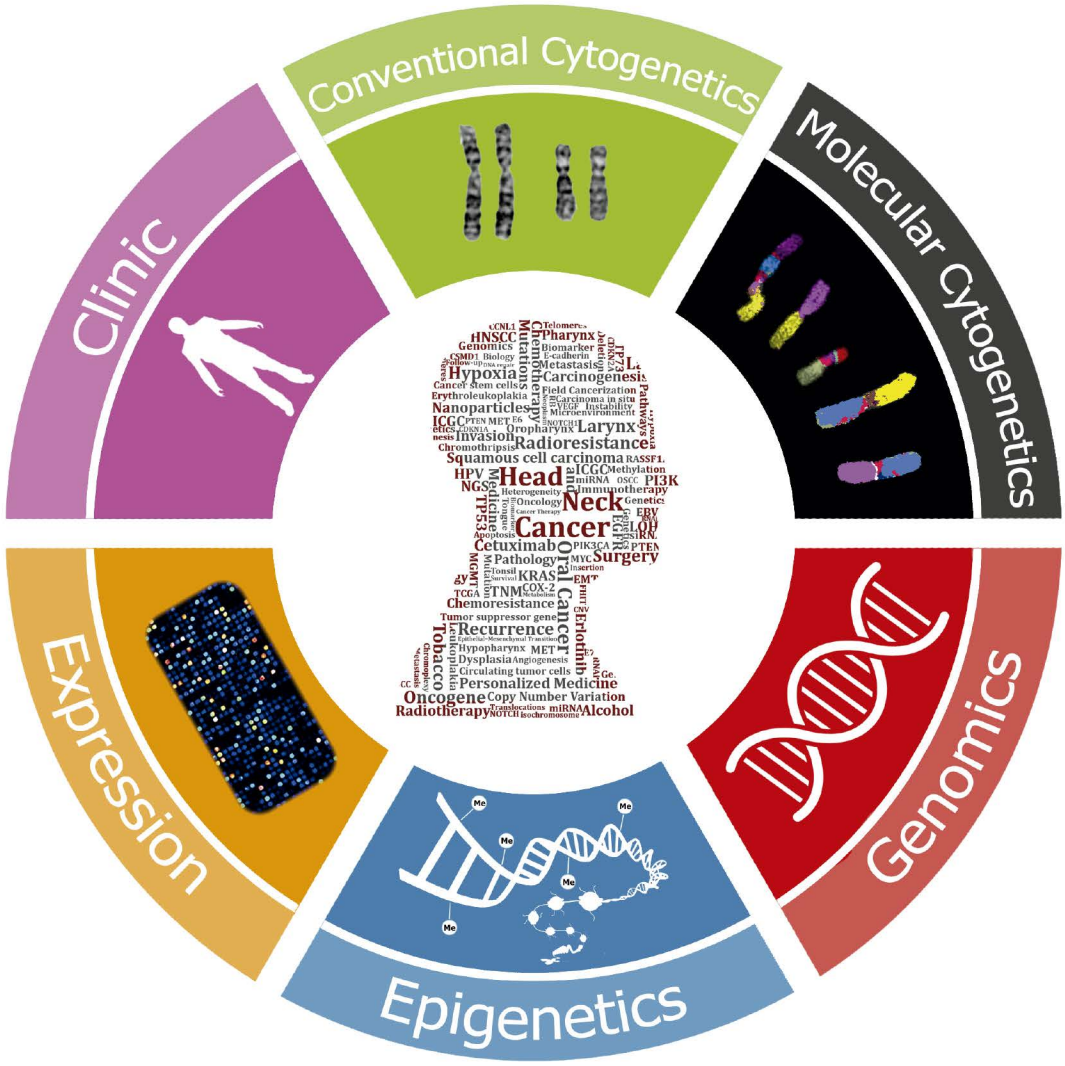


Ilda Patrícia Tavares da Silva Ribeiro

Head and Neck Squamous Cell Carcinoma:
integrating genomic, epigenetic and transcriptomic data - from bench to clinical applications

Tese de doutoramento do Programa Interuniversitário de Doutoramento em Envelhecimento e Doenças Crónicas,
orientada pela Senhora Professora Isabel Marques Carreira e pelo Senhor Professor Rui Manuel Reis,
o grau é conferido pelas Faculdade de Medicina da Universidade de Coimbra, Faculdade de Ciências Médicas da Universidade Nova de Lisboa
e Escola de Ciências da Saúde da Universidade do Minho, e é apresentada à Faculdade de Medicina da Universidade de Coimbra.

Julho 2017

· U C ