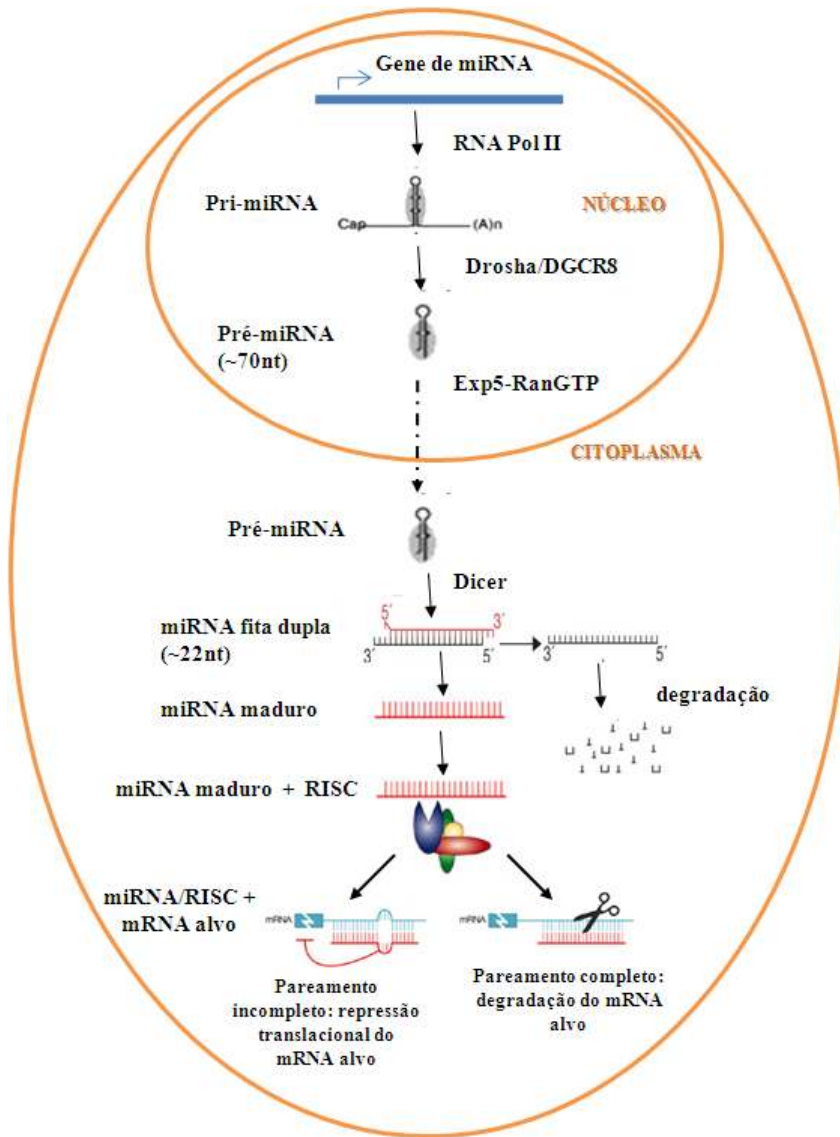
### *1.1. Biogênese e Função dos microRNAs*

Pequenos RNAs (sRNA) abrangem uma classe de RNAs não-codificadores de proteínas (ncRNAs), recentemente identificada. Estes sRNAs podem ser divididos em pelo menos seis categorias, as quais incluem microRNAs (miRNAs), RNA anão (tncRNAs), RNAs de interferência (siRNAs), RNAs com repetições associadas (rasiRNAs), RNAs modulatórios (smRNAs) e RNAs de interação com Piwi (*P-element induced wimpy testis*) (piRNA)[170-171].

Os miRNAs têm sido a classe de sRNA mais amplamente estudada[172-173]. MiRNAs são moléculas reguladoras altamente conservadas que se expressam nas células dos organismos eucarióticos e desempenham um papel importante na regulação da expressão gênica[174].

A biogênese e a função dos miRNAs envolvem um conjunto complexo de proteínas[173, 175]. Os miRNAs são codificados por genes localizados tanto em regiões não codificadoras como intrônicas, ou éxons de genes não codificantes de proteínas[173]. As seqüências primárias de miRNAs são transcritas pela RNA-polimerase II/III, e logo em seguida são processadas no núcleo pela enzima Rnase edonuclease III (Drosha), juntamente com o complexo microprocessador DGCR8, para a formação de um miRNA precursor. Este miRNA precursor é então exportado ao citoplasma celular pela proteína Exportina-5[176]. Finalmente, no citoplasma, o miRNA precursor é processado pela RNase III (Dicer) em um miRNA maduro e

funcional (Figura 2.1). Este miRNA funcional é uma molécula de dupla fita, que se liga ao complexo RNA indutor do silenciamento (RISC), que contém Argonauta 2[177], uma proteína com a atividade endonuclease direcionada para fitas de mRNA contendo seqüências complementares ao miRNA[178]. Dessa maneira, esse complexo de proteínas é responsável pelo fenômeno de silenciando gênico, e parcialmente responsável pela seleção da fita guia[179]. Posteriormente, uma das fitas é removida da fita dupla dos miRNAs e degradada; a fita restante guia o RISC para se ligar aos mRNAs complementares[180]. O complexo miRNA-RISC interage com sítios ligantes da região 3'UTR do mRNA, devido à complementaridade de uma seqüência de 5-7 nucleotídeos do miRNA maduro e do mRNA alvo[180-181]. Neste contexto, o miRNA direciona a regulação pós-transcricional, baseada nas características do mRNA alvo reconhecido pelo miRNA. Este mRNA alvo pode ser clivado em uma posição fixa, complementar à seqüência do miRNA, ou pode ter a tradução de sua proteína inibida, sem necessariamente ocorrer a degradação do mRNA (Figura 2.1)[179]. Desta forma, quando o miRNA possui uma complementaridade de bases absoluta ao mRNA alvo, esse miRNA causa a degradação da fita de mRNA. Porém, se a ligação das seqüências forem parcialmente complementares, ocorre uma repressão da tradução do mRNA, sem a degradação da sua fita.



**Figura 2.1. Via de biogênese dos miRNAs em eucariotos.** O gene de miRNA é transcrito pela RNA polimerase II ou III em um miRNA primário (pri-miRNA), e então clivado pelo complexo de enzimas Drosha-DGCR8 em miRNA precursor (pré-miRNA). Este miRNA precursor é exportado do núcleo para o citoplasma pela Exportina-5-Ran-GTP. No citoplasma, a Dicer RNase, em associação com a proteína TRBP, cliva o miRNA precursor, originando um miRNA maduro e funcional composto de ~22 nucleotídeos de comprimento. A fita passageira do miRNA de dupla fita (ilustrada em preto) é degradada, e o miRNA maduro se liga ao complexo de proteínas de silenciamento gênico, RISC. O então complexo miRNA+RISC é direcionado ao mRNA alvo, para controlar a expressão pós-transcricional de genes-alvo. Este processo pode ocorrer através de dois mecanismos: pareamento completo de bases, pelo qual ocorre degradação do mRNA alvo; e pareamento incompleto de bases, pelo qual ocorre uma repressão da tradução do mRNA alvo para proteínas.

Até o presente, aproximadamente 700 genes que codificam miRNAs foram identificados no genoma humano (<http://microrna.sanger.ac.uk/sequences/>; Release 3.0 of March 2009). Estima-se que cada miRNA possa se ligar a muitos mRNAs, e que mRNAs podem ter a sua tradução ou estabilidade regulada por mais de um miRNA[182-185]. Na maioria dos casos, os miRNAs impedem ou reduzem a eficiência da tradução dos mRNAs alvo; contudo, os miRNAs podem também regular positivamente a expressão de genes-alvo por uma complementaridade de seqüência parcial ou incompleta.

Considerando que o significado biológico da desregulação dos miRNAs depende dos seus genes alvo codificadores de proteínas, é importante entender os mecanismos pelos quais os miRNAs interagem com seus mRNAs alvo. Por exemplo, têm sido demonstrado que mutações podem diminuir a acessibilidade do mRNA alvo, e assim reduzir a repressão translacional mediada pelo miRNA[181]. Estes estudos são importantes para determinar a complementariedade exata dos miRNAs com mRNAs alvo, desde que isto é um fator crítico na função de miRNAs como reguladores da expressão gênica.

### *1.2.miRNAs e Câncer*

No câncer, as alterações genéticas e epigenéticas estão geralmente associadas com a inativação de genes supressores tumorais e com a ativação de proto-oncogenes, tanto em genes codificadores como em não-codificadores de proteínas. Como consequência, as seqüências de DNA alteradas são transcritas em mRNA, e mRNA é traduzido a proteína funcionalmente aberrante. A expressão genética também é regulada por vários mecanismos, tais como edição do RNA, *splicing* alternativo, degradação da fita do RNA e miRNA[186]. Dentre estes mecanismos,

a regulação da expressão gênica mediada por miRNAs tem sido relevantemente implicada na patogênese do câncer[187-190].

No genoma humano, muitos dos genes que codificam miRNAs estão localizados em sítios frágeis cromossômicos[191], e estes genes parecem ser independentemente transcritos[192-193]. A primeira evidência do envolvimento de miRNAs no câncer ocorreu com a caracterização molecular da deleção em 13q14 em pacientes com leucemia linfocítica crônica (LLC)[194]. Neste estudo dois miRNAs descritos na literatura, miR-15a e miR-16a, foram mapeados no cromossomo 13q14, uma região que se encontra deletada em mais de 50% dos casos de leucemia linfocítica crônica de células B (B-LLC). Adicionalmente, estes dois miRNAs foram os únicos dois encontrados em uma pequena região cromossômica comum entre pacientes com LLC. De maneira interessante, uma análise de expressão mostrou que miR-15a e miR-16a estavam ausentes ou com expressão diminuída na maioria (~68%) dos pacientes com LLC[194].

Desde então, vários estudos têm mostrado que os miRNAs representam uma classe importante de genes implicados na tumorigênese, uma vez que se encontram aberrantemente expressos, deletados, amplificados ou mutados em câncer[191, 194-198]. A expressão desregulada de certos miRNAs foi associada a doenças proliferativas, tais como câncer de mama [199], glioblastoma primário[200], carcinoma hepatocelular[201], carcinoma papilar da tireóide[202], câncer de pulmão[203-204], carcinomas gástrico e de cólon[197], e tumores endócrinos do pâncreas[203]; todos estes estudos forneceram evidência do papel dos miRNAs como agentes oncogênicos ou supressores tumorais. Como por exemplo, foi demonstrado que o gene *BCL2* interage com os miRNAs miR-15a e miR-16-1 em LLC[205]. A expressão destes miRNAs estava inversamente correlacionada à expressão do gene *BCL2*, e se demonstrou que ambos os miRNAs regulavam negativamente *BCL2* em nível pós-transcricional[205]. Estes

achados foram comprovados em estudos *in vitro* através do uso de linhagens de células leucêmicas, também sugerindo o uso terapêutico potencial do miR-15a e miR-16-1 em casos de LLC que possuam níveis de expressão aumentada do gene *BCL2*[205].

Em tumores sólidos, tais como o câncer de pulmão, se sugere que o miRNA miR-let-7 desempenha um papel crítico no desenvolvimento e patogênese do tumor. Takamizawa e colaboradores[198] observaram que miR-let-7 possuía expressão frequentemente diminuída em carcinomas de pulmão, tanto *in vivo* como *in vitro*; e esse baixo nível de expressão se associou significativamente com a redução da sobrevida dos pacientes após a cirurgia. Johnson e colaboradores[206] mostraram que miR-let-7 regula negativamente a expressão de dois oncogenes importantes para a tumorigênese do câncer de pulmão, os genes *RAS* e *MYC*, através da inibição da tradução de seus mRNAs alvo[206]. Ao contrário do miR-let-7, miRNAs do *cluster* miR-17-92 são altamente expressos no câncer de pulmão, especialmente nos de pequenas células, que é o tipo mais agressivo[207]. Adicionalmente, Hayashita e colaboradores[207] encontraram que a expressão aumentada dos miRNAs do grupo miR-17-92 estava associada com a regulação positiva de pelo menos um dos membros da família do gene *myc*, que muitas vezes está amplificado e/ou com expressão aumentada no câncer de pulmão de pequenas células[208].

Um estudo avaliando o perfil de expressão de miRNAs em carcinomas de mama mostrou que os padrões de expressão dos miRNAs eram significativamente diferentes entre os tecidos normal e neoplásico[199]. Entre estes miRs, miR-125b, miR-145 e miR-155 tiveram expressão significativamente diminuída nos tumores de mama. Este estudo também comparou a expressão dos miRNAs com as características histopatológicas do câncer de mama, e detectou que os estágios tumorais (pT1 ou pT2-3), os índices de proliferação (baixo ou alto), a expressão dos receptores de estrogênio (ER) e progesterona (PR) (positivo ou negativo) e a invasão vascular

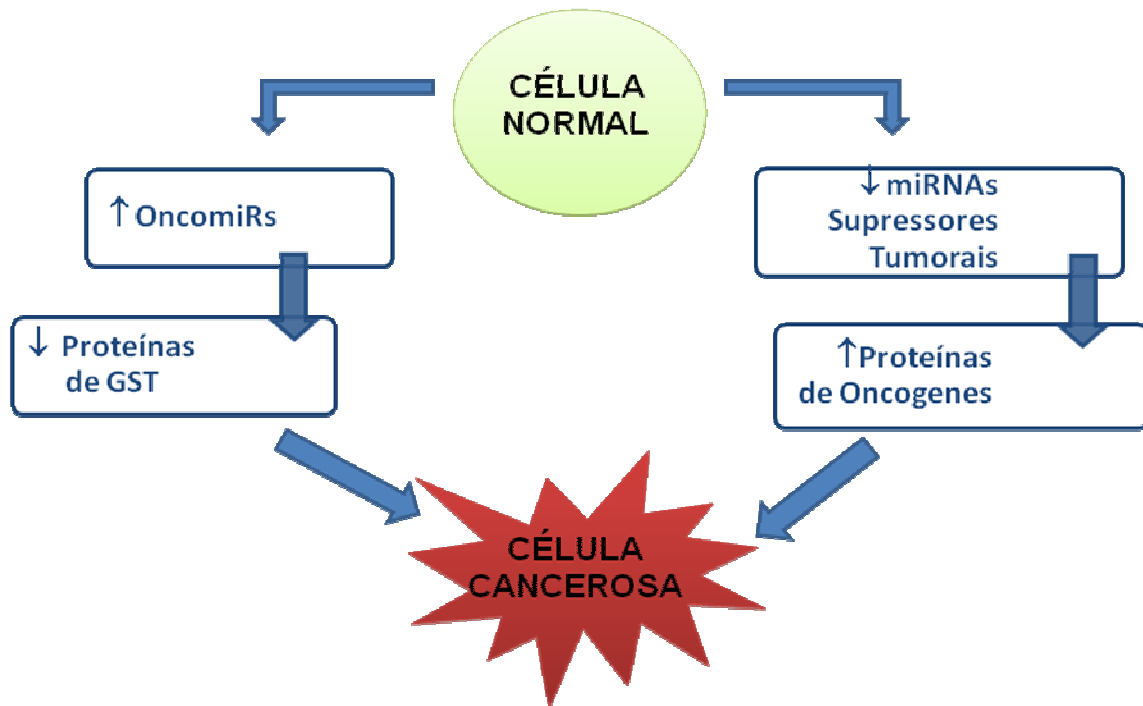
(ausente ou presente) associaram-se com alguns miRNAs diferencialmente expressos ( $p < 0.05$ )[199]. Por exemplo, a expressão da família miR-30 estava diminuída tanto em tumores ER como PR negativos, o que indica um possível mecanismo de regulação por estes hormônios. Outra associação interessante foi o baixo nível de expressão de vários miRNAs da família let-7 em amostras de câncer de mama com altos índices de proliferação, um achado que sugere que este miRNA pode ser um indicativo para um pobre prognóstico no câncer de mama.

A desregulação da expressão de miRNAs também foi encontrada em câncer de cólon. Um estudo por Michael e colaboradores[197] identificou 28 miRNAs diferencialmente expressos entre adenocarcinoma de cólon e mucosa normal; eles encontraram que os miRNAs maduros miR-143 e miR-145 possuíam expressão reduzida constante em polipose adenomatosa e em câncer de cólon avançado. Neste estudo, os autores também examinaram o nível transcricional da forma precursora dos miRNAs miR-143 e miR-145; estes precursores não estavam alterados nos tecidos coloreticais pré-canceroso e neoplástico, sugerindo que a transcrição alterada não é responsável pelos níveis reduzidos de miRNA maduros[197].

Recentemente, He e colaboradores[202] demonstraram que, entre vários miRNAs desregulados, miR-221, miR-222 e miR-146 tiveram expressão muito elevada (11 a 19-*folds*) em carcinomas papilares da tireóide (CPT), comparado com tecidos de tireóide normais. CPT está normalmente associado com níveis de expressão aberrantes do proto-oncogene *KIT*, que foi anteriormente apontado como um alvo putativo destes miRNAs[182-183, 209]. He e colaboradores[202] verificaram assim, que os níveis do mRNA e da proteína do gene *KIT* estavam consideravelmente diminuídos, e correlacionaram esses achados com os níveis aumentados de miR-221, miR-222 e miR-146 em CPT, sugerindo um mecanismo de regulação negativa do *KIT* por estes miRNAs.

A expressão desregulada de miRNAs em cânceres humanos, como nos exemplos descritos acima, provavelmente representará dois caminhos distintos pelos quais os genes supressores tumorais e oncogenes podem ser, respectivamente, suprimidos ou ativados (Figura 2.2). Por fim, a avaliação da expressão de miRNAs em tumores, em comparação com os respectivos tecidos normais, pode ajudar a identificar miRNAs que estejam implicados no câncer, bem como estabelecer o seu papel patogénico na doença[199].





**Figura 2.2.** miRNAs e o desenvolvimento do câncer. A expressão aumentada dos miRNAs oncogênicos (oncomirs) pode diminuir os produtos das proteína expressas por genes supressores tumorais (GSTs). Por outro lado, a expressão reduzida dos miRNAs supressores tumorais pode gerar níveis elevados das proteínas oncogênicas. O resultado final desse impareamento genético é a disrupção de mecanismos tais como a apoptose e a diferenciação celular, o que contribui para a proliferação celular aumentada e metástase.

### 1.2.1. miRNAs com Efeito de Oncogenes

Os miRNAs que possuem expressão significativamente aumentada no câncer podem ser considerados como oncogenes. Eles também são conhecidos como “*oncomirs*”, uma vez que funcionam promovendo o desenvolvimento de tumores através da inibição negativa de genes supressores tumorais e/ou genes que controlam a diferenciação celular ou apoptose.

Embora os miRNAs com expressão aumentada pareçam atuar como oncogenes, apenas poucos deles já foram bem caracterizados. O miR-17-92 é um bom exemplo de *oncomiR*; este grupo de miRNAs está localizado no cromossomo 13q31, uma região genômica frequentemente amplificada em câncer de pulmão e vários tipos de linfomas[194, 207]. Os níveis de expressão do miR-17-92 foram detectados significativamente aumentados em câncer de pulmão e linfomas, em comparação com os respectivos tecidos normais, especialmente nas formas mais agressivas desses tumores, tais como o câncer de pulmão de pequenas células e os linfomas de célula B[194, 207]. Além disso, estudos *in silico* indicam que o miR-17-92 pode ter vários genes alvo; mais de 600 genes são preditos para miR-19a e miR-20 (dois membros do grupo miR-17-92). Entre estes genes alvo, dois são genes supressores tumorais bem caracterizados: *PTEN* (fosfatase e tensina homólogos deletados no cromossomo 10), que promove apoptose via P13K-Akt-PKB[210]; e *RB2*[185]. Entretanto, não há dados experimentais funcionais confirmando que os genes *PTEN* e *RB2* são alvos de qualquer um dos membros de grupo miR-17-92.

Um estudo mostrou que a expressão do grupo miR-17-92 estava relacionada à expressão do gene *c-Myc*; tanto o miR-17-92 como *c-Myc* regulam a expressão de um terceiro gene, o *E2F1*, que é um fator de transcrição do ciclo celular[211]. O gene *c-Myc* é um dos proto-oncogenes mais bem caracterizados; ele atua como um fator de transcrição que regula a proliferação celular e o crescimento e a morte celular mediada por apoptose. É bem estipulado

que a expressão desregulada ou a disfunção deste gene leva à transformação maligna de muitos tumores humanos[212]. O'Donnell e colaboradores[211] demonstraram que o *c-Myc* ativa simultaneamente a transcrição tanto do *E2F1* como do miR-17-92.

O miR-21 é um dos miRNAs mais proeminentes implicados na gênese e na progressão do câncer que têm sido descrito na literatura[203]. Não só foi implicado na promoção do crescimento tumoral[213], proliferação[214], anti-apoptose[215] e resposta à quimioterapia[216], como também têm sido mostrado com expressão aumentada em diferentes tipos de câncer. Demonstrou-se também que o miR-21 desempenha um papel relevante na proliferação e apoptose celular; este miRNA foi implicado como um importante fator anti-apoptótico, uma vez que silenciando a sua expressão foi observado um aumento na morte celular por apoptose em células de glioblastomas[215]. Si e colaboradores[213] examinaram os níveis de expressão 157 miRNAs em cinco tumores de mama de estágio avançado, comparado com tecidos normais, e encontraram a expressão aumentada do miR-21 nesses tumores. Confirmando um estudo anterior[215], Iorio e colaboradores[199] observaram uma expressão aumentada do miR-21 em tumores mamários; e Meng e colaboradores[216] também demonstraram níveis de expressão aumentada do miR-21 em colangiócitos malignos, além de outros miRNAs tais como miR-let7a-1 e miR-16-1. Também, oligonucleotídeos anti-miR-21 demonstraram suprimir tanto o crescimento celular *in vitro* como o crescimento tumoral em modelos animais *in vivo*. Tais achados sugerem que o miR-21 se associa com o aumento da apoptose celular e com a diminuição dos níveis de expressão da proteína anti-apoptótica Bcl-2 em tumores de mama[213].

Um relato recente indicou que o gene *PDCD4* (*Programmed Cell Death 4*) é um dos alvos do miR-21[217]. *PDCD4* é uma proteína pro-apoptótica, que se encontra com expressão frequentemente reduzida em carcinomas hepatocelulares (CHCs)[218]. Notavelmente, a

expressão do miR-21 estava aumentada nestes CHCs, sugerindo que esse miRNA possa inibir a apoptose através da regulação negativa tanto do *PDCD4* como *PTEN*, um outro gene que também é alvo do miR-21[213]. Um estudo recente forneceu dados experimentais mostrando que o gene *PDCD4* é negativamente regulado pelo miR-21, e que tal miRNA está associado com invasão e metástase no câncer coloretal[219]. Zhu S e colaboradores[220] também descreveram o papel do miR-21 em invasão tumoral e metástase, através da sua ligação com os mRNAs dos genes *PDCD4*, *PTEN* e do gene tropomiosina 1 (*TPMI*), um outro importante gene supressor tumoral. Este estudo mostrou que a expressão aumentada de *TPMI* suprime a capacidade invasiva de linhagens celulares de câncer mamário (MDA-MB-231), e sugeriu que o miR-21 afeta o crescimento tumoral, a invasão e metástase, mediados provavelmente pela repressão simultânea dos genes supressores tumorais *TPMI*, *PDCD4* e *PTEN*. Estudos prévios em nosso laboratório mostraram que o gene *PDCD4* possui expressão diminuída em CCEOs, tanto em níveis de mRNA como de proteína, em comparação com tecidos orais normais; a expressão reduzida desse gene foi associada com metástase em linfonodos e com o aumento do potencial invasivo das células tumorais (Reis e colaboradores, em preparo para publicação). Além disso, nosso grupo confirmou que a região 3'UTR do gene *PDCD4* é alvo do miR-21, e que os níveis desse miRNA estavam aumentados em tumores apresentando níveis reduzidos de expressão do gene *PDCD4* (Reis e colaboradores, em preparo para publicação). Em conjunto, estes estudos sugerem que o miR-21 possui um forte potencial oncogênico devido à sua capacidade de regular múltiplos genes, afetando assim vários mecanismos associados ao câncer.

### 1.2.2. miRNAs com Efeito de Genes Supressores Tumorais (GSTs)

Os miRNAs que possuem um efeito de genes supressores tumorais (GSTs), também conhecidos como miRNA supressores tumorais[221], agem inibindo negativamente oncogenes[213, 215, 218]. Atualmente, mostrou-se que vários miRNAs possuem comportamento de genes supressores tumorais, cuja expressão diminuída se encontra associada ao câncer. Na literatura, a família do miR-let-7 foi extensivamente estudada como uma das família de miRNAs com efeitos mais similares a um GST. Os miRNAs desta família estão mapeados em regiões cromossômicas diferentes (9p22.31, 11q24.1, 22q13.31, 21q21.1 e 19q13.43), que estão freqüentemente deletadas em vários cânceres humanos[191]. Utilizando o câncer de ovário como modelo, Shell S e colaboradores[222] detectaram a expressão diminuída do miR-let-7 e a expressão aumentada do gene *HMGA2* como um marcador para o câncer ovariano menos diferenciado. Os autores sugeriram que o miR-let-7 e o gene *HMGA2* seriam melhores marcadores prognósticos do que os clássicos marcadores conhecidos, tais como a E-caderina, a vimentina e o *Snail*.

Takamizawa e colaboradores[198] observaram que o nível de expressão do miR-let-7 se encontrava reduzido em câncer de pulmão e estava significativamente associado com a sobrevida mais curta dos pacientes após a cirurgia. Neste estudo, eles fizeram experimentos de transfecção, *in vitro*, para recuperar a expressão do miR-let-7 em linhagens celulares de adenocarcinoma de pulmão, e observaram que este miRNA inibiu o crescimento das células tumorais de pulmão. Também em câncer de pulmão, se demonstrou que o miR-let-7 regula negativamente a expressão do oncogene *RAS*, através do seu ligamento direto na região 3' UTR do mRNA desse gene[206]. Os níveis do miR-let-7 estavam significativamente reduzidos em tecidos tumorais do pulmão, quando comparados com os tecidos normais, e foram associados com níveis significativamente aumentados da proteína RAS. Estes resultados sugerem que a regulação de *RAS* pelo miR-let-7

possa ser um mecanismo pelo qual o miR-let-7 atue como um gene supressor tumoral na carcinogênese de pulmão[206].

### *1.3.Papel dos miRNAs em CCEOs*

Como descrito acima, existem evidências do papel dos miRNAs no câncer humano. O câncer de cabeça e pescoço, inclusive CCEO, é um câncer que tem recebido atenção no campo de perfil genômico. Enquanto se sabe que os principais fatores de risco dos CCECPs são o fumo e o álcool, as vias moleculares exatas envolvidas na carcinogênese dos CCEOs ainda não são totalmente entendidas. Como tal, algumas prioridades neste campo é a necessidade de identificar novo(s) gene(s) e via(s) que pode(m) ajudar a elucidar os mecanismos desta doença. Os miRNAs são uma fonte interessante de pesquisa, uma vez que estas moléculas desempenham um papel importante em nível de regulação genética pós-transcricional, ainda não completamente caracterizado no câncer.

Vários estudos associando a desregulação de miRNA com o desenvolvimento, invasão e metástase do câncer foram previamente descritos neste capítulo. Juntos, estes estudos mostraram que o aumento ou diminuição da expressão dos miRNAs desempenha um papel crítico na biologia do câncer. Baseado neste conceito, os pesquisadores têm tentado associar o perfil de expressão dos miRNA e os seus prováveis genes alvo, com um papel na carcinogênese oral. Em CCECP, a proteína do gene homólogo do grupo de alta mobilidade A2 (*HMG A2*), foi considerada associada com o aumento da quimiosensibilidade seletiva direcionada ao inibidor da topoisomerase II em células cancerosas. Esta proteína desempenha um papel importante na regulação das respostas genotóxicas, e foi demonstrada ser regulada pelo miR-98 nas células do câncer de cabeça e pescoço[223].

Outro estudo [224] utilizou *microarrays* para investigar a expressão de 261 miRNAs maduros em nove linhagens celulares de CCECP de quatro sítios anatômicos diferentes, incluindo carcinomas da cavidade oral. Estes autores relataram 33 miRNAs com expressão aumentada e 22 com expressão diminuída em CCECP. Entre os miRNAs com expressão alterada, o miR-21 apresentou expressão significativamente aumentada, achado que concorda com estudos anteriores em glioblastoma[215], câncer de pulmão[204] e tumores de mama[220]. Outro estudo utilizando linhagens celulares de CCECP, incluindo CCEOs, mostrou a expressão aumentada do miR-205, que contrariamente, se encontrava com expressão diminuída em outros estudos com linhagens celulares de outros cânceres, tais como pulmão, mama, colorectal, próstata e pâncreas[225]. Estudos adicionais, incluindo aqueles com o uso de amostras de pacientes, ainda se fazem necessários para o melhor delineamento do papel dos miRNAs na biologia dos carcinomas de cavidade oral.

Uma limitação importante da maior parte dos estudos é o uso de linhagens celulares tumorais para a investigação de perfis de expressão de miRNAs, além da falta de inclusão de células normais da cavidade oral para comparação dos perfis de expressão. A expressão dos miRNAs dessas linhagens celulares tumorais pode não reproduzir exatamente o mesmo perfil que os tumores originais, uma vez que o crescimento celular e o acúmulo de alterações, *in vitro*, pode afetar a expressão dos microRNAs[226]. Isto também pode ser uma explicação do porque muitos estudos utilizando tais linhagens celulares têm achado resultados contraditórios com os descritos na literatura.

Finalmente, a expressão anormal de miRNAs foi encontrada em células pré-malignas de adenoma e em células malignas de carcinoma de cólon[197]; desta forma, é interessante investigar a expressão de miRNA em leucoplasia oral, que é uma das OPMLs mais prevalentes

com risco de progressão a CCEO. Uma vez que a ativação e/ou inativação de genes está implicada na transformação de leucoplasias orais a carcinomas, a desregulação da expressão de miRNAs pode desempenhar um papel importante na progressão desta doença.

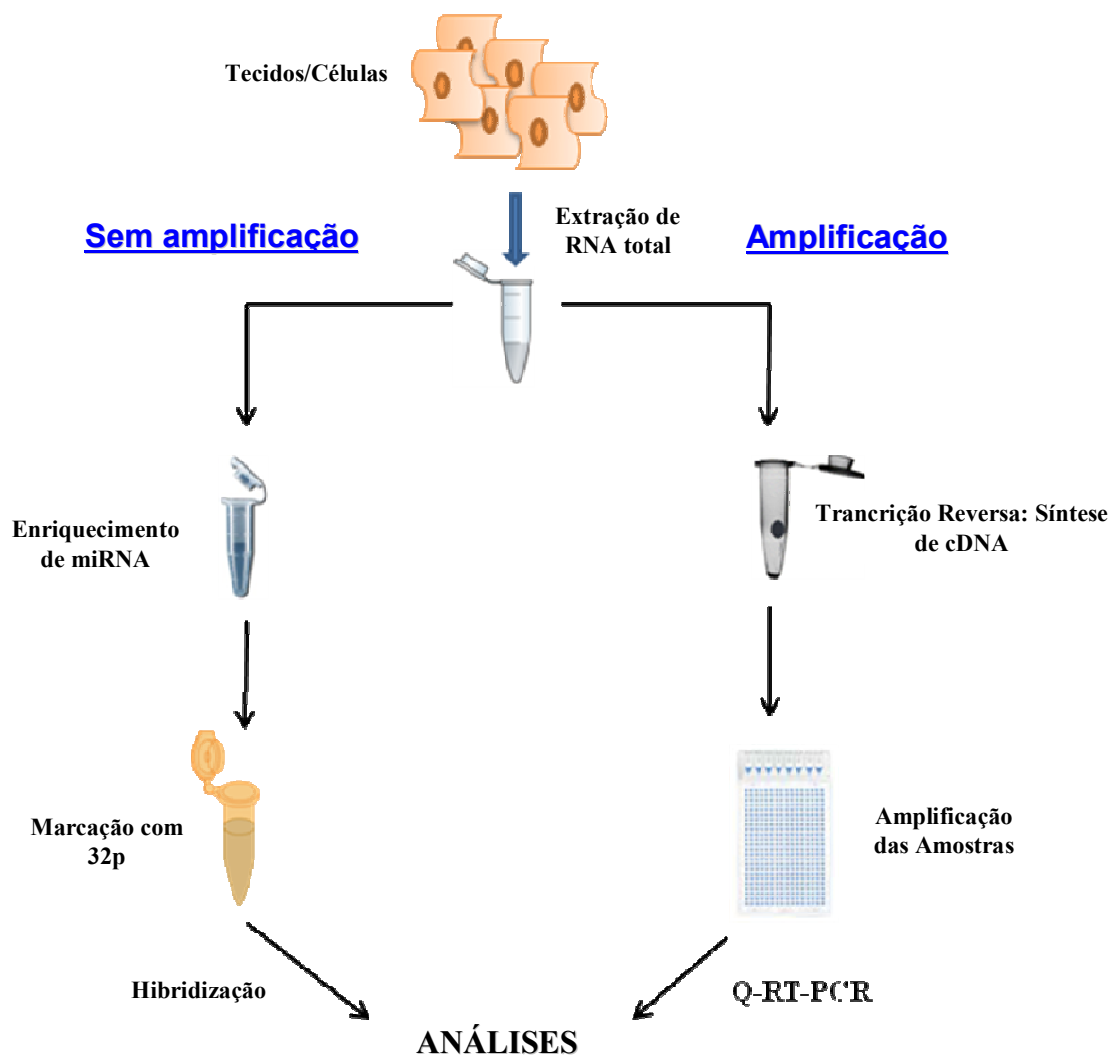
#### 1.4. Métodos de Análise do Perfil de Expressão de miRNAs

Durante os últimos anos, várias estratégias têm sido utilizadas para a análise de expressão e determinação de miRNAs envolvidos em doenças tais como o câncer. Destas estratégias, as análises em ampla escala têm sido as mais utilizadas na investigação dos perfis de expressão dos miRNAs. A tecnologia de *microarrays*, que foi originalmente desenvolvida para a análise dos perfis de expressão de cDNA (mRNA), também tem sido aplicada à expressão de miRNAs[227]. Outras metodologias também foram desenvolvidas adaptando tecnologia existentes e que são baseadas em hibridação e amplificação. Atualmente, as tecnologias mais frequentemente utilizadas para a análise de expressão de miRNA são o *microarray* e a PCR quantitativa em tempo real (Q-RT-PCR). Embora menos utilizados, outros métodos tais como *northern blot*, *cloning-mirage*, *rake* e nanotecnologia (Tabela 2.1) também fazem parte das técnicas atualmente desenvolvidas para a análise do perfil de expressão de miRNAs.

No geral, os métodos de análise existentes podem ser separados em duas categorias principais: (1) a que requer a amplificação das amostra de RNA (PCR, *Polymerase Chain Reaction*), e (2) o que utiliza oligonucleotídeos para a hibridação direta da amostra, sem necessidade de amplificação (Figura 2.3). Os protocolos de hibridização geralmente necessitam de uma quantidade inicial relativamente maior de RNA total, enquanto o procolo baseado em PCR (amplificação) necessitam de uma quantidade muito menor de material inicial. Por isso, os protocolos de amplificação são bem mais convenientes quando se utiliza amostras de RNA



extraído de amostras de arquivo (FFPE). As duas principais técnicas que representam cada uma dessas duas categorias acima mencionadas, e as mais utilizadas na detecção de miRNAs, são o *microarrays* para miRNAs (ensaio baseado em hibridização), que contém seqüências correspondentes de todos os miRNAs descritos no genoma humano, e a PCR em tempo real (ensaio baseado em amplificação), que é uma técnica bastante específica e sensível na análise de expressão de microRNAs[228-229]. Estas técnicas são discutidas nas seções seguintes.



**Figure 2.3.** Ilustração das duas categorias que representam as técnicas de análise do perfil de expressão dos miRNAs. No painel da direita, as técnicas que requerem a amplificação do RNA total. O lado esquerdo representa as metodologias que não necessitam da amplificação do RNA, e que utilizam a hibridização direta das amostras.

#### 1.4.1. *Microarranjos (microarrays) de miRNA*

A maior parte dos experimentos de *microarrays* objetivam determinar os perfis de expressão de RNA mensageiro (mRNA), através da análise simultânea dos níveis de expressão de centenas a milhares de genes, que estão diferencialmente expressos entre tecidos normal e tumoral. O perfil de expressão de miRNAs utilizando a tecnologia de *microarrays* é um campo relativamente novo, e ainda possui alguns desafios, tais como a quantidade de RNA necessária para análise, o tipo de tecido e o estágio da doença em que se pretende determinar os níveis de expressão dos miRNAs.

Atualmente, as plataformas de *microarrays* disponíveis para análise de miRNAs se concentram na exploração da biogênese dos miRNAs, na especificidade dos miRNAs em um determinado tecido e na expressão diferencial de miRNAs entre tecidos normais e tumorais. Portanto, a utilização desta tecnologia permite a identificação de vias genéticas e mecanismos relacionados à doença[230-232].

Os *microarrays* de miRNAs possuem sondas projetadas para a identificação de miRNA maduro e de seus precursores ou pre-miRNAs[227, 233], o que faz o uso desses *arrays* uma vantagem em relação a outras técnicas utilizadas para a análise de miRNAs[234-235]. Outra vantagem consiste no fato desse método possuir um custo relativamente baixo e regularmente disponível em laboratórios sem robótica e equipamentos especializados para a fabricação desses *microarrays*.

De fato, estudos recentes obtiveram sucesso na utilização de *microarrays* para a análise do perfil de expressão global dessas pequenas moléculas de RNA. Estudos utilizando *microarrays* identificaram assinaturas de miRNAs que podem ser úteis como marcadores diagnósticos ou terapêuticos em doenças como o câncer[232]. Estes estudos têm proporcionado

contribuições importantes tanto na pesquisa básica como na pesquisa clínica aplicada, e possuem o potencial de modificar a prática da medicina para um enfoque mais personalizado[236-237].

#### 1.4.2. PCR Quantitativa (Q-RT-PCR)

A hibridação de amostras de RNA em uma plataforma de *microarray* normalmente requer uma grande quantidade de RNA total; contudo, é comum no campo da pesquisa a disponibilidade de quantidades limitadas de RNA, tais como RNA isolado a partir de material arquivado (FFPE) e microdissectado. Para resolver este problema, desenvolveu-se a tecnologia baseada em PCR. Utilizando os princípios da PCR, a PCR quantitativa em tempo real têm sido empregada para quantificar os níveis de expressão de miRNAs.

Na quantificação de miRNAs, a PCR quantitativa em tempo real utilizando sondas *TaqMan* (Applied Biosystem) mostraram uma alta especificidade, sensibilidade e reproducibilidade[238-240]. A especificidade dos ensaios da PCR utilizando sondas *TaqMan* se baseia na ação combinada de dois *primers* e uma sonda *TaqMan* de seqüência específica para a detecção e quantificação de miRNAs[238]. Este método de quantificação de miRNA utiliza a tecnologia de transcrição reversa (*stem-loop RT*) seguida pela análise por PCR em tempo real. Os *primers* *stem-loop RT* são melhores do que os convencionais quanto a eficiência e especificidade da transcrição reversa de miRNA a cDNA. Os ensaios *TaqMan* para miRNA são específicos para miRNAs maduros, e também podem discriminar os miRNAs que se diferenciam por apenas um nucleotídeo[238].

Além disso, este método tem a vantagem de não ser afetado pela contaminação com DNA genômico. A quantificação exata dos miRNAs é realizada com um mínimo de até 25pg de RNA total, pertinente para amostras microdissectadas. A Q-RT-PCR também se mostrou ser um

método viável para a análise de expressão de miRNA extraído de tecido FFPE[241-242], que são espécimes bastante valiosas para estudos retrospectivos de doenças humanas, incluindo o câncer. O nosso laboratório avaliou anteriormente o uso de qRT-PCR para análise em larga escala da expressão de miRNAs em amostras clínicas de FFPE (Goswami e colaboradores, em preparação). Examinamos a utilidade da plataforma de microfluidos *TaqMan Low Density Arrays* (TLDA) v1.0 (Applied Biosystems) para o estudo da expressão de miRNA usando amostras FFPE e comparando-as com as amostras frescas correspondentes de linfomas de célula do manto (LCM). Os nossos achados mostraram que a plataforma TLDA é útil para gerar perfis de expressão de miRNAs a partir de amostras FFPE, uma vez que os perfis obtidos desses *arrays* foram reproduzíveis quando comparamos amostras coletadas a fresco e congeladas com amostras FFPE.

De fato, a alta sensibilidade, especificidade e precisão deste método levam em conta a análise direta de uma única célula sem a necessidade da purificação do ácido nucléico[238]. A Q-RT-PCR permite a expressão rápida, exata e sensível de miRNAs, possuindo o potencial de identificar possíveis biomarcadores que sejam específicos a determinado tecido ou doença.

Além das tecnologias principais anteriormente apresentadas, existe também outros métodos para a análise de miRNAs (Table 2.1). Estas técnicas enfocam características e necessidades específicas, tais como o custo do experimento, o tamanho da mostra, a quantidade de mostra, a velocidade para a realização do experimento e o objetivo de identificação de novos miRNAs. Entretanto, estas novas metodologias ainda estão em desenvolvimento e requerem otimização para aplicações futuras em nível de diagnóstico e prognóstico de pacientes.

**Table 2.1.** Métodos para análise do perfil de expressão dos miRNAs: vantagens e desvantagens.

| <b>Método</b>              | <b>Vantagens</b>   | <b>Limitações</b>   | <b>Referência</b> |
|----------------------------|--|---|-------------------|
| <b>Microarray de miRNA</b> | Baixo custo; larga escala  | Isótopos radioativos na fabricação das sondas                           | [243]             |
| <b>RT real time PCR</b>    | Detecção rápida de miRNA e seus precursores                          | Alto custo  | [187, 228]        |
| <b>Northern blot</b>       | Metodologia mais confiável – “the gold standard”                     | Utilização de isótopos radiotivos, pequena escala e baixa sensibilidade | [244]             |
| <b>Rake</b>                | Alta especificidade  | Baixa sensibilidade   | [245]             |
| <b>Bead-based método</b>   | Múltiplas amostras podem ser testadas, alta velocidade e baixo custo | Condições muito específicas   | [246]             |
| <b>Cloning-mirage</b>      | Para descobrimento de novos miRNAs                                   | Alto custo para seqüenciamento  | [247]             |
| <b>Nanotechnology</b>      | Técnica sensível (500 fmol)  | Alto custo, requer instrumento especial                                 | [248]             |

### 1.5. Análise de miRNAs com Amostras FFPE

Os laboratórios de patologia no mundo inteiro contêm um vasto arquivo de amostras que podem estar utilizadas para a análise genética[249-250]. Estas amostras são fixadas na formalina, e embebidas em parafina (FFPE) e são guardadas durante até 20 anos, e possivelmente por mais tempo, dependendo do Hospital e/ou leis governamentais. Considerando o tempo de armazenamento destas amostras, análises retrospectivas podem ser executadas em estudos que contenham dados clínico-patológicos de pacientes monitorados por períodos significantes.

O processo de fixação em formalina e embebedamento de amostras em parafina é rotineiro nos laboratórios de patologia[241]. Isto se faz um problema para estudos de expressão gênica porque a fixação em formalina e o processamento subsequente em etanol resultam na formação de *crosslinks* entre as moléculas de RNA e proteínas, levando a uma redução significativa da recuperação de mRNA dos tecidos FFPE. Além disso, a degradação enzimática antes e durante o processo de fixação em formalina resulta na redução da integridade e rendimento do mRNA[241]. Além disso, a fixação em formalina e o processamento em etanol levam à produção de mono-metil e aductos etolixados com as bases dos ácidos nucléicos, bem como fragmentos de depuração[251-254], reduzindo a eficiência da transcrição reversa durante a qRT-PCR, e afetando outras aplicações após a PCR[241].

Apesar destes problemas com a fixação e o processamento dos tecidos, a extração de miRNAs a partir de tecidos FFPE não parece ser um problema, já que o pequeno tamanho e a estrutura secundária dos miRNAs contribuem para a sua estabilidade durante esse processo de fixação[255]. Também acredita-se que os miRNAs sejam moléculas protegidas pelo complexo *RISC*, e por isso devem ser menos suscetíveis à degradação quando comparados com mRNAs[256]. Os miRNAs são assim mais estáveis e mais facilmente recuperados do tecido

FFPE do que os mRNAs[241], o que fazem deles uma melhor opção para as análises de expressão gênica utilizando amostras FFPE[255].



## 2. Alteração Global do Número de Cópias do DNA (CNAs)

### 2.1. Variação no Número de Cópias (CNV) e Alterações no Número de Cópias (CNA) do DNA

A variação do número de cópia (CNV) é definida quando um determinado segmento no genoma humano apresenta ganhos ou perdas do material genético. Neste caso, as alterações que representam  $>2$  cópias são consideradas ganhos; aquelas alterações que representam  $<2$  cópias determinam as perdas. Esta última alteração pode ser heterozigota ou uma deleção homozigota, dependendo do número de cópias perdidas, se somente 1 cópia ou ambas as cópias, respectivamente. Os CNVs possuem  $\sim 1\text{Kb}$  de comprimento, e são alterações polimórficas normalmente presentes presentes na população em geral[257-258]. Eles podem ser somáticos, quando ocorrem *de novo*, ou herdados, se presentes no DNA da linhagem germinativa. Os CNVs também se associam com várias regiões genômicas, tais como genes, duplicações de segmentos genéticos e elementos funcionais[259]. Quando esses CNVs ocorrem em *loci* associado a doenças, tais como o câncer, que é resultado de alterações genéticas em muitas destas regiões[260], denominamos esse evento como alterações do número de cópias do DNA (CNAs).

As CNAs são de particular interesse na biologia tumoral, e têm tido importância crescente na pesquisa do câncer devido à sua associação com a desregulação da expressão genética, como demonstrado em estudos do câncer de cólon[261] e mama[262]. Estudos como estes associaram ganhos ou perdas do DNA com alterações na expressão gênica de 1.5 vezes em carcinomas comparado com tecidos normais[262]. Adicionalmente, as regiões de CNAs são de particular interesse no câncer, desde que podem conter oncogenes mapeados dentro de regiões amplificadas, e abrigar genes supressores tumorais nas regiões de perda[263].

A análise do genoma em larga escala tem sido extensivamente aplicada ao estudo do câncer e outras doenças. Em consequência, estes estudos contribuíram para a identificação de genes e vias genéticas associados com etiologia e progressão de doenças. Especialmente, a análise genômica global do DNA tumoral levou à identificação de alterações na sequência e no número de cópias, associadas com o diagnóstico, prognóstico e resposta ao tratamento de vários tipos de câncer[264-266]. Têm-se demonstrado que as alterações no DNA, tais como amplificações ou deleções, podem resultar em expressão aumentada ou diminuída de genes-chave. A expressão genética anormal pode ter várias consequências, dependendo dos genes afetados, que favorecem o crescimento tumoral; por exemplo, pode conferir uma vantagem ao crescimento de células malignas, ou suprimir mecanismos associados com a morte celular[267]. Os genes desregulados podem ser, por isso, alvos importantes no prognóstico e intervenção terapêutica, e a identificação de tais alterações (CNAs) pode ajudar a elucidar mecanismos potenciais implicados na tumorigênese.

## *2.2. Análise no número de cópias do DNA em Câncer*

A análise genômica do câncer se beneficiou muito com a introdução da técnica de hibridação genômica comparativa (CGH), que permitiu fazer o mapa dos ganhos e perdas em nível cromossômico de todo o genoma[268-269]. Subseqüentemente ao CGH, a técnica de CGH *array* (aCGH) foi introduzida em 1998, permitindo automação aumentada, reproducibilidade melhorada e maior precisão devido ao mapeamento mais exato das aberrações cromossômicas. A análise de aCGH é uma técnica conduzida pela marcação diferencial com fluorocromos do DNA genômico total de uma amostra teste (tumor) e uma amostra referência (normal), seguida pela co-hibridização no *array* composto de sondas de DNA que incluem representações de todo o

genoma. A razão entre as intensidades de fluorescência permite a identificação de CNA entre amostra normal e tumoral. A tecnologia de aCGH tem sido aplicada com sucesso para caracterizar anormalidades congênitas[268-270], caracterizar e classificar tumores[250], além de ser útil para uma melhor compreensão do processo da instabilidade cromossômica no câncer humano[271].

As aberrações do número de cópias do DNA podem variar e expor certa especificidade em tipos tumorais diferentes, uma vez que tais alterações ocorrem pela seleção clônica das células que se multiplicam durante o desenvolvimento do tumor[272]. Esta especificidade pode ser empregada para a classificação da origem e/ou sub-tipo destes tumores. Por exemplo, perfis de CNAs de tumores gástricos foram associados com diferentes aspectos prognósticos e de sobrevivência dos pacientes[273] e perfis específicos de CNAs em tumores gástricos foram relatados em pacientes jovens comparados com pacientes com idade mais avançada[274].

Jong K e colaboradores[275] realizaram uma meta-análise utilizando dados disponíveis de perfis de aCGH de 373 tumores primários de várias localizações anatômicas. Os resultados mostraram que os tecidos que possuem uma origem em comum se agruparam em um mesmo *cluster*; interessante, tumores de origem hematopoética e mesenquimal estavam agrupados separadamente dos tumores de origem epitelial. O *clustering* não-supervisionado dos perfis de aCGH também podem distinguir subclasses dentro de tumores com a mesma origem tecidual, como foi demonstrado com uma série de tumores mamários[276]. Os estudos utilizando a análise supervisionada podem determinar as regiões cromossômicas que distinguem uma classe de tumor de outro, baseado nos dados aCGH[277].

Enfim, o aCGH tem sido amplamente utilizado em estudos que examinam CNAs. Inicialmente, a técnica de aCGH foi aplicada com sucesso utilizando DNA de alto peso

molecular, extraído de espécimes de tecido coletados a fresco e congelados ou de linhagens celulares[278]. Recentemente, estudos têm demonstrado que a análise de aCGH também é possível através da utilização de DNA extraído de tecidos arquivados (FFPE)[279-283]. O uso de tecidos FFPE para aplicações em aCGH tem muitas vantagens, tais como as exemplificadas abaixo.

### *2.3. aCGH em Amostras FFPE*

Embora o aCGH seja um método eficiente para identificar e classificar oncogenes e genes supressores tumorais em tumores sólidos, os métodos atualmente utilizados para obter dados de CNAs a partir de amostras arquivadas (FFPE) ainda não estão totalmente otimizados, com muitas variações em protocolos utilizados e no grau de sucesso das hibridações.

Como a qualidade de DNA dos tecidos FFPE é variável, existem algumas perguntas que ainda são pertinentes quanto ao uso do material FFPE para a análise de aCGH: Primeiramente, é possível confiantemente identificar CNAs em DNA obtido de tecidos FFPE, utilizando aCGH? Em segundo lugar, o aCGH pode ser aplicado quando se obtém um pequeno número de células microdissecadas de amostras de tecido FFPE? Em terceiro lugar, é possível distinguir verdadeiras alterações de artefatos experimentais nos tumores analisados? Já se sabe que a maior parte da variabilidade na qualidade do DNA é devido ao tempo de fixação da mostra e também a duração do armazenamento dessas amostras em arquivos. As melhoras no manejo dessas amostras em muitos laboratórios de patologia, incluindo a redução do tempo de fixação para 24-48 h e a utilização de formalina tamponizada, contribuíram para a melhoria da qualidade do DNA extraído destes tecidos[284]. Para poder prever o sucesso das hibridações realizadas nos

experimentos de aCGH, alguns laboratórios avaliam a qualidade do DNA em gel de eletroforese. Esta análise pode fornecer informação sobre o tamanho, quantidade e distribuição dos tamanhos de fragmentos do DNA, mas não prediz necessariamente quais amostras terão resultados bem sucedidos na análise de aCGH. Neste contexto, aCGH para espécimes FFPE seria mais eficiente quando a qualidade do DNA pudesse ser avaliada, de uma maneira rápida, barata e viável. Van Beers EH e colaboradores[283] otimizaram procedimentos de isolamento do DNA genômico de amostras FFPE em combinação com um controle de qualidade do DNA baseado em PCR *multiplex*. Neste estudo, os autores partiram da suposição de que os *crosslinks* do DNA causados pela fixação das amostras eram a principal causa da má qualidade das hibridizações. Dessa maneira, eles se concentraram no desenvolvimento de um método melhorado de extração de DNA para reduzir a ocorrência de *crosslinks*, combinado com a PCR *multiplex* e uma marcação química específica do DNA teste com o fluorescente cianina cis-platina.

Como descrito anteriormente, a capacidade da tecnologia de aCGH de identificar pequenos CNAs depende criticamente da qualidade do DNA e da heterogeneidade da amostra. Por isso, a seleção de um DNA com qualidade suficiente, mesmo quando utilizando material FFPE, tem grande importância na análise de aCGH. A amplificação do genoma total (WGA) pode ser realizada quando as quantidades de DNA obtidas de determinadas amostras, incluindo tecidos FFPE, são insuficientes[279, 281]. A utilização de WGA para DNA isolado de amostras FFPE será discutida nas seções seguintes deste capítulo. Outro fator que pode afetar a análise de aCGH é a utilização de protocolos de marcação enzimática das amostras, que acabam reduzindo o tamanho médio do DNA, e também aumentando problemas com a hibridização devido à ligação não-específica[285], sobretudo quando o tamanho médio dos fragmentos são inferiores a 200 bp[283]. Como alternativa, protocolos de marcação química do DNA utilizando cianina cis-

platina têm substituído a marcação enzimática, e gerado bons resultados nas análises de aCGH, mesmo quando utilizando amostras FFPE[283, 286].

#### 2.4 . Amplificação Genômica Global (WGA) para a Análise de aCGH

É possível utilizar pequenas quantidades de DNA extraído de tecidos congelados ou FFPE para aCGH, através da amplificação genômica global (WGA) do DNA[281, 287]. Este DNA necessita um sistema robusto e imparcial de WGA, que seja tolerante com o DNA de má qualidade altamente degradado extraído de amostras do tipo FFPE[281]. O método mais comumente utilizado para WGA é a PCR de oligonucleotídeo degenerado (DOP-PCR). Este método também tem sido utilizado para DNA submetido à análise de CGH cromossômico extraído de amostras FFPE. A DOP-PCR utiliza um *primer* de oligonucleotídeo universal degenerado e com baixa temperatura de anelamento para a amplificação global do DNA[288].

Uma das principais preocupações para a amplificação do DNA é se a técnica de amplificação escolhida é viável, reproduzível e não introduz erros na análise de aCGH; por exemplo, um estudo anterior mostrou que a DOP-PCR introduz *vieses* na amplificação[289]. Na análise de CGH cromossômico (resolução em metáfase), estes *vieses* tem sido aceitáveis; uma vez que os investigadores utilizam estratégias tais como a exclusão, da análise, de regiões cromossômicas que podem ser inadequadamente representadas para alterações no número de cópias, tais como nos cromossomos 1p, 16p e 19[290]. Alterações ocorrem frequentemente nestas regiões, pela sua característica de serem enriquecidas em guanina e citosina, e, portanto, propensas para produzir resultados falso-positivos na análise CGH[289, 291]. Contudo, analisando modificações do número de cópias ao nível de aCGH, tais *vieses* podem ser mais

pronunciados e as limitações menos aceitáveis. Estes problemas podem ser resolvidos através da realização de experimentos de CGH inverso (i.e., a marcação recíproca das amostras teste e referência com cy3 e cy5, respectivamente); neste caso, as aberrações em 1p, 16p e 19 são consideradas reais se presentes em ambas as hibridizações. No presente estudo, realizamos estes experimentos de marcação recíproca para todas as amostras analisadas.

Little SE e colaboradores[281] compararam alguns métodos de WGA, tais como o *GenomePlex WGA Kit*, para amplificar tecidos congelados e de FFPE para o uso em aCGH. A técnica do *GenomePlex WGA* funciona através da digestão randômica do DNA em uma biblioteca fragmentada de moléculas de DNA, tornando-o potencialmente útil quando se trabalha com material FFPE. Neste método, o DNA fragmentado se torna facilmente amplificável por PCR quando flanqueado por *primers* com sítios universais, gerando uma WGA. Os autores compararam o DNA amplificado com esta técnica e o amplificado com o DOP-PCR padrão, utilizando-os em *arrays* com um espaçamento aproximado de 0.9Mb de todo o genoma. A análise de amostras congeladas *versus* FFPE demonstrou a eficácia do protocolo utilizado pelo *GenomePlex Kit* e melhor confiabilidade na sua amplificação quando comparada com o método de DOP-PCR[281].

Um outro método recente de WGA inclui o uso de uma DNA polimerase isolada do bacteriófago *Phi29* para a amplificação com primer randômicos, baseada em uma reação de deslocamento isotérmico da fita de DNA[292-293]. Esta técnica utiliza *primers* hexâmeros randômicos que se anelam em múltiplos sítios do genoma, e então a reação de amplificação se inicia pela polimerase *Phi29* nestes sítios, após a denaturação da fita linear do DNA. Durante a síntese, o deslocamento da fita do DNA complementar gera o novo DNA de fita simples que está então disponível para o anelamento por *primers* adicionais. A exigência de DNA genômico de

alta qualidade faz esta técnica não conveniente para o uso com amostras FFPE; entretanto, experimentos de aCGH com amostras do tipo FFPE utilizando este método de amplificação foram publicados[287]. As conseqüências deste estudo são surpreendentes, considerando a natureza fragmentada padrão do DNA isolado de FFPE; como os dados informaram, os produtos amplificados tinham dezenas de kilobases de comprimento.

Stoecklein e colaboradores[294] também descreveram um método de WGA denominado “SCOMP” (*single cell comparative genomic hybridization*, hibridação genômica comparativa de célula única) utilizando tecidos FFPE com diferentes tempos de armazenamento. A utilização deste método para a amplificação do DNA genômico obtido de uma única célula foi um sucesso para a técnica de CGH clássico, para a as análises de perda de heterozigose e para o seqüenciamento de DNA[295]. Em contraste com as técnicas de DOP-PCR e de pré-amplificação de extensão de *primers*, o método SCOMP é baseado em digestão do DNA utilizando a enzima de restrição *MseI*, seguido por ligação dos fragmentos de DNA através do uso de um único *primer* randômico.

Assim, como o DNA de tecidos FFPE é substancialmente degradado[296-297], a fragmentação extensa introduzida pela técnica de SCOMP pode resultar em seqüências de DNA muito curtas, que contenham só um ou nenhum sítio de restrição da *MseI*, e, que dessa maneira, não seriam eficientemente amplificadas por SCOMP. Além disso, o DNA fixado em formalina pode ser resistente à restrição por *MseI*, também resultando em uma amplificação ineficiente. Stoecklein e colaboradores[294] mostraram que não foi possível amplificar as seqüências selecionadas (318bp fragmentos) de quantidades diferentes de DNA fixado em formalina, *MseI*-digerido, em contraste com um DNA controle não-fragmentado. Em nosso estudo, o método de SCOMP foi inicialmente usado para WGA, e não resultou em uma amplificação bem sucedida



do DNA isolado de amostras FFPE. Adicionalmente, os tecidos de FFPE utilizados neste estudo foram obtidos entre 1991-2005, e foi demonstrado que tecidos que envelhecem mais de 7 anos possuem uma redução da eficiência de amplificação por SCOMP[294]. Mesmo em estudos de aCGH que mostraram uma amplificação bem sucedida de DNA isolado de amostras FFPE utilizando a técnica de SCOMP[294], a idade de armazenamento das espécimes FFPE parecia interferir no sinal da intensidade das hibridizações.

Outros relatórios demonstraram o uso de uma técnica de WGA, utilizando a enzima *sequenase* para amplificar DNA genômico em dois estágios ([298], adaptado de [299]). Este método de amplificação foi utilizado com sucesso para amplificar uma quantidade inicial inferior a 10ng de DNA, e foi útil na comparação do enriquecimento genômico relativo entre duas amostras[299]. O protocolo inclui dois estágios de reações enzimáticas; no estágio I, a enzima *sequenase* é utilizada para estender os *primers* que se anelaram randomicamente (*primer A*) e gerar uma biblioteca de fragmentos que serão subseqüentemente submetidos a um número limitado de ciclos de PCR. Durante o estágio II, o *primer B* específico (com seqüência parcialmente igual ao *primer A*) é utilizado para amplificar os produtos anteriormente gerados pela incorporação dos amino-alquil dUTPs. Esta técnica gera DNA suficiente para o uso subseqüente por outras técnicas; e a necessidade de um DNA fragmentado para realização desse protocolo faz esta técnica ideal para o uso com amostras do tipo FFPE. Em nosso estudo, utilizamos esta técnica com sucesso para amplificar o DNA genômico extraído de tecidos FFPE. Estes resultados são mostrados no Capítulo 5.

## 2.5. Comparação de Diferentes Plataformas de a-CGH – BACs vs. Arrays de Oligonucleotídeos.

Ficou evidente que as variações no número de cópias do DNA são muito comuns no genoma humano e se associam com muitas doenças genéticas, incluindo o câncer. Por conseguinte, muitos *arrays* foram projetados para identificar estas alterações em nível de DNA, bem como descobrir amplificações e deleções de oncogenes e genes supressores tumorais[300-301]. Identificar as alterações genômicas segmentárias específicas, e os genes que elas contêm, poderá produzir marcadores moleculares para o diagnóstico e terapia. Para obter tratamentos eficazes, é essencial que a identificação dessas alterações genéticas associadas com a doença sejam acuradas.

Como descrito anteriormente neste capítulo, o aCGH oferece a possibilidade de uma análise precisa de alterações no número de cópias do DNA. O aCGH surgiu de uma transformação do CGH clássico, através do uso de clones de sondas genômicas com seqüências verificadas e mapeadas no genoma, impressas em microarranjos (*arrays*) para servir como alvo de hibridação para amostras biológicas[279]. Os métodos de análise de dados de aCGH continuam em desenvolvimento, e muitas vezes são otimizados para determinadas características da plataforma em uso, ou das aplicações a que servem. Diferentes plataformas de aCGH estão disponíveis e têm sido freqüentemente utilizadas no estudo do câncer. Estas plataformas se diferenciam em vários aspectos, tais como a resolução, que é determinada pelo tamanho e/ou espaçamento entre as sondas genômicas utilizadas para a construção do *array*.

Existem plataformas cujo objetivo é hibridizar o DNA num *array* que contenha seqüências de cromossomos bacterianos artificiais, conhecidas como BACs, e por isso são chamados BAC *arrays*. Estes *arrays* estão atualmente disponíveis para determinar regiões genômicas com ganhos ou perdas com níveis de resolução de megabases (1Mb) ou submegabases (100KBs)[302-303]. Ao contrário, as plataformas recentes de companhias como a

*Agilent*, *Affymetrix*, *NimbleGen* e *Illumina* oferecem *arrays* de oligonucleotídeos que representam 500.000 ou mais regiões genômicas ou seqüências representadas nos *arrays*, e que avaliam as diferenças no número de cópias do DNA de amostras alvo comparadas com uma amostra de referência. Plataformas como as produzidas pela *Affymetrix* e *Illumina* apresentam a vantagem de também permitirem a análise de polimorfismos de base única (SNPs), e por isso têm sido preferencialmente adotadas nestes estudos.

Vários relatos da literatura compararam o uso de diferentes plataformas para DNA (p. ex. *BACs vs. oligos*) no estudo do câncer, e constataram que dependendo do *array* utilizado, existem particulares vantagens e desvantagens para cada plataforma (p. ex. resolução mais alta e proporção de sinal específico para inespecífico)[300, 304]. A capacidade de aCGH de determinar alterações em regiões do genoma depende do tamanho e do posicionamento dos clones *array*. Os *BAC arrays*, com uma resolução média de 150KBs, possuem tipicamente grandes segmentos de 100-160KBs, ao passo que os *arrays* de oligonucleotídeos contêm tamanhos de segmento mais curtos da ordem de 50-100KBs. Dessa maneira, as plataformas de oligonucleotídeos fornecem uma detecção melhorada de ganhos e perdas em comparação com os *BAC arrays*, devido à sua resolução mais alta (em média de 35KB), e devido ao maior número de sondas por *array*. Contudo, uma desvantagem da plataforma de oligos é que o tamanho pequeno das sondas desse *array* pode resultar em diminuição da intensidades dos sinais de hibridação, e conseqüentemente, gerar um número maior de achados falso-positivos. Tipicamente, 20–50 oligonucleotídeos adjacentes são necessários para uma chamada acurada (isto é, > 90%). Assim, a identificação das regiões de CNA deve ser compensada pelo uso de instrumentos estatísticos e algoritmos mais complexos, que possam excluir qualquer *vies* experimental.

Um estudo realizado pelo Grupo de Pesquisa de Microarrays (MARG) em 2007, demonstrou como tanto as plataformas de BAC como as de oligos aCGH podem determinar ganhos e perdas do número de cópias do DNA. Este grupo analisou os CNAs do DNA da linhagem celular HL60 de leucemia (que possui amplificação conhecida em 8q24, e deleções em 5q11.2-q31, 6q12, 9p21.3-p22, 10p12-p15, 14q22-q31, 17p12-p13.3)[305], comparando 5 plataformas comercialmente disponíveis: *RPCI BAC 19K arrays*, *Agilent 44K*, *Illumina HAP 550K*, *Affymetrix 500K* e *Affymetrix U133 array* de expressão, que foi recentemente mostrado ser útil na identificação de alterações no número de cópias do DNA[306]. Eles encontraram boa precisão dos resultados em todas as plataformas, com um coeficiente de variação (CV) dos experimentos duplicados entre 2.1 a 10.7%. Contudo, a plataforma *Illumina* e as plataformas da *Affymetrix 500KB* identificaram 9 dos 9 (100%) de todas as regiões de perda e ganhos conhecidas no DNA da linhagem celular HL60, enquanto as plataformas da *Agilent*, *Affymetrix U133* e *RPCI BAC arrays* identificaram 8 das 9 (89%) regiões conhecidas. Este grupo também analisou a variação do número de cópias usando um método chamado segmentação binária circular (CBS); neste método, a razão do log do sinal de intensidade de HL60/referência de quatro experimentos independentes para cada plataforma, foi utilizada para avaliar a resolução das plataformas. Esta análise mostrou que todas as plataformas testadas foram capazes de determinar ganhos e perdas conhecidos, sugerindo que essas plataformas possuem boa resolução para a determinação de CNAs.

Mais recentemente, Hester e colaboradores[304] compararam as plataformas da *Agilent*, *Illumina 550K*, *Affymetrix 500 KB* e *Affymetrix U133 v2.0* e uma plataforma de BAC array produzida no *Roswell Park Cancer Institute*, Buffalo, NY. Este estudo confirmou que essas plataformas possuem uma boa capacidade para determinar as regiões de ganho e perdas

conhecidas na mesma linhagem celular acima descrita (HL-60). Além disso, novas regiões de deleções foram identificadas nesta linhagem celular, sugerindo diferenças relacionadas à resolução de cada plataforma e/ou aos parâmetros de análise(s) utilizada(s). Dessa maneira, algumas plataformas de aCGH são melhor ajustadas para o campo de pesquisa, enquanto outras são preferencialmente utilizadas para aplicações diagnósticas.

Estes estudos demonstram resultados encorajadores na utilização destas plataformas de aCGH para a determinação de CNAs. De acordo com estes relatos, na ausência de diferenças técnicas substanciais, a seleção da plataforma a ser utilizada deve ser baseada em considerações práticas e na aplicação destinada. Em nosso estudo, utilizamos uma plataforma de oligo aCGH, 385K *Whole-Genome Tiling array*, disponível pela *NimbleGen* (Roche). O formato deste *array* utiliza 385.000 sondas que cobrem todo o genoma humano e são projetadas para detectar CNAs a uma resolução tão baixa quanto ~5-10Kb. Selecionamos esta plataforma, uma vez que ela possui características importantes, tais como ultra-alta densidade e sondas longas de oligos que permitem uma alta resolução e análise abrangente do genoma inteiro. Também, a plataforma da *NimbleGen* foi aquela que demonstrou resultados de hibridização mais viáveis e consistentes, uma qualidade maior dos dados gerados para produtos de DNA amplificados de amostras FFPE (isto é, com um número significativo e satisfatório de sinais específicos e um menor número de sinais inespecíficos), quando em comparação com outras plataformas que testamos (*BAC arrays*, *Affymetrix UI33* e *Agilent 244K*). Adicionalmente, a cobertura destas sondas de regiões de baixa cópia de repetição (p. ex., duplicações segmentárias) permite a detecção aumentada de alterações do número de cópia do DNA associadas com rearranjos patogênicos, que podem passar despercebidos em outras plataformas. A *Roche NimbleGen* também oferece instrumentos de software de análises e processamento de dados, ao mesmo tempo que permite a utilização de

outros softwares comercialmente disponíveis para a análise de CNAs, tais como PARTEK e NEXUS, que são compatíveis com os arquivos de dados gerados pelo NimbleScan.

## 2.6. Análises de CNAs em CCECP

Várias regiões cromossômicas têm sido implicadas em perdas (2p, 3p, 4, 5q, 7q, 8p, 9p, 11, 13, 14q, 15p, 16p e 18q) e ganhos (3q, 5p, 7q, 8q24, 9q22-34, 11q13, 14q24, 15q, 16p, 19, 20q24, e 22q) nos cânceres de cabeça e pescoço, e determinam as alterações genéticas mais comuns neste tipo tumoral[307-309]. Estas alterações também foram descritas em outros estudos; através da combinação de técnicas, tais como CGH clássico, cariótipo espectral (SKY) e *microarrays* de expressão gênica. Squire JA e colaboradores[310] examinaram as alterações no número de cópias do DNA de sete linhagens celulares e quatro CCECP primários. Os ganhos mais freqüentes foram observados em 3q24-27, 8q e 6q22-qter, assim como perda em 18q22-qter; além disso, a análise de SKY detectou a presença de rearranjos complexos e heterogeneidade nestas amostras.

Em um estudo recente, Lin M e colaboradores[311] mostraram resultados semelhantes para alterações freqüentemente observadas em CCECP; relataram amplificação em 3q26.3 e 11q13.3, além dos ganhos em 3q29, 8q13.1, 8q22.3, 9q32, 10q24.32, 14q32.32, 17q25.1 e 20q13.33. Estes autores também sugeriram que múltiplos sítios de amplificação possam surgir como resultado de amplificações em 3p26.3 e 11q13.3, e relatou que vários genes incluindo *MYCL1*, *CCNL1*, *PIK3CA*, *EGFR*, *FGFR1*, *EDD1*, *LRP12*, *MYCC*, *CCND1*, *EMSI*, *MDM2*, *ERBB2*, estão presentes nestas regiões[311].

A análises de CNA e de perda de heterozigose foram combinadas com técnicas de pequena escala (p. ex., FISH e métodos baseados em PCR) e de ampla escala (p. ex., aCGH) em estudos de CCECP, e correlacionadas com recorrência de tumores e sobrevida dos pacientes[312]. A amplificação do receptor do fator de crescimento epidermal (*EGFR*) é um dos eventos de CNAs mais conhecidos em CCECP, e está correlacionado com a baixa sobrevida dos pacientes[313]. Além disso, estudos têm mostrado que CNAs em 3q, 7p, 11q e 22q, e LOH em 3p, 8p, 9p e 13q se associam muitas vezes com a sobrevida dos pacientes, e que CNA em 11q e LOH em 9p e 17p são alterações genéticas associadas com a recorrência tumoral[312]. Outras anormalidades cromossômicas, tais como ganhos em 3q21-29 e 12q24 e perdas em 5q11-15, 6q14-21, 8p21-22, 18q22 e 21q11-21, também foram correlacionadas com a sobrevida[314-316]. Uma região de deleção no cromossomo 22q13, envolvendo especificamente o gene *DI1*, que codifica uma proteína do metabolismo de drogas, foi associada com um pior prognóstico de pacientes com CCECP e história familiar de câncer[317].

Num estudo recente, Bauer VL e colaboradores[318] identificaram regiões significantes de ganhos em 1q43 e 16q23-24, e de perda em 18q22 que foram associadas com um pior prognóstico em 117 pacientes com CCECP de estadiamento avançado, utilizando a combinação das análises de aCGH (BAC *array* - *Humarray* 3.2) e FISH.

Outro estudo mostrou aberrações do número de cópias cromossômicas em amostras de CCECP hrHPV (pacientes com papiloma vírus de alto risco) positivas e negativas[319]. Neste estudo, perdas em 3p11.2-26.3, 5q11.2-35.2, 9p21.1-24 e 18q12.1-23, e ganhos em 11q12.1-13.4 estavam presentes em tumores HPV negativos e ausentes em tumores HPV positivos. Esta observação apoia a hipótese de que podem existir dois caminhos diferentes para o desenvolvimento dos CCECP, categorizando-os em dois grupos: um determinado pelo papiloma

vírus de alto risco, HPV16, e outro determinado por carcinógenos ambientais do tabaco e álcool[319-320]. Ganhos em 3q22.2-qter, 5p15.2-pter, 8p11.2-qter, 9q22 - 34.1 e 20p-20q, e perdas em 11q14.1-qter e 13q11-33, foram freqüentes (> 33% dos casos) em ambos os grupos de CCECP[319]. Estas alterações genéticas estavam presentes em tumores HPV positivos e negativos, indicando que possam ser eventos mais importantes em estágios tardios da progressão dos CCECP, nos quais os tumores, que se originam de fatores etiológicos diferentes (viral e não-viral), acabam compartilhando as mesmas alterações genéticas.

### *2.7. CNAs Aplicados à Progressão do Câncer Oral*

Até o presente momento, são poucos os estudos que identificaram alterações genômicas globais em lesões orais pré-malignas, principalmente devido à raridade das displasias de baixo grau com dados clínicos longitudinais detalhados. Estes estudos têm se concentrado na identificação de genes/mecanismos genéticos envolvidos na progressão de CCEOs.

Noutomi Y e colaboradores[321] utilizaram CGH clássico para determinar alterações genéticas tanto em CCEOs primários, como em lesões displásicas adjacentes biopsiadas de 35 pacientes. Eles identificaram que o ganho em 8q22-23 foi a alteração mais freqüentemente observada tanto em displasia leve como em CCEO, e sugeriram este evento como um dos que ocorre primeiro durante a tumorigênese. Adicionalmente, eles mostraram que o número médio de aberrações do número de cópias das sequências do DNA (DSCNAs) aumentava com a progressão de displasia leve à carcinoma invasivo. Interessantemente, este estudo demonstrou que os CCEOs com um grande (L) número de DSCNAs se segregavam dos CCEOs com pequeno (S) número de DSCNAs (CCEOs-L se associaram com o câncer de estágios avançados,



e os CCEOs-S estavam associados com os estágios iniciais da doença). Os ganhos em 3q26-qter, 8q, 11q13, 14q e 20q, e as perdas em 4q, 5q12-22, 6q, 8p, 13q e 18q22-qter, foram alterações comuns entre CCEO-L e CCEO-S. Contudo, ganhos em 5p15, 7p, 17q11-22 e 18p, e perdas em 3p14-21, 4p e 9p, foram detectados exclusivamente em CCEOs-L, e nas suas lesões displásticas adjacentes. Estes achados sugerem que duas categorias diferentes de CCEOs possam se originar de tipos diferentes de displasias, através de diferentes vias genéticas[321].

Em um estudo mais recente, Garnis C e colaboradores[322] realizaram uma análise genômica global utilizando o *tiling BAC* aCGH (British Columbia Cancer Research Centre, Vancouver, Canada) para avaliar as alterações do número de cópias do DNA em um total de 62 lesões orais pré-malignas de estágios iniciais vs. avançado (32 displasias de alto grau (HGD) e carcinoma *in situ* (CIS), 9 displasias de baixo grau (LGDs) e 21 LGDs não-progressivas), em comparação com um grupo diferente de CCEOs (n=24). Estes autores também compararam os CNAs dos tecidos pré-malignos (OPMLs) de diferentes graus durante as etapas da progressão do câncer oral, a fim de identificar alterações no DNA específicas para a transformação maligna. Os resultados mostraram que as lesões HGD, também classificadas como pré-invasivas, possuíam um grau de instabilidade genômica também observada em tumores invasivos. A alteração mais comumente observada em lesões HGD/CIS estavam presentes em 1p, 2q, 3q, 5q, 7q e 8p, enquanto os CNAs mais comuns em CCEO invasivo ocorreram em 5p, 9q, 11q e 19p[322]. Tais resultados sugerem que as alterações genéticas responsáveis pela iniciação e progressão da doença são diferentes daquelas que levam a invasão tumoral. É interessante notar que não houve nenhuma alteração comumente compartilhada entre as lesões LGD progressivas. O cromossomo 9p foi o mais freqüentemente alterado (78%), seguido pelos cromossomos 8q, 20p e 20q (56%);

estes achados indicam que o grau de instabilidade genômica em uma dada OPML pode se útil para prever a probabilidade dessas lesões progredirem à carcinoma oral.

Outro estudo mostrou não só CNAs associados com a progressão de CCEO, mas também com outros sítios anatômicos do CCECP, através da análise de aberrações cromossômicas em amostras pré-malignas e malignas de pacientes egípcios[323]. Este estudo foi conduzido com 41 pacientes com lesões pré-malignas, 79 pacientes com lesões malignas de laringe, esôfago, nasofaringe, cavidade nasal e cavidade oral, e 50 indivíduos controles sem câncer. Os autores mostraram que 41% das lesões pré-cancerosas possuíam perda de 3p, e ganhos em 3q, 8q e 11q13 foram observados em 51%, 29%, e 22% dos casos, respectivamente. Em lesões malignas, 63% possuíam perdas em 3p13-p24, 59.5% em 5q12-23, 49.5% em 8p22-p23, 45.5% em 9p21-p24, 40.5% em 18q22-q23, e os ganhos foram observados em 3q para 66% dos casos, 8q para 39% e em 11q13 para 16% dos casos. Estes resultados sugerem que a análise de aberrações cromossômicas, em 3p, 3q, 8q e 11q13, nas lesões orais suspeitas em grupos de alto risco (fumantes e alcoólicos), possa ser importante para o primeiro diagnóstico de CCECP. A existência de uma alta percentagem de alterações genéticas nestas regiões poderia levar a uma intervenção mais ativa junto ao paciente, tais como químio-prevenção e o monitoramento mais frequente destes pacientes para a detecção precoce da doença.

Neste presente estudo, realizamos uma análise genômica global para a detecção de CNAs comuns entre amostras seqüenciais de pacientes com leucoplasias progressivas e CCEOs correspondentes, do mesmo sítio anatômico, através da comparação dos genomas destes tecidos orais durante a progressão ao câncer. A inclusão de amostras seqüenciais de um mesmo paciente permitiu uma comparação mais acurada das alterações genéticas que ocorrem nas lesões orais durante a transformação maligna. Até o presente, não há estudos publicados, de outros grupos de

pesquisa, utilizando amostras seqüenciais para delinear alterações genéticas (por exemplo, CNAs) associadas com a progressão do câncer oral. Recentemente, o nosso grupo publicou o primeiro estudo utilizando tais amostras seqüenciais de leucoplasias e CCEOs para a análise de expressão de miRNAs, e determinou uma assinatura de miRNAs associada com a progressão de CCEO[324]. Com o uso deste grupo amostral que inclui espécimes seqüenciais dos mesmos pacientes, o nosso estudo contribui para uma melhor compreensão das alterações genômicas que ocorrem durante os diferentes estágios da progressão tumoral dos CCEOs. Além disso, os nossos resultados oferecem a possibilidade do delineamento de um modelo genético de progressão, baseado em alterações do número de cópias do DNA e de expressão de miRNAs, associados com a transformação maligna desta doença. Os achados deste estudo tem o potencial de contribuir para o desenvolvimento de estratégias precoces de intervenção, ajudando em nível prognóstico e conseqüentemente resultando em uma melhor sobrevida dos pacientes.

## Referências

1. Evans PhR, M.P., Gullane PJ, *Principles and practice of head and neck oncology* 2003, London, UK: Martin Dunitz. 633.
2. Ragin, C.C., F. Modugno, and S.M. Gollin, *The epidemiology and risk factors of head and neck cancer: a focus on human papillomavirus*. J Dent Res, 2007. **86**(2): p. 104-14.
3. Parkin, D.M., et al., *Global cancer statistics, 2002*. CA Cancer J Clin, 2005. **55**(2): p. 74-108.
4. Jemal, A., et al., *Cancer statistics, 2007*. CA Cancer J Clin, 2007. **57**(1): p. 43-66.
5. Wynder, E.L. and I.J. Bross, *Aetiological factors in mouth cancer; an approach to its prevention*. Br Med J, 1957. **1**(5028): p. 1137-43.
6. Brennan, P., *Behind the scenes at Baltimore*. Br Dent J, 1995. **178**(1): p. 33-4.
7. Choi, S.Y. and H. Kahyo, *Effect of cigarette smoking and alcohol consumption in the aetiology of cancer of the oral cavity, pharynx and larynx*. Int J Epidemiol, 1991. **20**(4): p. 878-85.
8. Franceschi, S., et al., *Comparison of the effect of smoking and alcohol drinking between oral and pharyngeal cancer*. Int J Cancer, 1999. **83**(1): p. 1-4.
9. Macfarlane, G.J., et al., *Alcohol, tobacco, diet and the risk of oral cancer: a pooled analysis of three case-control studies*. Eur J Cancer B Oral Oncol, 1995. **31B**(3): p. 181-7.
10. Morse, D.E., et al., *Smoking and drinking in relation to oral epithelial dysplasia*. Cancer Epidemiol Biomarkers Prev, 1996. **5**(10): p. 769-77.
11. Blot, W.J., et al., *Smoking and drinking in relation to oral and pharyngeal cancer*. Cancer Res, 1988. **48**(11): p. 3282-7.
12. Castellsague, X., et al., *The role of type of tobacco and type of alcoholic beverage in oral carcinogenesis*. Int J Cancer, 2004. **108**(5): p. 741-9.
13. Dal Maso, L., et al., *Alcohol drinking outside meals and cancers of the upper aerodigestive tract*. Int J Cancer, 2002. **102**(4): p. 435-7.
14. Lewin, F., et al., *Smoking tobacco, oral snuff, and alcohol in the etiology of squamous cell carcinoma of the head and neck: a population-based case-referent study in Sweden*. Cancer, 1998. **82**(7): p. 1367-75.
15. Goldstein, D.P. and J.C. Irish, *Head and neck squamous cell carcinoma in the young patient*. Curr Opin Otolaryngol Head Neck Surg, 2005. **13**(4): p. 207-11.
16. Wiseman, S.M., et al., *Squamous cell carcinoma of the head and neck in nonsmokers and nondrinkers: an analysis of clinicopathologic characteristics and treatment outcomes*. Ann Surg Oncol, 2003. **10**(5): p. 551-7.
17. Koch, W.M., et al., *Head and neck cancer in nonsmokers: a distinct clinical and molecular entity*. Laryngoscope, 1999. **109**(10): p. 1544-51.
18. Llewellyn, C.D., et al., *Squamous cell carcinoma of the oral cavity in patients aged 45 years and under: a descriptive analysis of 116 cases diagnosed in the South East of England from 1990 to 1997*. Oral Oncol, 2003. **39**(2): p. 106-14.
19. Greer, R.O., Jr., L.R. Eversole, and L.K. Crosby, *Detection of human papillomavirus-genomic DNA in oral epithelial dysplasias, oral smokeless tobacco-associated*

- leukoplakias, and epithelial malignancies.* J Oral Maxillofac Surg, 1990. **48**(11): p. 1201-5.
20. Park, N.H., et al., *In vitro and animal studies of the role of viruses in oral carcinogenesis.* Eur J Cancer B Oral Oncol, 1992. **28B**(2): p. 145-52.
  21. Lo, Y.M., et al., *Quantitative and temporal correlation between circulating cell-free Epstein-Barr virus DNA and tumor recurrence in nasopharyngeal carcinoma.* Cancer Res, 1999. **59**(21): p. 5452-5.
  22. Makitie, A.A., et al., *Epstein-Barr virus DNA measured in nasopharyngeal brushings in patients with nasopharyngeal carcinoma: pilot study.* J Otolaryngol, 2004. **33**(5): p. 299-303.
  23. Vokes, E.E., D.N. Liebowitz, and R.R. Weichselbaum, *Nasopharyngeal carcinoma.* Lancet, 1997. **350**(9084): p. 1087-91.
  24. D'Souza, G., et al., *Case-control study of human papillomavirus and oropharyngeal cancer.* N Engl J Med, 2007. **356**(19): p. 1944-56.
  25. zur Hausen, H. and E.M. de Villiers, *Human papillomaviruses.* Annu Rev Microbiol, 1994. **48**: p. 427-47.
  26. Huang, B., et al., *High-capacity adenoviral vector-mediated reduction of huntingtin aggregate load in vitro and in vivo.* Hum Gene Ther, 2007. **18**(4): p. 303-11.
  27. Boyer, S.N., D.E. Wazer, and V. Band, *E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway.* Cancer Res, 1996. **56**(20): p. 4620-4.
  28. Hubbert, N.L., S.A. Sedman, and J.T. Schiller, *Human papillomavirus type 16 E6 increases the degradation rate of p53 in human keratinocytes.* J Virol, 1992. **66**(10): p. 6237-41.
  29. Munger, K., et al., *The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes.* J Virol, 1989. **63**(10): p. 4417-21.
  30. Munoz, N., et al., *Epidemiologic classification of human papillomavirus types associated with cervical cancer.* N Engl J Med, 2003. **348**(6): p. 518-27.
  31. zur Hausen, H., *Papillomavirus infections--a major cause of human cancers.* Biochim Biophys Acta, 1996. **1288**(2): p. F55-78.
  32. Adams, V., et al., *Prevalence of human papilloma virus DNA in head and neck cancers carrying wild-type or mutant p53 tumor suppressor genes.* Anticancer Res, 1999. **19**(1A): p. 1-6.
  33. Badaracco, G., et al., *Human papillomavirus in head and neck carcinomas: prevalence, physical status and relationship with clinical/pathological parameters.* Anticancer Res, 2000. **20**(2B): p. 1301-5.
  34. Blot, W.J., *Esophageal cancer trends and risk factors.* Semin Oncol, 1994. **21**(4): p. 403-10.
  35. Gillison, M.L., et al., *Evidence for a causal association between human papillomavirus and a subset of head and neck cancers.* J Natl Cancer Inst, 2000. **92**(9): p. 709-20.
  36. Haraf, D.J., et al., *Human papilloma virus and p53 in head and neck cancer: clinical correlates and survival.* Clin Cancer Res, 1996. **2**(4): p. 755-62.
  37. Hodge, K.M., M.B. Flynn, and T. Drury, *Squamous cell carcinoma of the upper aerodigestive tract in nonusers of tobacco.* Cancer, 1985. **55**(6): p. 1232-5.

38. Hoshikawa, T., et al., *Detection of human papillomavirus DNA in laryngeal squamous cell carcinomas by polymerase chain reaction*. Laryngoscope, 1990. **100**(6): p. 647-50.
39. Klusmann, J.P., S. Weissenborn, and P.G. Fuchs, *Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck*. N Engl J Med, 2001. **345**(5): p. 376; author reply 377.
40. Lindel, K., et al., *Human papillomavirus positive squamous cell carcinoma of the oropharynx: a radiosensitive subgroup of head and neck carcinoma*. Cancer, 2001. **92**(4): p. 805-13.
41. Mellin, H., et al., *Human papillomavirus (HPV) DNA in tonsillar cancer: clinical correlates, risk of relapse, and survival*. Int J Cancer, 2000. **89**(3): p. 300-4.
42. Mineta, H., et al., *Human papilloma virus (HPV) type 16 and 18 detected in head and neck squamous cell carcinoma*. Anticancer Res, 1998. **18**(6B): p. 4765-8.
43. Mork, J., et al., *Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck*. N Engl J Med, 2001. **344**(15): p. 1125-31.
44. Niedobitek, G., et al., *Detection of human papillomavirus type 16 DNA in carcinomas of the palatine tonsil*. J Clin Pathol, 1990. **43**(11): p. 918-21.
45. Nishioka, S., et al., *Human papillomavirus as a risk factor for head and neck cancers--a case-control study*. Acta Otolaryngol Suppl, 1999. **540**: p. 77-80.
46. Ostwald, C., et al., *Human papillomavirus DNA in oral squamous cell carcinomas and normal mucosa*. J Oral Pathol Med, 1994. **23**(5): p. 220-5.
47. Paz, I.B., et al., *Human papillomavirus (HPV) in head and neck cancer. An association of HPV 16 with squamous cell carcinoma of Waldeyer's tonsillar ring*. Cancer, 1997. **79**(3): p. 595-604.
48. Shindoh, M., et al., *Detection of human papillomavirus DNA sequences in oral squamous cell carcinomas and their relation to p53 and proliferating cell nuclear antigen expression*. Cancer, 1995. **76**(9): p. 1513-21.
49. Sisk, E.A., et al., *Human papillomavirus infection in "young" versus "old" patients with squamous cell carcinoma of the head and neck*. Head Neck, 2000. **22**(7): p. 649-57.
50. Venuti, A., et al., *Physical state and expression of human papillomavirus in laryngeal carcinoma and surrounding normal mucosa*. J Med Virol, 2000. **60**(4): p. 396-402.
51. Hansson, B.G., et al., *Strong association between infection with human papillomavirus and oral and oropharyngeal squamous cell carcinoma: a population-based case-control study in southern Sweden*. Acta Otolaryngol, 2005. **125**(12): p. 1337-44.
52. Herrero, R., *Chapter 7: Human papillomavirus and cancer of the upper aerodigestive tract*. J Natl Cancer Inst Monogr, 2003(31): p. 47-51.
53. Schwartz, S.M., et al., *Oral cancer risk in relation to sexual history and evidence of human papillomavirus infection*. J Natl Cancer Inst, 1998. **90**(21): p. 1626-36.
54. Smith, E.M., et al., *Age, sexual behavior and human papillomavirus infection in oral cavity and oropharyngeal cancers*. Int J Cancer, 2004. **108**(5): p. 766-72.
55. Gillison, M.L., *Human papillomavirus-related diseases: oropharynx cancers and potential implications for adolescent HPV vaccination*. J Adolesc Health, 2008. **43**(4 Suppl): p. S52-60.
56. Hunter, K.D., E.K. Parkinson, and P.R. Harrison, *Profiling early head and neck cancer*. Nat Rev Cancer, 2005. **5**(2): p. 127-35.
57. Jemal, A., et al., *Cancer statistics, 2008*. CA Cancer J Clin, 2008. **58**(2): p. 71-96.

58. Parkin, D.M., *Global cancer statistics in the year 2000*. Lancet Oncol, 2001. **2**(9): p. 533-43.
59. Edwards, B.K., et al., *Annual report to the nation on the status of cancer, 1975-2002, featuring population-based trends in cancer treatment*. J Natl Cancer Inst, 2005. **97**(19): p. 1407-27.
60. Murdock, J.M. and J.L. Gluckman, *African-American and white head and neck carcinoma patients in a university medical center setting. Are treatments provided and are outcomes similar or disparate?* Cancer, 2001. **91**(1 Suppl): p. 279-83.
61. Pereira, M.C., et al., *Histologic subtypes of oral squamous cell carcinoma: prognostic relevance*. J Can Dent Assoc, 2007. **73**(4): p. 339-44.
62. Atula, T., et al., *The evaluation and treatment of the neck in carcinoma of the oral cavity*. Acta Otolaryngol Suppl, 1997. **529**: p. 223-5.
63. Brunin, F., et al., *Cancer of the base of the tongue: past and future*. Head Neck, 1999. **21**(8): p. 751-9.
64. Silverman, S., Jr. and P.B. Sugeran, *Oral premalignancies and squamous cell carcinoma*. Clin Dermatol, 2000. **18**(5): p. 563-8.
65. Forastiere, A., et al., *Head and neck cancer*. N Engl J Med, 2001. **345**(26): p. 1890-900.
66. Chambers, A.F., A.C. Groom, and I.C. MacDonald, *Dissemination and growth of cancer cells in metastatic sites*. Nat Rev Cancer, 2002. **2**(8): p. 563-72.
67. Ramaswamy, S., et al., *A molecular signature of metastasis in primary solid tumors*. Nat Genet, 2003. **33**(1): p. 49-54.
68. Chung, C.H., et al., *Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression*. Cancer Cell, 2004. **5**(5): p. 489-500.
69. Katayama, A., et al., *Expressions of matrix metalloproteinases in early-stage oral squamous cell carcinoma as predictive indicators for tumor metastases and prognosis*. Clin Cancer Res, 2004. **10**(2): p. 634-40.
70. O'Donnell, M.A., et al., *The essential role of interferon-gamma during interleukin-12 therapy for murine transitional cell carcinoma of the bladder*. J Urol, 2004. **171**(3): p. 1336-42.
71. Roepman, P., et al., *An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas*. Nat Genet, 2005. **37**(2): p. 182-6.
72. Thompson, E.W., D.F. Newgreen, and D. Tarin, *Carcinoma invasion and metastasis: a role for epithelial-mesenchymal transition?* Cancer Res, 2005. **65**(14): p. 5991-5; discussion 5995.
73. Sawair, F.A., et al., *Invasive front grading: reliability and usefulness in the management of oral squamous cell carcinoma*. J Oral Pathol Med, 2003. **32**(1): p. 1-9.
74. Haddad, R.I. and D.M. Shin, *Recent advances in head and neck cancer*. N Engl J Med, 2008. **359**(11): p. 1143-54.
75. Cohen, E.E., et al., *Response of some head and neck cancers to epidermal growth factor receptor tyrosine kinase inhibitors may be linked to mutation of ERBB2 rather than EGFR*. Clin Cancer Res, 2005. **11**(22): p. 8105-8.
76. Li, M., et al., *Induction of retinoic acid receptor-beta suppresses cyclooxygenase-2 expression in esophageal cancer cells*. Oncogene, 2002. **21**(3): p. 411-8.
77. Tseng, J.E., et al., *Phase II study of the antiangiogenesis agent thalidomide in recurrent or metastatic squamous cell carcinoma of the head and neck*. Cancer, 2001. **92**(9): p. 2364-73.

78. Slaughter, D.P., H.W. Southwick, and W. Smejkal, *Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin*. *Cancer*, 1953. **6**(5): p. 963-8.
79. Braakhuis, B.J., et al., *A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications*. *Cancer Res*, 2003. **63**(8): p. 1727-30.
80. Ha, P.K. and J.A. Califano, *The molecular biology of mucosal field cancerization of the head and neck*. *Crit Rev Oral Biol Med*, 2003. **14**(5): p. 363-9.
81. Izzo, J.G., et al., *Dysregulated cyclin D1 expression early in head and neck tumorigenesis: in vivo evidence for an association with subsequent gene amplification*. *Oncogene*, 1998. **17**(18): p. 2313-22.
82. Mao, L., et al., *Frequent microsatellite alterations at chromosomes 9p21 and 3p14 in oral premalignant lesions and their value in cancer risk assessment*. *Nat Med*, 1996. **2**(6): p. 682-5.
83. Partridge, M., et al., *A case-control study confirms that microsatellite assay can identify patients at risk of developing oral squamous cell carcinoma within a field of cancerization*. *Cancer Res*, 2000. **60**(14): p. 3893-8.
84. Rosin, M.P., et al., *Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia*. *Clin Cancer Res*, 2000. **6**(2): p. 357-62.
85. Fearon, E.R. and B. Vogelstein, *A genetic model for colorectal tumorigenesis*. *Cell*, 1990. **61**(5): p. 759-67.
86. Lee, J.J., et al., *Predicting cancer development in oral leukoplakia: ten years of translational research*. *Clin Cancer Res*, 2000. **6**(5): p. 1702-10.
87. Petersen, P.E., *Strengthening the prevention of oral cancer: the WHO perspective*. *Community Dent Oral Epidemiol*, 2005. **33**(6): p. 397-9.
88. Jin, Y.T., et al., *Genetic alterations in oral squamous cell carcinoma of young adults*. *Oral Oncol*, 1999. **35**(3): p. 251-6.
89. Warnakulasuriya, S., N.W. Johnson, and I. van der Waal, *Nomenclature and classification of potentially malignant disorders of the oral mucosa*. *J Oral Pathol Med*, 2007. **36**(10): p. 575-80.
90. Axell, T., et al., *Oral white lesions with special reference to precancerous and tobacco-related lesions: conclusions of an international symposium held in Uppsala, Sweden, May 18-21 1994. International Collaborative Group on Oral White Lesions*. *J Oral Pathol Med*, 1996. **25**(2): p. 49-54.
91. Kramer, I.R., et al., *Definition of leukoplakia and related lesions: an aid to studies on oral precancer*. *Oral Surg Oral Med Oral Pathol*, 1978. **46**(4): p. 518-39.
92. Axell T, H.P., Kramer IRH, Pindborg JJ, Shear M., *International seminar on oral leukoplakia and associated lesions related to tobacco habits*. *Community Dent Oral Epidemiol*, 1984. **12**: p. 145-54.
93. Pindborg JJ, R.P., Smith CJ, Van der Waal I., *World Health Organization International Histological Classification of Tumours*. Histological typing of cancer and precancer of the oral mucosa. 1997, Berlin: Springer.
94. Barnes L, E.J., Reichart P, Sidransky D, *World Health Organization. World Health Organization Classification of Tumours*, ed. P.G.H.a.N. Tumours. 2005, Lyon: International Agency for Research on Cancer (IARC).
95. Petti, S., *Pooled estimate of world leukoplakia prevalence: a systematic review*. *Oral Oncol*, 2003. **39**(8): p. 770-80.



96. van der Waal, I. and P.A. Reichart, *Oral proliferative verrucous leukoplakia revisited*. Oral Oncol, 2008. **44**(8): p. 719-21.
97. Lind, P.O., *Malignant transformation in oral leukoplakia*. Scand J Dent Res, 1987. **95**(6): p. 449-55.
98. Schepman, K.P., et al., *Malignant transformation of oral leukoplakia: a follow-up study of a hospital-based population of 166 patients with oral leukoplakia from The Netherlands*. Oral Oncol, 1998. **34**(4): p. 270-5.
99. Shiu, M.N., et al., *Risk factors for leukoplakia and malignant transformation to oral carcinoma: a leukoplakia cohort in Taiwan*. Br J Cancer, 2000. **82**(11): p. 1871-4.
100. Cowan, C.G., et al., *Potentially malignant oral lesions in northern Ireland: a 20-year population-based perspective of malignant transformation*. Oral Dis, 2001. **7**(1): p. 18-24.
101. Scheifele, C. and P.A. Reichart, *Is there a natural limit of the transformation rate of oral leukoplakia?* Oral Oncol, 2003. **39**(5): p. 470-5.
102. Schepman, K., et al., *Concomitant leukoplakia in patients with oral squamous cell carcinoma*. Oral Dis, 1999. **5**(3): p. 206-9.
103. Reichart, P.A. and H.P. Philipsen, *Oral erythroplakia--a review*. Oral Oncol, 2005. **41**(6): p. 551-61.
104. Shafer, W.G. and C.A. Waldron, *Erythroplakia of the oral cavity*. Cancer, 1975. **36**(3): p. 1021-8.
105. Shear, M., *Erythroplakia of the mouth*. Int Dent J, 1972. **22**(4): p. 460-73.
106. Sciubba, J.J., *Oral leukoplakia*. Crit Rev Oral Biol Med, 1995. **6**(2): p. 147-60.
107. Zakrzewska, J.M., et al., *Proliferative verrucous leukoplakia: a report of ten cases*. Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 1996. **82**(4): p. 396-401.
108. Reibel, J., *Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics*. Crit Rev Oral Biol Med, 2003. **14**(1): p. 47-62.
109. van der Waal, I., et al., *Oral leukoplakia: a clinicopathological review*. Oral Oncol, 1997. **33**(5): p. 291-301.
110. Mithani, S.K., et al., *Molecular genetics of premalignant oral lesions*. Oral Dis, 2007. **13**(2): p. 126-33.
111. Zhang, L., et al., *Increased genetic damage in oral leukoplakia from high risk sites: potential impact on staging and clinical management*. Cancer, 2001. **91**(11): p. 2148-55.
112. EW, B.J.a.S., *Oral Cavity Cancer*, in *Cancer of Head and Neck*, J.P. Sah, Editor. 2001, American Cancer Society: Hamilton-London.
113. Gupta, P.C., et al., *Incidence rates of oral cancer and natural history of oral precancerous lesions in a 10-year follow-up study of Indian villagers*. Community Dent Oral Epidemiol, 1980. **8**(6): p. 283-333.
114. Waldron, C.A. and W.G. Shafer, *Leukoplakia revisited. A clinicopathologic study 3256 oral leukoplakias*. Cancer, 1975. **36**(4): p. 1386-92.
115. Schepman, K.P., et al., *Tobacco usage in relation to the anatomical site of oral leukoplakia*. Oral Dis, 2001. **7**(1): p. 25-7.
116. Banoczy, J. and L. Sugar, *Longitudinal studies in oral leukoplakias*. J Oral Pathol, 1972. **1**(6): p. 265-72.
117. O'Grady, J.F. and P.C. Reade, *Candida albicans as a promoter of oral mucosal neoplasia*. Carcinogenesis, 1992. **13**(5): p. 783-6.

118. Barrett, A.W., V.J. Kingsmill, and P.M. Speight, *The frequency of fungal infection in biopsies of oral mucosal lesions*. Oral Dis, 1998. **4**(1): p. 26-31.
119. Renstrup, G., *Occurrence of candida in oral leukoplakias*. Acta Pathol Microbiol Scand B Microbiol Immunol, 1970. **78**(4): p. 421-4.
120. Allen, C.M., *Toluidine blue: proceed with caution?* Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 1998. **86**(3): p. 255.
121. Schepman, K.P. and I. van der Waal, *A proposal for a classification and staging system for oral leukoplakia: a preliminary study*. Eur J Cancer B Oral Oncol, 1995. **31B**(6): p. 396-8.
122. van der Waal, I. and T. Axell, *Oral leukoplakia: a proposal for uniform reporting*. Oral Oncol, 2002. **38**(6): p. 521-6.
123. Greenberg, M.S., *The "brush" controversy*. Oral Surg Oral Med Oral Pathol Oral Radiol Endod, 2002. **93**(3): p. 217-8.
124. Kujan, O., et al., *Potential applications of oral brush cytology with liquid-based technology: results from a cohort of normal oral mucosa*. Oral Oncol, 2006. **42**(8): p. 810-8.
125. Gandolfo, S., et al., *Toluidine blue uptake in potentially malignant oral lesions in vivo: clinical and histological assessment*. Oral Oncol, 2006. **42**(1): p. 89-95.
126. Zhang, L., et al., *Toluidine blue staining identifies high-risk primary oral premalignant lesions with poor outcome*. Cancer Res, 2005. **65**(17): p. 8017-21.
127. Epstein, J.B., et al., *Utility of toluidine blue in oral premalignant lesions and squamous cell carcinoma: continuing research and implications for clinical practice*. Head Neck, 2007. **29**(10): p. 948-58.
128. Sharwani, A., et al., *Assessment of oral premalignancy using elastic scattering spectroscopy*. Oral Oncol, 2006. **42**(4): p. 343-9.
129. Poh, C.F., et al., *Fluorescence visualization detection of field alterations in tumor margins of oral cancer patients*. Clin Cancer Res, 2006. **12**(22): p. 6716-22.
130. Westra, W.H. and D. Sidransky, *Fluorescence visualization in oral neoplasia: shedding light on an old problem*. Clin Cancer Res, 2006. **12**(22): p. 6594-7.
131. Epstein, J.B., et al., *The efficacy of oral lumenoscopy (ViziLite) in visualizing oral mucosal lesions*. Spec Care Dentist, 2006. **26**(4): p. 171-4.
132. White, J.M., et al., *Nd:YAG and CO2 laser therapy of oral mucosal lesions*. J Clin Laser Med Surg, 1998. **16**(6): p. 299-304.
133. Johnson, N.W., A.W. Ranasinghe, and K.A. Warnakulasuriya, *Potentially malignant lesions and conditions of the mouth and oropharynx: natural history--cellular and molecular markers of risk*. Eur J Cancer Prev, 1993. **2 Suppl 2**: p. 31-51.
134. Jones, A.S., *Prognosis in mouth cancer: tumour factors*. Eur J Cancer B Oral Oncol, 1994. **30B**(1): p. 8-15.
135. Schantz, S.P. and J.S. Ostroff, *Novel approaches to the prevention of head and neck cancer*. Proc Soc Exp Biol Med, 1997. **216**(2): p. 275-82.
136. Vogelstein, B. and K.W. Kinzler, *Cancer genes and the pathways they control*. Nat Med, 2004. **10**(8): p. 789-99.
137. Kim, M.M. and J.A. Califano, *Molecular pathology of head-and-neck cancer*. Int J Cancer, 2004. **112**(4): p. 545-53.

138. Choi, H.R., et al., *Differential expressions of cyclin-dependent kinase inhibitors (p27 and p21) and their relation to p53 and Ki-67 in oral squamous tumorigenesis*. *Int J Oncol*, 2003. **22**(2): p. 409-14.
139. Diwakar, N., et al., *Heterogeneity, histological features and DNA ploidy in oral carcinoma by image-based analysis*. *Oral Oncol*, 2005. **41**(4): p. 416-22.
140. Janot, F., et al., *Prognostic value of clinicopathological parameters in head and neck squamous cell carcinoma: a prospective analysis*. *Br J Cancer*, 1996. **73**(4): p. 531-8.
141. Wang, X., et al., *Intratumor genomic heterogeneity correlates with histological grade of advanced oral squamous cell carcinoma*. *Oral Oncol*, 2006. **42**(7): p. 740-4.
142. Califano, J., et al., *Genetic progression and clonal relationship of recurrent premalignant head and neck lesions*. *Clin Cancer Res*, 2000. **6**(2): p. 347-52.
143. Copper, M.P., et al., *Role of genetic factors in the etiology of squamous cell carcinoma of the head and neck*. *Arch Otolaryngol Head Neck Surg*, 1995. **121**(2): p. 157-60.
144. Jin, Y., et al., *Clonal chromosome abnormalities in premalignant lesions of the skin*. *Cancer Genet Cytogenet*, 2002. **136**(1): p. 48-52.
145. Califano, J., et al., *Genetic progression model for head and neck cancer: implications for field cancerization*. *Cancer Res*, 1996. **56**(11): p. 2488-92.
146. Ha, P.K., et al., *A transcriptional progression model for head and neck cancer*. *Clin Cancer Res*, 2003. **9**(8): p. 3058-64.
147. Karabulut, A., et al., *Observer variability in the histologic assessment of oral premalignant lesions*. *J Oral Pathol Med*, 1995. **24**(5): p. 198-200.
148. Warnakulasuriya, S., *Lack of molecular markers to predict malignant potential of oral precancer*. *J Pathol*, 2000. **190**(4): p. 407-9.
149. Spafford, M.F., et al., *Detection of head and neck squamous cell carcinoma among exfoliated oral mucosal cells by microsatellite analysis*. *Clin Cancer Res*, 2001. **7**(3): p. 607-12.
150. Partridge, M., et al., *Allelic imbalance at chromosomal loci implicated in the pathogenesis of oral precancer, cumulative loss and its relationship with progression to cancer*. *Oral Oncol*, 1998. **34**(2): p. 77-83.
151. Hogmo, A., et al., *Preneoplastic oral lesions: the clinical value of image cytometry DNA analysis, p53 and p21/WAF1 expression*. *Anticancer Res*, 1998. **18**(5B): p. 3645-50.
152. Zhou, X., et al., *Frequent allelic imbalances at 8p and 11q22 in oral and oropharyngeal epithelial dysplastic lesions*. *Cancer Genet Cytogenet*, 2005. **161**(1): p. 86-9.
153. Zhang, L. and M.P. Rosin, *Loss of heterozygosity: a potential tool in management of oral premalignant lesions?* *J Oral Pathol Med*, 2001. **30**(9): p. 513-20.
154. Smith, J., et al., *Biomarkers in dysplasia of the oral cavity: a systematic review*. *Oral Oncol*, 2009. **45**(8): p. 647-53.
155. Chen, Y.K., S.S. Hsue, and L.M. Lin, *p73 expression for human buccal epithelial dysplasia and squamous cell carcinoma: does it correlate with nodal status of carcinoma and is there a relationship with malignant change of epithelial dysplasia?* *Head Neck*, 2004. **26**(11): p. 945-52.
156. Jordan, R.C., et al., *Overexpression of matrix metalloproteinase-1 and -9 mRNA is associated with progression of oral dysplasia to cancer*. *Clin Cancer Res*, 2004. **10**(19): p. 6460-5.
157. Lo Muzio, L., et al., *Survivin, a potential early predictor of tumor progression in the oral mucosa*. *J Dent Res*, 2003. **82**(11): p. 923-8.

158. Rich, A.M., D. Kerdpon, and P.C. Reade, *p53 expression in oral precancer and cancer*. Aust Dent J, 1999. **44**(2): p. 103-5.
159. Cruz, I.B., et al., *p53 expression above the basal cell layer in oral mucosa is an early event of malignant transformation and has predictive value for developing oral squamous cell carcinoma*. J Pathol, 1998. **184**(4): p. 360-8.
160. Shahnavaz, S.A., et al., *p53 gene mutations in sequential oral epithelial dysplasias and squamous cell carcinomas*. J Pathol, 2000. **190**(4): p. 417-22.
161. Regezi, J.A., et al., *p53 protein expression in sequential biopsies of oral dysplasias and in situ carcinomas*. J Oral Pathol Med, 1995. **24**(1): p. 18-22.
162. Hogmo, A., et al., *Nuclear DNA content and p53 overexpression in stage I squamous cell carcinoma of the tongue compared with advanced tongue carcinomas*. Mol Pathol, 1998. **51**(5): p. 268-72.
163. Zhou, X., et al., *Allelic imbalance analysis of oral tongue squamous cell carcinoma by high-density single nucleotide polymorphism arrays using whole-genome amplified DNA*. Hum Genet, 2005. **118**(3-4): p. 504-7.
164. el-Naggar, A.K., et al., *Genotypic alterations in benign and malignant salivary gland tumors: histogenetic and clinical implications*. Am J Surg Pathol, 1997. **21**(6): p. 691-7.
165. Li, G., et al., *Association between the V109G polymorphism of the p27 gene and the risk and progression of oral squamous cell carcinoma*. Clin Cancer Res, 2004. **10**(12 Pt 1): p. 3996-4002.
166. Callender, T., et al., *PRAD-1 (CCND1)/cyclin D1 oncogene amplification in primary head and neck squamous cell carcinoma*. Cancer, 1994. **74**(1): p. 152-8.
167. Weber, A., C. Wittekind, and A. Tannapfel, *Genetic and epigenetic alterations of 9p21 gene products in benign and malignant tumors of the head and neck*. Pathol Res Pract, 2003. **199**(6): p. 391-7.
168. Gasco, M. and T. Crook, *The p53 network in head and neck cancer*. Oral Oncol, 2003. **39**(3): p. 222-31.
169. Papadimitrakopoulou, V.A., et al., *Cyclin D1 and p16 alterations in advanced premalignant lesions of the upper aerodigestive tract: role in response to chemoprevention and cancer development*. Clin Cancer Res, 2001. **7**(10): p. 3127-34.
170. Aravin, A. and T. Tuschl, *Identification and characterization of small RNAs involved in RNA silencing*. FEBS Lett, 2005. **579**(26): p. 5830-40.
171. Lee, R.C. and V. Ambros, *An extensive class of small RNAs in Caenorhabditis elegans*. Science, 2001. **294**(5543): p. 862-4.
172. Griffiths-Jones, S., et al., *miRBase: tools for microRNA genomics*. Nucleic Acids Res, 2008. **36**(Database issue): p. D154-8.
173. Grosshans, H. and F.J. Slack, *Micro-RNAs: small is plentiful*. J Cell Biol, 2002. **156**(1): p. 17-21.
174. Valencia-Sanchez, M.A., et al., *Control of translation and mRNA degradation by miRNAs and siRNAs*. Genes Dev, 2006. **20**(5): p. 515-24.
175. Schramke, V. and R. Allshire, *Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing*. Science, 2003. **301**(5636): p. 1069-74.
176. Carninci, P., et al., *The transcriptional landscape of the mammalian genome*. Science, 2005. **309**(5740): p. 1559-63.
177. Sen, G.L. and H.M. Blau, *Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies*. Nat Cell Biol, 2005. **7**(6): p. 633-6.

178. Rand, T.A., et al., *Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation*. Cell, 2005. **123**(4): p. 621-9.
179. Engels, B.M. and G. Hutvagner, *Principles and effects of microRNA-mediated post-transcriptional gene regulation*. Oncogene, 2006. **25**(46): p. 6163-9.
180. Doench, J.G. and P.A. Sharp, *Specificity of microRNA target selection in translational repression*. Genes Dev, 2004. **18**(5): p. 504-11.
181. Kertesz, M., et al., *The role of site accessibility in microRNA target recognition*. Nat Genet, 2007. **39**(10): p. 1278-84.
182. John, B., et al., *Human MicroRNA targets*. PLoS Biol, 2004. **2**(11): p. e363.
183. Krek, A., et al., *Combinatorial microRNA target predictions*. Nat Genet, 2005. **37**(5): p. 495-500.
184. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets*. Cell, 2005. **120**(1): p. 15-20.
185. Lewis, B.P., et al., *Prediction of mammalian microRNA targets*. Cell, 2003. **115**(7): p. 787-98.
186. Scholzova, E., et al., *RNA regulation and cancer development*. Cancer Lett, 2007. **246**(1-2): p. 12-23.
187. Chen, C.Z., *MicroRNAs as oncogenes and tumor suppressors*. N Engl J Med, 2005. **353**(17): p. 1768-71.
188. Esquela-Kerscher, A. and F.J. Slack, *Oncomirs - microRNAs with a role in cancer*. Nat Rev Cancer, 2006. **6**(4): p. 259-69.
189. Kent, O.A. and J.T. Mendell, *A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes*. Oncogene, 2006. **25**(46): p. 6188-96.
190. Wiemer, E.A., *The role of microRNAs in cancer: no small matter*. Eur J Cancer, 2007. **43**(10): p. 1529-44.
191. Calin, G.A., et al., *Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers*. Proc Natl Acad Sci U S A, 2004. **101**(9): p. 2999-3004.
192. Lim, L.P., et al., *Vertebrate microRNA genes*. Science, 2003. **299**(5612): p. 1540.
193. Sarnow, P., et al., *MicroRNAs: expression, avoidance and subversion by vertebrate viruses*. Nat Rev Microbiol, 2006. **4**(9): p. 651-9.
194. Calin, G.A., et al., *Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia*. Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15524-9.
195. Calin, G.A., et al., *A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia*. N Engl J Med, 2005. **353**(17): p. 1793-801.
196. Metzler, M., et al., *High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma*. Genes Chromosomes Cancer, 2004. **39**(2): p. 167-9.
197. Michael, M.Z., et al., *Reduced accumulation of specific microRNAs in colorectal neoplasia*. Mol Cancer Res, 2003. **1**(12): p. 882-91.
198. Takamizawa, J., et al., *Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival*. Cancer Res, 2004. **64**(11): p. 3753-6.
199. Iorio, M.V., et al., *MicroRNA gene expression deregulation in human breast cancer*. Cancer Res, 2005. **65**(16): p. 7065-70.

200. Ciafre, S.A., et al., *Extensive modulation of a set of microRNAs in primary glioblastoma*. *Biochem Biophys Res Commun*, 2005. **334**(4): p. 1351-8.
201. Murakami, Y., et al., *Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues*. *Oncogene*, 2006. **25**(17): p. 2537-45.
202. He, H., et al., *The role of microRNA genes in papillary thyroid carcinoma*. *Proc Natl Acad Sci U S A*, 2005. **102**(52): p. 19075-80.
203. Volinia, S., et al., *A microRNA expression signature of human solid tumors defines cancer gene targets*. *Proc Natl Acad Sci U S A*, 2006. **103**(7): p. 2257-61.
204. Yanaihara, N., et al., *Unique microRNA molecular profiles in lung cancer diagnosis and prognosis*. *Cancer Cell*, 2006. **9**(3): p. 189-98.
205. Cimmino, A., et al., *miR-15 and miR-16 induce apoptosis by targeting BCL2*. *Proc Natl Acad Sci U S A*, 2005. **102**(39): p. 13944-9.
206. Johnson, S.M., et al., *RAS is regulated by the let-7 microRNA family*. *Cell*, 2005. **120**(5): p. 635-47.
207. Hayashita, Y., et al., *A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation*. *Cancer Res*, 2005. **65**(21): p. 9628-32.
208. O'Donnell, E.P., et al., *Quantitative analysis of early chemically-induced pulmonary lesions in mice of varying susceptibilities to lung tumorigenesis*. *Cancer Lett*, 2006. **241**(2): p. 197-202.
209. Rehmsmeier, M., et al., *Fast and effective prediction of microRNA/target duplexes*. *RNA*, 2004. **10**(10): p. 1507-17.
210. Hammond, E.M. and A.J. Giaccia, *Hypoxia-inducible factor-1 and p53: friends, acquaintances, or strangers?* *Clin Cancer Res*, 2006. **12**(17): p. 5007-9.
211. O'Donnell, K.A., et al., *c-Myc-regulated microRNAs modulate E2F1 expression*. *Nature*, 2005. **435**(7043): p. 839-43.
212. Cole, M.D. and S.B. McMahon, *The Myc oncoprotein: a critical evaluation of transactivation and target gene regulation*. *Oncogene*, 1999. **18**(19): p. 2916-24.
213. Si, M.L., et al., *miR-21-mediated tumor growth*. *Oncogene*, 2007. **26**(19): p. 2799-803.
214. Roldo, C., et al., *MicroRNA expression abnormalities in pancreatic endocrine and acinar tumors are associated with distinctive pathologic features and clinical behavior*. *J Clin Oncol*, 2006. **24**(29): p. 4677-84.
215. Chan, J.A., A.M. Krichevsky, and K.S. Kosik, *MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells*. *Cancer Res*, 2005. **65**(14): p. 6029-33.
216. Meng, F., et al., *Involvement of human micro-RNA in growth and response to chemotherapy in human cholangiocarcinoma cell lines*. *Gastroenterology*, 2006. **130**(7): p. 2113-29.
217. Lu, Z., et al., *MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene*. *Oncogene*, 2008. **27**(31): p. 4373-9.
218. Zhang, H., et al., *Involvement of programmed cell death 4 in transforming growth factor-beta1-induced apoptosis in human hepatocellular carcinoma*. *Oncogene*, 2006. **25**(45): p. 6101-12.
219. Asangani, I.A., et al., *MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer*. *Oncogene*, 2008. **27**(15): p. 2128-36.

220. Zhu, S., et al., *MicroRNA-21 targets tumor suppressor genes in invasion and metastasis*. Cell Res, 2008. **18**(3): p. 350-9.
221. Long, X.B., et al., *Let-7a microRNA functions as a potential tumor suppressor in human laryngeal cancer*. Oncol Rep, 2009. **22**(5): p. 1189-95.
222. Shell, S., et al., *Let-7 expression defines two differentiation stages of cancer*. Proc Natl Acad Sci U S A, 2007. **104**(27): p. 11400-5.
223. Hebert, C., et al., *High mobility group A2 is a target for miRNA-98 in head and neck squamous cell carcinoma*. Mol Cancer, 2007. **6**: p. 5.
224. Tran, N., et al., *MicroRNA expression profiles in head and neck cancer cell lines*. Biochem Biophys Res Commun, 2007. **358**(1): p. 12-7.
225. Jiang, J., et al., *Real-time expression profiling of microRNA precursors in human cancer cell lines*. Nucleic Acids Res, 2005. **33**(17): p. 5394-403.
226. Gaur, A., et al., *Characterization of microRNA expression levels and their biological correlates in human cancer cell lines*. Cancer Res, 2007. **67**(6): p. 2456-68.
227. Liu, C.G., et al., *Expression profiling of microRNA using oligo DNA arrays*. Methods, 2008. **44**(1): p. 22-30.
228. Schmittgen, T.D., et al., *A high-throughput method to monitor the expression of microRNA precursors*. Nucleic Acids Res, 2004. **32**(4): p. e43.
229. Shi, R. and V.L. Chiang, *Facile means for quantifying microRNA expression by real-time PCR*. Biotechniques, 2005. **39**(4): p. 519-25.
230. Bushati, N. and S.M. Cohen, *microRNA functions*. Annu Rev Cell Dev Biol, 2007. **23**: p. 175-205.
231. Calin, G.A. and C.M. Croce, *MicroRNA signatures in human cancers*. Nat Rev Cancer, 2006. **6**(11): p. 857-66.
232. Yin, J.Q. and R.C. Zhao, *Identifying expression of new small RNAs by microarrays*. Methods, 2007. **43**(2): p. 123-30.
233. Shingara, J., et al., *An optimized isolation and labeling platform for accurate microRNA expression profiling*. RNA, 2005. **11**(9): p. 1461-70.
234. Liu, C.G., et al., *An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues*. Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9740-4.
235. Zhao, J.J., et al., *Genome-wide microRNA profiling in human fetal nervous tissues by oligonucleotide microarray*. Childs Nerv Syst, 2006. **22**(11): p. 1419-25.
236. Centola, M., et al., *Genome-scale assessment of molecular pathology in systemic autoimmune diseases using microarray technology: a potential breakthrough diagnostic and individualized therapy-design tool*. Scand J Immunol, 2006. **64**(3): p. 236-42.
237. Southern, E.M., *DNA microarrays. History and overview*. Methods Mol Biol, 2001. **170**: p. 1-15.
238. Chen, C., et al., *Real-time quantification of microRNAs by stem-loop RT-PCR*. Nucleic Acids Res, 2005. **33**(20): p. e179.
239. Liang, Y., et al., *Characterization of microRNA expression profiles in normal human tissues*. BMC Genomics, 2007. **8**: p. 166.
240. Mattie, M.D., et al., *Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies*. Mol Cancer, 2006. **5**: p. 24.

241. Doleshal, M., et al., *Evaluation and validation of total RNA extraction methods for microRNA expression analyses in formalin-fixed, paraffin-embedded tissues*. J Mol Diagn, 2008. **10**(3): p. 203-11.
242. Hui, A.B., et al., *Robust global micro-RNA profiling with formalin-fixed paraffin-embedded breast cancer tissues*. Lab Invest, 2009. **89**(5): p. 597-606.
243. Krichevsky, A.M., et al., *A microRNA array reveals extensive regulation of microRNAs during brain development*. RNA, 2003. **9**(10): p. 1274-81.
244. Valoczi, A., et al., *Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes*. Nucleic Acids Res, 2004. **32**(22): p. e175.
245. Nelson, P.T., et al., *Microarray-based, high-throughput gene expression profiling of microRNAs*. Nat Methods, 2004. **1**(2): p. 155-61.
246. Lu, J., et al., *MicroRNA expression profiles classify human cancers*. Nature, 2005. **435**(7043): p. 834-8.
247. Lau, N.C., et al., *An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans*. Science, 2001. **294**(5543): p. 858-62.
248. Neely, L.A., et al., *A single-molecule method for the quantitation of microRNA gene expression*. Nat Methods, 2006. **3**(1): p. 41-6.
249. van Beers, E.H., et al., *Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations*. Cancer Res, 2005. **65**(3): p. 822-7.
250. Wessels, L.F., et al., *Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors*. Cancer Res, 2002. **62**(23): p. 7110-7.
251. Cronin, M., et al., *Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay*. Am J Pathol, 2004. **164**(1): p. 35-42.
252. Dapson, R.W., *Macromolecular changes caused by formalin fixation and antigen retrieval*. Biotech Histochem, 2007. **82**(3): p. 133-40.
253. Korga, A., K. Wilkolaska, and E. Korobowicz, *[Difficulties in using archival paraffin-embedded tissues for RNA expression analysis]*. Postepy Hig Med Dosw (Online), 2007. **61**: p. 151-5.
254. Masuda, N., et al., *Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples*. Nucleic Acids Res, 1999. **27**(22): p. 4436-43.
255. Xi, Y., et al., *Systematic analysis of microRNA expression of RNA extracted from fresh frozen and formalin-fixed paraffin-embedded samples*. RNA, 2007. **13**(10): p. 1668-74.
256. Li, J., et al., *Comparison of miRNA expression patterns using total RNA extracted from matched samples of formalin-fixed paraffin-embedded (FFPE) cells and snap frozen cells*. BMC Biotechnol, 2007. **7**: p. 36.
257. Iafrate, A.J., et al., *Detection of large-scale variation in the human genome*. Nat Genet, 2004. **36**(9): p. 949-51.
258. Sebat, J., et al., *Large-scale copy number polymorphism in the human genome*. Science, 2004. **305**(5683): p. 525-8.
259. Redon, R., et al., *Global variation in copy number in the human genome*. Nature, 2006. **444**(7118): p. 444-54.



260. LaFramboise, T., et al., *Allele-specific amplification in cancer revealed by SNP array analysis*. PLoS Comput Biol, 2005. **1**(6): p. e65.
261. Tsafirir, D., et al., *Relationship of gene expression and chromosomal abnormalities in colorectal cancer*. Cancer Res, 2006. **66**(4): p. 2129-37.
262. Pollack, J.R., et al., *Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors*. Proc Natl Acad Sci U S A, 2002. **99**(20): p. 12963-8.
263. Lasko, D., W. Cavenee, and M. Nordenskjold, *Loss of constitutional heterozygosity in human cancer*. Annu Rev Genet, 1991. **25**: p. 281-314.
264. Albertson, D.G. and D. Pinkel, *Genomic microarrays in human genetic disease and cancer*. Hum Mol Genet, 2003. **12 Spec No 2**: p. R145-52.
265. Forozan, F., et al., *Comparative genomic hybridization analysis of 38 breast cancer cell lines: a basis for interpreting complementary DNA microarray data*. Cancer Res, 2000. **60**(16): p. 4519-25.
266. Lucito, R., et al., *Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation*. Genome Res, 2003. **13**(10): p. 2291-305.
267. Schwab, M., *Oncogene amplification in neoplastic development and progression of human cancers*. Crit Rev Oncog, 1990. **2**(1): p. 35-51.
268. Gray, J.W., et al., *Molecular cytogenetics of human breast cancer*. Cold Spring Harb Symp Quant Biol, 1994. **59**: p. 645-52.
269. Kallioniemi, O.P., et al., *Comparative genomic hybridization: a rapid new method for detecting and mapping DNA amplification in tumors*. Semin Cancer Biol, 1993. **4**(1): p. 41-6.
270. Veltman, I., et al., *A novel case of infantile sacral teratoma and a constitutional t(12;15)(q13;q25) pat*. Cancer Genet Cytogenet, 2002. **136**(1): p. 17-22.
271. Nessling, M., et al., *Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue*. Cancer Res, 2005. **65**(2): p. 439-47.
272. Albertson, D.G., et al., *Quantitative mapping of amplicon structure by array CGH identifies CYP24 as a candidate oncogene*. Nat Genet, 2000. **25**(2): p. 144-6.
273. Weiss, M.M., et al., *Genomic profiling of gastric cancer predicts lymph node status and survival*. Oncogene, 2003. **22**(12): p. 1872-9.
274. Buffart, T.E., et al., *Gastric cancers in young and elderly patients show different genomic profiles*. J Pathol, 2007. **211**(1): p. 45-51.
275. Jong, K., et al., *Cross-platform array comparative genomic hybridization meta-analysis separates hematopoietic and mesenchymal from epithelial tumors*. Oncogene, 2007. **26**(10): p. 1499-506.
276. Fridlyand, J., et al., *Breast tumor copy number aberration phenotypes and genomic instability*. BMC Cancer, 2006. **6**: p. 96.
277. van de Wiel, M.A., et al., *Expression microarray analysis and oligo array comparative genomic hybridization of acquired gemcitabine resistance in mouse colon reveals selection for chromosomal aberrations*. Cancer Res, 2005. **65**(22): p. 10208-13.
278. Pollack, J.R., et al., *Genome-wide analysis of DNA copy-number changes using cDNA microarrays*. Nat Genet, 1999. **23**(1): p. 41-6.

279. Johnson, N.A., et al., *Application of array CGH on archival formalin-fixed paraffin-embedded tissues including small numbers of microdissected cells*. Lab Invest, 2006. **86**(9): p. 968-78.
280. Lambros, M.B., et al., *Unlocking pathology archives for molecular genetic studies: a reliable method to generate probes for chromogenic and fluorescent in situ hybridization*. Lab Invest, 2006. **86**(4): p. 398-408.
281. Little, S.E., et al., *Array CGH using whole genome amplification of fresh-frozen and formalin-fixed, paraffin-embedded tumor DNA*. Genomics, 2006. **87**(2): p. 298-306.
282. Reis-Filho, J.S., et al., *The molecular genetics of breast cancer: the contribution of comparative genomic hybridization*. Pathol Res Pract, 2005. **201**(11): p. 713-25.
283. van Beers, E.H., et al., *A multiplex PCR predictor for aCGH success of FFPE samples*. Br J Cancer, 2006. **94**(2): p. 333-7.
284. Legrand, B., et al., *DNA genotyping of unbuffered formalin fixed paraffin embedded tissues*. Forensic Sci Int, 2002. **125**(2-3): p. 205-11.
285. Alers, J.C., et al., *Universal linkage system: an improved method for labeling archival DNA for comparative genomic hybridization*. Genes Chromosomes Cancer, 1999. **25**(3): p. 301-5.
286. Raap, A.K., et al., *Array comparative genomic hybridization with cyanin cis-platinum-labeled DNAs*. Biotechniques, 2004. **37**(1): p. 130-4.
287. Bredel, M., et al., *Amplification of whole tumor genomes and gene-by-gene mapping of genomic aberrations from limited sources of fresh-frozen and paraffin-embedded DNA*. J Mol Diagn, 2005. **7**(2): p. 171-82.
288. Telenius, H., et al., *Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer*. Genomics, 1992. **13**(3): p. 718-25.
289. Isola, J.J., et al., *Genetic aberrations detected by comparative genomic hybridization predict outcome in node-negative breast cancer*. Am J Pathol, 1995. **147**(4): p. 905-11.
290. Kirchhoff, M., et al., *Detection of chromosomal gains and losses in comparative genomic hybridization analysis based on standard reference intervals*. Cytometry, 1998. **31**(3): p. 163-73.
291. Kallioniemi, O.P., et al., *Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors*. Genes Chromosomes Cancer, 1994. **10**(4): p. 231-43.
292. Dean, F.B., et al., *Comprehensive human genome amplification using multiple displacement amplification*. Proc Natl Acad Sci U S A, 2002. **99**(8): p. 5261-6.
293. Lizardi, P.M., et al., *Mutation detection and single-molecule counting using isothermal rolling-circle amplification*. Nat Genet, 1998. **19**(3): p. 225-32.
294. Stoecklein, N.H., et al., *SCOMP is superior to degenerated oligonucleotide primed-polymerase chain reaction for global amplification of minute amounts of DNA from microdissected archival tissue samples*. Am J Pathol, 2002. **161**(1): p. 43-51.
295. Klein, C.A., et al., *Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells*. Proc Natl Acad Sci U S A, 1999. **96**(8): p. 4494-9.
296. Goelz, S.E., S.R. Hamilton, and B. Vogelstein, *Purification of DNA from formaldehyde fixed and paraffin embedded human tissue*. Biochem Biophys Res Commun, 1985. **130**(1): p. 118-26.
297. Yagi, N., et al., *The role of DNase and EDTA on DNA degradation in formaldehyde fixed tissues*. Biotech Histochem, 1996. **71**(3): p. 123-9.

298. Sadikovic, B., et al., *Genome-wide H3K9 histone acetylation profiles are altered in benzopyrene-treated MCF7 breast cancer cells*. J Biol Chem, 2008. **283**(7): p. 4051-60.
299. Bohlander, S.K., et al., *A method for the rapid sequence-independent amplification of microdissected chromosomal material*. Genomics, 1992. **13**(4): p. 1322-4.
300. Lockwood, W.W., et al., *Recent advances in array comparative genomic hybridization technologies and their applications in human genetics*. Eur J Hum Genet, 2006. **14**(2): p. 139-48.
301. Stranger, B.E., et al., *Relative impact of nucleotide and copy number variation on gene expression phenotypes*. Science, 2007. **315**(5813): p. 848-53.
302. Fiegler, H., et al., *DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones*. Genes Chromosomes Cancer, 2003. **36**(4): p. 361-74.
303. Ishkanian, A.S., et al., *A tiling resolution DNA microarray with complete coverage of the human genome*. Nat Genet, 2004. **36**(3): p. 299-303.
304. Hester, S.D., et al., *Comparison of comparative genomic hybridization technologies across microarray platforms*. J Biomol Tech, 2009. **20**(2): p. 135-51.
305. Peiffer, D.A., et al., *High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping*. Genome Res, 2006. **16**(9): p. 1136-48.
306. Auer, H., et al., *Gene-resolution analysis of DNA copy number variation using oligonucleotide expression microarrays*. BMC Genomics, 2007. **8**: p. 111.
307. Bergamo, N.A., et al., *Classic and molecular cytogenetic analyses reveal chromosomal gains and losses correlated with survival in head and neck cancer patients*. Clin Cancer Res, 2005. **11**(2 Pt 1): p. 621-31.
308. Bockmuhl, U., et al., *[Genetic screening of head-neck carcinomas using comparative genomic hybridization (CGH)]*. Laryngorhinotologie, 1996. **75**(7): p. 408-14.
309. Singh, B., et al., *Molecular cytogenetic characterization of head and neck squamous cell carcinoma and refinement of 3q amplification*. Cancer Res, 2001. **61**(11): p. 4506-13.
310. Squire, J.A., et al., *Molecular cytogenetic analysis of head and neck squamous cell carcinoma: By comparative genomic hybridization, spectral karyotyping, and expression array analysis*. Head Neck, 2002. **24**(9): p. 874-87.
311. Lin, M., et al., *DNA copy number gains in head and neck squamous cell carcinoma*. Oncogene, 2006. **25**(9): p. 1424-33.
312. Chen, Y. and C. Chen, *DNA copy number variation and loss of heterozygosity in relation to recurrence of and survival from head and neck squamous cell carcinoma: a review*. Head Neck, 2008. **30**(10): p. 1361-83.
313. Temam, S., et al., *Epidermal growth factor receptor copy number alterations correlate with poor clinical outcome in patients with head and neck squamous cancer*. J Clin Oncol, 2007. **25**(16): p. 2164-70.
314. Bockmuhl, U., et al., *Association of 8p23 deletions with poor survival in head and neck cancer*. Otolaryngol Head Neck Surg, 2001. **124**(4): p. 451-5.
315. Pearlstein, R.P., et al., *Loss of 18q predicts poor survival of patients with squamous cell carcinoma of the head and neck*. Genes Chromosomes Cancer, 1998. **21**(4): p. 333-9.
316. Wreesmann, V.B., et al., *Genetic abnormalities associated with nodal metastasis in head and neck cancer*. Head Neck, 2004. **26**(1): p. 10-5.
317. Reis, P.P., et al., *Quantitative real-time PCR identifies a critical region of deletion on 22q13 related to prognosis in oral cancer*. Oncogene, 2002. **21**(42): p. 6480-7.

318. Bauer, V.L., et al., *Chromosomal changes characterize head and neck cancer with poor prognosis*. J Mol Med, 2008. **86**(12): p. 1353-65.
319. Smeets, S.J., et al., *Genome-wide DNA copy number alterations in head and neck squamous cell carcinomas with or without oncogene-expressing human papillomavirus*. Oncogene, 2006. **25**(17): p. 2558-64.
320. Westra, W.H., et al., *Inverse relationship between human papillomavirus-16 infection and disruptive p53 gene mutations in squamous cell carcinoma of the head and neck*. Clin Cancer Res, 2008. **14**(2): p. 366-9.
321. Noutomi, Y., et al., *Comparative genomic hybridization reveals genetic progression of oral squamous cell carcinoma from dysplasia via two different tumorigenic pathways*. J Pathol, 2006. **210**(1): p. 67-74.
322. Garnis, C., et al., *Genomic imbalances in precancerous tissues signal oral cancer risk*. Mol Cancer, 2009. **8**: p. 50.
323. Abou-Elhamd, K.E. and T.N. Habib, *The flow cytometric analysis of premalignant and malignant lesions in head and neck squamous cell carcinoma*. Oral Oncol, 2007. **43**(4): p. 366-72.
324. Cervigne, N.K., et al., *Identification of a microRNA signature associated with progression of leukoplakia to oral carcinoma*. Hum Mol Genet, 2009.

### **3. Hipótese, Objetivos e Material e Métodos**

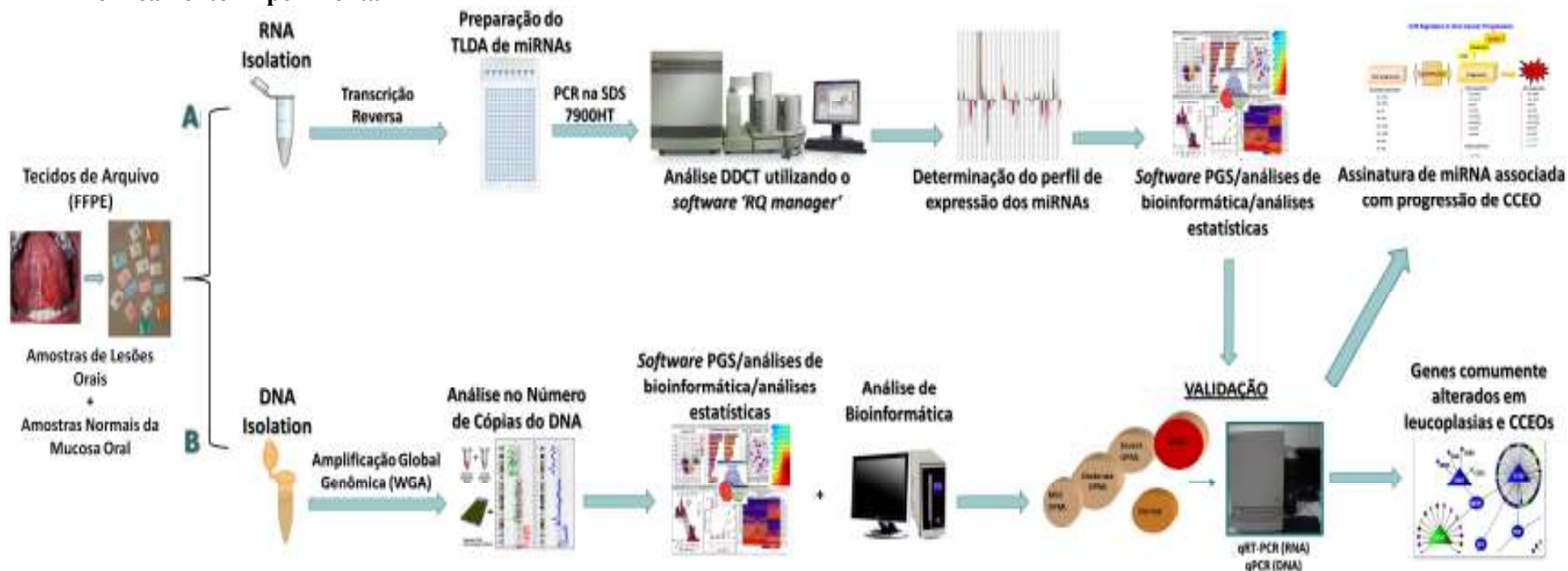
#### **Hipótese**

Alterações genéticas comuns entre lesões orais potencialmente malignas (leucoplasia) e carcinomas de células escamosas orais são biomarcadores de transformação maligna.

#### **Objetivos**

1. Investigar as alterações nos níveis de expressão de microRNAs em leucoplasias progressivas e CCEOs de mesmo sítio anatômico.
2. Identificar ampliações e deleções genômicas globais, comuns entre leucoplasias progressivas e CCEOs correspondentes, e ausentes em leucoplasias que não progrediram, utilizando a análise de CGH em *arrays*.

## Delineamento Experimental



Neste estudo, amostras de arquivo (FFPE) de lesões orais (OPMLs e CCEOs) e amostras normais da mucosa oral foram coletadas de 92 pacientes (24 pacientes do grupo experimental e 68 pacientes do grupo da validação dos experimentos). A) o painel superior mostra a análise de expressão de miRNAs; RNA total foi isolado dos tecidos, convertido em cDNA através de transcrição reversa, e em seguida aplicado ao *TaqMan Low Density Array* (TLDA) para a análise de expressão global de 365 miRNAs. Para comparar os perfis de expressão entre tecidos normal, OPMLs e CCEOs correspondentes, os dados de expressão foram submetidos à análise estatística e de bioinformática utilizando o *software* Partek Genomics Suite (PGS), e outros testes estatísticos apropriados. A validação destes resultados foi efetuada em um grupo independente de amostras utilizando qRT-PCR; tais análises determinaram uma assinatura de miRNAs associada com a progressão de carcinomas orais. B) o painel inferior mostra a análise de aCGH; o DNA foi isolado, amplificado pela técnica *WGA sequenase-based*, e subsequentemente submetida à análise com *385K NimbleGen tiling array* para a identificação de alterações no número de cópias do DNA em OPMLs e CCEOs, em comparação com amostras normais. Os dados resultantes desta análise foram submetidos ao *software* PGS, e as regiões de ganho e perda associadas com progressão oral foram determinadas. Os genes mapeados dentro destas regiões desempenham um papel potencialmente relevante a biologia tumoral, incluindo carcinomas orais. Estes genes foram selecionados para validação por PCR quantitativa (qPCR), utilizando *SyBR Green I*. O programa *I2D* foi utilizado para delinear as vias genéticas envolvendo estes genes.

## Material e Métodos

### *Grupo Amostral*

#### *A) Amostras utilizadas para os experimentos (expressão de miRNAs e CGH array)*

O grupo amostral deste estudo inclui 15 pacientes, com material fixado e embebido em parafina (FFPE) de um total de 54 amostras (37 leucoplasias, 17 carcinomas). No grupo de 37 leucoplasias, 4 são não-displásicas (todas progrediram à carcinoma) e 33 são lesões displásicas (32 progrediram à carcinoma) (Figura 3.1). Todos os carcinomas têm pelo menos uma leucoplasia correspondente precedendo o tumor. Tais amostras seqüenciais são muito difíceis de se obter devido a múltiplos fatores, incluindo a sua natureza dinâmica. Adicionalmente, 9 leucoplasias não-progressivas (5 pacientes da análise aCGH e 4 pacientes da análise de expressão de miRNA), e 7 tecidos FFPE de mucosa oral normal foram incluídos neste estudo. Todas as amostras são espécimes de arquivos e foram coletadas entre 1991 e 2005, do *Toronto General Hospital (TGH)* e da *School of Dentistry*, Universidade de Toronto, Canadá. Os registros médicos foram usados para obter informações clínicas dos pacientes, tais como idade, sexo, sítio e laudo da biópsia, isto é, pacientes com leucoplasia que progrediram ou não à carcinoma; neste último caso, os pacientes selecionados tiveram um seguimento médio de sete anos. Este estudo foi aprovado pelo Conselho de Ética em Pesquisa da *University Health Network (UHN)*.

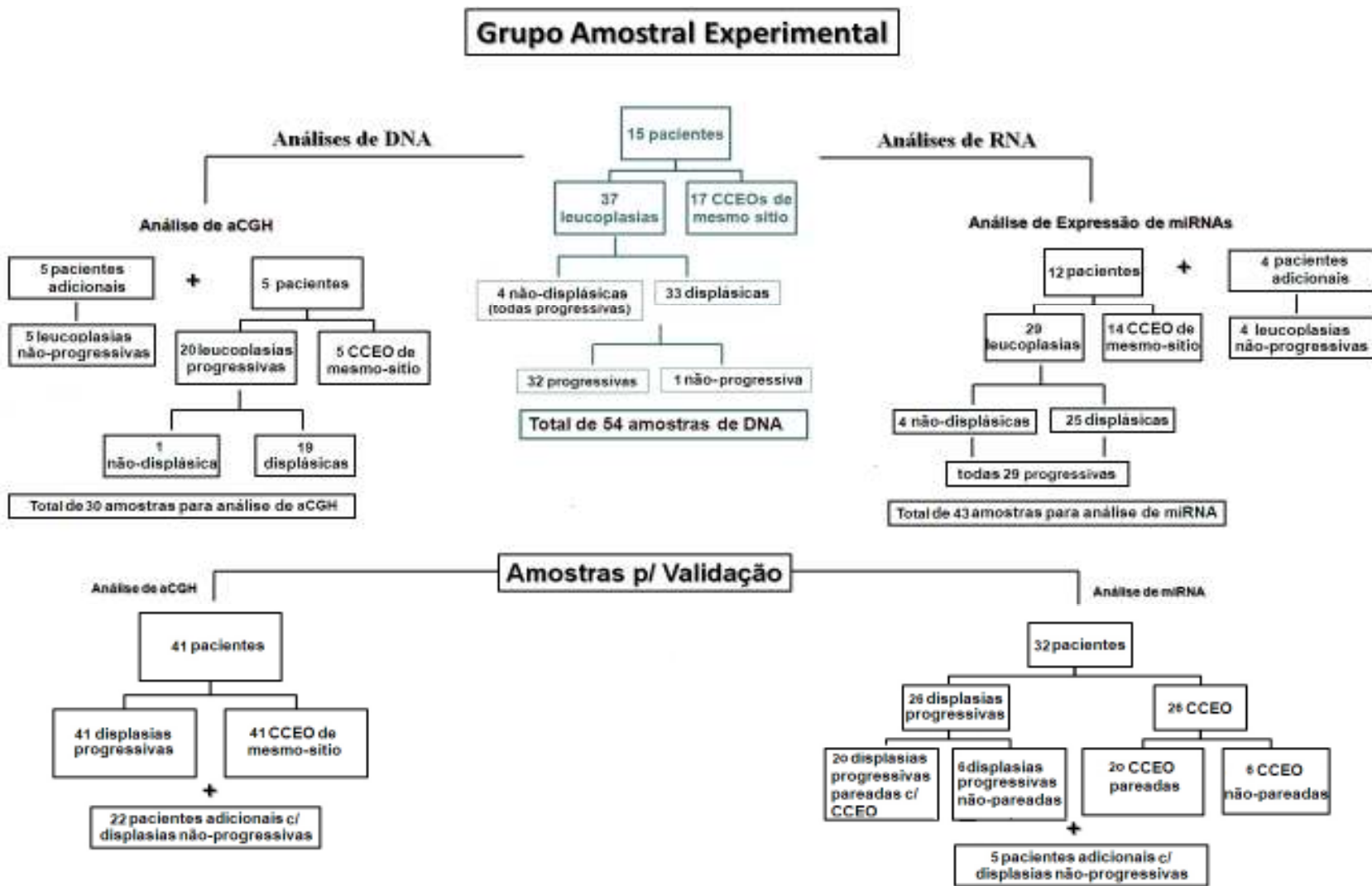
#### *B) Amostras utilizadas para validação dos experimentos*

O grupo amostral utilizado para a validação dos experimentos inclui 41 pacientes, dos quais foram coletadas um total de 82 amostras (41 displasias progressivas e 41 CCEOs correspondentes) (Figura 3.1). Todas as amostras de carcinomas têm pelo menos um displasia progressiva correspondente precedendo o tumor. O DNA foi obtido dos 82 tecidos para

validação dos dados de aCGH por PCR quantitativa. Contudo, apenas 32 dos 41 pacientes possuíam amostra em quantidade suficiente para a validação dos resultados de expressão de miRNA (Figura 3.1). Estas também são espécimes de arquivos (FFPE) e foram coletadas entre 1991 e 2005, do *Toronto General Hospital (TGH)*, Toronto, Canadá. As características detalhadas destas amostras estão descritas nas Tabelas suplementares 3.1 e 3.2.



Figura 3.1. Detalhes do grupo amostral experimental e do grupo amostral de validação, utilizados neste estudo.



#### **4. Manuscrito 1: Análise de expressão de miRNAs**

A desregulação da expressão de miRNAs em diferentes tipos de tumores sólidos indica o seu envolvimento em vias genéticas fundamentais, assim como sua interação com importante genes codificadores de proteínas associadas com a tumorigênese. Apesar dos relatos da literatura que mostram a associação de miRNAs com a biologia e progressão tumoral, até o presente não há outros estudos que tenham identificado uma assinatura de miRNAs, câncer-específica, associada com a progressão de CCEOs. As neoplasias da cavidade oral representam um bom modelo para determinar como os miRNAs estão envolvidos em tumorigênese, e para avaliar se eles estão associados com os eventos envolvidos na progressão epitelial da mucosa oral.

## **Identification of a microRNA signature associated with progression of leukoplakia to oral carcinoma**

Nilva K. Cervigne<sup>1,2</sup>, Patricia P. Reis<sup>1</sup>, Jerry Machado<sup>1,3</sup>, Bekim Sadikovic<sup>4</sup>, Grace Bradley<sup>1,5</sup>, Natalie N. Galloni<sup>6</sup>, Melania Pintilie<sup>7,8</sup>, Igor Jurisica<sup>9,10</sup>, Ralph Gilbert<sup>11</sup>, Patrick Gullane<sup>11</sup>, Jonathan Irish<sup>11</sup>, Suzanne Kamel-Reid<sup>1,3,12\*</sup>

<sup>1</sup>Division of Applied Molecular Oncology, Ontario Cancer Institute, the University Health Network, Toronto, M5G 2M9 ON Canada

<sup>2</sup>Department of Genetics, Bioscience Institute, Sao Paulo State University, Botucatu, SP Brazil

<sup>3</sup>Department of Laboratory Medicine and Pathobiology, University of Toronto, M5G 1L5 ON Canada

<sup>4</sup>Department of Pediatric Laboratory Medicine, and the Genetics & Genome Biology Program, the Hospital for Sick Children, Toronto, ON M5G 1X8 Canada

<sup>5</sup>Faculty of Dentistry, University of Toronto, Toronto, M5G 1G6 ON Canada

<sup>6</sup>Department of Otolaryngology, Hospital Calderon Guardia, San Jose, Costa Rica

<sup>7</sup>Department of Biostatistics, Ontario Cancer Institute, the University Health Network, Toronto, M5G 2M9 ON Canada

<sup>8</sup>Dalla Lana School of Public Health Sciences, University of Toronto, M5T 3M7 ON Canada

<sup>9</sup>Department of Computer Science, Ontario Cancer Institute, the University Health Network, Toronto, M5G 2M9 ON Canada

<sup>10</sup>Department of Signaling Biology, Ontario Cancer Institute, the University Health Network, Toronto, M5G 2M9 ON Canada

<sup>11</sup>Department of Otolaryngology/Surgical Oncology, Princess Margaret Hospital, The University of Toronto and the University Health Network, Toronto, M5G 2M9 ON Canada

<sup>12</sup>Department of Pathology, Toronto General Hospital, Ontario Cancer Institute, the University Health Network, Toronto, M5G 2C4 ON Canada

\*Correspondence to:

Dr. Suzanne Kamel-Reid, PhD, FACMG  
610 University Avenue, Rm. 9-622  
Princess Margaret Hospital, Ontario Cancer Institute and University Health Network  
Toronto, Ontario, M5G 2M9  
CANADA  
Tel. +1 340-4800 x 5739

Fax. +1 (416) 340-3596

E-mail: [Suzanne.Kamel-Reid@uhn.on.ca](mailto:Suzanne.Kamel-Reid@uhn.on.ca)

## **Abstract**

MicroRNAs (miRs) are non-coding RNA molecules involved in cancer initiation and progression. Deregulated miR expression has been implicated in cancer; however there are no studies implicating a miR signature associated with progression in Oral Squamous Cell Carcinoma (OSCC). Although OSCC may develop from oral leukoplakia, clinical and histological assessments have limited prognostic value in predicting which leukoplakic lesions will progress. Our aim was to quantify miR expression changes in leukoplakia and same-site OSCC, and to identify a miR signature associated with progression. We examined miR expression changes in 43 sequential progressive samples from 12 patients and 4 non-progressive leukoplakias from 4 different patients, using TaqMan Low Density Arrays (TLDA). The findings were validated using quantitative PCR in an independent cohort of 52 progressive dysplasias and OSCCs, and 5 non-progressive dysplasias. Global miR expression profiles distinguished progressive leukoplakia/OSCC from non-progressive leukoplakias/normal tissues. 109 miRs were highly expressed exclusively in progressive leukoplakia and invasive OSCC. miR-21, miR-181b and miR-345 expression was consistently increased and associated with increases in lesion severity during progression. Over-expression of miR-21, miR-181b and miR-345 may play an important role in malignant transformation. Our study provides the first evidence of a miR signature potentially useful for identifying leukoplakias at risk of malignant transformation.

## **Introduction**

Oral squamous cell carcinomas (OSCCs) are among the most common types of head and neck cancers, representing the sixth leading cause of cancer death worldwide [2]. Reports indicate that 16-62% of OSCCs develop from pre-malignant oral leukoplakia [3]. Oral leukoplakia are lesions presenting as a “white patch” in the oral mucosa [4, 5] and are currently classified based on clinical and histopathological assessment. Clinically, they are classified as homogeneous or non-homogeneous; the latter group is considered to have a higher risk of malignant transformation [6, 7]. Histologically, these lesions are classified as non-dysplastic or dysplastic (mild, moderate or severe)[7]. Dysplastic leukoplakias are associated with increased risk of transformation [6, 7]. However, clinical and histological characteristics have limited prognostic value, making it difficult to predict which leukoplakias will progress. There is therefore an urgent need for predictive biomarkers that can aid in the analysis of oral leukoplakia. The search for genetic biomarkers of tumor progression is one potential approach; such biomarkers can be used to identify which leukoplakias have an increased likelihood of malignant transformation.

Previous studies have attempted to identify genetic alterations in oral pre-malignant lesions, including leukoplakia, and to correlate these with progression. For example, Califano and colleagues [8] reported loss of heterozygosity (LOH) at 9p21, 3p, and 17p13 in squamous hyperplasia, as well as LOH at 13q11, 13q21, and 14q31 in dysplasias. Similarly, Ha and colleagues [9] reported transcriptional changes occurring from normal tissue to pre-malignant lesions to invasive head and neck carcinomas. However, consistent genetic changes involved in the progression of oral leukoplakia remain unknown. The inability to predict progression may be, in part, because previous studies were based on candidate-gene approaches using unpaired dysplasias and carcinomas. Our study includes the analysis of sequential leukoplakias and same-

site OSCC from the same patients. The analysis of such sequential samples is extremely valuable for the identification of biomarkers that can accurately predict, at the time of diagnosis, which leukoplakia will most likely transform to carcinoma.

Although gene expression changes are informative, there is increasing evidence for the causal involvement of microRNAs (miRs) in cancer initiation and progression [10-12]. miRs are small (18-22 nucleotides in length), non-coding RNA molecules [13] predicted to regulate at least 30% of mRNA transcripts [14-16]. miRs function in concert with tumor suppressors and oncoproteins to regulate key cellular pathways [17, 18] involved in development, differentiation and tumorigenesis. In this study, we examined global miR expression changes during progression of oral malignancy. We identified a specific miR signature that is aberrantly expressed in sequential samples of oral leukoplakia and same-site OSCC. These miRs represent the first non-coding genes implicated in oral carcinoma progression and may be able to predict which leukoplakia have a higher risk of malignant transformation.

## **Material and Methods**

### Patients and samples

This study was approved by the University Health Network (UHN) Research Ethics Board. Our sample group consisted of 43 Formalin-Fixed Paraffin Embedded (FFPE) specimens from 12 patients, collected between 1991-2005, from UHN and the Faculty of Dentistry, University of Toronto. Of these, 29 were leukoplakias and 14 carcinomas (Table 4.1). All carcinomas had at least one preceding leukoplakia. Of the 29 progressive leukoplakias, 6 were non-dysplastic and 23 dysplastic. Four non-progressive leukoplakia samples were also collected (Table 4.2); these were from patients who did not develop OSCC within 5-9 years of follow-up. We were able to

collect seven normal oral mucosa samples (controls) from dental procedures (Table 4.2), representing 14.9% of the training set. As these controls are difficult to obtain, we based the number of required normal samples on the literature (10% of the sample size) [41].

An additional set of 52 FFPE samples was used for validation; 40 were paired dysplasia and OSCC samples from 20 patients. The remaining were six dysplasias and six OSCCs, unpaired, from different 12 patients. In order to ensure that miR changes detected in progressive tissues were not present in non-progressive tissues, we included a group of five additional non-progressive samples in the validation set to be analyzed by real-time PCR.

Formalin-fixed, paraffin-embedded tissue is the most ample tissue source for retrospective studies. The use of FFPE tissues allows retrospective data analysis and increases the feasibility of long term follow-up to enable clinical associations. The extraction of total RNA (including miRs) from FFPE tissue for molecular analysis can be undertaken using commercially available nucleic acid extraction kits. Analysis of their utility has shown that miRs are more stable and more easily recovered from FFPE tissue than mRNAs [42]. Hoefig et al [43] compared miR expression levels in frozen and FFPE tissues of tonsil and liver samples and showed that miR expression was barely influenced by FFPE treatment.

### RNA Isolation

Histopathological assessment was performed by an oral pathologist (GB) to ensure dysplasia or carcinoma in >80% of tissue sections. Samples were needle micro-dissected, according to standard protocols. Total RNA was isolated and purified using the RecoverAll™ Total Nucleic Acid Isolation Kit (Ambion, Austin, TX, USA), according to the manufacturer's protocol. This method allows isolation of total RNA, enriching for miRs from FFPE specimens. The protease digestion conditions of the RecoverAll Kit are designed to release a maximal amount of RNA

fragments of all sizes, including microRNAs, in a relatively short amount of time. RNA quality and quantity were evaluated by spectrophotometry.

#### MicroRNA Expression Quantification

The TaqMan Low Density Array (TLDA) Human MicroRNA Panel and Multiplex RT pools of eight 48-microRNA sets (Applied Biosystems, Streetsville, ON, Canada) were used for global miR expression analysis by quantitative real-time PCR (QRT-PCR). Briefly, total RNA was converted into specific cDNA derived from mature miRs, and quantified using the TaqMan microRNA assay. The TLDA card contained 365 lyophilized human TaqMan miR sequences plus three small nucleolar RNAs (RU6B, RNU4 and RNU44) as endogenous controls. Data were quantified and analyzed using Sequence Detection System (version 2.3) (Applied Biosystems). miR relative expression was normalized against endogenous controls and pooled normal oral mucosa samples using the equation  $2^{-\Delta CT}$ , where  $\Delta CT = (CT_{miRNA} - CT_{snoRNAs})$  [44].

#### MicroRNA Expression Profiling Analysis

Raw expression data were log transformed and the .xls files with relative enrichment values (relative to normal) were imported into Partek Genome Suite (PGS) (Partek Inc, St Louis, MO) [45]. We performed unsupervised hierarchical clustering analysis based on the Euclidean, average linkage, agglomerative method (PGS), blinded to the identity of the samples. Principal Component Analysis (PCA) was performed to visualize high-dimensional data generated based on the sample groupings. Samples were grouped as normal, non-progressive, progressive, and malignant.

Significant differences in miR expression were detected using the 1-way ANOVA test, by comparing miR expression values in non-progressive, progressive, and malignant samples vs. normal. The miR expression data were assigned p-value and average fold-change (relative to



normal samples), and individual miRs satisfying  $p < 0.01$  were compared using Venn analysis. Data were filtered based on fold change;  $\geq 2$ -fold over- or under-expression.

#### Mixed model of linear regression analysis

A statistical linear mixed effects regression model was performed, taking into account that the sequential samples came from the same patient. The type of histology was the explanatory variable. This analysis allowed us to identify miRs that suggested a monotonic change (decrease or increase) with increased severity (non-progressive, keratosis, mild, moderate, severe dysplasia, and OSCC).

#### Validation of deregulated microRNA expression

The RecoverAll Total Nucleic Acid Isolation kit (Ambion) was used for isolation of total RNA from an additional set of 52 progressive and 5 non-progressive FFPE samples. miR expression analysis was performed for the 15 most significantly over-expressed miRs in progressive leukoplakias and same-site OSCCs (miR-21, -345, -181b, -184, -520g, -133a, -133b, -196a, -1, -206, -106b, -649, -518b, -146a, and miR-17-5p) by QRT-PCR, according to manufacturer's protocols (Applied Biosystems). Briefly, total RNA was converted into cDNA using specific microRNA primer-probes. The RT-PCR product was diluted 10-fold, mixed with TaqMan Universal PCR Master Mix (Applied Biosystems), and added to 384-well custom plates using a liquid-handling robot. Each miR was spotted in triplicate; expression data was normalized against RNU44; pooled normal oral mucosa was used as the reference for  $\Delta\Delta C_t$  analysis.

MicroRNA expression was compared to normal using the one-sided Wilcoxon signed rank test; validated miRs would be over-expressed in both progressive dysplasias and OSCC. A linear mixed effect regression analysis was performed to test the difference between progressive dysplasias and OSCCs for each miR.

## Results

### Lesion-specific miR expression profiles

Our goal was to identify deregulated miR expression associated with progressive leukoplakia and OSCC, compared to non-progressive leukoplakia and normal oral mucosa. Since miR expression is tissue specific [19], we determined baseline miR expression in normal oral mucosa samples, alongside normal tongue RNA (Stratagene). Of the 365 miRs analyzed, 286 were consistently expressed in all normal samples and were used for further analyses. Samples clustered into three distinct groups: 22/29 progressive leukoplakia and 14/14 invasive OSCCs clustered together (Figure 4.1A), indicating that they shared a common miR expression profile. Interestingly, 4/29 progressive leukoplakia clustered separately; all had histological characteristics of mild hyperplasia (with or without mild atypia) and 3/4 samples were from the same patient. All non-progressive leukoplakia and normal oral mucosa grouped together; 3/29 progressive leukoplakias (two hyperplasias and one moderate dysplasia) were found within this cluster.

We found a large number of commonly over-expressed miRs in progressive leukoplakia and OSCC, a rare finding in non-progressive leukoplakia. Median expression values were calculated for every miR in each cluster, and compared with normal tissue, in order to identify deregulated miRs. 125/286 miRs were deregulated in progressive leukoplakias and OSCCs that clustered together. 109/125 were significantly over-expressed ( $p < 0.01$ ). 103/286 miRs were deregulated in non-progressive leukoplakia compared to normal. A higher frequency (55/103) of under-expressed miRs was observed in non-progressive leukoplakia; only two miRs were significantly over-expressed.

A group of four over-expressed miRs (miR-21, miR-181b, miR-345, miR-146a) were found in progressive leukoplakia and OSCCs that clustered together, but not in normal oral mucosa or

non-progressive leukoplakia (Figure 4.2). The marked over-expression of these miRs in progressive leukoplakias and OSCCs may indicate their involvement in molecular pathways associated with progression.

Principal Component Analysis (PCA) was used to analyze 3-dimensional clustering of miR expression in non-progressive, progressive leukoplakia and OSCC. As with our hierarchical clustering analysis, progressive leukoplakia and OSCC grouped together, whereas normal and non-progressive samples showed proximal clustering, segregating away from the first group (figure 1B). Strikingly, PCA identified the same three groupings of progressive leukoplakia uncovered by our initial Euclidean clustering analysis (Figure 4.1B).

We next analyzed miR expression using the ANOVA test ( $p < 0.01$ ). As expected for carcinoma tissues containing an accumulation of genetic changes, OSCC samples showed the largest number ( $n=53$ ) of deregulated miRs compared to normal. 35/53 were highly expressed, exclusively in OSCC; 16/53 were significantly over-expressed in both progressive leukoplakia and OSCC. Eleven miRs were exclusively deregulated (ten over- and one under-expressed) in progressive leukoplakia. Three miRs were exclusively deregulated in non-progressive leukoplakia (two significantly over- and one under-expressed). Two miRs were commonly deregulated in all three sample groups, compared to normal tissue (Figure 4.3).

#### Deregulated miR expression during progression

We used a mixed linear regression analysis model to determine whether deregulated miR expression was associated with different stages of progression. hsa-mir-21, hsa-mir-181b and hsa-mir-345 (over-expressed in all progressive leukoplakias and OSCCs that clustered together) demonstrated consistent increases in expression levels, associated with increased lesion severity. Thus, the gradual over-expression of these three miRs may play a role in malignant

transformation. In direct contrast, these miRs were unchanged in the group of non-progressive leukoplakias. Average miR expression based on this model is shown in Figure 4.4.

Our data was also adjusted for multiple comparisons using Bonferroni correction, and the 3 miRs (miR-21, miR-181b and miR-345) identified using this model remained statistically significant ( $p < 0.00017$ ).

#### miRs as biomarkers of oral cancer progression

Based on significance, 15 deregulated miRs were selected for validation. These miRs demonstrated significant over-expression in both progressive leukoplakias and same-site OSCCs compared to normal; among these were hsa-mir-21, hsa-mir-181b and hsa-mir-345, which demonstrated consistent increases in expression levels, associated with lesion severity (Figure 4.4). We initially confirmed the expression of these miRs by QRT-PCR in a subset of leukoplakia and OSCCs. The expression patterns of these 15 miRs were consistent between the TLDA and QRT-PCR assays ( $r = 0.94$ , data not shown).

Validation by QRT-PCR was performed to verify whether these 15 miRs were significantly over-expressed in both progressive dysplasias and same-site OSCCs, but not in non progressive dysplasias, when compared to normal. The miR expression data were subjected to the one-side Wilcoxon test, which enabled us to validate the over-expression of 8/15 miRs (miR-146a, miR-181b, miR-184, miR-21, miR-345, miR-518b, miR-520g and miR-649) in progressive dysplasias and OSCCs. These 8 miRs were found unchanged or under-expressed in non-progressive samples compared to normal. A fold-change larger than 1 was considered over-expressed; the associated p values are shown in Table 4.3.

We also applied linear mixed effect regression analysis to compare miR expression levels between OSCC and same-site dysplasias, and confirmed that the same eight miRs had similar over-expression levels in progressive dysplasias and OSCCs, with  $p > 0.01$ .

Five miRs (miR-1, miR-133a, miR-133b, miR-196a and miR-206) had differential expression levels between progressive dysplasia and OSCC ( $p < 0.005$ ) and did not validate our initial TLDA findings of over-expression in both progressive leukoplakia and OSCC. miR-196a and miR-206 were significantly over-expressed in OSCCs, but under-expressed in progressive dysplasias. Additionally, miR-1, miR-133a and miR-133b were under-expressed in both groups at significantly different levels.

## **Discussion**

MicroRNA expression profiling has been able to identify signatures associated with cancer initiation and progression. Ours is the first study to examine miR expression in sequential oral lesions, with the aim of identifying a signature as a potential predictor of malignant transformation. The miR signature identified was able to distinguish between progressive and non-progressive leukoplakia, providing the first evidence of miRs as potential biomarkers in oral carcinoma progression.

In our study, the inclusion of sequential progressive samples of oral leukoplakia and same-site OSCC, from the same patients, allowed us to follow the miR changes as the tissue progressed from a benign to a transformed state, and, finally, invasive carcinoma. The non-progressive samples were included for comparison, to ensure that miR changes detected in progressive tissues were not present in non-progressive tissues. The inclusion of miR expression data from the fully progressed OSCC allowed us to verify that miR changes detected early in leukoplakia

were also present in the invasive carcinoma. When we compared miR expression profiles of histologically different tissues, we successfully demonstrated a clear segregation between normal/non-progressive leukoplakia and progressive leukoplakia/invasive OSCC. The miR expression pattern showed little difference between pre-malignant lesions that progressed and malignant tissues, in contrast to the differences observed between normal oral mucosa and pre-malignant tissues that progressed. Thus, our data demonstrate that pre-malignant lesions that progressed possess many of the genetic alterations present in invasive cancers. This is consistent with the hypothesis that the majority of genetic alterations may occur early during head and neck cancer progression [9].

About 10% of the progressive leukoplakias (3/29) were found within the non-progressive leukoplakia samples and normal oral mucosa tissues that grouped together. Two of these leukoplakias (hyperplasias) were non-dysplastic; the other leukoplakia showed evidence of moderate dysplasia. The clustering of these samples together with the non-progressive and normal samples may be due to sampling, or misclassification due to genetic heterogeneity. Such biological variation can be often seen in global genomic analysis.

Altogether, 64 miRs were significantly deregulated in progressive leukoplakia and OSCCs (Figure 4.3), whereas only five miRs were altered in non-progressive leukoplakia and OSCC. This is not surprising, as comparison between groups with the greatest genetic and phenotypic differences across the continuum from normal oral tissue to OSCC is more likely to be significant, in contrast with groups that are more phenotypically and genetically similar.

In Head and Neck Cancer, a previous microarray study assessed miR expression in nine cancer cell lines from four different sites [20]. As in our study, over-expressed miRs, such as hsa-miR-16 and hsa-miR-21, were identified. The deregulation of these miRs is likely to reflect primary

tumor biology, suggesting that they may have important roles in oral carcinogenesis. However, they also identified several under-expressed miRs across all cell lines, while we identified mostly miR over-expression. These differing findings may be due to differences between sample types, such as established cell lines *vs.* primary patient material, as well as the comparison with normal oral tissue in our study.

Interestingly, three miRs; hsa-mir-21, hsa-mir-181b and hsa-mir-345 had a consistent and gradual increase in expression significantly associated with progression (Figure 4.4). This is the first evidence of a consistent process of differential alteration in miR profiles occurring during oral cancer progression.

A large-scale analysis including lung, breast, stomach, prostate, colon, and pancreatic tumors, identified a miR signature composed of a large number of over-expressed miRs [1]. Among these miRs are some with well characterized cancer association, such as miR-21. This miR is frequently over-expressed in solid and hematological tumors, and it has been shown to play roles in cell proliferation and apoptosis [21]. miR-21 has been implicated as an anti-apoptotic factor, as knocking it down increases apoptotic cell death in human glioblastoma cells [22]. This miR is also over-expressed in hepatocellular carcinoma (HCC) and may inhibit apoptosis through regulation of PTEN and Programmed Cell Death 4 (PDCD4), which are frequently down-regulated in HCC [23]. Tumor suppressor genes, such as tropomyosin 1 (TPM1) [24] and serpin peptidase inhibitor, clade B (ovalbumin), member 5 (SERPINB5) [25], have been shown to be targets of miR-21, suggesting that this miR may also play a role in tumor invasion and metastasis.

miR-181b is frequently up-regulated in B cell chronic lymphocytic leukemias [26], brain tumors [27], thyroid [28] and gastro-intestinal carcinomas [29]. miR-181b is known to down-regulate

the homeobox protein, HOXA11 (a repressor of differentiation), thus establishing that miR-181 is functional associated with the process of mammalian skeletal muscle differentiation [30].

miR-345, has not been well characterized in the literature yet, however some studies have shown its over-expression in prostate carcinomas [31], and under-expression in thyroid [32] and pancreatic carcinoma [33]. We suggest that hsa-miR-21, hsa-miR-181b and hsa-miR-345 represent a unique group of miRs that may be potentially involved in oral cancer progression.

A recent report showed that a miR expression ratio (miR-221:miR-375) has high selective potential in distinguishing head and neck tumor from normal tissue, and suggests that this ratio may be used as diagnostic tool [34]. This study also detected an increase in miR-21 and miR-181b expression in tumors compared to normal tissue. These two miRs were also over-expressed in our data set and furthermore strengthen our prognostic signature of progression.

MicroRNAs over-expressed in tumors may be considered oncogenes, as their over-expression accelerates tumorigenesis [35]. These oncogenic miRs, also called “oncomirs”, usually promote tumor development by inhibiting tumor suppressor genes and/or genes that control cell differentiation or apoptosis. In this study, we found that miR over-expression appears to have a higher correlation with oral cancer progression. These findings were corroborated by our validation results, showing the potential relevance of miR over-expression in oral carcinogenesis. In the additional cohort of samples used for validation, we showed over-expression of 8 miRs (miR-21, -345, -181b, -184, -520g, -649, - 518b, -146a) in progressive dysplasias and OSCCs. Such widespread over-expression may occur because the identified miRs are located in cancer-associated genomic regions, and could be deregulated by epigenetic mechanisms or alterations in the miR processing machinery [36]. Of the deregulated miRs, two putative oncomirs were significantly over-expressed in our progressive dysplasias and OSCCs. Recently, miR-184 was



shown to be associated with antiapoptotic and proliferative processes in tongue carcinoma [37] and miR-21 was shown to play a role in HNSCC tumorigenesis [38]. Together with our data, these findings suggest that miR over-expression may be an early event during oral cancer progression.

Two miRs, miR-206 and miR-196a, were over-expressed in OSCC, but under-expressed in dysplasias. In our initial analysis, these miRs were over-expressed in leukoplakias that progressed, as well as same-site OSCCs. Both training and validation sample sets were composed of samples representing the different grades of progressive dysplasias, however the training set had a higher number of severe dysplasias (n=8) compared to the validation set (n=3). Interestingly, when we looked only at the severe dysplasias from both the training and the validation sample sets, severe dysplasias showed over-expression of these two miRs. These findings may suggest that under-expression of these miRs in mild to moderate progressive dysplasias, and their subsequent over-expression in severe dysplasias and invasive carcinoma, may be involved in later stages of OSCC progression.

A proposed model of miR changes associated with progression of leukoplakia to OSCC is shown in figure 5. We show that a subset of over-expressed miRs (miR-146b, miR-181b, miR-21, miR-345, miR-518b, miR-520g, miR-649 and miR-184) may be considered early detectable events in oral tumor progression, since they were commonly deregulated in progressive and malignant lesions, and were unchanged or under-expressed in non-progressive lesions. In particular, miR-345, miR-21 and miR-181b are strongly associated with increased lesion severity during the transition from histologically pre-malignant to malignant lesions. According to this model, the final stages of carcinogenesis may also involve changes in miR-196a and miR-206, which appear under-expressed in premalignant lesions and over-expressed in carcinomas. Interestingly, these

miRs are predicted to target cancer-associated genes, eg, *TNFRSF10B*, *ACAT1*, *NFIB*, *CCLI*, *MSH2*, *ACYPI*, and *PCBP2* (miR-21); *GRMI*, *MAP3K10*, *CCDC42*, *SNHG5* (miR-181b); *TSPO*, *RFXDC1*, *ZNF133*, *MORN3* (miR-345). Alterations in miR expression associated with genomic aberration and transcriptional changes may lead to deregulation of key cellular processes (eg, differentiation, proliferation, and apoptosis) that underlie oral cancer development and progression. The combination of miR expression profiles and mRNA target prediction may provide a snapshot of genes/pathways deregulated in cancer.

MicroRNA signatures may be valuable as surrogate markers of tumor progression. Recent transcriptional profiling studies have enabled the discovery of numerous miRs differentially expressed in normal tissues vs. tumors, and associated with cancer development, diagnosis, and prognosis [39, 40]. miR signatures may be used to detect and classify cancer and predict the severity of disease, with certain profiles of miR expression linked to more aggressive cancers. Our data support the use of a miR signature to segregate lesions for risk of malignant transformation and suggest that over-expression of a specific subset of miRs may constitute a signature for progression. The miR over-expression signature we identified may be applied in the future as a “multi-miR expression analysis tool”, which, together with clinical and histological analyses, will help assess which leukoplakias have a higher risk of progression to malignancy. Such information would ultimately lead to improved treatment and, consequently, improved patient survival.

Our study identified the first non-coding set of RNAs that may be relevant in oral cancer progression and suggests that these miRs may be potential biomarkers to assess which leukoplakias have a risk of malignant transformation. Our study highlights the potential of miR

expression analysis in early detection of oral cancer, ultimately improving diagnosis and patient outcome.

### **Funding**

This work was supported by the Galloway Fund [RG and SKR] and the Cancer Research Society [SKR].

### **Acknowledgments**

We appreciate the contribution of Dr. Mahadeo Sukhai, for carefully reading and editing this manuscript and Dr. Rashmi Goswami for helping with sample assessment.

### **Conflicts of interest**

All Authors declare no financial or conflict of interest.

**Table 4.1.** Sample set of sequential oral leukoplakia and same-site OSCCs.

| Patient | Sample ID | Sample Site          | Histopathological Diagnosis                          | Biopsy date (d/m/yr) |
|---------|-----------|----------------------|--|----------------------|
| 1       | 1a        | Tonsil               | Focal keratosis <sup>1</sup>                         | 02.10.2001           |
|         | 1b        | Alveolus + FOM       | Mild squamous hyperplasia <sup>1</sup>               | 15.07.2003           |
|         | 1c        | Anterior FOM         | Severe dysplasia                                     | 07.05.2004           |
|         | 1d        | Anterior FOM         | Invasive moderately differentiated keratinizing OSCC | 07.05.2004           |
| 2       | 2a        | Tongue               | Moderate dysplasia                                   | 23.02.1996           |
|         | 2b        | Tongue               | Invasive moderately differentiated OSCC              | 13.08.2003           |
| 3       | 3a        | Tongue               | Keratosis with moderate dysplasia                    | 11.06.1999           |
|         | 3b        | Tongue               | Invasive moderately differentiated OSCC              | 12.12.2000           |
| 4       | 4a        | Tongue               | Severe squamous dysplasia                            | 21.11.1997           |
|         | 4b        | Tongue               | Mild squamous dysplasia                              | 08.10.1998           |
|         | 4c        | Tongue               | Severe squamous dysplasia                            | 03.11.1998           |
|         | 4d        | Tongue               | Severe squamous dysplasia+SCC micro foci             | 26.11.1998           |
|         | 4e        | Tongue               | Invasive moderately differentiated OSCC              | 28.02.2000           |
| 5       | 5a        | Left lateral tongue  | Keratosis <sup>1</sup>                               | 08.01.1993           |
|         | 5b        | Right lateral tongue | moderate dysplasia                                   | 08.03.1993           |
|         | 5c        | Right lateral tongue | Keratosis  | 08.03.1993           |
|         | 5d        | FOM                  | Severe dysplasia                                     | 27.10.1994           |
|         | 5e        | Buccal mucosa        | Hyperkeratotic squamous papilloma+dysplasia          | 06.02.1997           |
|         | 5f        | Left buccal mucosa   | Invasive OSCC  | 28.10.1997           |
| 6       | 6a        | Right tongue         | Moderate to severe dysplasia                         | 14.05.2003           |
|         | 6b        | Right tongue         | Invasive moderately differentiated OSCC              | 23.06.2003           |
|         | 6c        | Right tongue         | Invasive moderately differentiated OSCC              | 12.09.2003           |
| 7       | 7a        | Left tongue          | Moderate to severe dysplasia                         | 13.05.1999           |
|         | 7b        | Left tongue          | Severe dysplasia                                     | 25.04.2000           |
|         | 7c        | Left tongue          | Microinvasive OSCC                                   | 17.05.2000           |
| 8       | 8a        | Left buccal mucosa   | Keratosis <sup>1</sup>                               | 28.03.1991           |
|         | 8b        | Left buccal mucosa   | Moderate dysplasia                                   | 15.09.1995           |
|         | 8c        | Left buccal mucosa   | Mild dysplasia                                       | 26.09.1996           |
|         | 8d        | Left buccal mucosa   | Severe dysplasia                                     | 09.10.2003           |
|         | 8e        | Left buccal mucosa   | SCC micro foci                                       | 20.11.2003           |
|         | 8f        | Left buccal mucosa   | Invasive moderately differentiated OSCC              | 25.06.2004           |
| 9       | 9a        | Left lateral tongue  | Moderate dysplasia                                   | 04.02.2002           |
|         | 9b        | Left lateral tongue  | Invasive moderately differentiated OSCC              | 31.05.2005           |
| 10      | 10a       | Right tongue         | Moderate dysplasia                                   | 19.04.2001           |
|         | 10b       | Right tongue         | Poorly differentiated OSCC                           | 12.09.2005           |
| 11      | 11a       | Right tongue         | Mild squamous hyperplasia <sup>1</sup>               | 06.11.2003           |
|         | 11b       | Right tongue         | Moderately differentiated OSCC                       | 15.04.2005           |
| 12      | 12a       | Right tongue         | Mild dysplasia                                       | 07.04.1994           |
|         | 12b       | Right tongue         | Moderate dysplasia                                   | 07.04.1994           |
|         | 12c       | Right tongue         | Keratosis mild atypia                                | 23.02.2001           |
|         | 12d       | Right tongue         | Atypical squamous epithelium                         | 10.10.2003           |
|         | 12e       | Right tongue         | Moderately differentiated OSCC                       | 30.10.2003           |
|         | 12f       | Right tongue         | Mild squamous hyperplasia <sup>1</sup>               | 08.03.2004           |

<sup>1</sup>Non-dysplastic lesions; OSCC: oral squamous cell carcinoma; FOM: floor of mouth; WD: well differentiated, MD: moderately differentiated, PD: poorly differentiated.

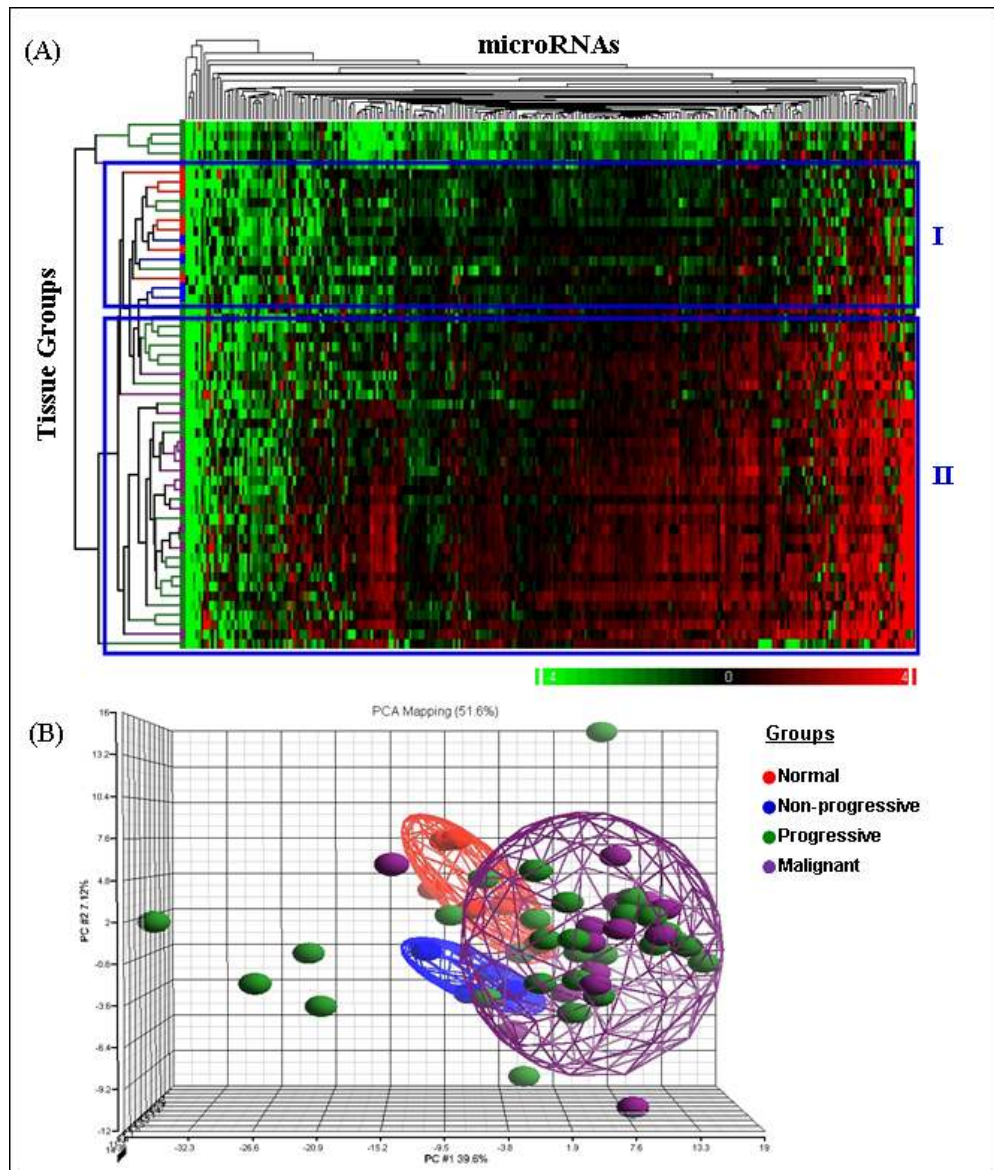
**Table 4.2.** Characteristics of the non-progressive leukoplakia and normal oral mucosa samples.

| Patient                             | Sample site          | Histopathological<br>Diagnosis       | Age (yrs) | Gender | Follow-up<br>(non-progressive leukoplakias) |
|-------------------------------------|----------------------|--------------------------------------|-----------|--------|---|
| <b>Non-Progressive Leukoplakias</b> |                      |                                      |           |        |   |
| <b>Training sample set</b>          |                      |                                      |           |        |   |
| Pt16 (1NP)                          | Floor of mouth       | Severe dysplasia                     | 63        | F      | 6 years                                     |
| Pt 17 (2NP)                         | Floor of mouth       | Moderate dysplasia                   | 60        | M      | 5 years                                     |
| Pt18 (3NP)                          | Mandibular mucosa    | Mild dysplasia                       | 58        | M      | 7 years                                     |
| Pt19 (4NP)                          | Mandibular mucosa    | Mild dysplasia                       | 32        | M      | 9 years                                     |
| <b>Validation sample set</b>        |                      |                                      |           |        |   |
| Pt1 (1NP)                           | Right ventral tongue | Hyperkeratosis w/ mild dysplasia     | 56        | F      | 5 years 10m                                 |
| Pt2 (2NP)                           | Mandibular mucosa    | Hyperkeratosis w/ mild dysplasia     | 50        | M      | 9 years 2m                                  |
| Pt3 (3NP)                           | Tongue               | Hyperkeratosis w/ moderate dysplasia | 40        | M      | 6years 9m                                   |
| Pt4 (4NP)                           | Floor of mouth       | Hyperkeratosis w/ mild dysplasia     | 55        | F      | 5 years 2m                                  |
| Pt5 (5NP)                           | Flour of mouth       | Hyperkeratosis w/ severe dysplasia   | 78        | F      | 5 years 11m                                 |
| <b>Normal Oral Mucosa</b>           |                      |                                      |           |        |   |
| 1N                                  | Tongue               | Normal                               | 37        | M      | -   |
| 2N                                  | Buccal mucosa        | Normal                               | 24        | M      | -   |
| 3N                                  | Buccal Mucosa        | Normal                               | 40        | M      | -   |
| 4N                                  | Lip mucosa           | Normal                               | 56        | M      | -   |
| 5N                                  | Tongue               | Normal                               | 36        | F      | -   |
| 6N                                  | Tongue               | Normal                               | 41        | F      | -   |
| 7N                                  | Lower lip mucosa     | Normal                               | 59        | M      | -   |

**Table 4.3.** Quantitative expression of a subset of 15 miRs in progressive dysplasias and OSCCs relative to normal oral mucosa.

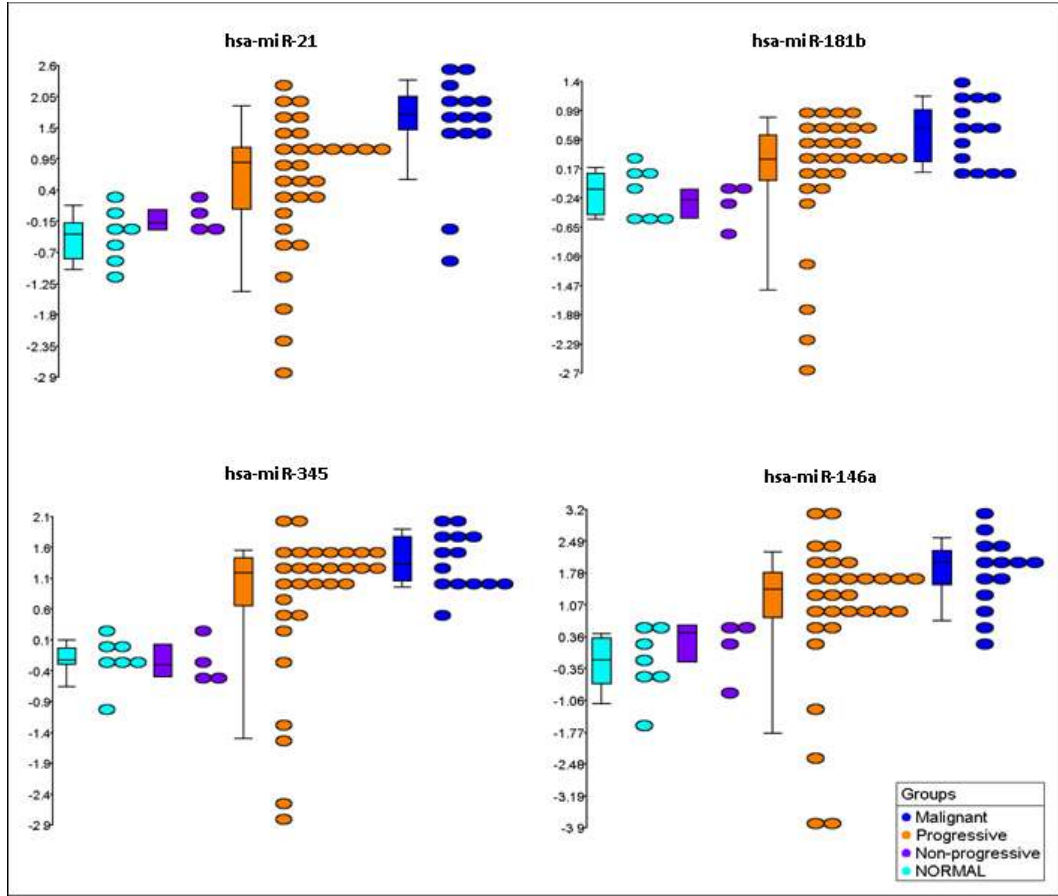
| miR ID          | Dysplasia                   |                 | OSCC                        |                   |
|-----------------|-----------------------------|-----------------|-----------------------------|-------------------|
|                 | <i>fold change (median)</i> | <i>p-value</i>  | <i>fold change (median)</i> | <i>p-value</i>    |
| miR-1           | 0.008                       | 1               | 0.123                       | 0.99              |
| miR-106b        | 1.157                       | 0.067           | 1.134                       | 0.054             |
| miR-133a        | 0.016                       | 1               | 0.259                       | 0.99              |
| miR-133b        | 0.013                       | 1               | 0.319                       | 1                 |
| <b>miR-146a</b> | 5.031                       | <b>0.00019</b>  | 7.547                       | <b>0.0000082</b>  |
| miR-17-5p       | 0.351                       | 0.98            | 0.644                       | 0.97              |
| <b>miR-181b</b> | 2.190                       | <b>0.00032</b>  | 3.539                       | <b>0.0000015</b>  |
| <b>miR-184</b>  | 1.860                       | <b>0.0094</b>   | 2.388                       | <b>0.000054</b>   |
| miR-196a        | 0.519                       | 0.16            | 4.185                       | 0.0016            |
| miR-206         | 0.074                       | 0.99            | 4.286                       | 0.00014           |
| <b>miR-21</b>   | 2.846                       | <b>0.0025</b>   | 3.988                       | <b>0.0000003</b>  |
| <b>miR-345</b>  | 2.073                       | <b>0.00032</b>  | 3.528                       | <b>0.0000033</b>  |
| <b>miR-518b</b> | 2.557                       | <b>0.00007</b>  | 3.076                       | <b>0.0000082</b>  |
| <b>miR-520g</b> | 2.570                       | <b>0.0028</b>   | 9.932                       | <b>0.00000012</b> |
| <b>miR-649</b>  | 3.770                       | <b>0.000005</b> | 3.464                       | <b>0.0000003</b>  |

This analysis was performed using quantitative real-time PCR. The median fold changes and p-values for each miR were determined using Wilcoxon one-sided test; here we considered a fold change larger than 1 as over-expressed in Dysplasias and OSCCs when compared to normal. miRs that had  $p < 0.01$ , in both dysplasia and OSCC, were considered validated by this analysis (miRs and respective p values highlighted in bold).



**Figure 4.1. Clustering of miR expression in oral lesions**

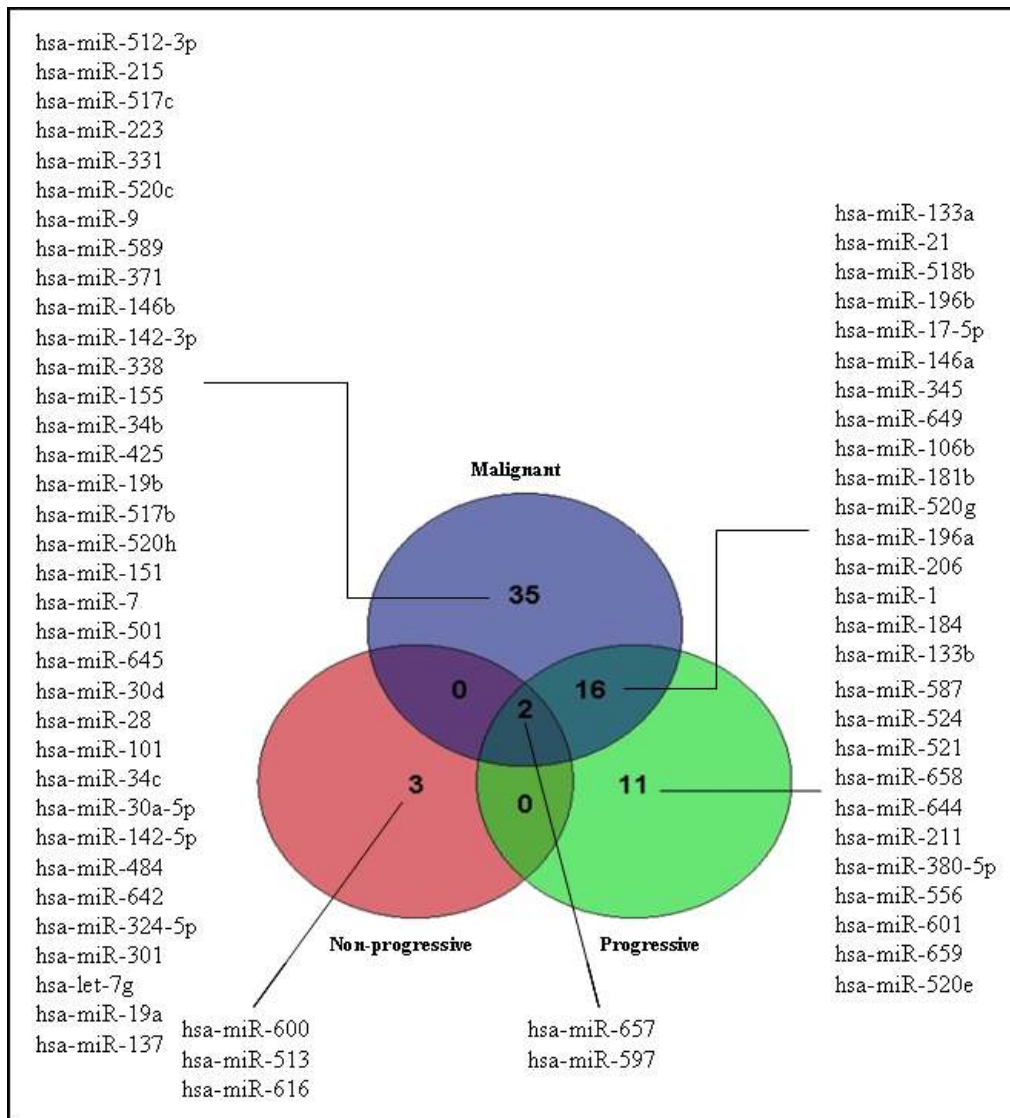
Partek Genome Suite analysis: (A) Hierarchical clustering of miR expression levels. Two distinct clusters were observed: normal/non-progressive leukoplakias (I) and progressive leukoplakias/OSCCs (II). (B) Principal Component Analysis confirming the similarity between malignant and progressive lesions (purple and green, respectively), and showing how normal and non-progressive leukoplakias (red and blue, respectively) are grouped separately.



**Figure 4.2. Highly expressed miRs identified by hierarchical clustering**

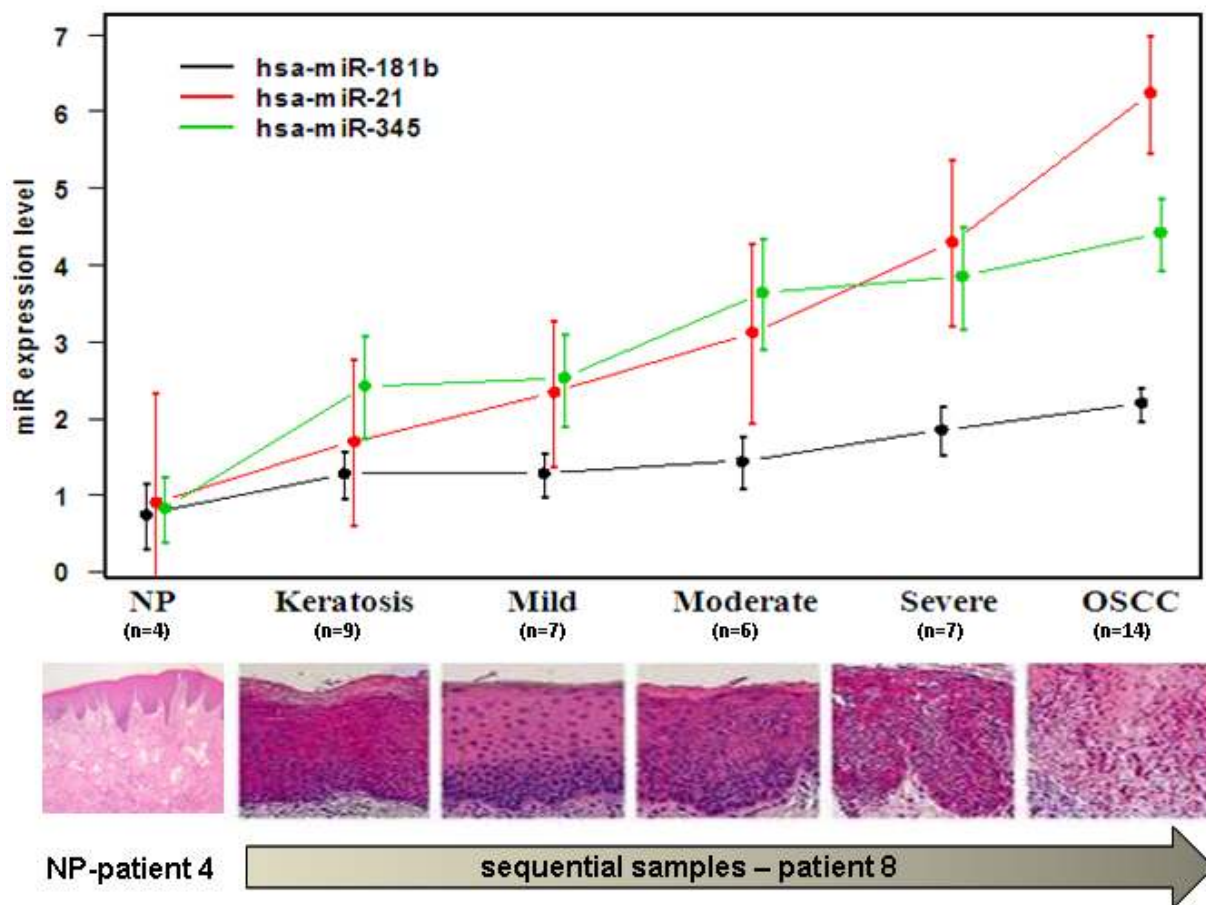
Box Plots representing four miRs highly over-expressed across progressive lesions and OSCCs. Bars show whiskers extending to minimum and maximum observations with median marked in the centre. The standard-bar represents the miR expression level for each group. Circles refer to the number of samples in each group (Malignant, n=14; Progressive, n=23; Non-progressive, n=4; Normal, n=7).





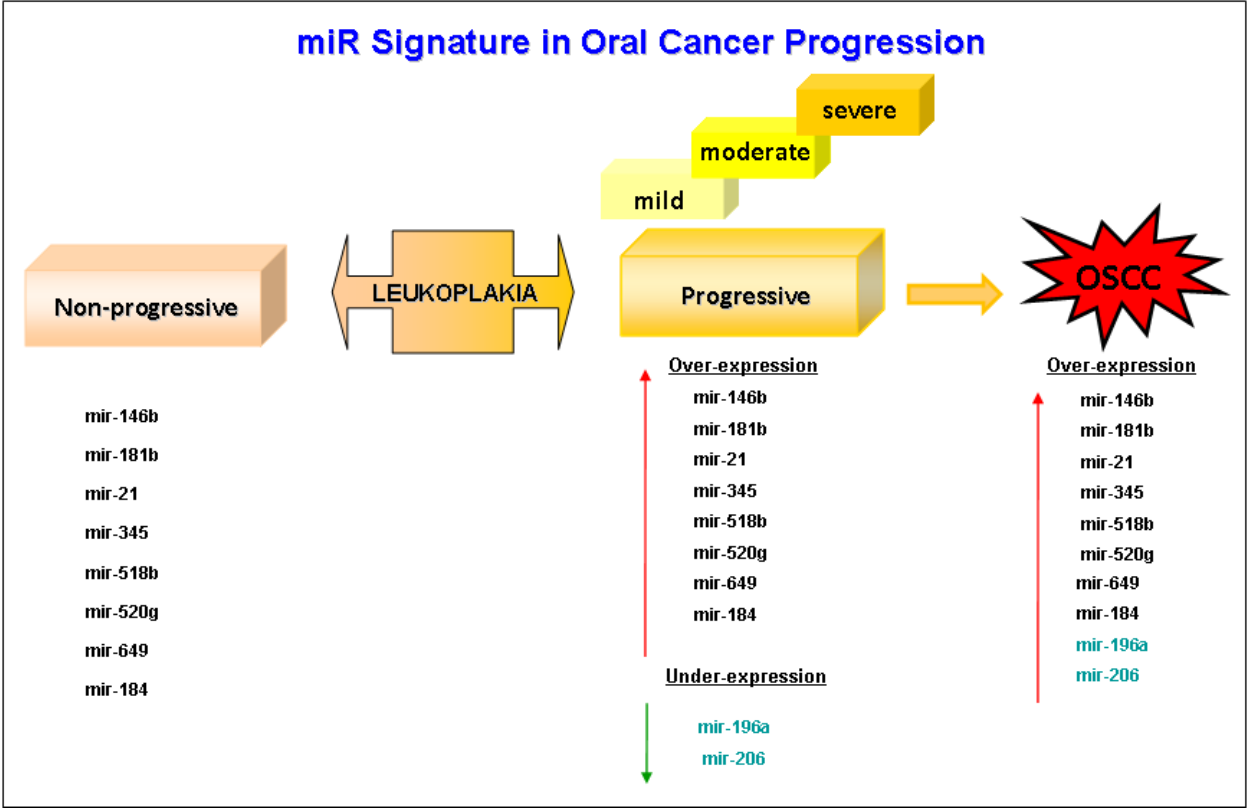
**Figure 4.3. VENN diagram showing miR signatures in oral lesions**

Significantly deregulated miRs ( $p < 0.01$ ) within the 3 groups compared to normal oral mucosa. Eleven miRs were specific to progressive leukoplakias, 35 specific to malignant tumors and 3 to non-progressive leukoplakias. Eighteen miRs were commonly deregulated between progressive leukoplakias and malignant tumors; 2/18 miRs were commonly deregulated in non-progressive, progressive and malignant groups.



**Figure 4.4. miR over-expression associated with increased lesion severity**

Statistical linear regression model analysis showing that miR-21, miR-181b, and miR-345 expression increases in oral tissue lesions during cancer progression. Bottom illustrates H&E-stained tissue sections showing a representative example of sequential samples from one patient, and its association with miR deregulation during progression. Y axis: mean miR expression level; x axis: type of oral tissue specimens analyzed (NP: Non Progressive samples; followed by the progressive samples - keratosis, mild, moderate and severe dysplasias, and oral squamous cell carcinoma).



**Figure 4.5.** miR signature of oral cancer progression

Signature of deregulated miRNAs during oral cancer progression. The model shows a subset of 9/11 early changed miRNAs that were over-expressed in progressive dysplasias and same-site OSCCs. 2/11 miRNAs (highlighted in blue) were under-expressed in progressive leukoplakias, but over-expressed in carcinomas. All 11 miRNAs were unchanged in non-progressive and normal oral mucosa.

## References

1. Volinia, S., et al., *A microRNA expression signature of human solid tumors defines cancer gene targets*. Proc Natl Acad Sci U S A, 2006. **103**(7): p. 2257-61.
2. Parkin, D.M., P. Pisani, and J. Ferlay, *Global cancer statistics*. CA Cancer J Clin, 1999. **49**(1): p. 33-64, 1.
3. Lee, J.J., et al., *Predicting cancer development in oral leukoplakia: ten years of translational research*. Clin Cancer Res, 2000. **6**(5): p. 1702-10.
4. Pindborg, J.J., et al., *Histological typing of cancer and precancer of the oral mucosa*. 2. ed. ed. International histological classification of tumours, ed. W.H. Organization. 1997: Berlin, Germany : Springer. 87.
5. Warnakulasuriya, S., N.W. Johnson, and I. van der Waal, *Nomenclature and classification of potentially malignant disorders of the oral mucosa*. J Oral Pathol Med, 2007. **36**(10): p. 575-80.
6. Holmstrup, P., et al., *Long-term treatment outcome of oral premalignant lesions*. Oral Oncol, 2006. **42**(5): p. 461-74.
7. Reibel, J., *Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics*. Crit Rev Oral Biol Med, 2003. **14**(1): p. 47-62.
8. Califano, J., et al., *Genetic progression model for head and neck cancer: implications for field cancerization*. Cancer Res, 1996. **56**(11): p. 2488-92.
9. Ha, P.K., et al., *A transcriptional progression model for head and neck cancer*. Clin Cancer Res, 2003. **9**(8): p. 3058-64.
10. Calin, G.A., et al., *Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers*. Proc Natl Acad Sci U S A, 2004. **101**(9): p. 2999-3004.
11. Lu, J., et al., *MicroRNA expression profiles classify human cancers*. Nature, 2005. **435**(7043): p. 834-8.
12. Tsuchiya, S., Y. Okuno, and G. Tsujimoto, *MicroRNA: biogenetic and functional mechanisms and involvements in cell differentiation and cancer*. J Pharmacol Sci, 2006. **101**(4): p. 267-70.
13. Ambros, V., *The functions of animal microRNAs*. Nature, 2004. **431**(7006): p. 350-5.
14. Bentwich, I., *Prediction and validation of microRNAs and their targets*. FEBS Lett, 2005. **579**(26): p. 5904-10.
15. Griffiths-Jones, S., et al., *miRBase: microRNA sequences, targets and gene nomenclature*. Nucleic Acids Res, 2006. **34**(Database issue): p. D140-4.
16. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets*. Cell, 2005. **120**(1): p. 15-20.
17. Brennecke, J., et al., *bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila*. Cell, 2003. **113**(1): p. 25-36.
18. Hatfield, S.D., et al., *Stem cell division is regulated by the microRNA pathway*. Nature, 2005. **435**(7044): p. 974-8.
19. Babak, T., et al., *Probing microRNAs with microarrays: tissue specificity and functional inference*. RNA, 2004. **10**(11): p. 1813-9.

20. Tran, N., et al., *MicroRNA expression profiles in head and neck cancer cell lines*. Biochem Biophys Res Commun, 2007. **358**(1): p. 12-7.
21. Krichevsky, A.M. and G. Gabriely, *miR-21: a small multi-faceted RNA*. J Cell Mol Med, 2009. **13**(1): p. 39-53.
22. Chan, J.A., A.M. Krichevsky, and K.S. Kosik, *MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells*. Cancer Res, 2005. **65**(14): p. 6029-33.
23. Lu, Z., et al., *MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene*. Oncogene, 2008. **27**(31): p. 4373-9.
24. Zhu, S., et al., *MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1)*. J Biol Chem, 2007. **282**(19): p. 14328-36.
25. Zhu, S., et al., *MicroRNA-21 targets tumor suppressor genes in invasion and metastasis*. Cell Res, 2008. **18**(3): p. 350-9.
26. Garzon, R., et al., *MicroRNA gene expression during retinoic acid-induced differentiation of human acute promyelocytic leukemia*. Oncogene, 2007. **26**(28): p. 4148-57.
27. Ciafre, S.A., et al., *Extensive modulation of a set of microRNAs in primary glioblastoma*. Biochem Biophys Res Commun, 2005. **334**(4): p. 1351-8.
28. Schetter, A.J., et al., *MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma*. JAMA, 2008. **299**(4): p. 425-36.
29. Nikiforova, M.N., et al., *MicroRNA expression profiling of thyroid tumors: biological significance and diagnostic utility*. J Clin Endocrinol Metab, 2008. **93**(5): p. 1600-8.
30. Naguibneva, I., et al., *The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation*. Nat Cell Biol, 2006. **8**(3): p. 278-84.
31. Porkka, K.P., et al., *MicroRNA expression profiling in prostate cancer*. Cancer Res, 2007. **67**(13): p. 6130-5.
32. Pallante, P., et al., *MicroRNA deregulation in human thyroid papillary carcinomas*. Endocr Relat Cancer, 2006. **13**(2): p. 497-508.
33. Lee, E.J., et al., *Expression profiling identifies microRNA signature in pancreatic cancer*. Int J Cancer, 2007. **120**(5): p. 1046-54.
34. Avissar, M., et al., *MicroRNA expression ratio is predictive of head and neck squamous cell carcinoma*. Clin Cancer Res, 2009. **15**(8): p. 2850-5.
35. Cho, W.C., *OncomiRs: the discovery and progress of microRNAs in cancers*. Mol Cancer, 2007. **6**: p. 60.
36. Calin, G.A. and C.M. Croce, *MicroRNA signatures in human cancers*. Nat Rev Cancer, 2006. **6**(11): p. 857-66.
37. Wong, T.S., et al., *Mature miR-184 as Potential Oncogenic microRNA of Squamous Cell Carcinoma of Tongue*. Clin Cancer Res, 2008. **14**(9): p. 2588-92.
38. Chang, S.S., et al., *MicroRNA alterations in head and neck squamous cell carcinoma*. Int J Cancer, 2008. **123**(12): p. 2791-7.
39. Chen, Y. and R.L. Stallings, *Differential patterns of microRNA expression in neuroblastoma are correlated with prognosis, differentiation, and apoptosis*. Cancer Res, 2007. **67**(3): p. 976-83.
40. Yanaihara, N., et al., *Unique microRNA molecular profiles in lung cancer diagnosis and prognosis*. Cancer Cell, 2006. **9**(3): p. 189-98.
41. Calin, G.A. and C.M. Croce, *Investigation of microRNA alterations in leukemias and lymphomas*. Methods Enzymol, 2007. **427**: p. 193-213.

42. Doleshal, M., et al., *Evaluation and validation of total RNA extraction methods for microRNA expression analyses in formalin-fixed, paraffin-embedded tissues*. J Mol Diagn, 2008. **10**(3): p. 203-11.
43. Hoefig, K.P., et al., *Unlocking pathology archives for microRNA-profiling*. Anticancer Res, 2008. **28**(1A): p. 119-23.
44. Schmittgen, T.D., et al., *A high-throughput method to monitor the expression of microRNA precursors*. Nucleic Acids Res, 2004. **32**(4): p. e43.
45. Sadikovic, B., et al., *In vitro analysis of integrated global high-resolution DNA methylation profiling with genomic imbalance and gene expression in osteosarcoma*. PLoS ONE, 2008. **3**(7): p. e2834.

## **5. Manuscrito 2: Análise do número de cópias do DNA using aCGH**

A aplicação da tecnologia de CGH em *arrays* para a identificação de alterações no número de cópias do DNA contribuiu para a caracterização de genes e vias genéticas associadas com a etiologia e progressão desta doença. No nosso estudo, realizamos uma análise genômica global utilizando Hibridização Genômica Comparativa (aCGH)c de amostras seqüenciais progressivas, de leucoplasias leve, moderada e severa, a CCEOs de mesmo-sítio, para identificar marcadores genéticos associados com a progressão tumoral. Os genes necessários para a transformação maligna estarão provavelmente alterados em leucoplasias progressivas e nos carcinomas correspondentes.

## Copy number profiling of sequential progressive leukoplakia and oral squamous cell carcinomas by aCGH analysis

Nilva K. Cervigne<sup>1,2</sup>, Jerry Machado<sup>3</sup>, Bekim Sadikovic<sup>4</sup>, Grace Bradley<sup>1,5</sup>, Rashmi S. Goswami<sup>3</sup>, Bayardo Perez-Ordonez<sup>6</sup>, Natalie Naranjo Galloni<sup>7</sup>, Ralph Gilbert<sup>8</sup>, Patrick Gullane<sup>8</sup>, Jonathan C. Irish<sup>8</sup>, Patricia P. Reis<sup>1,2</sup>, Suzanne Kamel-Reid<sup>1,3,6\*</sup>

<sup>1</sup>Division of Applied Molecular Oncology, Ontario Cancer Institute, the University Health Network, Toronto, ON Canada

<sup>2</sup>Department of Genetics, Bioscience Institute, Sao Paulo State University, Botucatu, SP, Brazil

<sup>3</sup>Department of Laboratory Medicine and Pathobiology, University of Toronto, ON, Canada

<sup>4</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

<sup>5</sup>Faculty of Dentistry, University of Toronto, Toronto, ON Canada

<sup>6</sup>Department of Pathology, Toronto General Hospital, Ontario Cancer Institute, the University Health Network, Toronto, Ontario, Canada

<sup>7</sup>Department of Otolaryngology, Hospital Calderon Guardia, San Jose, Costa Rica

<sup>8</sup>Department of Otolaryngology/Surgical Oncology, Princess Margaret Hospital, The University of Toronto and the University Health Network, Toronto, Ontario, Canada

\*Correspondence to:

Dr. Suzanne Kamel-Reid, PhD, FACMG

610 University Avenue, Rm. 9-622

Princess Margaret Hospital, Ontario Cancer Institute and University Health Network

Toronto, Ontario, M5G 2M9

CANADA

Tel. +1 (416) 340-4800 x 5739

Fax. +1 (416) 340-3596

E-mail: [Suzanne.Kamel-Reid@uhn.on.ca](mailto:Suzanne.Kamel-Reid@uhn.on.ca)



**KEYWORDS:** oral leukoplakia, oral carcinoma, biomarkers, progression.

**Running Title:** Copy number alterations and oral cancer progression

### **Abstract**

Oral squamous cell carcinomas (OSCCs) are the sixth leading cause of cancer death worldwide. A significant proportion of OSCCs (16-62%) arise from oral potential malignant lesions (OPML), such as leukoplakia. Since cancer progression is due to genetic damage over time, the detection of genetic changes in progressive OPMLs is important for the identification of lesions at risk for malignant transformation, thus improving patient outcome. In this study, aCGH was performed on 25 progressive samples of oral squamous cell carcinomas. DNA losses were observed in approximately 20% of all samples and were mainly seen on chromosomes 5q31.2 (35%), 16p13.2 (30%), 9q33.1, 9q33.2 9 (25%), and 17q11.2, 3p26.2, 18q21.1, 4q34.1 and 8p23.2 (20%). DNA copy number increases were identified on chromosome 1p in 20/25 cases (80%) with high-level amplification detected at 1p35 and 1p36. Other regions of gain were observed on 11q13.4 (68%), 9q34.13 (64%), 21q22.3 (60%), 6p21 and 6q25 (56%), and 10q24, 19q13.2, 22q12, 5q31.2, 7p13, 10q24, and 14q22 (48%). These copy number alterations (CNAs) were detected in all grades of dysplasia that progressed, as well as their corresponding OSCCs, in 70% of patients analyzed in our study, suggesting they may represent CNAs associated with progression of oral leukoplakia to OSCC. Using strict criteria, we annotated and identified 16 altered genes within these regions; 14/16 genes, including *KHDRBS1*, *CAMSAP1L1*, *PARP1*, *RAB1A*, *HBEGF*, *PAIP2*, *TBRG4*, *NPM3*, *CHRD12*, *IRF9*, *REC8*, *GMPK2*, *BTBD7*, *FBXO7* were amplified, and *CSMD1* and *MYO5B* were deleted in progressive leukoplakias and OSCCs, compared to normal. In summary, unlike other genome-wide studies in oral cancer progression

that show a consistent pattern of changes on selected chromosome arms, but do not identify specific genes involved in OSCC progression. Our study highlighted potentially useful genes on chromosomes 1p, 2p, 5q, and 14q that might serve as biomarkers of oral cancer progression. Future studies will combine these results with functional analyses, to determine how these changes are associated with genetic progression of OSCC.

## **Introduction**

Oral squamous cell carcinomas (OSCCs) are the sixth leading cause of cancer death worldwide[1-2], with 28,260 expected new cases and 7,230 deaths every year in the United States[3]. Patients with OSCC have benefited from the latest advances in surgical techniques, and treatment by radiation and chemotherapy, which may enhance quality of life and improve survival. However, despite these advances, the 5-year survival rate of OSCC patients remains at approximately 50%[4-6]. These low survival rates are mainly due to the presence of late-stage disease at the time of diagnosis and to disease recurrence. In order to improve patient outcome, better methods of oral cancer detection and a better understanding of the genetics events associated with disease progression are important.

Since malignant transformation is due to genetic damage over time[7], the identification of genetic changes in sequential progressive lesions may be useful for predicting lesions at risk for malignant transformation. It is known that a significant proportion of OSCCs (16-62%, [8-9]) arise from oral potential malignant lesions (OPML), such as leukoplakia. Oral leukoplakia is a lesion that presents as a “white patch” in the oral mucosa[4]. Currently, these lesions are classified based on clinical and histopathological assessment; clinically, leukoplakia are homogeneous or non-homogeneous, the latter has a higher risk of transformation. Histologically, they are classified as non-dysplastic or dysplastic[10], and the presence of epithelial dysplasia is associated with an increased risk for transformation of up to 31%[11]. However, clinical and histological characteristics have limited prognostic value for predicting which leukoplakia will progress to malignancy.

The search for genetic biomarkers is an approach for potentially identifying which leukoplakias have an increased risk of malignant transformation. Previous studies identified large

chromosomal regions and LOH events associated with progression in dysplasias and OSCCs from different patients. As these studies identify fairly large genomic regions, specific genes involved in progression remain unknown.

High resolution global genomic profiling analysis allowed us to identify copy number gains and losses, and to narrow down regions containing genes that are likely to be involved in progression of leukoplakias. Genes identified herein have the potential to be used as diagnostic markers, for prediction of which leukoplakia may have a higher risk of progression. Such biomarkers can be used to initiate early intervention, ultimately improving patient survival.

## **Material and Methods**

### Patient samples

#### *Training Sample Set*

We collected 20 progressive leukoplakias (sequential samples) and 5 OSCCs (N=25) from 5 patients. All samples were formalin fixed paraffin-embedded (FFPE) archival tissues. Of the 20 leukoplakias, 4 were non-dysplastic and 16 were dysplastic (mild, moderate or severe). All carcinomas had at least one corresponding premalignant leukoplakia. An additional 5 non-progressive leukoplakia samples, from 5 different patients, were included in this analysis. The characteristics of the training sample set are described in Table 5.1.

### DNA Isolation from FFPE samples

All samples underwent histopathological analysis by an oral pathologist (GB) to ensure the presence of dysplasia or carcinoma in at least 80% of each tissue section. Samples were needle micro-dissected, according to standard protocols, to select the target cell population for DNA extraction and genomic analysis. In short, genomic DNA was isolated from 5-10, 10  $\mu$ m-thick

FFPE tissue sections. After xylene deparaffinization, tissues were incubated in Cell Lysis Solution buffer (5 PRIME, Gaithersburg MD, USA) and Proteinase K solution (20mg/ml) for 2 days at 56°C (fresh aliquots of proteinase K were added at 17 and 24 hours). Genomic DNA was isolated and purified using the ArchivePure DNA Cell/Tissue Kit-4g (5 PRIME, Mat#2900269, Gaithersburg MD, USA), with the final elution of DNA into water. All DNA samples were quantified using a NanoDrop Spectrophotometer, and checked by agarose gel electrophoresis for quality. All samples yielded DNA of sufficient quality for analysis. All 30 samples (training set) were subjected to whole genomic amplification (WGA) using a sequenase-based approach (modified from the Affymetrix Chromatin Immunoprecipitation Assay, as per Sadikovic et al[12]) in order to yield enough DNA quantity for aCGH analysis (~2ug). High quality normal male genomic DNA (Promega, P/Ns G1471) was used as the reference sample; this reference DNA is widely used in other aCGH studies[13-15]. Male genomic DNA (Promega) was heat fragmented for 10 minutes at 95°C, subsequently subjected to WGA, and hybridized against each test sample.

#### Sequenase-based WGA

Two rounds of WGA were used to randomly amplify 30 FFPE DNA samples (minimum of 10 and maximum of 200ng of DNA). This amplification protocol was successfully used by others to amplify less than 10ng of DNA and was successfully used in the comparison of relative enrichment between two samples[16]. The protocol consists of two sets of enzymatic reactions (Table 5.2); in Round I, sequenase enzyme is used to extend randomly annealed primers (Primer A) and to generate templates for subsequent PCR. During Round II, the specific primer B (sequence is partially the same as Primer A), was used to amplify the templates previously generated by dNTPs (10mM) incorporation. Following each amplification round, the DNA was

purified using the QIAquick® PCR Purification Kit (Qiagen), according to the manufacturer's protocol. The final purified PCR product was eluted into 50 µl of Sigma water, and five µL of product was used to run a 1% agarose gel, to verify the presence of a 200bp –1kb DNA “smear” for successfully amplified samples.

To verify the fidelity of the WGA, we first sought to determine the correlation between amplified and unamplified template DNA from FFPE samples, by analyzing matched pairs of samples (29T and 201T). In all experiments, the WGA protocol was used for both test (tumor) and reference (Promega DNA) samples. A summary of the correlation data is given in Supplemental table 5.1. All WGA samples displayed adequate signal-to-background ratios. CGH profiles of paired samples did not display any chromosomal gains or losses due to WGA. Overall, we obtained representative data when comparing amplified vs. unamplified FFPE samples, resulting in a Person's correlation coefficient ranging between  $R^2=0.80-0.97$  for sample for the majority of probes referring to both samples 29T and 201T, which is fairly consistent with the limits of experimental variation. These values were reflected in the mean absolute deviations of the log<sub>2</sub> ratios, calculated for all probes across the genome on the array.

#### Whole Genome Tiling Array-CGH

We used the NimbleGen 385K whole genome tiling v2.0 array, which contains over 385,000 oligonucleotides probes (60-mer with a median probe spacing of ~7kb) providing genome wide coverage. Array-CGH experiments including quality control, DNA labeling, hybridization, scanning, and data extraction were performed by NimbleGen Systems core facility (Reykjavik, Iceland). The complete experimental protocol is provided at NimbleGen Arrays User's guide ([https://projects.cgb.indiana.edu/download/attachments/5363/NimbleGen\\_CGH\\_Users\\_Guide\\_v3p1.pdf?version=2](https://projects.cgb.indiana.edu/download/attachments/5363/NimbleGen_CGH_Users_Guide_v3p1.pdf?version=2)). Briefly, 1µg of genomic DNA was used for dual color labeling (inverse

Cy3/Cy5). All 30 samples were successfully labeled, meeting the quality control criteria. Following hybridization, washing and scanning were performed according to the the manufacturer's protocol (NimbleGen-Roche). Array CGH data generation was performed using commercially available software (SignalMap version 1.8, Nimblegen).

#### aCGH Copy number Data analysis

Partek Genomic Suite (PGS) software was used to identify copy number alterations (CNAs). First, the *pair* data files were inputted in the PGS software, which automatically loaded log<sub>2</sub> ratio intensities for all probes across the tiling array. We performed unsupervised hierarchical clustering analysis based on the Euclidean, average linkage, agglomerative method (PGS), blinded to the identity of the samples. We first sought to identify CNAs associated with oral cancer progression. For this, we performed copy number analysis across all samples, which were categorized into two groups: progressive leukoplakias with corresponding OSCCs (n=25), and non-progressive leukoplakias (n=5). We determined CNAs present in progressive leukoplakia and corresponding OSCCs, and absent in non-progressive leukoplakia. We then compared these data against the copy number variation (CNV) frequency data available from the general control population (2,115 predominantly European background individuals; half from Ontario[17], and half from Germany[18]). This analysis allowed filtering out any CNVs that were present in the general population which are not relevant to disease biology/tumorigenesis. Additionally, in order to map the genetic alterations occurring during progression, we assessed the CNAs within the progressive samples from each patient.

CNAs were detected using the genomic segmentation algorithm in PGS. Genomic aberrations were assessed with a segmentation stringency of 10 consecutive genomic markers utilizing  $p < 0.001$  as cut off, and signal to noise ratio cut off 0.3 for amplifications and deletions. We used

a copy number cut-off of 2 copies to identify gains and losses; ratios  $<0.85$  were considered regions of loss, whereas  $>1.15$  represented regions of gain. This analysis excluded genes mapped on sex chromosomes, and regions with no known genes.

## Results

Currently, the most widely used technique to study copy number alterations is array comparative genomic hybridization (aCGH). This technique has also been applied to DNA extracted from archival formalin-fixed paraffin-embedded (FFPE) clinical specimens, such as lung and prostate cancer, to elucidate key genes involved in disease development and progression[15, 19]. We have successfully applied a WGA protocol for amplification of low yield DNA from FFPE oral samples, to accurately assess DNA copy number gains and losses. Array CGH using amplified FFPE samples allowed the identification of global copy number gains and losses, with similar results when compared to DNA from unamplified FFPE samples; data were thus consistent within the limits of experimental variation ( $R^2 = 0.804$ ).

Copy number alterations were analyzed blinded to the histology of the samples, and the unsupervised Euclidean hierarchical clustering analysis (PGS) showed that the majority of progressive leukoplakias (16/20) and OSCCs clustered together, and separately from normal and non-progressive leukoplakias (Figure 5.1), indicating that they share common CNAs.

The genomic segmentation algorithm used to detect amplifications and deletions showed a total of 8,409 change calls in the group of progressive leukoplakia and OSCCs, and 2,170 change calls in non-progressive samples. These results were then filtered for CNAs only found in both groups that were very rare or absent among the CNVs found in the general population. This analysis showed that out of the 8,409 change calls, 4,081 (48.5%) were unique to the group of progressive



leukoplakia and same-site OSCCs; and that 1,146 out of 2,170 were CNAs present in the group of non-progressive leukoplakia. These 1,146 change calls present in non-progressive samples were then subtracted from our data containing 4,081 CNAs. Such an approach was used to finally determine the genetic changes that may be involved in oral cancer progression, since we selected CNAs specific present in progressive leukoplakia and corresponding OSCCs, and absent in non-progressive leukoplakias. This analysis showed that a total of 2,935 CNA calls were present in all progressive leukoplakias and OSCCs, but not in non-progressive leukoplakias. A larger number of gains were commonly found in progressive leukoplakia and OSCC (80%), in contrast to a small number of losses (20%) (Figure 5.2). DNA losses were represented mainly on chromosomes 5q31.2 (35%), 16p13.2 (30%), 9q33.1-9q33.2 (25%), and 17q11.2, 3p26.2, 18q21.1, 4q34.1 and 8p23.2 (20%) (Figure 5.3). DNA copy number gains were identified on chromosome 1p in 20/25 cases (80%) with high-level amplifications at 1p35 and 1p36. Over-representations were found at 11q13.4 (68%), 9q34.13 (64%), 21q22.3 (60%), 6p21 and 6q25 (56%), and 10q24, 19q13.2, 22q12, 5q31.2, 7p13, and 14q22 (48%) (Figure 5.3).

We observed an average of 113, 61, 153, and 178 significant change calls ( $p < 0.001$ ) in sequential progressive samples of mild ( $n=4$ ), moderate ( $n=3$ ), severe ( $n=6$ ) leukoplakia and OSCCs ( $n=6$ ), across all patients. Figure 4 shows a representative example of CNAs found in the sequential progressive samples of patient 4. Note that samples containing foci of cells with two different grades (samples 4d and 10d), were included in the group of higher grade of dysplasia. Also, as severe dysplasias and carcinoma *in situ* (CIS) have very similar histology and do not differ in biologically, they were considered as a unique group (severe dysplasias) in our analysis. We finally detected a total of 696 different chromosomal regions commonly altered in progressive leukoplakia and OSCC samples; 144/696 were regions of loss, and 552/696 were

regions of gain. In order to find the alterations involved in the progression of OPML to invasive carcinoma, we focused our analysis on CNAs from these 696 regions that were present in at least one OSCC and one preceding OPML from the same patient. This analysis revealed 193 different regions of gains and 15 of loss. Interestingly, 38/193 and 5/15 regions were commonly changed in all sequential samples (OSCC and preceding leukoplakias) from at least one patient. Even after all the filtering analysis, we still observed CNA losses mapped to 3p26.2, 8p23.2, 9q33.1-9q33.2, 17q11.2, and 18q21, and gains to 1p35-36, 1q32, 2p14, 5q31, 6p21, 6q25, 7p13, 10q24, 11q13.4, 12p13, 14q22, 19q13, and 22q12.3.

These CNAs were detected in low to high grade dysplasias, and their corresponding OSCCs, for the majority (70%) of patients analyzed in our study. Since these regions may contain genes that are relevant for the process of neoplastic transformation of leukoplakia to OSCC, we annotated the 263 genes (255 amplified and 8 deleted) detected within the identified regions. We then determined their potential relevance and involvement in cancer biology using two freely available databases, UCSC Genome browser (<http://genome.ucsc.edu/>) and NCBI (<http://www.ncbi.nlm.nih.gov/>). From this analysis we selected a list of 79 top candidate genes (Supplemental Table 5.2), which were then subjected to further evaluation using the ONCOMINE v.4 cancer profiling database (Research edition), which is a cancer microarray database and web-based data-mining platform aimed at facilitating discovery from genome-wide expression analyses[20]. This analysis verified whether deregulated mRNA expression of these candidate genes had been previously reported in head and neck cancer studies. Due to sample availability, we selected 16 out of the 79 genes (Table 5.3) for validation of the identified genes and their corresponding CNAs. Validation analysis will be performed in a separate cohort of

progressive dysplasias and their corresponding OSCCs, and non-progressive leukoplakias, using quantitative real-time PCR (see Chapter 6 of this Thesis).

## **Discussion**

The application of high-throughput molecular techniques to the study of human disease has contributed immensely to the identification of genes and pathways associated with disease etiology and progression. In particular, the genomic analysis of tumor DNA has identified alterations in sequence and copy number associated with diagnosis, prognosis, and treatment response in a variety of cancer types[21-23]. Amplification or deletion of distinct chromosomal regions can lead to deregulated gene expression, thus conferring a growth advantage to malignant cells[24]. Amplified or deleted genes are, therefore, important targets for therapeutic invention, and identification of such copy number alterations can help elucidate potential mechanisms involved in tumour development and progression.

Ours is the first study to examine CNAs in sequential oral lesions from the same patient, with the aim of identifying copy number changes associated with malignant transformation, as the tissue progresses from benign to invasive carcinoma. Non-progressive samples were also included to ensure that CNAs detected in progressive leukoplakia were not present in non-progressive lesions, ensuring that such changes would likely be involved in malignant transformation.

When we compared CNA profiles of histologically different tissues using unsupervised hierarchical clustering analysis, we were able to show that normal/non-progressive leukoplakia segregated away from progressive leukoplakia/invasive OSCCs. Overall, the number of copy number changes increased from the lowest to the highest grade of dysplasia, culminating with largest changes in invasive OSCCs. Our data demonstrate that leukoplakia lesions that progress

already possess many of the genetic alterations present in invasive cancers. This is consistent with the hypothesis that the majority of genetic alterations may occur early during head and neck cancer progression[25].

In our study, we identified deletions in about 20% of cases, with deletions on chromosomes 3p, 9q, and 18q being preferentially found in progressive leukoplakias and same-site OSCC. This is in agreement with other studies showing association of allelic loss of 9p and 3p with HNSCC progression[26-27]. 18q loss was also previously reported and associated with patient poor prognosis and metastasis in HNSCC[28-29].

In this study, the over-representations found at 1p35-36, 11q13, 19q, and 22q12, in low to high grade sequential progressive dysplasias and OSCCs, agrees with a previous study showing that these regions were correlated with tumor progression in HNSCC[26]. In particular, 11q13 gains or amplification were associated with poor prognosis of HNSCC[30]. Interestingly, a recent study implicated gains at 11q with high risk for esophageal squamous cell carcinoma development in a population in China[31]. This study showed that one of the most gene-rich CNA regions with gains was 11q13.1-13.4, with gains significantly correlated with increased RNA expression in over 80% of these genes. Additionally, gains at 11q13 were described as associated with poor prognosis for other cancer types such as prostate[32] and thyroid[33], associated with larger tumor sizes in hepatocellular carcinomas[34], and predictive markers of systemic recurrence in breast cancer[35].

Using strict criteria, we identified 16 altered genes within the regions containing CNAs associated with progression of oral leukoplakia to OSCC. We showed that 14/16 genes, including *KHDRBS1*, *CAMSAP1L1*, *PARP1*, *RAB1A*, *HBEGF*, *PAIP2*, *TBRG4*, *NPM3*, *CHRD12*, *IRF9*, *REC8*, *GMPK2*, *BTBD7*, *FBXO7* were amplified, and *CSMD1* and *MYO5B*

were deleted in progressive leukoplakias and OSCCs, as compared to normal tissues. The Gene Ontology (GO) categories for those genes includes mainly functions associated with protein binding, cell cycle, cell differentiation, cell proliferation, transcription factors, and cell surface receptors linked to signal transduction (CancerGenes website <http://cbio.mskcc.org/CancerGenes/Select.action>; [36]). The most representative GO category was that of the protein binding genes, which included *BTBD7*, *FBXO7*, *IRF9*, *PAIP2*, *PARP1*, *REC8*, and *MYO5B*. The other deleted gene, *CSMD1*, was included in membrane-membrane interactions group. *RAB1A* represented a protein transportation and transduction signaling, and *KHDRS1* represented, together with *PARP1* and *IRF9*, regulators of transcription factor categories. Additionally, *TBRG4* represented the apoptosis/protein kinase activity, *CHRDL2* represented the cell differentiation, and *HBEGF* represented the epidermal growth factor receptor/transduction signaling genes categories.

Genetic alterations in OPMLs and OSCCs may represent key changes in disease initiation and development. In particular, DNA amplification can occur at an earlier stage of cancer development, in which oncogenic pathways may be critically disrupted[37]. One of the most frequently (80%) amplified genes in our data set was *KHDRBS1* (also known as *p62*), a cell proliferation and cell surface receptor of signal transduction located at 1p35.1. A recent study found an interesting association between *Ras* and *p62*; they showed that the signaling adaptor *p62* is induced by *Ras*, its levels are increased in human tumors, and it is required for *Ras*-induced survival and cellular transformation[38]. Considering that the balance between cell death and survival are important aspects of oncogenic transformation, further investigations of *KHDRBS1-RAS* involvement in oral cancer progression may be of relevance to determine the outcome of oral tumorigenesis.

Interestingly, *RAB1A*, a member of Ras oncogene family which maps to 2p14, was also amplified in our data set. Its aberrant expression was previously identified in human tongue squamous cell cancer (TSCC)[39]. These authors also found a high prevalence of protein Rab1A-overexpression not only in TSCCs (98%) but also in premalignant lesions (93%), suggesting that the Rab1A may be a potential biomarker of tongue carcinogenesis.

Another gene of particular interest is *HBEGF*, located at 5q31.3. This gene encodes a protein that is an EGFR ligand and it is up-regulated in a wide variety of pathological conditions, including cancer[34, 40-42]. A recent study performed gene expression analysis of clinical samples and cells that preferentially infiltrate the brain isolated from patients with advanced breast cancer; they identified the epidermal growth factor receptor (*EGFR*) ligand *HBEGF*, together with other two genes, *COX2* and *ST6GALNAC5*, as mediators of cancer cell passage through the blood-brain barrier[40]. In the literature, *COX-2* is extensively expressed in oral cancer and oral premalignant lesions and seems to be enhanced in high-risk oral lesions[43]. It has also been reported that the over-expression of *EGFR* is associated with a poor prognosis of head and neck cancer patients, and its inhibition improves the survival of patients[44-45]. Also, *EGFR* was recently showed to be the most frequently amplified and highly expressed gene in both human and mouse oral tumors ( $p < 0.01$ )[46]. Of Note, a member of Ras family (*RAB1A*), which is one of the downstream effectors of *EGFR*, was also found to be amplified in our sample set. Interestingly, a FISH analysis performed by[37] showed that amplification of *EGFR* and *CCND1* can coexist in single cells of an oral dysplasia, suggesting the dependence on multiple oncogenes for OPML progression.

*CHRD2*, located at 11q13.4, is another gene that deserves special consideration. This gene encodes a putative extracellular matrix protein and it was shown to be over-expressed in breast,

lung and colon tumors, compared to the corresponding normal tissues[47]. The *REC8* gene, in 14q12, is part of cohesin gene family. The REC8 protein participates in the segregation of homologs at the first meiotic division[48], and it has been shown that the up-regulation of meiosis-specific genes, such as *REC8*, regulates the extent of arrested mitoses and polyploidy in lymphoma cell lines[49]. *REC8*, together with *PAIP2*, have been described as differently expressed in OSCCs[50]; these genes were, respectively, representative of two clusters enriched in respect to mitosis-related and translation GO terms with this study. Translation initiation is regulated in response to mitogenic stimulation, and thus associated with cell cycle progression and cell growth. Many alterations in translational control can occur in cancer, but further investigations need to be done in order to discover their role in cancer development and progression[51]. Changes in the expression of the components of the translational machinery can lead to global changes, such as an increase in protein expression and translational activation of the RNA (mRNAs and miRNAs) molecules involved in cell growth and proliferation.

Of note, other studies identified large chromosomal regions and the presence of LOH events in dysplasias and OSCCs, but from different patients[37, 52-53]. In a previous study by us [54], we have identified a miRNA signature associated with progression of oral leukoplakia to same-site carcinomas. We showed that increased expression levels of miR-21, miR-181b, and miR-345 associated with increased lesion grade severity during progression. It is known that changes in miRNA expression are complex and, likely transcriptional and post-transcriptional regulation, changes in the expression of miRNA biogenesis enzymes can be partially due to genomic loss or gain. In this study, we used another global genomic approach to investigate how DNA CNAs contribute to malignant transformation, in addition to the previously identified miRNA changes. Interestingly, miR-345, located at 14q32.2 and miR-181b, located at 1q32.1 map to regions of

DNA gain in our sample set, suggesting that over-expression of these two miRs may be driven by gains at these genomic loci. As we do not observe any CNA in the chromosomal region where miR-21 is mapped, its expression may be regulated by other transcriptional and/or post-transcriptional changes rather than CNA.

In summary, genome-wide studies in oral cancer progression show a consistent pattern of changes on selected chromosome arms, and are useful to identify specific genes involved in OSCC progression. Our study has also identified potential genomic markers of interest on chromosomes 1p, 2p, 5q, 8p, 11q, 14q, 18q, and 22q in oral cancer. Tumors showed a broad variety in the patterns of alterations, suggesting that alternative mechanisms of genomic instability may play a role in disease progression. Also, our findings show the potential utility of combining CNA and miRNA expression analysis to identify genes involved in oral carcinogenesis. Future studies should combine these results with functional analyses, and to further determine how these genetic changes drive progression to OSCC.



**Table 5.1: Training sample set of sequential oral leukoplakia and same-site OSCCs.**

| <b>Patient</b> | <b>Sample ID</b> | <b>Sample Site</b>   | <b>Histopathological Diagnosis</b>          | <b>Biopsy date (d/m/yr)</b> |
|----------------|------------------|----------------------|---|-----------------------------|
| <b>1</b>       | 1a               | Tonsil               | Focal keratosis <sup>1</sup>                | 02.10.2001                  |
|                | 1b               | Alveolus + FOM       | Mild squamous hyperplasia <sup>1</sup>      | 15.07.2003                  |
|                | 1c               | Anterior FOM         | Carcinoma <i>in situ</i>                    | 07.05.2004                  |
|                | 1d               | Anterior FOM         | Invasive moderately differentiated OSCC     | 07.05.2004                  |
| <b>4</b>       | 4a               | Tongue               | Severe squamous dysplasia                   | 21.11.1997                  |
|                | 4b               | Tongue               | Mild squamous dysplasia                     | 08.10.1998                  |
|                | 4c               | Tongue               | Severe squamous dysplasia                   | 03.11.1998                  |
|                | 4d               | Tongue               | Severe dysplasia + SCC micro foci           | 26.11.1998                  |
|                | 4e               | Tongue               | Invasive moderately differentiated OSCC     | 28.02.2000                  |
| <b>5</b>       | 5a               | Right lateral tongue | Keratosis <sup>1</sup>                      | 08.03.1993                  |
|                | 5b               | Right lateral tongue | Moderate dysplasia                          | 08.03.1993                  |
|                | 5c               | FOM                  | Severe dysplasia                            | 27.10.1994                  |
|                | 5d               | Buccal mucosa        | Mild dysplasia                              | 06.02.1997                  |
|                | 5e               | Left buccal mucosa   | Invasive OSCC                               | 28.10.1997                  |
| <b>10</b>      | 10a              | Left buccal mucosa   | Keratosis <sup>1</sup>                      | 28.03.1991                  |
|                | 10b              | Left buccal mucosa   | Moderate dysplasia                          | 15.09.1995                  |
|                | 10c              | Left buccal mucosa   | Mild dysplasia                              | 26.09.1996                  |
|                | 10d              | Left buccal mucosa   | Moderate + Severe dysplasia                 | 09.10.2003                  |
|                | 10e              | Left buccal mucosa   | Severe dysplasia + carcinoma <i>in situ</i> | 20.11.2003                  |
|                | 10f              | Left buccal mucosa   | Invasive moderately differentiated OSCC     | 25.06.2004                  |
| <b>15</b>      | 15a              | Right tongue         | Mild dysplasia                              | 07.04.1994                  |
|                | 15b              | Right tongue         | Moderate dysplasia                          | 07.04.1994                  |
|                | 15c              | Right tongue         | Keratosis mild atypia                       | 23.02.2001                  |
|                | 15d              | Right tongue         | Moderately differentiated OSCC              | 30.10.2003                  |
|                | 15e              | Right tongue         | Mild squamous hyperplasia <sup>1</sup>      | 08.03.2004                  |

<sup>1</sup>Non-dysplastic lesions; OSCC: oral squamous cell carcinoma; FOM: floor of mouth; WD: well differentiated, MD: moderately differentiated, PD: poorly differentiated.

**Table 5.2.** PCR conditions for the two round whole genomic DNA amplification (modified from Affymetrix Chromatin Immunoprecipitation Assay, as per Sadikovic et al.[12]).

| <b>Round I</b>            |                        |   |
|---------------------------|------------------------|---|
| <b>Components</b>         | <b>Volume for 1Rxn</b> | <b>Cycling conditions</b>   |
| DNA (10-200ng)            | 10 $\mu$ l             | 1. Heat 2 min at 94°C   |
| 5x Sequenase Buffer       | 4 $\mu$ l              | 2. Rapid cool to 10°C and hold 5 min at 10°C  |
| Primer A(200 $\mu$ M)     | 4 $\mu$ l              | 3. Add reaction mixture* to sample  |
| Total volume              | 18 $\mu$ l             | 4. Ramp from 10°C to 37°C over 8 min  |
|                           |                        | 5. Hold at 37°C for 8 min; rapid ramp to 94°C and hold for 2 min.                                     |
|                           |                        | 6. Rapid ramp to 10°C and hold for 5 min at 10°C while adding 1.2 $\mu$ l of diluted Sequenase (1:10) |
|                           |                        | 7. Go back to step 4 (3 times)  |
| <b>Round II</b>           |                        |   |
| Purified Round I template | 8 $\mu$ l              | 1. 2 min at 94°C  |
| MgCl <sub>2</sub> (50mM)  | 2.5 $\mu$ l            | 2. 1 min at 94°C  |
| 10x PCR Buffer            | 5 $\mu$ l              | 3. 2 min at 52°C  |
| dNTP (10mM)               | 2 $\mu$ l              | 4. 3 min 72°C   |
| Primer B (100 $\mu$ M)    | 2 $\mu$ l              | 5. Go to 2, 4 times   |
| Taq Pol (5U/ $\mu$ l)     | 1 $\mu$ l              | 6. 30 sec 94°C  |
| ddH <sub>2</sub> O        | 29.5 $\mu$ l           | 7. 2 min at 52°C  |
|                           |                        | 8. 1,5min at 72°C   |
|                           |                        | 9. Go to 6, 24 times  |
|                           |                        | 10. 5 min at 72°C   |
|                           |                        | 11. $\infty$ at 4°C   |

30 cycles

\*Reaction Mixture: diluted sequenase (1/10 from 13U/ $\mu$ l stock), 0.5  $\mu$ l 25mM dNTPs, 1  $\mu$ l 0.1M DTT and 0.1  $\mu$ l 20mg/ml BSA (total volume = 2.6  $\mu$ l).

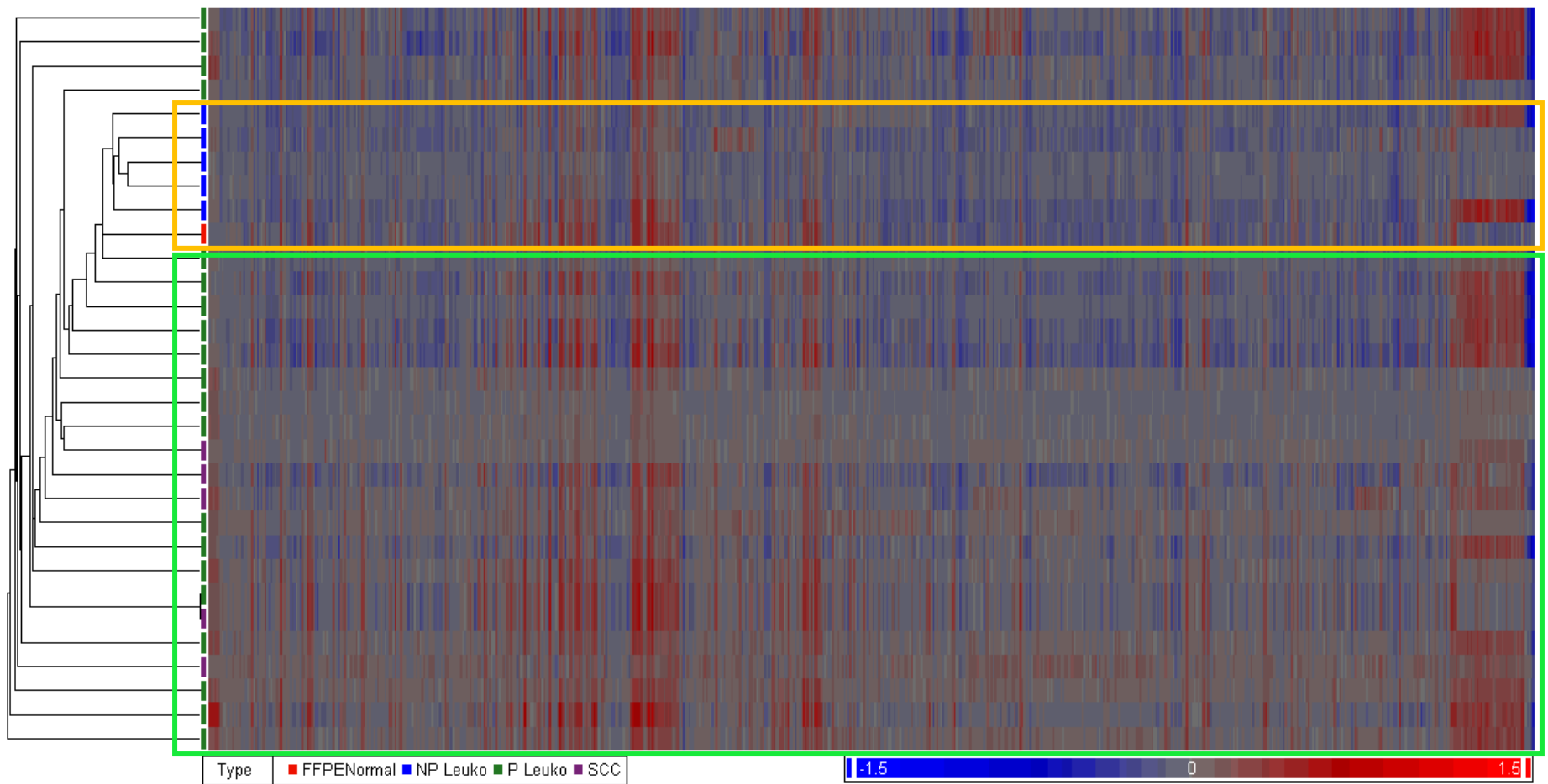
**Table 5.3.** Genes selected for Validation by quantitative Real-time PCR

| CYTOBAND       | RefSeq_genes     | GENE NAME  | FUNCTION  | INVOLVEMENT IN CANCER   |
|----------------|------------------|--|---|---|
| <b>GAINS</b>   | <b>AMPLIFIED</b> |  |   |   |
| <b>1p35.1</b>  | KHDRBS1          | KH domain containing, RNA binding, signal transduction associated 1(p62,Sam68) | DNA, RNA, and protein binding, SH3/SH2 adaptor activity, transcription repressor activity   | YES: Duran A et al 2008 Cancer Cell (The signaling adaptor p62 is an important NF-kappaB mediator in tumorigenesis). Lazer G et al 2007 Cell Signal.  |
| <b>1q32.1</b>  | CAMSAP1L1        | calmodulin regulated spectrin-associated protein (KIAA1078)                    | Unknown   | YES:KIF14 (same region) is a candidate oncogene in the 1q minimal region of genomic gain in multiple cancers. Corson TW et al 2005 Oncogene.  |
| <b>1q42.12</b> | PARP1            | poly (ADP-ribose) polymerase family, member 1                                  | Involved in the base excision repair (BER) pathway, by catalyzing the poly(ADP-ribosyl)ation of a limited number of acceptor proteins involved in chromatin architecture and in DNA metabolism. This modification follows DNA damages and appears as an obligatory step in a detection/signaling pathway leading to the reparation of DNA strand breaks | YES:Zaremba T et al 2009.   |
| <b>2p14</b>    | RAB1A            | RAB1A, member RAS oncogene family isoform 1                                    | Probably required for transit of protein from the ER through Golgi compartment. Binds GTP and GDP and possesses intrinsic GTPase activity   | YES: Tongue; Shimada K et al 2005 BJ Cancer (found a high prevalence of Rab1A-overexpression not only in TSCCs (98%) but also in premalignant lesions (93%).  |
| <b>5q31.3</b>  | HBEGF            | heparin-binding EGF-like growth factor   | epidermal growth factor receptor binding.growth factor activity (EGFR ligand)   | YES: Chechlinska M et al 2009;Bos PD et al 2009 Nature.   |
| <b>5q31.2</b>  | PAIP2            | poly(A) binding protein interacting protein 2                                  | translation repressor activity (interacts with VEGF mRNA - prognostic value in H&N SCC )  | YES: Onesto C et al 2006 (H&N cancer; Vascular endothelial growth factor-A and Poly(A) binding protein-interacting protein 2 expression in human head and neck carcinomas: correlation and prognostic significance). Onesto C et al 2004. |

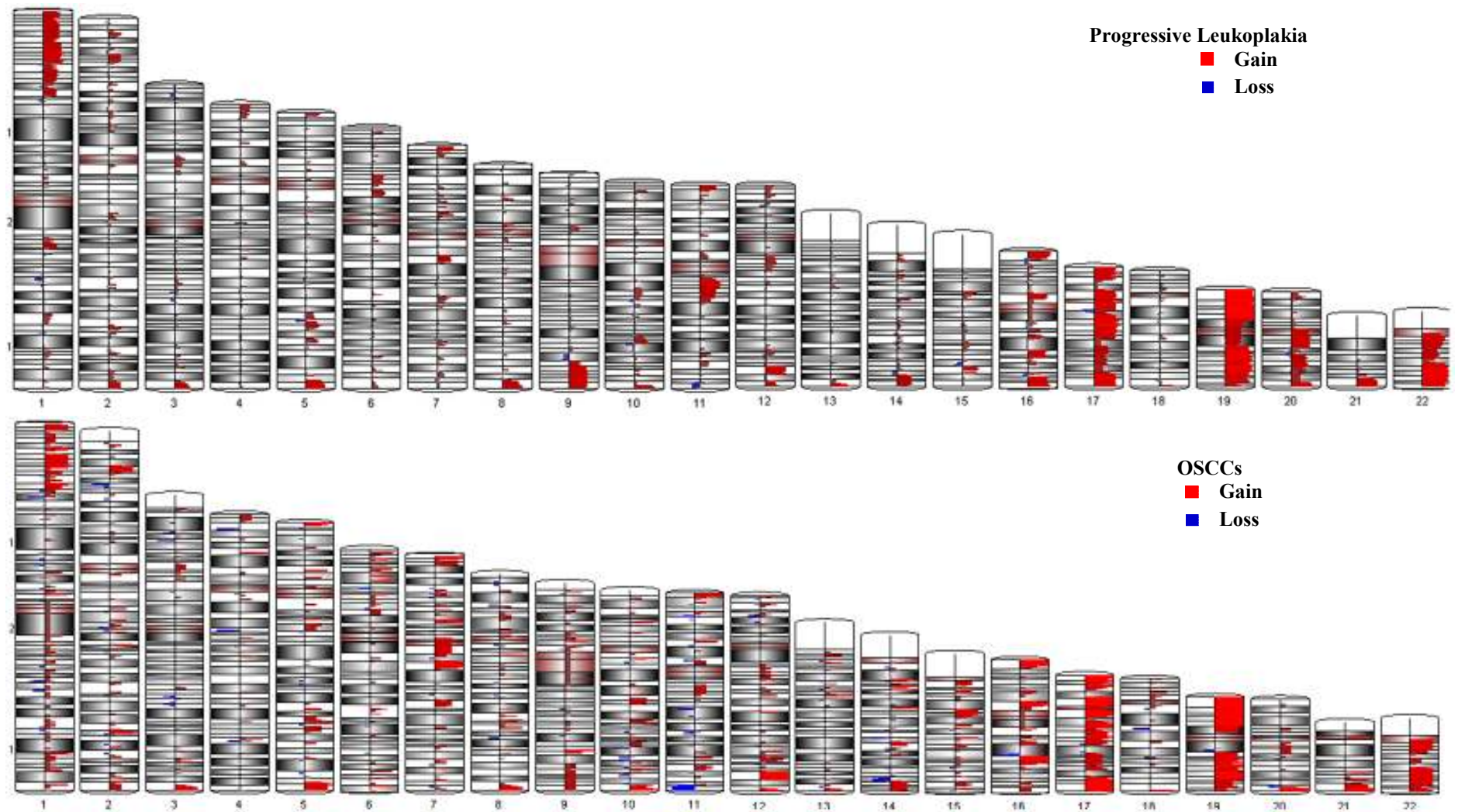
|                 |       |   |   |  |
|-----------------|-------|---|---|--|
| <b>7p13</b>     | TBRG4 | transforming growth factor beta regulator 4 | protein kinase activity   | No reports   |
| <b>10q24.32</b> | NPM3  | nucleophosmin/nucleoplasmin 3               | nucleic acid binding, protein binding, rRNA processing and transcription          | Two genetic pathways, t(1;10) and amplification of 3p11-12, in myxoinflammatory fibroblastic sarcoma, haemosiderotic fibrolipomatous tumour, and morphologically similar lesions - J. Pathol. 2009 Apr;217(5):716-27.  |
| <b>11q13.4</b>  | CHRD2 | chordin-like 2, breast tumor novel factor 1 | cell differentiation, cartilage development, multicellular organismal development | BNF-1, a novel gene encoding a putative extracellular matrix protein, is overexpressed in tumor tissues. Gene 2003 Jun 5;311:105-10.   |
| <b>14q12</b>    | IRF9  | interferon regulatory factor 9              | protein binding, transcription factor activity                                    | Expression of protein mediators of type I interferon signaling in human squamous cell carcinoma of the skin. Cancer Epidemiol. Biomarkers Prevention 2000 Sep;9(9):993-7. PTF1alpha/p48 transcription factor couples proliferation and differentiation in the exocrine pancreas [corrected]. Gastroenterology. 2004 Sep;127(3):937-49. |
| <b>14q12</b>    | REC8  | REC8 homolog (yeast)                        | protein binding, male meiosis I, spermatogenesis                                  | Upregulation of meiosis-specific genes in lymphoma cell lines following genotoxic insult and induction of mitotic catastrophe. BMC Cancer 2006 Jan 9;6:6.  |
| <b>14q12</b>    | GMPR2 | guanosine monophosphate reductase 2         | GMP reductase activity, metal ion binding, oxidoreductase activity                | Cloning and functional characterization of GMPR2, a novel human guanosine monophosphate reductase, which promotes the monocytic differentiation of HL-60 leukemia cells.   |
| <b>14q32.12</b> | BTBD7 | BTB (POZ) domain containing 7               | protein binding   | FUP1, a gene associated with hepatocellular carcinoma, stimulates NIH3T3 cell proliferation and tumor formation in nude mice. Biochem Biophys Res Commun. 2001 Sep 7;286(5):1033-8   |

|                |                |                                  |  |  |
|----------------|----------------|----------------------------------|--|--|
| <b>22q12.3</b> | FBXO7          | F-box protein 7                  | Protein binding, ubiquitin-protein ligase activity. This gene encodes a member of the F-box protein family which is characterized by an approximately 40 amino acid motif, the F-box. The F-box proteins constitute one of the four subunits of the ubiquitin protein ligase complex called SCFs (SKP1-cullin-F-box), which function in phosphorylation-dependent ubiquitination. The F-box proteins are divided into 3 classes: Fbws containing WD-40 domains, Fbls containing leucine-rich repeats, and Fbxs containing either different protein-protein interaction modules or no recognizable motifs. The protein encoded by this gene belongs to the Fbxs class and it may play a role in regulation of hematopoiesis | Fbxo7 gets proactive with cyclin D/cdk6. Cell Cycle. 2006 Feb;5(3):279-82. Epub 2006 Feb 8 |
| <b>LOSSES</b>  | <b>DELETED</b> |                                  |  |  |
| <b>8p23.2</b>  | CSMD1          | CUB and Sushi multiple domains 1 | protein binding  | YES: Beuten J et al 2009; Sheffer M et al 2009; Kuo KT et al 2009.                         |
| <b>18q21.1</b> | MYO5B          | MYO5B                            | ATP, actin, calmodulin, nucleotide binding   | YES:Kuang SQ et al 2008.   |

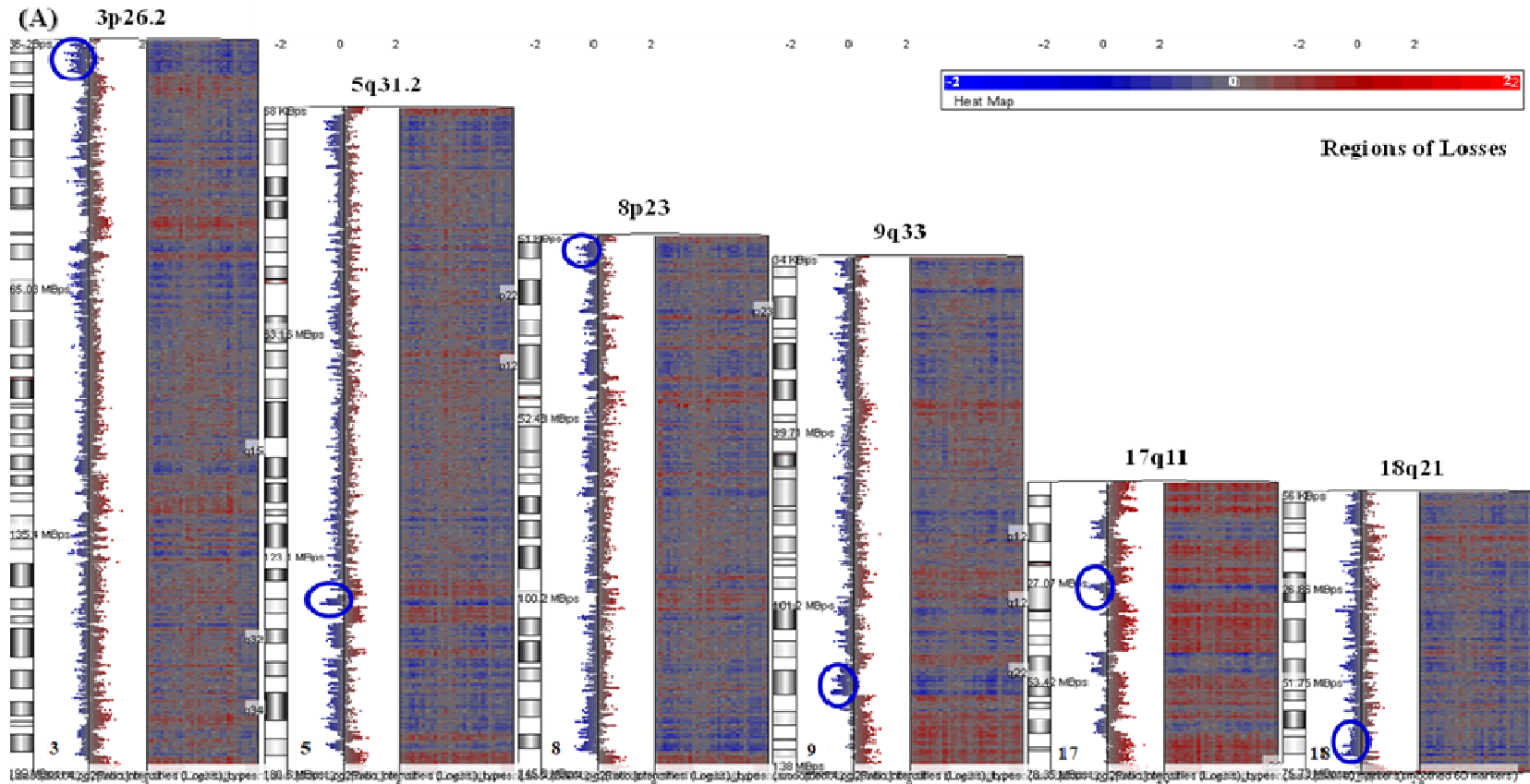
**Figure 5.1.** The panel shows the Hierarchical clustering of CNAs generated by Partek Genome Suite software. Two main distinct clusters were observed: normal/non-progressive leukoplakias (yellow box) and progressive leukoplakias/OSCCs (green box), suggesting similarity between malignant and progressive lesions (purple and green, respectively), and showing how normal and non-progressive leukoplakias (red and blue, respectively) are grouped separately.



**Figure 5.2.** Panel showing similar copy number changes profile for progressive leukoplakias (upper panel) and same-site OSCCs (bottom panel). Regions of gain (in red) over-represented the CNAs common between progressive leukoplakia and OSCC samples, compared to the regions of loss (in blue).

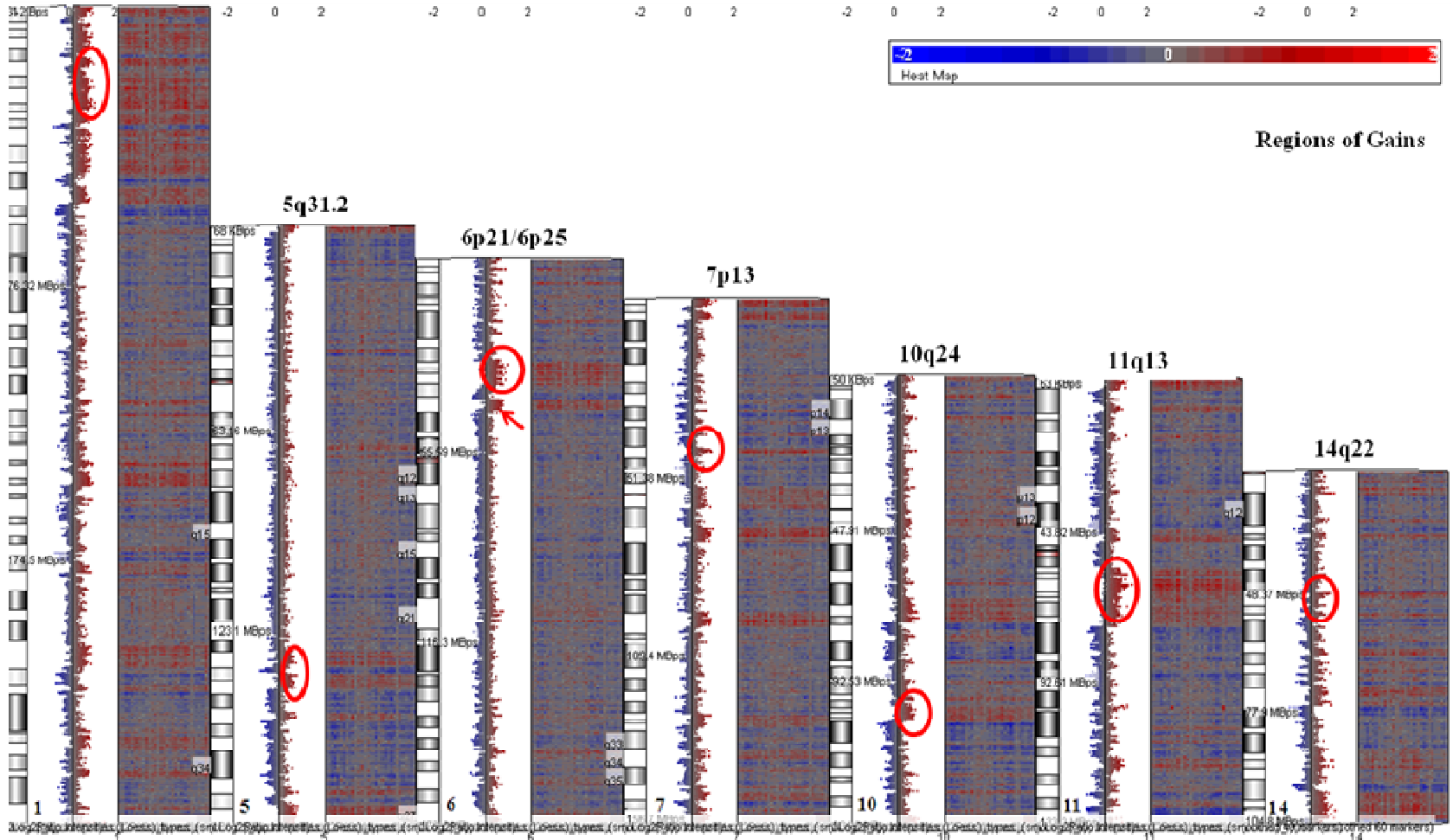


**Figure 5.3. (A)** Chromosomal view segmentation for chromosomes 3, 5, 8, 9, 17, and 18 with significant regions of loss (blue circles) common between progressive leukoplakia and OSCCs. **(B)** Panel showing chromosomal view segmentation for chromosomes 1, 5, 6, 7, 10, 11, and 14 with significant regions of gains (red circles) across all samples. Heat Map indicates CNA < 2 copies in blue for loss, and > 2 copies in red for gains.

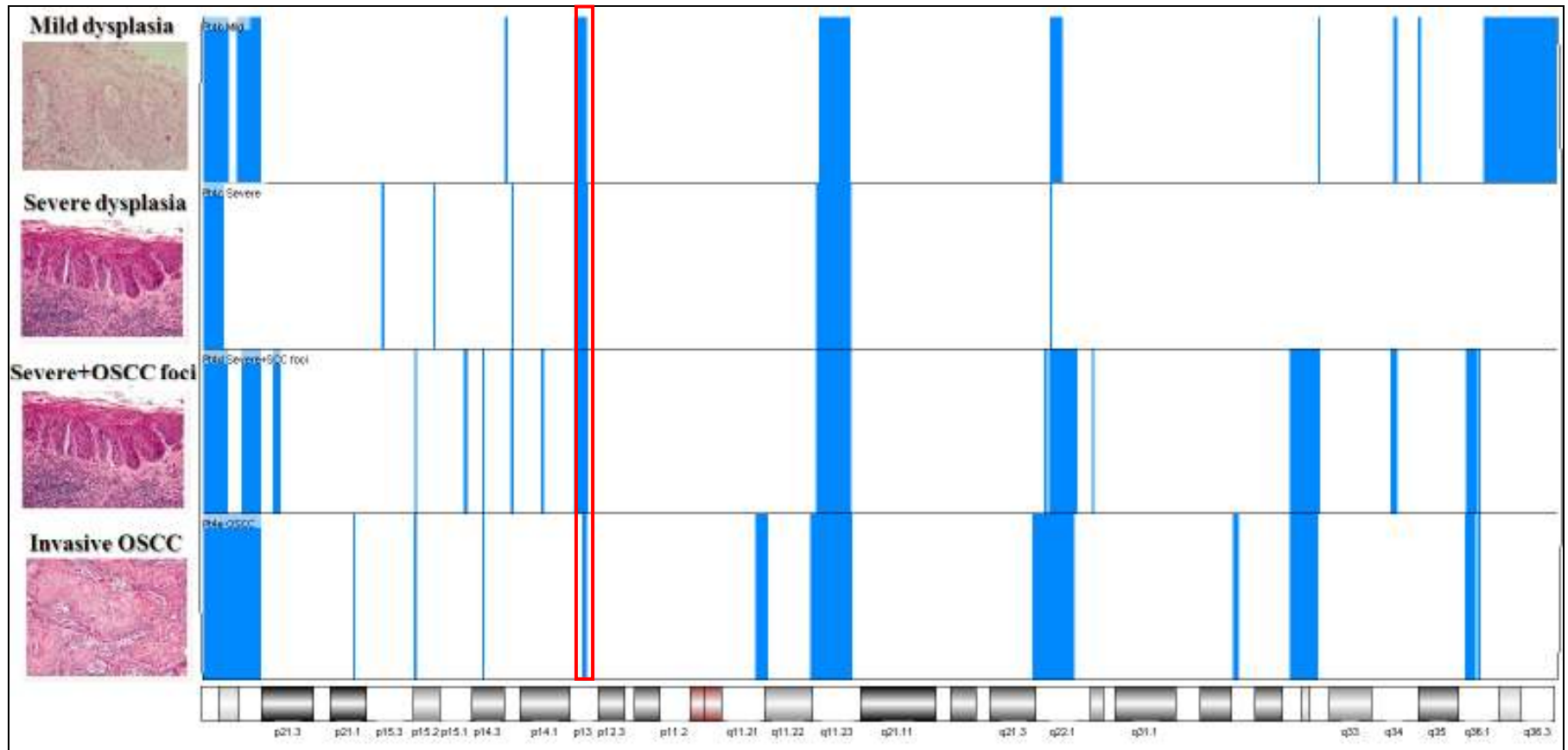




**(B) 1p35-36**



**Figure 5.4.** Showing a representative example of CNA (gain at 7p13; red box) found in sequential progressive samples of patient 4. Overall, we observed such changes present in low to high grade sequential progressive dysplasias and OSCCs from a same patient, suggesting these CNAs may be correlated with oral tumor progression.



## References

1. Hunter, K.D., E.K. Parkinson, and P.R. Harrison, *Profiling early head and neck cancer*. Nat Rev Cancer, 2005. **5**(2): p. 127-35.
2. Parkin, D.M., *Global cancer statistics in the year 2000*. Lancet Oncol, 2001. **2**(9): p. 533-43.
3. Carvalho, A.L., et al., *Trends in incidence and prognosis for head and neck cancer in the United States: a site-specific analysis of the SEER database*. Int J Cancer, 2005. **114**(5): p. 806-16.
4. Atula, T., et al., *The evaluation and treatment of the neck in carcinoma of the oral cavity*. Acta Otolaryngol Suppl, 1997. **529**: p. 223-5.
5. Brunin, F., et al., *Cancer of the base of the tongue: past and future*. Head Neck, 1999. **21**(8): p. 751-9.
6. Forastiere, A., et al., *Head and neck cancer*. N Engl J Med, 2001. **345**(26): p. 1890-900.
7. Silverman, S., Jr. and P.B. Sugeran, *Oral premalignancies and squamous cell carcinoma*. Clin Dermatol, 2000. **18**(5): p. 563-8.
8. Lee, J.J., et al., *Predicting cancer development in oral leukoplakia: ten years of translational research*. Clin Cancer Res, 2000. **6**(5): p. 1702-10.
9. Schepman, K., et al., *Concomitant leukoplakia in patients with oral squamous cell carcinoma*. Oral Dis, 1999. **5**(3): p. 206-9.
10. Reibel, J., *Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics*. Crit Rev Oral Biol Med, 2003. **14**(1): p. 47-62.
11. Mithani, S.K., et al., *Molecular genetics of premalignant oral lesions*. Oral Dis, 2007. **13**(2): p. 126-33.
12. Sadikovic, B., et al., *Genome-wide H3K9 histone acetylation profiles are altered in benzo(a)pyrene-treated MCF7 breast cancer cells*. J Biol Chem, 2008. **283**(7): p. 4051-60.
13. Devries, S., et al., *Array-based comparative genomic hybridization from formalin-fixed, paraffin-embedded breast tumors*. J Mol Diagn, 2005. **7**(1): p. 65-71.
14. Paris, P.L., et al., *High-resolution analysis of paraffin-embedded and formalin-fixed prostate tumors using comparative genomic hybridization to genomic microarrays*. Am J Pathol, 2003. **162**(3): p. 763-70.
15. Paris, P.L., et al., *High resolution oligonucleotide CGH using DNA from archived prostate tissue*. Prostate, 2007. **67**(13): p. 1447-55.
16. Bohlander, S.K., et al., *A method for the rapid sequence-independent amplification of microdissected chromosomal material*. Genomics, 1992. **13**(4): p. 1322-4.
17. Roberts R, S.A., Wells GA, Williams KA, Kavaslar N, McPherson R, *Identifying genes for coronary heart disease: An idea whose time has come*. Canadian Journal of Cardiology, 2007. **23**(A): p. 7A-15A.
18. Krawczak, M., et al., *PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships*. Community Genet, 2006. **9**(1): p. 55-61.
19. Johnson, N.A., et al., *Application of array CGH on archival formalin-fixed paraffin-embedded tissues including small numbers of microdissected cells*. Lab Invest, 2006. **86**(9): p. 968-78.
20. Rhodes, D.R., et al., *ONCOMINE: a cancer microarray database and integrated data-mining platform*. Neoplasia, 2004. **6**(1): p. 1-6.
21. Albertson, D.G. and D. Pinkel, *Genomic microarrays in human genetic disease and cancer*. Hum Mol Genet, 2003. **12 Spec No 2**: p. R145-52.

22. Forozan, F., et al., *Comparative genomic hybridization analysis of 38 breast cancer cell lines: a basis for interpreting complementary DNA microarray data*. *Cancer Res*, 2000. **60**(16): p. 4519-25.
23. Lucito, R., et al., *Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation*. *Genome Res*, 2003. **13**(10): p. 2291-305.
24. Schwab, M., *Amplification of the MYCN oncogene and deletion of putative tumour suppressor gene in human neuroblastomas*. *Brain Pathol*, 1990. **1**(1): p. 41-6.
25. Ha, P.K., et al., *A transcriptional progression model for head and neck cancer*. *Clin Cancer Res*, 2003. **9**(8): p. 3058-64.
26. Bockmuhl, U., et al., *Distinct patterns of chromosomal alterations in high- and low-grade head and neck squamous cell carcinomas*. *Cancer Res*, 1996. **56**(23): p. 5325-9.
27. Califano, J., et al., *Genetic progression model for head and neck cancer: implications for field cancerization*. *Cancer Res*, 1996. **56**(11): p. 2488-92.
28. Soder, A.I., et al., *Distinct nonrandom patterns of chromosomal aberrations in the progression of squamous cell carcinomas of the head and neck*. *Cancer Res*, 1995. **55**(21): p. 5030-7.
29. Van Dyke, D.L., et al., *Recurrent cytogenetic abnormalities in squamous cell carcinomas of the head and neck region*. *Genes Chromosomes Cancer*, 1994. **9**(3): p. 192-206.
30. Akervall, J.A., et al., *Chromosomal abnormalities involving 11q13 are associated with poor prognosis in patients with squamous cell carcinoma of the head and neck*. *Cancer*, 1995. **76**(5): p. 853-9.
31. Hu, N., et al., *Genomic characterization of esophageal squamous cell carcinoma from a high-risk population in China*. *Cancer Res*, 2009. **69**(14): p. 5908-17.
32. Paris, P.L., et al., *An oncogenic role for the multiple endocrine neoplasia type 1 gene in prostate cancer*. *Prostate Cancer Prostatic Dis*, 2009. **12**(2): p. 184-91.
33. Lee, J.J., et al., *Array-CGH identifies cyclin D1 and UBCH10 amplicons in anaplastic thyroid carcinoma*. *Endocr Relat Cancer*, 2008. **15**(3): p. 801-15.
34. Tanaka, Y., et al., *Gain of GRHL2 is associated with early recurrence of hepatocellular carcinoma*. *J Hepatol*, 2008. **49**(5): p. 746-57.
35. Hwang, K.T., et al., *Genomic copy number alterations as predictive markers of systemic recurrence in breast cancer*. *Int J Cancer*, 2008. **123**(8): p. 1807-15.
36. Higgins, M.E., et al., *CancerGenes: a gene selection resource for cancer genome projects*. *Nucleic Acids Res*, 2007. **35**(Database issue): p. D721-6.
37. Tsui, I.F., et al., *Multiple pathways in the FGF signaling network are frequently deregulated by gene amplification in oral dysplasias*. *Int J Cancer*, 2009. **125**(9): p. 2219-28.
38. Duran, A., et al., *The signaling adaptor p62 is an important NF-kappaB mediator in tumorigenesis*. *Cancer Cell*, 2008. **13**(4): p. 343-54.
39. Shimada, K., et al., *Aberrant expression of RAB1A in human tongue cancer*. *Br J Cancer*, 2005. **92**(10): p. 1915-21.
40. Bos, P.D., et al., *Genes that mediate breast cancer metastasis to the brain*. *Nature*, 2009. **459**(7249): p. 1005-9.
41. Chechlinska, M. and R. Nowak, *The sensitivity of BRCA1 mutation carriers to ionising radiation: questions of methodology*. *Breast Cancer Res Treat*, 2009. **115**(2): p. 433.
42. Teramoto, T., H. Kawamata, and T. Fujimori, *How does adenosquamous carcinoma arise in the small intestine?* *J Gastroenterol*, 2003. **38**(8): p. 810-1.
43. Wang, Z., *The role of COX-2 in oral cancer development, and chemoprevention/ treatment of oral cancer by selective COX-2 inhibitors*. *Curr Pharm Des*, 2005. **11**(14): p. 1771-7.

44. Cohen, E.E., et al., *Protein kinase C zeta mediates epidermal growth factor-induced growth of head and neck tumor cells by regulating mitogen-activated protein kinase*. *Cancer Res*, 2006. **66**(12): p. 6296-303.
45. Kalyankrishna, S. and J.R. Grandis, *Epidermal growth factor receptor biology in head and neck cancer*. *J Clin Oncol*, 2006. **24**(17): p. 2666-72.
46. Sheu, J.J., et al., *Functional genomic analysis identified epidermal growth factor receptor activation as the most common genetic event in oral squamous cell carcinoma*. *Cancer Res*, 2009. **69**(6): p. 2568-76.
47. Wu, I. and M.A. Moses, *BNF-1, a novel gene encoding a putative extracellular matrix protein, is overexpressed in tumor tissues*. *Gene*, 2003. **311**: p. 105-10.
48. Ellermeier, C. and G.R. Smith, *Cohesins are required for meiotic DNA breakage and recombination in Schizosaccharomyces pombe*. *Proc Natl Acad Sci U S A*, 2005. **102**(31): p. 10952-7.
49. Kalejs, M., et al., *Upregulation of meiosis-specific genes in lymphoma cell lines following genotoxic insult and induction of mitotic catastrophe*. *BMC Cancer*, 2006. **6**: p. 6.
50. Severino, P., et al., *Global gene expression profiling of oral cavity cancers suggests molecular heterogeneity within anatomic subsites*. *BMC Res Notes*, 2008. **1**: p. 113.
51. Meric, F. and K.K. Hunt, *Translation initiation in cancer: a novel target for therapy*. *Mol Cancer Ther*, 2002. **1**(11): p. 971-9.
52. Garnis, C., et al., *Genomic imbalances in precancerous tissues signal oral cancer risk*. *Mol Cancer*, 2009. **8**: p. 50.
53. Weber, A., C. Wittekind, and A. Tannapfel, *Genetic and epigenetic alterations of 9p21 gene products in benign and malignant tumors of the head and neck*. *Pathol Res Pract*, 2003. **199**(6): p. 391-7.
54. Cervigne, N.K., et al., *Identification of a microRNA signature associated with progression of leukoplakia to oral carcinoma*. *Hum Mol Genet*, 2009.

## 6. Conclusões e Direções Futuras

Apesar de estudos tentarem identificar genes/vias genéticas implicadas na progressão do câncer oral, os genes envolvidos na transformação maligna de leucoplasia à carcinoma permanecem desconhecidos, o que torna difícil prever qual destas lesões orais transformarão. Isto acontece principalmente porque a maior parte dos estudos se concentraram na identificação de alterações em regiões cromossômicas grandes, bem como em amostras de lesões orais de pacientes diferentes. No presente estudo, mostramos que as análises genômicas globais, tais como a análise de expressão de miRNAs e a análise do número de cópias do DNA, possuem o potencial para superar este problema, uma vez que permitem a identificação de genes e vias genéticas e regulatórias específicas envolvidos na progressão da doença. A inclusão de amostras seqüentes de leucoplasias e CCEOs de mesmo sítio, do mesmo paciente, permitiu uma comparação exata das alterações genéticas que ocorrem nestas lesões durante a progressão de graus leve à severo à carcinoma invasivo.

Utilizando análises genômicas em ampla escala, fomos capazes de identificar um conjunto de biomarcadores potenciais associados com a progressão oral. O nosso estudo se concentrou em dois objetivos principais: a análise de expressão de microRNAs e a análise do número de cópias do DNA, que têm sido amplamente utilizadas para a identificação de alterações genéticas associadas com o desenvolvimento e progressão de doenças, tais como o câncer. Estas técnicas foram aqui utilizadas como instrumentos para identificar modificações genéticas comumente presentes em lesões pré-malignas e CCEOs correspondentes.

Neste estudo, propusemos um modelo de progressão para miRNAs diferencialmente desregulados durante a transformação maligna de leucoplasia a CCEO. Mostramos que um subconjunto de miRNAs com expressão aumentada (miR-146b, miR-181b, miR-21, miR-345, miR-518b, miR-520g, miR-649 e miR-184) são considerados eventos detectáveis em estágios iniciais da progressão tumoral,

uma vez que estes miRNAs estavam comumente desregulados em lesões que progrediram de grau displásico leve e em lesões malignas, enquanto estão inalterados ou com expressão diminuída em lesões não-progressivas. O nosso estudo identificou o primeiro grupo de RNAs não-codificantes que podem ser valiosos como marcadores da progressão do câncer oral, sugerindo que estes miRs possam ser úteis para ajudar a avaliar quais leucoplasias possuem risco aumentado de transformação maligna. Os nossos dados apoiam a utilização da análise de expressão de miRNAs para segregar lesões com risco de transformação maligna, e sugere que a expressão aumentada de específicos miRNAs possa constituir uma assinatura de progressão para carcinomas orais. Tal assinatura poderia ser aplicada no futuro como um instrumento de "análise de expressão de multi-miRs", que, em conjunto com análises clínica e histológica, ajudará a avaliar quais leucoplasias possuem risco aumentado de progressão à malignidade. Também destacamos o potencial da análise de expressão de miRNAs para a detecção de estágios iniciais do câncer oral, o que levaria a métodos melhorados de diagnóstico e prognóstico do paciente.

Outra análise genômica global realizada neste estudo foi a análise do número de cópias do DNA para determinar o perfil de alterações do DNA associadas com a progressão do câncer oral. Aplicamos com sucesso um novo método de WGA, para produzir perfis de aCGH acurados para amostras arquivadas (FFPE). O nosso estudo também identificou potenciais marcadores genômicos associados com carcinomas orais, nos cromossomos 1p, 2p, 5q e 14q, que serão futuramente validados em um grupo amostral maior independente de lesões orais progressivas e carcinomas correspondentes. Para tal, reunimos 82 amostras orais que progrediram (41 displasias e 41 CCEOs) e 22 amostras que não progrediram à carcinoma. As amostras dos pacientes mostraram variedade nos perfis de CNA, sugerindo que os mecanismos alternativos de instabilidade genômica podem desempenhar um papel relevante na carcinogênese oral. Também exemplificamos o potencial da combinação das análises de CNA e expressão de miRNA para identificar genes envolvidos em carcinogênese oral. As amostras incluídas na

análise de aCGH também fizeram parte do grupo amostral utilizado para avaliar as alterações na expressão de miRNAs que ocorrem durante a progressão de carcinomas orais[1]. Dessa maneira, poderemos integrar os dados resultantes da análise do número de cópias do DNA com o perfil de expressão de miRNA determinado para essas amostras, utilizando o software Partek Genomics Suite (PGS). Esta análise também será útil para explicar se deregulação na expressão de miRNAs seria possivelmente causada por ganhos/amplificação ou perdas do número de cópias do DNA.

Os resultados das análises de expressão de miRNA e do número de cópias do DNA são úteis para caracterizar biomarcadores que possam auxiliar a prever quais leucoplasias possuem um risco aumentado de transformação maligna. Isto resultaria em um melhor prognóstico e sobrevida dos pacientes. Os estudos futuros, como os delineados abaixo, devem combinar estes resultados, bem como incluir análises funcionais para melhor caracterizar e determinar como essas alterações genéticas se associam com a progressão a CCEO, e se elas podem servir como marcadores de estágios iniciais da doença.

### *Direções Futuras*

Sabe-se que os microRNAs possuem o potencial de regular mRNA de diversos genes, incluindo genes supressores tumorais e oncogenes. Contudo, devido ao pareamento imperfeito entre a maioria dos miRNAs e seus alvos, cada miRNA pode regular centenas de genes. A identificação desses alvos tem sido acelerada pelo desenvolvimento de vários bancos de dados que podem ser acessados para compilar os potenciais mRNAs alvo dos miRNAs. As alterações na expressão de miRNAs, associada com aberração genômica e alterações transcricionais, podem levar à desregulação de processos celulares importantes (diferenciação, proliferação e apoptose), que são a base para o desenvolvimento e



progressão do câncer oral. A combinação de perfis de expressão de miRNA e a predição dos seus mRNAs alvo possui a capacidade de fornecer informações sobre os genes e vias gênicas desregulados na doença. Os nossos resultados da análise dos miRNAs possuem aplicações potenciais neste campo, para a identificação de genes/vias desregulados que estejam envolvidos na progressão do câncer. Esta análise pode ser realizada utilizando ferramentas de bioinformática para a predição dos genes alvo, seguida pela validação desses resultados. A análise de bioinformática é baseada em bancos de dados miRNA disponíveis *online*, tal como o *miRanda*[2], que pode estar utilizado para prever os mRNA alvos dos 10 miRNAs (miR-146a, miR-181b, miR-184, miR-21, miR-345, miR-518b, miR-520g, miR-649, miR-196a, e miR-206) identificados no nosso estudo como associados com a progressão ao câncer oral. Futuramente, os genes alvo serão determinados para cada um destes miRs, considerando os valores maiores dos *scores* da análise de predição efetuada pelo banco de dados para cada miRNA e respectivo mRNA. Adicionalmente, esses dados serão submetidos a uma análise para a construção de uma rede de interação genética entre os mRNAs alvo preditos para cada um destes miRNAs, utilizando o Banco de dados de Interação *Interlogous* (I2D)[3], que também verificará se estes miRNAs possuem alguma via/interação genética em comum. Os genes serão anotados utilizando os bancos de dados públicos disponíveis *online* (*NCBI*, *UCSC*), e o seu envolvimento na doença e possível função na carcinogênese será avaliado através de uma revisão da literatura. Os genes e as vias que possuem os critérios estabelecidos serão selecionados para experimentos funcionais.

Estes estudos são necessários para a validação funcional dos alvos preditos e para determinar a relação de regulação entre os miRNAs e seus mRNAs alvo. O delineamento experimental de tais estudos deve incluir: (1) a identificação e a validação dos alvos preditos dos miRNAs, (2) *screening* para miRNAs que regulam a expressão de um gene, e (3) *screening* de miRNAs que afetam um processo celular. Aumentando ou diminuindo a expressão de um miRNA específico permite a investigação dos

papéis específicos desse miRNA na iniciação e desenvolvimento de câncer. Atualmente existem vários métodos, tais como inibidores *antisense*, transgênicos, promotores específicos e mutações em ponto, utilizados para aumentar ou diminuir os níveis de expressão de miRNAs conhecidos, para estudar a regulação genética celular endógena ou investigar alguma resposta fenotípica. A utilização de inibidores *antisense* para bloquear a função de miRNA tem sido aplicada com sucesso por outros grupos, e pode ser um bom enfoque para os nossos estudos futuros. Através dessa estratégia, um RNA *antisense* artificial compete com o mRNA celular no pareamento com o respectivo miRNA. O RNA *antisense* é pareado com seu miRNA e inibi a função desse miRNA. Essa técnica foi adotada por dois grupos de pesquisa independentes, para inibir seqüências de miRNAs e reduzir o silenciamento de RNA por siRNA[4-5], e para inibir quatro miRNAs, *in vivo*, através da modificação de RNAs *antisense*[6]. Ensaio de mutagênese para inserir mutações nos sítios de pareamento de miRNAs com mRNAs também podem ser empregados para estudar a regulação genética dos miRNAs em câncer. Uma vantagem óbvia desta metodologia é a possibilidade de estudar a interação direta dos miRNAs e seus genes alvo. Vários estudos mostraram que “a seqüência original” é importante para os miRNAs reconhecerem seus alvos, e o aumento da má combinação nas seqüências originais reduziria significativamente a função de regulação genética dos miRNAs[7-8]. Assim, também podemos usar este princípio da inserção de mutações em ponto nas seqüências dos miRNAs ou de seus alvos. Uma ou duas alterações de nucleotídeos “na região original” de um miRNA específico reduzirão dramaticamente a possibilidade do miRNA parer seu mRNA alvo, resultando na desregulação da expressão desses mRNAs. Se esses miRNAs e/ou seus alvos estiverem envolvidos em determinados processos celulares, tais como transcrição, proliferação celular, etc, uma mutação em ponto como esta pode resultar em desregulação celular e, conseqüentemente, transformação maligna.

Para a realização de tais experimentos, é importante a existência de modelos animais que reflitam exatamente as alterações celulares e moleculares associadas com a iniciação e a progressão do câncer oral, para entender o desenvolvimento desses carcinomas e avaliar a eficiência de novos marcadores relevantes para a doença. Um dos melhores modelos animais caracterizados para o estudo de CCEO é o hamster (conhecido como *cheek pouch system*) de carcinogênese oral, que correlaciona os eventos moleculares seqüenciais comuns que ocorrem durante a progressão de lesões orais pré-malignas a lesões invasivas malignas[9-10]. Os eventos deste modelo incluem mutações e alterações na expressão de oncogenes e genes supressores tumorais, tais como p53, H-ras e p16[11-13], a expressão de marcadores de proliferação celular[14], resposta a citocinas imune-relacionadas[15] e aumento do crescimento celular em resposta a vários fatores, tais como o fator de crescimento epidermal e fator de crescimento de transformação[16]. As principais vantagens deste modelo são a semelhança entre a mucosa buccal queratinizante desse hamster e a mucosa oral humana, a ausência de tumores espontâneos e a suscetibilidade de influências sistêmicas, tais como hormônios, micronutrientes e outros[17-18]. Apesar de tais modelos terem melhorado a compreensão dos estágios seqüenciais implicados na oncogênese oral, estudos adicionais das vias relacionadas a estes eventos são necessários para esclarecer os mecanismos envolvidos na progressão de carcinomas orais.

Nesta tese, também realizamos análise genômica global do número de cópias do DNA, utilizando a técnica de aCGH, que nos permitiu identificar CNAs de regiões cromossômicas/genes específicos associadas com a progressão do câncer oral. A análise de bioinformática dos dados gerados pela análise de aCGH resultou na identificação de 16 genes candidatos que estavam significativamente amplificados ou deletados tanto em leucoplasias progressivas como em CCEOs; tais genes serão provavelmente requisitados para a ocorrência da transformação maligna. Para confirmar a relevância destes genes como biomarcadores de progressão, validaremos estes genes em um grupo adicional de 104

lesões orais (41 displasias progressivas, 41 CCEOs de mesmo sítio e 22 displasias não-progressivas, com um mínimo de 5 anos de seguimento; Tabelas Suplementares 3 e 4). O DNA genômico isolado destas amostras será submetido a PCR quantitativa em tempo real (qPCR), que é um método de validação mais rápido e de baixo custo, comparado com a análise de hibridização fluorescente *in situ* (FISH), também utilizada para validação de dados gerados pela análise de aCGH. As reações de PCR serão realizadas utilizando o *ABI Prism 7900 Sequence Detection System* (Applied Biosystems). As alterações do número de cópias do DNA nesta análise serão determinadas utilizando dois controle internos, *GAPDH* e B-globin (*HBB*), para considerar a possível variação relacionada com as quantidades de DNA aplicadas ou com a presença de reagentes inibidores na PCR. Ao invés de utilizar o método de quantificação absoluta através da curva padrão de análise, estudos tem demonstrado que é possível calcular o número de cópias haplóide baseado nos valores dos Ct observados[19-20]. No nosso estudo, a quantificação será realizada utilizando este método de Ct comparativo:  $2^{-\Delta\Delta Ct} = (1+E)^{-\Delta Ct \text{ gene alvo}} / (1+E)^{-\Delta Ct \text{ gene referência}}$ , com E = eficiência da reação de PCR (dado como 0.95). Uma amostra do DNA genômico de referência, de alta qualidade (Promega), será utilizada como o calibrador da reação, para a obtenção da quantidade relativa do DNA presente nas amostras alvo. Em um passo adicional, os genes validados por qPCR serão submetidos à análise de FISH para confirmar a sua amplificação em uma pequeno grupo amostral de leucoplasias progressivas seqüenciais e CCEOs correspondentes. Os genes finalmente validados serão submetidos à análises que identifiquem vias gênicas associadas ao câncer com o fim de afunilar e determinar as possíveis via(s) envolvida(s) na progressão de câncer oral.

Enfim, a regulação da expressão dos miRNAs ainda é pouco entendida, mas alguns estudos sugeriram que um mecanismo possível da expressão diferencial de miRNAs pode ser a alteração do número de cópias dos genes transcritos em miRNA[21]. Um dos nossos objetivos futuros é o de integrar os nossos resultados de expressão de miRNAs aos nossos dados do número de cópias do DNA

utilizando o software Partek Geomics Suite (PGS). Neste aspecto, tentaremos determinar a correlação entre a expressão desses miRNAs e as regiões de CNAs determinados pela nossa análise de aCGH; esta análise indicará se as ampliações ou as perdas do DNA estariam presentes nas regiões nas quais os miRNAs estão mapeados. Com o enfoque descrito, seremos capazes de descobrir e elucidar mecanisticamente as bases da heterogeneidade molecular adicionais da carcinogênese oral, e confirmar a utilidade/importância destas alterações genéticas como novos marcadores prognósticos para a progressão do câncer oral.

## Referências

1. Cervigne, N.K., et al., *Identification of a microRNA signature associated with progression of leukoplakia to oral carcinoma*. Hum Mol Genet, 2009.
2. John, B., et al., *Human MicroRNA targets*. PLoS Biol, 2004. 2(11): p. e363.
3. Sultan, M., et al., *Binary tree-structured vector quantization approach to clustering and visualizing microarray data*. Bioinformatics, 2002. 18 Suppl 1: p. S111-9.
4. Hutvagner, G., et al., *Sequence-specific inhibition of small RNA function*. PLoS Biol, 2004. 2(4): p. E98.
5. Meister, G., et al., *Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing*. RNA, 2004. 10(3): p. 544-50.
6. Krutzfeldt, J., et al., *Silencing of microRNAs in vivo with 'antagomirs'*. Nature, 2005. 438(7068): p. 685-9.
7. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets*. Cell, 2005. 120(1): p. 15-20.
8. Lewis, B.P., et al., *Prediction of mammalian microRNA targets*. Cell, 2003. 115(7): p. 787-98.
9. Morris, A.L., D.B. Scott, and A.B. Reiskin, *Carcinogenesis in the hamster cheek pouch. I. Correlation of histopathology with soluble sulfhydryl groups*. Cancer Res, 1961. 21: p. 1352-9.
10. Salley, J.J., *Experimental carcinogenesis in the cheek pouch of the Syrian hamster*. J Dent Res, 1954. 33(2): p. 253-62.
11. Husain, Z., et al., *Sequential expression and cooperative interaction of c-Ha-ras and c-erbB genes in in vivo chemical carcinogenesis*. Proc Natl Acad Sci U S A, 1989. 86(4): p. 1264-8.
12. Muscarella, P., et al., *Identification and sequencing of the Syrian Golden hamster (Mesocricetus auratus) p16(INK4a) and p15(INK4b) cDNAs and their homozygous gene deletion in cheek pouch and pancreatic tumor cells*. Gene, 2001. 278(1-2): p. 235-43.
13. Schwartz, J.L. and G. Shklar, *Verification in syngeneic hamsters of in vitro transformation of hamster oral mucosa by 7,12-dimethylbenz(a)anthracene*. Oral Oncol, 1997. 33(6): p. 431-8.
14. Schwartz, J.L., et al., *Different sensitivity to cell killing and chromosome mutation induction by gamma rays in two human lymphoblastoid cell lines derived from a single donor: possible role of apoptosis*. Mutagenesis, 1995. 10(3): p. 227-33.
15. Schwartz, J.L., et al., *The administration of beta carotene to prevent and regress oral carcinoma in the hamster cheek pouch and the associated enhancement of the immune response*. Adv Exp Med Biol, 1990. 262: p. 77-93.
16. Wong, D.T., et al., *Sequential expression of transforming growth factors alpha and beta 1 by eosinophils during cutaneous wound healing in the hamster*. Am J Pathol, 1993. 143(1): p. 130-42.
17. Dipple, A. and C.A. Bigger, *Metabolic properties of in vitro systems*. Ann N Y Acad Sci, 1983. 407: p. 26-33.
18. Shklar, G., E. Eisenberg, and E. Flynn, *Immunoenhancing agents and experimental leukoplakia and carcinoma of the hamster buccal pouch*. Prog Exp Tumor Res, 1979. 24: p. 269-82.
19. De Preter, K., et al., *Quantification of MYCN, DDX1, and NAG gene copy number in neuroblastoma using a real-time quantitative PCR assay*. Mod Pathol, 2002. 15(2): p. 159-66.
20. Ferreira, I.D., V.E. Rosario, and P.V. Cravo, *Real-time quantitative PCR with SYBR Green I detection for estimating copy numbers of nine drug resistance candidate genes in Plasmodium falciparum*. Malar J, 2006. 5: p. 1.

21. Porkka, K.P., et al., *MicroRNA expression profiling in prostate cancer*. *Cancer Res*, 2007. 67(13): p. 6130-5.

## 7. Apêndice

### 7.1. Material Suplementar

**Tabela Suplementar 3.1:** Grupo amostral incluindo displasias progressivas e CCEOs correspondentes de um mesmo paciente, para a validação dos dados de aCGH.

| Paciente | Amostra ID | Sítio Anatômico | Diagnóstico Histopatológico |
|----------|------------|-----------------|-----------------------------|
| 1 VS     | 1A         | Tonsila         | Displasia                   |
|          | 1B         | Tonsila         | CCEO                        |
| 2 VS     | 2A         | Língua          | Displasia                   |
|          | 2B         | Língua          | CCEO                        |
| 3 VS     | 3A         | FOM             | Displasia                   |
|          | 3B         | FOM             | CCEO                        |
| 4 VS     | 4A         | Mucosa bucal    | Displasia                   |
|          | 4B         | Mucosa bucal    | CCEO                        |
| 5 VS     | 5A         | Língua          | Displasia                   |
|          | 5B         | Língua          | CCEO                        |
| 6 VS     | 6A         | FOM             | Displasia                   |
|          | 6B         | FOM             | CCEO                        |
| 7 VS     | 7A         | Tonsila         | Displasia                   |
|          | 7B         | Tonsila         | CCEO                        |
| 8 VS     | 8A         | Mucosa bucal    | Displasia                   |
|          | 8B         | Mucosa bucal    | CCEO                        |
| 9 VS     | 9A         | FOM             | Displasia                   |
|          | 9B         | FOM             | CCEO                        |
| 10 VS    | 10A        | Língua          | Displasia                   |
|          | 10B        | Língua          | CCEO                        |
| 11 VS    | 11A        | Língua          | Displasia                   |
|          | 11B        | Língua          | CCEO                        |
| 12 VS    | 12A        | Mucosa bucal    | Displasia                   |
|          | 12B        | Mucosa bucal    | CCEO                        |
| 13 VS    | 13A        | Língua          | Displasia                   |
|          | 13B        | Língua          | CCEO                        |
| 14 VS    | 14A        | Mucosa bucal    | Displasia                   |
|          | 14B        | Mucosa bucal    | CCEO                        |
| 15 VS    | 15A        | Mucosa bucal    | Displasia                   |
|          | 15B        | Mucosa bucal    | CCEO                        |
| 16 VS    | 16A        | FOM             | Displasia                   |
|          | 16B        | FOM             | CCEO                        |
| 17 VS    | 17A        | ND              | Displasia                   |
|          | 17B        | ND              | CCEO                        |
| 18 VS    | 18A        | Língua          | Displasia                   |
|          | 18B        | Língua          | CCEO                        |
| 19 VS    | 19A        | Mucosa bucal    | Displasia                   |
|          | 19B        | Mucosa bucal    | CCEO                        |
| 20 VS    | 20A        | FOM             | Displasia                   |
|          | 20B        | FOM             | CCEO                        |
| 21 VS    | 21A        | Língua          | Displasia                   |
|          | 21B        | Língua          | CCEO                        |
| 22 VS    | 22A        | FOM             | Displasia                   |



|              |     |                    |           |
|--------------|-----|--------------------|-----------|
|              | 22B | FOM                | CCEO      |
| <b>23 VS</b> | 23A | ND                 | Displasia |
|              | 23B | ND                 | CCEO      |
| <b>24 VS</b> | 24A | Tonsila            | Displasia |
|              | 24B | Tonsila            | CCEO      |
| <b>25 VS</b> | 25A | Língua             | Displasia |
|              | 25B | Língua             | CCEO      |
| <b>26 VS</b> | 26A | Língua             | Displasia |
|              | 26B | Língua             | CCEO      |
| <b>27 VS</b> | 27A | FOM <sup>198</sup> | Displasia |
|              | 27B | FOM                | CCEO      |
| <b>28 VS</b> | 28A | Mucosa bucal       | Displasia |
|              | 28B | Mucosa bucal       | CCEO      |
| <b>29 VS</b> | 29A | Língua             | Displasia |
|              | 29B | Língua             | CCEO      |
| <b>30 VS</b> | 30A | FOM                | Displasia |
|              | 30B | FOM                | CCEO      |
| <b>31 VS</b> | 31A | Língua             | Displasia |
|              | 31B | Língua             | CCEO      |
| <b>32 VS</b> | 32A | Tonsila            | Displasia |
|              | 32B | Tonsila            | CCEO      |
| <b>33 VS</b> | 33A | FOM                | Displasia |
|              | 33B | FOM                | CCEO      |
| <b>34 VS</b> | 34A | ND                 | Displasia |
|              | 34B | ND                 | CCEO      |
| <b>35 VS</b> | 35A | FOM                | Displasia |
|              | 35B | FOM                | CCEO      |
| <b>36 VS</b> | 36A | FOM                | Displasia |
|              | 36B | FOM                | CCEO      |
| <b>37 VS</b> | 37A | FOM                | Displasia |
|              | 37B | FOM                | CCEO      |
| <b>38 VS</b> | 38A | Língua             | Displasia |
|              | 38B | Língua             | CCEO      |
| <b>39 VS</b> | 39A | Língua             | Displasia |
|              | 39B | Língua             | CCEO      |
| <b>40 VS</b> | 40A | ND                 | Displasia |
|              | 40B | ND                 | CCEO      |
| <b>41 VS</b> | 41A | Mucosa bucal       | Displasia |
|              | 41B | Mucosa bucal       | CCEO      |

**CCEO: Carcinoma de células escamosas oral; FOM: Assoalho da boca; ND: Não-determinado; VS: amostras do grupo da validação (*validation set*).**

**Tabela Suplementar 3.2:** Grupo de leucoplasias orais que não-progrediram a carcinoma, utilizadas para a validação dos resultados da análise de aCGH.

| <b>Paciente ID</b> | <b>Sítio Anatômico</b> | <b>Diagnóstico Histopatológico</b> | <b>Idade</b> | <b>Sexo</b> |
|--------------------|------------------------|------------------------------------|--------------|-------------|
| NP 1               | Palato mole            | Moderate Displasia                 | 59           | M           |
| NP 2               | Língua                 | Moderate Displasia+hyperkeratosis  | 37           | M           |
| NP 3               | Língua                 | Moderate Displasia                 | 66           | F           |
| NP 4               | FOM                    | Moderate Displasia+hyperkeratosis  | 33           | F           |
| NP 5               | Mucosa bucal           | Moderate Displasia                 | 44           | M           |
| NP 6               | Mucosa bucal           | Mild Displasia+hyperkeratosis      | 50           | F           |
| NP 7               | Mucosa bucal           | Mild Displasia+hyperkeratosis      | 75           | M           |
| NP 8               | Mucosa mandibular      | Moderate Displasia                 | 76           | F           |
| NP 9               | Língua                 | Moderate Displasia                 | 81           | F           |
| NP 10              | Língua                 | Mild Displasia                     | 62           | M           |
| NP 11              | Mucosa bucal           | Moderate Displasia+hyperkeratosis  | 61           | F           |
| NP 12              | ND                     | Severe Displasia+hyperkeratosis    | 66           | M           |
| NP 13              | FOM                    | Severe Displasia                   | 49           | M           |
| NP 14              | FOM                    | Mild Displasia+hyperkeratosis      | 53           | F           |
| NP 15              | ND                     | Mild Displasia                     | 67           | M           |
| NP 16              | Língua                 | Mild Displasia+hyperkeratosis      | 51           | M           |
| NP 17              | FOM                    | Moderate Displasia                 | 50           | F           |
| NP 18              | Mucosa mandibular      | Moderate Displasia+hyperkeratosis  | ND           | M           |
| NP 19              | Mucosa mandibular      | Moderate Displasia+hyperkeratosis  | 65           | M           |
| NP 20              | Língua                 | Mild Displasia+hyperkeratosis      | 56           | M           |
| NP 21              | Mucosa bucal           | Mild Displasia+hyperkeratosis      | 39           | F           |
| NP 22              | Mucosa bucal           | Mild Displasia+hyperkeratosis      | 69           | M           |

**FOM: assoalho da boca; M: masculino; F: feminino; ND: não-determinado.**

**Supplemental Table 5.1.** Correlations between amplified and unamplified frozen and FFPE samples as measured by aCGH

| Chromosomes | HN00 29T (unamp) | HN02 201T (unamp) |
|-------------|------------------|-------------------|
|             | HN00 29T (amp)   | HN02 201T (amp)   |
| 1           | 0.8812           | 0.7163            |
| 2           | 0.8734           | 0.7831            |
| 3           | 0.9125           | 0.6154            |
| 4           | 0.8452           | 0.7546            |
| 5           | 0.8321           | 0.7287            |
| 6           | 0.7542           | 0.5835            |
| 7           | 0.7743           | 0.6036            |
| 8           | 0.8438           | 0.7845            |
| 9           | 0.8125           | 0.7524            |
| 10          | 0.7904           | 0.6785            |
| 11          | 0.8562           | 0.7247            |
| 12          | 0.8693           | 0.8015            |
| 13          | 0.7742           | 0.7248            |
| 14          | 0.6623           | 0.5863            |
| 15          | 0.6321           | 0.5740            |
| 16          | 0.7597           | 0.6378            |
| 17          | 0.8452           | 0.7356            |
| 18          | 0.8631           | 0.7025            |
| 19          | 0.9523           | 0.8672            |
| 20          | 0.9729           | 0.8965            |
| 21          | 0.8735           | 0.8472            |
| 22          | 0.7582           | 0.7242            |
| X           | 0.6834           | 0.6736            |
| Y           | 0.7129           | 0.6895            |

Pearson's correlation coefficients are given as a mean for all probes on the array, subdivided by chromosome.

**Supplemental Table 5.2.** List of 79 top candidate genes used to further evaluation on the ONCOMINE v.4 cancer profiling database.

| Chr | Cytoband        | RefSeq_genes                                    | GENE NAME  | FUNCTION   |
|-----|-----------------|---|--|--|
| 1   | 1p35.1          | KHDRBS1   | KH domain containing, RNA binding, signal transduction associated 1(p62,Sam68) | DNA, RNA, and protein binding. SH3/SH2 adaptor activity. transcription repressor activity.   |
| 1   | 1p35.3          | PTPRU   | protein tyrosine phosphatase, receptor type, U (PCP-2)                         | PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation.  |
| 1   | 1p36.23         | PARK7   | Parkinson disease protein 7 (Protein DJ-1) (ONCOGENE DJ1)                      | Acts as a positive regulator of androgen receptor- dependent transcription. Has cell-growth promoting activity and transforming activity.  |
| 1   | 1p36.23         | TNFRSF9   | Tumor necrosis factor receptor superfamily member 9                            | This receptor contributes to the clonal expansion, survival, and development of T cells. It can also induce proliferation in peripheral monocytes, enhance T cell apoptosis induced by TCR/CD3 triggered activation, and regulate CD28 co-stimulation to promote Th1 cell responses. |
| 1   | 1p36.13         | AKR7A3,AKR7L,AKR7A2                             | aldo-keto reductase family 7, member A3 A2 (AKR7A3; AKR7A2)                    | Can reduce the dialdehyde protein-binding form of aflatoxin B1 (AFB1) to the non-binding AFB1 dialcohol. May be involved in protection of liver against the toxic and carcinogenic effects of AFB1, a potent hepatocarcinogen (By similarity).                                       |
| 1   | 1p36.13         | UBR4  | ubiquitin protein ligase E3 component n-recognin 4 (p600)                      | Ubiquitin-protein ligase activity; calmodulin binding  |
| 1   | 1p36.12,1p36.13 | PLA2G2A, PLA2G2C,PLA2G5,PLA2G2E,PLA2G2D,PLA2G2F | phospholipase A2, group IIA (PLA2G2A)  | participate in the regulation of the phospholipid metabolism in biomembranes including eicosanoid biosynthesis. Catalyzes the calcium-dependent hydrolysis of the 2- acyl groups in 3-sn-phosphoglycerides   |
| 1   | 1p36.12,1p36.13 | NBL1  | neuroblastoma, suppression of tumorigenicity 1 2 (DAN)                         | Possible candidate as a tumor suppressor gene of neuroblastoma. May play an important role in preventing cells from entering the final stage (G1/S) of the transformation process. Also, these proteins are secreted, and act as BMP (bone morphogenetic protein) antagonists by     |

binding to BMPs and preventing them from interacting with their receptors.

|   |         |           |   |  |
|---|---------|-----------|---|--|
| 1 | 1q32.1  | CAMSAP1L1 | calmodulin regulated spectrin-associated protein (KIAA1078)                           | unknown  |
| 1 | 1q32.1  | CSRP1     | Homo sapiens cysteine and glycine-rich protein 1 (CSRP1), transcript variant 1, mRNA. | This gene family includes a group of LIM domain proteins, which may be involved in regulatory processes important for development and cellular differentiation. The LIM/double zinc-finger motif found in this gene product occurs in proteins with critical functions in gene regulation, cell growth, and somatic differentiation.                     |
| 1 | 1q32.1  | PTPN7     | protein tyrosine phosphatase, non-receptor type (HePTP)                               | protein tyrosine phosphatase activity. Protein binding.  |
| 1 | 1q32.1  | RNPEP     | Uncharacterized protein RNPEP   | NA   |
| 1 | 1q32.1  | PPP1R12B  | protein phosphatase 1, regulatory (inhibitor)   | Regulates myosin phosphatase activity. Augments Ca(2+) sensitivity of the contractile apparatus  |
| 1 | 1q32.1  | ELF3      | E74-like factor 3, (Epithelium-specific Ets transcription factor 1) (ESE-1)           | Transcriptional activator that binds and transactivates ETS sequences containing the consensus nucleotide core sequence GGA[AT].   |
| 1 | 1q32.1  | LMOD1     | Homo sapiens leiomodin 1 (smooth muscle) (LMOD1), mRNA                                | tropomyosin binding  |
| 1 | 1q32.1  | LAD1      | ladinin 1   | structural molecule activity   |
| 1 | 1q42.12 | PARP1     | poly (ADP-ribose) polymerase family, member 1   | Involved in the base excision repair (BER) pathway, by catalyzing the poly(ADP-ribosylation) of a limited number of acceptor proteins involved in chromatin architecture and in DNA metabolism. This modification follows DNA damages and appears as an obligatory step in a detection/signaling pathway leading to the reparation of DNA strand breaks. |
| 1 | 1q42.13 | PSEN2     | presenilin 2 isoform 2  | protein binding  |
| 1 | 1q42.12 | MIXL1     | Homeobox protein MIXL1 (MIX1 homeobox-like protein 1)                                 | Transcription factor that play a central role in proper axial mesendoderm morphogenesis and endoderm formation. Required for efficient differentiation of  |

|   |               |         |   |   |
|---|---------------|---------|---|---|
|   |               |         |   | cells from the primitive streak stage to blood, by acting early in the recruitment and/or expansion of mesodermal progenitors to the hemangioblastic and hematopoietic lineages |
| 2 | 2p23.3        | GCKR    | glucokinase (hexokinase 4) regulator                                | protein coding. enzyme inhibitor activity   |
| 2 | 2p14          | SERTAD2 | SERTA domain containing 2   | protein coding.transcription coactivator activity.  |
| 2 | 2p14          | RAB1A   | RAB1A, member RAS oncogene family isoform 1                         | Probably required for transit of protein from the ER through Golgi compartment. Binds GTP and GDP and possesses intrinsic GTPase activity                                       |
| 2 | 2p14          | SPRED2  | sprouty-related protein with EVH-1 domain 2                         | regulate growth factor-induced activation of the MAP kinase cascade   |
| 2 | 2p16.1        | PNPT1   | polyribonucleotide nucleotidyltransferase 1                         | transferase activity. RNA binding.  |
| 2 | 2p15,2p14     | UGP2    | UDP-glucose pyrophosphorylase 2                                     | Plays a central role as a glucosyl donor in cellular metabolic pathways.nucleotidyltransferase activity, protein binding.   |
| 3 | 3p24.1        | SLC4A7  | solute carrier family 4, sodium bicarbonate cotransporter, member 7 | transporter activity. Protein binding   |
| 5 | 5q31.2        | ETF1    | eukaryotic translation termination factor 1                         | translation release factor activity, codon specific. Protein, RNA binding.  |
| 5 | 5q31.3        | HBEGF   | heparin-binding EGF-like growth factor                              | epidermal growth factor receptor binding.growth factor activity (EGFR ligand)   |
| 5 | 5q31.2,5q31.3 | NRG2    | neuregulin 2  | growth factor activity  |
| 5 | 5q31.2,5q31.3 | CXXC5   | CXXC finger 5   | DNA, metal ion binding ( WNT pathways )**   |
| 5 | 5q31.3        | PURA    | purine-rich element binding protein A                               | RNA polymerase II transcription factor activity, enhancer binding.SMAD binding. transcription factor binding.double-stranded telomeric DNA binding                              |
| 5 | 5q31.2        | PAIP2   | poly(A) binding protein interacting protein 2                       | translation repressor activity (interacts with VEGF mRNA - prognostic value in H&N SCC )  |
| 5 | 5q31.2        | SLC23A1 | solute carrier family 23 (nucleobase transporters), member 1        | nucleobase transmembrane transporter activity   |
| 5 | 5q31.2        | CTNNA1  | catenin (cadherin-associated protein),                              | structural molecule activity. Cadherin  |

|    |          |         |  |   |
|----|----------|---------|--|---|
|    |          |         | alpha 1, 102kDa  | biding  |
| 5  | 5q31.3   | DIAPH1  | diaphanous homolog 1 (Drosophila)                          | receptor biding. Actin biding   |
| 5  | 5q31.3   | HDAC3   | histone deacetylase 3                                      | transcription factor binding. histone deacetylase activity. transcription corepressor activity  |
| 5  | 5q35.1   | DUSP1   | dual specificity phosphatase 1                             | human cellular response to environmental stress as well as in the negative regulation of cellular proliferation   |
| 5  | 5q31.4   | WNT8A   | wingless-type MMTV integration site family, member 8       | cell signalling, development, regulation of cell fate, embriogenesis  |
| 6  | 6p21.33  | MSH5    | mutS homolog 5 (E. coli)                                   | This gene encodes a member of the mutS family of proteins that are involved in DNA mismatch repair or meiotic recombination processes. This protein is similar to a Saccharomyces cerevisiae protein that participates in meiotic segregation fidelity and crossing-over.   |
| 6  | 6q25.4   | EZR     | ezrin  | This protein plays a key role in cell surface structure adhesion, migration and organization, and it has been implicated in various human cancers.  |
| 7  | 7p13     | TBRG4   | transforming growth factor beta regulator 4                | protein kinase activity   |
| 7  | 7q32.3   | MKLN1   | muskelin 1, intracellular mediator containing kelch motifs | Muskelin is an intracellular protein that acts as a mediator of cell spreading and cytoskeletal responses to the extracellular matrix component thrombospondin I - protein binding, signal transduction.  |
| 9  | 9q34.13  | RAPGEF1 | Rap guanine nucleotide exchange factor (GEF) 1             | This gene encodes a human guanine nucleotide exchange factor. It transduces signals from CRK by binding the SH3 domain of CRK, and activating several members of the Ras family of GTPases. This signaling cascade that may be involved in apoptosis, integrin-mediated signal transduction, and cell transformation. |
| 10 | 10q24.32 | MGEA5   | meningioma expressed antigen 5 (hyaluronidase)             | cetyltransferase activity, metabolism   |

|           |          |         |   |   |
|-----------|----------|---------|---|---|
| <b>10</b> | 10q24.32 | NPM3    | nucleophosmin/nucleoplasmin 3                     | nucleic acid binding, protein binding, rRNA processing and transcription  |
| <b>10</b> | 10q24.32 | FGF8    | fibroblast growth factor 8 (androgen-induced)     | FGF family members possess broad mitogenic and cell survival activities, and are involved in a variety of biological processes, including embryonic development, cell growth, morphogenesis, tissue repair, tumor growth and invasion. This protein is known to be a factor that supports androgen and anchorage independent growth of mammary tumor cells. Overexpression of this gene has been shown to increase tumor growth and angiogenesis.   |
| <b>11</b> | 11q13.4  | CHRD2   | chordin-like 2, breast tumor novel factor 1       | cell differentiation, cartilage development, multicellular organismal development.  |
| <b>11</b> | 11q12.3  | LGALS12 | lectin, galactoside-binding, soluble, 12          | lactose binding, sugar binding, apoptosis   |
| <b>12</b> | 12q13.12 | MLL2    | myeloid/lymphoid or mixed-lineage leukemia 2      | DNA binding, metal ion binding, methyltransferase activity  |
| <b>12</b> | 12p13.31 | PTPN6   | protein tyrosine phosphatase, non-receptor type 6 | The protein encoded by this gene is a member of the protein tyrosine phosphatase (PTP) family. PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. N-terminal part of this PTP contains two tandem Src homolog (SH2) domains, which act as protein phosphotyrosine binding domains, and mediate the interaction of this PTP with its substrates. This PTP is expressed primarily in hematopoietic cells, and functions as an important regulator of multiple signaling pathways in hematopoietic cells. This PTP has been shown to interact with, and dephosphorylate a wide spectrum of phospho-proteins involved in hematopoietic cell signaling. |



|    |          |        |  |   |
|----|----------|--------|--|---|
| 12 | 12p13.31 | VWF    | von Willebrand factor preproprotein                              | Endothelial cell marker. The glycoprotein encoded by this gene functions as both an antihemophilic factor carrier and a platelet-vessel wall mediator in the blood coagulation system. It is crucial to the hemostasis process. Mutations in this gene or deficiencies in this protein result in von Willebrand's disease. An unprocessed pseudogene has been found on chromosome 22.   |
| 14 | 14q22.1  | ARF6   | ADP-ribosylation factor 6  | This gene encodes a member of the human ARF gene family, which is part of the RAS superfamily. The ARF genes encode small guanine nucleotide-binding proteins that stimulate the ADP-ribosyltransferase activity of cholera toxin and play a role in vesicular trafficking and as activators of phospholipase D. The product of this gene is localized to the plasma membrane, and regulates vesicular trafficking, remodelling of membrane lipids, and signaling pathways that lead to actin remodeling. |
| 14 | 14q12    | TINF2  | TERF1 (TRF1)-interacting nuclear factor 2                        | telomeric DNA binding   |
| 14 | 14q12    | IRF9   | interferon regulatory factor 9                                   | protein binding, transcription factor activity  |
| 14 | 14q12    | PSME1  | proteasome (prosome, macropain) activator subunit 1 (PA28 alpha) | proteasome activator  |
| 14 | 14q12    | REC8   | REC8 homolog (yeast)   | protein binding, male meiosis I, spermatogenesis  |
| 14 | 14q12    | GMMPR2 | guanosine monophosphate reductase 2                              | GMP reductase activity, metal ion binding, oxidoreductase activity.   |

|    |         |        |  |   |
|----|---------|--------|--|---|
| 14 | 14q12   | TGM1   | transglutaminase 1 (K polypeptide epidermal type I, protein-glutamine-gamma-glutamyltransferase) | Acetyltransferase activity, calcium ion binding activity, involved in keratinization. The protein encoded by this gene is a membrane protein that catalyzes the addition of an alkyl group from an alkylamine to a glutamine residue of a protein, forming an alkylglutamine in the protein. This protein alkylation leads to crosslinking of proteins and catenation of polyamines to proteins. This gene contains either one or two copies of a 22 nt repeat unit in its 3' UTR. Mutations in this gene have been associated with autosomal recessive lamellar ichthyosis (LI) and nonbullous congenital ichthyosiform erythroderma (NCIE). |
| 14 | 14q22.1 | CDKL1  | cyclin-dependent kinase-like 1 (CDC2-related kinase)   | ATP binding, nucleotide binding, cyclin-dependent protein kinase activity. This gene product is a member of a large family of CDC2-related serine/threonine protein kinases. It accumulates primarily in the nucleus.   |
| 14 | 14q22.1 | ATL1   | atlastin GTPase 1  | GTP binding, GTPase activity, protein binding, nucleotide binding. The protein encoded by this gene is a GTPase and a Golgi body transmembrane protein. The encoded protein can form a homotetramer and has been shown to interact with spastin and with mitogen-activated protein kinase kinase kinase 4. This protein may be involved in axonal maintenance as evidenced by the fact that defects in this gene are a cause of spastic paraplegia type 3.  |
| 14 | 14q22.1 | MAP4K5 | mitogen-activated protein kinase kinase kinase 5   | ATP, nucleotide and protein binding, protein serine/threonine kinase activity. This gene encodes a member of the serine/threonine protein kinase family, that is highly similar to yeast SPS1/STE20 kinase. Yeast SPS1/STE20 functions near the beginning of the MAP kinase signal cascades that is essential for yeast   |

|    |          |          |   |   |
|----|----------|----------|---|---|
|    |          |          |   | pheromone response. This kinase was shown to activate Jun kinase in mammalian cells, which suggested a role in stress response.   |
| 14 | 14q32.12 | BTBD7    | BTB (POZ) domain containing 7                         | protein binding   |
| 16 | 16p13.13 | TNFRSF17 | tumor necrosis factor receptor superfamily, member 17 | The protein encoded by this gene is a member of the TNF-receptor superfamily. This receptor is preferentially expressed in mature B lymphocytes, and may be important for B cell development and autoimmune response. This receptor has been shown to specifically bind to the tumor necrosis factor (ligand) superfamily, member 13b (TNFSF13B/TALL-1/BAFF), and to lead to NF-kappaB and MAPK8/JNK activation. This receptor also binds to various TRAF family members, and thus may transduce signals for cell survival and proliferation. |
| 16 | 16p13.13 | GSPT1    | G1 to S phase transition 1                            | GTP binding, nucleotide and protein binding, translation release factor activity. G1/S transition of mitotic cell cycle; nuclear-transcribed mRNA catabolic process, nonsense-mediated decay; protein amino acid methylation.   |
| 16 | 16p13.13 | GSPT1    | G1 to S phase transition 1                            | GTP binding, nucleotide and protein binding, translation release factor activity. G1/S transition of mitotic cell cycle; nuclear-transcribed mRNA catabolic process, nonsense-mediated decay; protein amino acid methylation.   |
| 19 | 19q13.2  | LEUTX    | leucine twenty homeobox                               | sequence-specific DNA binding, transcription factor activity. Regulation of transcription, DNA-dependent.   |

|    |         |         |  |   |
|----|---------|---------|--|---|
| 19 | 19q13.2 | CLC     | Charcot-Leyden crystal protein           | Lysophospholipase activity; hydrolase activity; sugar binding. Lysophospholipases are enzymes that act on biological membranes to regulate the multifunctional lysophospholipids. The protein encoded by this gene is a lysophospholipase expressed in eosinophils and basophils. It hydrolyzes lysophosphatidylcholine to glycerophosphocholine and a free fatty acid. This protein may possess carbohydrate or IgE-binding activities. It is both structurally and functionally related to the galectin family of beta-galactoside binding proteins. It may be associated with inflammation and some myeloid leukemias.   |
| 19 | 19q13.2 | LGALS14 | lectin, galactoside-binding, soluble, 14 | Sugar binding. This gene is predominantly expressed in placenta. The encoded protein belongs to the galectin (galaptin/S-lectin) family. The members of galectin family contain one or two carbohydrate recognition domains, which can bind beta-galactoside.   |
| 20 | 20q12   | EMILIN3 | elastin microfibril interfacier 3        | uncharacterized   |
| 20 | 20q12   | LPIN3   | lipin 3                                  | Phosphatidate phosphatase activity. Humans lipodystrophy is characterized by loss of body fat, fatty liver, hypertriglyceridemia, and insulin resistance. Mice carrying mutations in the fatty liver dystrophy (fld) gene have similar phenotypes. Through positional cloning, the mouse gene responsible for fatty liver dystrophy was isolated and designated Lpin1. The nuclear protein encoded by Lpin1 was named lipin. Lpin1 mRNA was expressed at high levels in adipose tissue and was induced during differentiation of preadipocytes. These results indicated that lipin is required for normal adipose tissue development and provided a candidate gene for human lipodystrophy. Through database searches, mouse and human EST and genomic sequences with similarities to |

Lpin1 were identified. These included two related mouse genes (Lpin2 and Lpin3) and three human homologs (LPIN1, LPIN2, and LPIN3). Human LPIN1 gene has been mapped to 2p25.; linkages of fat mass and serum leptin levels to this same region have been noted. Human LPIN2 and LPIN3 mapped to chromosomes 18p11 and 20q11-q12, respectively.

|           |               |        |                                  |   |
|-----------|---------------|--------|----------------------------------|---|
| <b>21</b> | 21q22.3       | ZNF295 | zinc finger protein 295          | DNA binding, metal ion binding, protein binding, zinc ion binding. ZNF295 may be involved in the bi-directional control of gene expression in concert with ZFP161.  |
| <b>21</b> | 21q22.3       | PRDM15 | PR domain containing 15          | DNA binding, metal ion binding, protein binding, zinc ion binding.  |
| <b>22</b> | 22q12.3       | FBXO7  | F-box protein 7                  | Protein binding, ubiquitin-protein ligase activity. This gene encodes a member of the F-box protein family which is characterized by an approximately 40 amino acid motif, the F-box. The F-box proteins constitute one of the four subunits of the ubiquitin protein ligase complex called SCFs (SKP1-cullin-F-box), which function in phosphorylation-dependent ubiquitination. The F-box proteins are divided into 3 classes: Fbws containing WD-40 domains, Fbfs containing leucine-rich repeats, and Fbxs containing either different protein-protein interaction modules or no recognizable motifs. The protein encoded by this gene belongs to the Fbxs class and it may play a role in regulation of hematopoiesis. |
| <b>16</b> | 16p13.2       | A2BP1  | ataxin 2-binding protein 1       | protein C-terminus binding. RNA, nucleotide binding   |
| <b>9</b>  | 9q33.1,9q33.2 | DBC1   | deleted in bladder cancer 1      | protein binding.  |
| <b>8</b>  | 8p23.2        | CSMD1  | CUB and Sushi multiple domains 1 | protein binding.  |

|           |         |       |  |  |
|-----------|---------|-------|--|--|
| <b>3</b>  | 3p26.2  | EDEM1 | ER degradation enhancer,<br>mannosidase alpha-like 1 | protein binding.                           |
| <b>18</b> | 18q21.1 | MYO5B | MYO5B  | ATP, actin, calmodulin, nucleotide binding |

## 7.2. Outros Manuscritos submetidos e/ou em Preparo para publicação

### 7.2.1. Submetidos:

#### 1. *microRNA evaluation of unknown primary lesions in the head and neck.*

Barker EV FRCS, PhD\*, **Cervigne NK MSc**\*, Goswami R.S. MD\*, Reis PP PhD\*, Weinreb I MD, Irish JC MSc, MD, Kamel-Reid S PhD.

\* Equal contributors

Journal: Molecular Cancer, **Aceito sob revisão**

Unknown primary malignancy in the head and neck is not an infrequent diagnosis for patients with metastatic cervical lymph nodes. Although linked with a relatively good prognosis following radiation treatment, widespread radiation is coupled with significant morbidity. Altered microRNA (miRNA) expression has been associated with both cancer progression and metastasis. We sought to determine whether miRNA expression analysis could be used as a diagnostic tool to discover the primary site of malignancy, within the head and neck. We used real-time PCR quantification to identify miRNA expression profiles of squamous cell carcinoma of the tonsil, base of tongue and post-nasal space, as well as their corresponding metastatic lymph nodes, from 6 patients. Our results revealed that each cancer maintained its expression profile between the primary site and the nodal metastasis ( $r= 0.82$ ,  $p<0.0001$ ). In addition, each anatomical sub-site maintained a distinct miRNA profile between individual patients ( $r=0.79$ ,  $p<0.0001$ ). Finally, between sub-sites, the miRNA profiles were distinct ( $p<0.0001$ ). As proof of principle, our study provides an indication that miRNA expression analysis may be useful to compare the primary lesion and local metastatic disease. This may be clinically relevant to predict the primary site of origin of metastatic disease, when the primary site remains obscure.

#### 2. *Expression profiling identifies genes associated with human osteosarcoma oncogenesis and correlation of RUNX2 overexpression with poor response to chemotherapy.*

Bekim Sadikovic<sup>1,2</sup>, Paul Thorner<sup>1</sup>, Susan Chilton-MacNeill<sup>1</sup>, Jeff W. Martin<sup>3</sup>, **Nilva K. Cervigne**<sup>4,5</sup>, Jeremy Squire<sup>3</sup>, Maria Zielenska<sup>1,2</sup>

<sup>1</sup>Department of Pediatric Laboratory Medicine, Pathology Division, Hospital for Sick Children, Toronto, Canada, M5G 1X8.

<sup>2</sup>Genetics and Genome Biology Program, Hospital for Sick Children, Toronto, ON, Canada.

<sup>3</sup>Department of Pathology and Molecular Medicine, Richardson Labs, Queen's University, Kingston, Canada, K7L 3N6

<sup>4</sup>Division of Applied Molecular Oncology, Ontario Cancer Institute, the University Health Network, Toronto, Canada, M5G 2M9.

<sup>5</sup>Dept. of Genetics, Biosciences Institute, UNESP, Botucatu, SP, Brazil

MS ID: 2117142176319882, Journal: BMC Cancer

Human osteosarcoma is the most common pediatric bone tumour. There is limited understanding of molecular mechanisms underlying osteosarcoma oncogenesis, and a lack of good diagnostic as well as prognostic clinical markers for this disease. The best prognostic marker related to the overall survival of osteosarcoma patients is the response to chemotherapy. Recent discoveries have highlighted a potential role of a number of genes including: *RECQL4*, *DOCK5*, *SPP1*, *RUNX2*, *RBI*, *CDKN1A*, *P53*, *IBSP*, *LSAMP*, *MYC*, *TNFRSF1B*, *BMP2*, *HISTH2BE*, *FOS*, *CCNB1*, and *CDC5L*. Our objective was to assess relative expression levels of these 16 genes as potential biomarkers of osteosarcoma oncogenesis and chemotherapy response in human tumours. We performed quantitative expression analysis in a panel of 22 human osteosarcoma tumours with differential response to chemotherapy, and 5 normal human osteoblasts. *RECQL4*, *SPP1*, *RUNX2*, and *IBSP* were significantly overexpressed, and *DOCK5*, *CDKN1A*, *RBI*, *P53*, and *LSAMP* showed significant loss of expression relative to normal osteoblasts. In addition to being overexpressed in osteosarcoma tumour samples relative to normal osteoblasts, *RUNX2* was the only gene of the 16 to show significant overexpression in tumours that had a poor response to chemotherapy relative to good responders. These data provide further evidence of the loss of tumour suppressive pathways, and activation of specific oncogenic mechanisms associated with osteosarcoma oncogenesis, while highlighting the role of *RUNX2* expression as a potential biomarker of chemotherapy failure in osteosarcoma.

**3. The tumor suppressor Programmed Cell Death 4 inhibits tumor cell invasion and it is under-expressed in Oral Squamous Cell Carcinoma.**

Patricia P. Reis<sup>1</sup>, Miranda Tomenson<sup>1,2</sup>, **Nilva K. Cervigne**<sup>1,3</sup>, Jerry Machado<sup>1,4</sup>, Mahadeo A. Sukhai<sup>5</sup>, Igor Jurisica<sup>6,7</sup>, Melania Pintilie<sup>8,9</sup>, Natalie Naranjo Galloni<sup>10</sup>, Bayardo Perez-Ordenez<sup>11</sup>, Colleen



Simpson<sup>12</sup>, David Goldstein<sup>12</sup>, Dale Brown<sup>12</sup>, Ralph Gilbert<sup>12</sup>, Patrick Gullane<sup>12</sup>, Jonathan Irish<sup>12</sup>, Suzanne Kamel-Reid<sup>1,2,4,11</sup>

<sup>1</sup>Div. of Applied Molecular Oncology, Princess Margaret Hospital, Ontario Cancer Institute, University Health Network, Toronto, Ontario, Canada

<sup>2</sup>Dept. of Medical Biophysics, University of Toronto, Toronto, ON, Canada

<sup>3</sup>Dept. of Genetics, Biosciences Institute, UNESP, Botucatu, SP, Brazil

<sup>4</sup>Dept. of Laboratory Medicine and Pathobiology, University of Toronto, ON, Canada

<sup>5</sup>Div. of Cancer Genomics and Proteomics, Ontario Cancer Institute, University Health Network, Toronto, Ontario, Canada

<sup>6</sup>Dept. of Computer Science, Ontario Cancer Institute, University Health Network, Toronto, Ontario, Canada

<sup>7</sup>Dept. of Signaling Biology, Ontario Cancer Institute, University Health Network, Toronto, Ontario, Canada

<sup>8</sup>Dalla Lana School of Public Health Sciences, University of Toronto, Toronto, ON, Canada

<sup>9</sup>Dept. of Biostatistics, Princess Margaret Hospital, University Health Network, Toronto, ON, Canada

<sup>10</sup>Dept. of Otolaryngology, Hospital Calderon Guardia, San Jose, Costa Rica

<sup>11</sup>Dept. of Pathology, Toronto General Hospital, Ontario Cancer Institute, University Health Network, Toronto, Ontario, Canada

<sup>12</sup>Dept. of Otolaryngology/Surgical Oncology, Princess Margaret Hospital, Ontario Cancer Institute and The University Health Network, Toronto, Ontario, Canada

Journal: Molecular Cancer

PDCD4 has been found under-expressed and associated with tumorigenesis processes, such as disease progression and metastasis. There are no current studies characterizing PDCD4 expression and its clinical relevance in Oral Squamous Cell Carcinoma (OSCC). Our goal was to determine whether PDCD4 expression was correlated with important patient clinicopathological data. Since nodal metastasis is a major prognostic factor in OSCC, we aimed to determine whether PDCD4 had functional relevance in OSCC invasion. PDCD4 mRNA expression levels were assessed in 50 OSCCs and 25 normal oral tissues. PDCD4 was under-expressed in 43/50 (86%) OSCCs, with significantly reduced mRNA levels in patients with nodal metastasis ( $p=0.0027$ ), and marginally associated with T3-T4 tumor stage ( $p=0.054$ ). PDCD4 protein expression was assessed, by immunohistochemistry (IHC), in 28/50 OSCCs and adjacent normal tissues; PDCD4 protein was absent/under-expressed in 25/28 (89%) of OSCCs, and marginally associated with nodal metastasis ( $p=0.059$ ). A matrigel invasion assay showed that PDCD4 expression suppressed invasion, and si-RNA mediated PDCD4 loss was associated with

increased invasive potential of oral carcinoma cells. We also showed that miR-21 levels were increased in tumors showing PDCD4 under-expression, both at the mRNA and protein levels, and that PDCD4 expression may be down-regulated in OSCC by direct binding of miR-21 to the 3'UTR PDCD4 mRNA. Our data showed a potential role for the loss of PDCD4 expression in OSCC tumorigenesis and invasion, and demonstrated a mechanism of PDCD4 down-regulation in OSCC. PDCD4 may represent a clinically relevant biomarker with prognostic value in OSCC, due to its association with nodal metastasis.

#### *4. Optimization and analysis of a quantitative real-time PCR-based technique to determine microRNA expression in formalin-fixed paraffin-embedded samples*

Goswami RS<sup>1,2</sup>, Machado J<sup>2</sup>, **Cervigne NK**<sup>2,3</sup>, Reis PP<sup>2,3</sup>, Kuruvilla J<sup>3</sup>, Crump M<sup>3</sup>, Bailey D<sup>4</sup>, Kamel-Reid S<sup>1,2,4</sup>.

<sup>1</sup>Laboratory Medicine & Pathobiology

<sup>2</sup>Division of Applied Molecular Oncology

<sup>3</sup>Dept. of Genetics, Biosciences Institute, UNESP, Botucatu, SP, Brazil

<sup>4</sup>Department of Hematology/Medical Oncology

<sup>5</sup>Department of Pathology, University Health Network, Toronto, ON, Canada.

Journal: RNA

MicroRNAs (miRs) are non-coding RNA molecules, 17-27 nucleotides in length, involved in post-transcriptional regulation. miRs play diverse roles in tissue development, differentiation, cell proliferation and apoptosis. mRNA expression studies may be limited to fresh, frozen samples due to degradation and cross-linking that occurs after formalin fixation. However, miRs may be less prone to such processes, facilitating their expression studies in formalin-fixed paraffin-embedded (FFPE) tissue. This presents an advantage as there is a large repository of FFPE samples within anatomical pathology laboratories, available for characterization and analysis combined with retrospective patient clinical data. miR expression levels can be quantified in FFPE tissues using quantitative real-time PCR techniques. Our study demonstrates that the TaqMan Low Density Arrays (TLDA) platform is suitable for miR expression profiling analysis, and outlines the optimal performance conditions of this platform for clinical FFPE samples. Additionally, we showed that miR expression profiles of FFPE samples are

comparable to those of fresh-frozen samples, opening up a realm of research possibilities for retrospective studies.

### 7.2.2. *Em preparação:*

#### 1. *Global Genomic Analysis of Oral Carcinomas from Young and Older Patients using Affymetrix SNP 6.0 Arrays.*

Jerry Machado<sup>a,b</sup>, **Nilva K. Cervigne**<sup>b,c</sup>, Christian R. Marshall<sup>d</sup>, Rashmi Goswami<sup>a,b</sup>, Bayardo Perez-Ordóñez<sup>e</sup>, Patrick Gullane<sup>f</sup>, Jonathan C. Irish<sup>f</sup>, Patricia P. Reis<sup>b</sup>, Suzanne Kamel-Reid<sup>a,b,g,h</sup>.

<sup>a</sup>Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

<sup>b</sup>Division of Applied Molecular Oncology, Princess Margaret Hospital, Ontario Cancer Institute and University Health Network, Toronto, Ontario, Canada

<sup>c</sup>Department of Genetics, Bioscience Institute, Sao Paulo State University, Botucatu, SP, Brazil

<sup>d</sup>The Centre for Applied Genomics, Hospital for Sick Children, Toronto, Ontario, Canada

<sup>e</sup>Department of Pathology, Toronto General Hospital, University Health Network, Toronto, Ontario, Canada.

<sup>f</sup>Department of Otolaryngology/Surgical Oncology, Princess Margaret Hospital, Ontario Cancer Institute and University Health Network, Toronto, Ontario, Canada

<sup>g</sup>Institute for Medical Sciences, University of Toronto, Toronto, Ontario, Canada

<sup>h</sup>Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

Oral squamous cell carcinoma (OSCC) most commonly develops in older patients ( $\geq 60$  years of age) with a history of tobacco and alcohol use. However, young individuals ( $\leq 40$  years of age) can also develop OSCC, often without common risk factors (e.g. smoking and alcohol). The absence or shortened time of carcinogen exposure and the development of carcinoma at an early age suggest aberrant genetic events in the development of OSCCs in young patients that are different than OSCCs from older patients. We used Affymetrix SNP 6.0 arrays to assess copy number, loss of heterozygosity (LOH), and copy number neutral LOH (CNNLOH) status within tumors and adjacent normal mucosa from 19 young and 20 old patients. Copy number profiles were similar between tumors of young and older patients, however focal regions of different copy number alterations (CNAs) were observed. In addition we observed an increase of LOH and CNNLOH in tumors from older patients compared to young patients,

suggesting that LOH and CNNLOH may have a more prominent role in oral tumors from older patients. However, similar to CNAs there were specific regions of LOH and CNNLOH that were only observed in young patient tumors. We are currently validating potential genes involved in these regions of CNAs, LOH and CNNLOH in tumors from young and older patients, to gain an understanding of the genetic changes involved in early-onset oral carcinoma.

## *2. Telomere length abnormalities occur early during oral cancer progression*

**Nilva K. Cervigne**<sup>1,2</sup>, Patricia P. Reis<sup>1</sup>, Grace Bradley<sup>1,3</sup>, Melania Pintilie<sup>4,5</sup>, Bayardo Perez-Ordenez<sup>7</sup>, Suzanne Kamel-Reid<sup>1,6,7\*</sup>

<sup>1</sup>Division of Applied Molecular Oncology, Ontario Cancer Institute, the University Health Network, Toronto, M5G 2M9 ON Canada

<sup>2</sup>Department of Genetics, Bioscience Institute, Sao Paulo State University, Botucatu, SP Brazil

<sup>3</sup>Faculty of Dentistry, University of Toronto, Toronto, M5G 1G6 ON Canada

<sup>4</sup>Department of Biostatistics, Ontario Cancer Institute, the University Health Network, Toronto, M5G 2M9 ON Canada

<sup>5</sup>Dalla Lana School of Public Health Sciences, University of Toronto, M5T 3M7 ON Canada

<sup>6</sup>Department of Laboratory Medicine and Pathobiology, University of Toronto, M5G 1L5 ON Canada

<sup>7</sup>Department of Pathology, Toronto General Hospital, Ontario Cancer Institute, the University Health Network, Toronto, M5G 2C4 ON Canada

Oral squamous cell carcinoma (OSCC) arises as a consequence of multiple genetic damage events induced in many chromosomes and genes, and it is the accumulation of these changes that appears to lead to carcinoma. Telomere maintenance by telomerase or, in its absence, alternative lengthening of telomeres, protects this acquired genetic changes ensuring immortality without losing eukaryotic linear DNA. When this maintenance does not occur, DNA is lost and end-replication problems arise. In our study we objective to verify if telomere length appear to be one of the earliest and most prevalent alterations acquired in the multistep process of tumor progression by comparison of oral malignant tissue with their corresponding leukoplakias. Real-Time quantification of relative telomere length (RQ-Telo Assay) was used to assess the telomere length of 35 oral sample (mild, moderate and severe dysplasias and their corresponding OSCCs). We found that all oral lesions examined showed telomere length abnormalities; with most lesions (94.7%) displaying abnormally short or very short telomeres (table 2). Altered telomere length was detected in the majority of samples: 27 samples (71%) contained very short telomeres and 9 (23.7%) had short telomere lengths. Two (5.3%) samples (severe dysplasia + carcinoma

and mincroinvasive carcinoma) presented long telomeres. Our preliminary results indicate that telomere erosion may be associated with malignant transformation of leukoplakias. We need to assess more distant normal tissues to ensure that this is the case. Future studies are thus necessary to determine if telomere length assessment could be used as a biomarker in oral cancer.

### 3. A Comprehensive Review of the Molecular Characterization of Oral Cancer.

**Nilva K. Cervigne**<sup>b,c\*</sup>, Jerry Machado<sup>a,b\*</sup>, Patricia P. Reis<sup>b</sup>, Suzanne Kamel-Reid<sup>a,b,d,e</sup>.

<sup>a</sup>Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

<sup>b</sup>Division of Applied Molecular Oncology, Princess Margaret Hospital, Ontario Cancer Institute and University Health Network, Toronto, Ontario, Canada

<sup>c</sup>Department of Genetics, Bioscience Institute, Sao Paulo State University, Botucatu, SP, Brazil

<sup>d</sup>Institute for Medical Sciences, University of Toronto, Toronto, Ontario, Canada

<sup>e</sup>Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

\* Equal contributors

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)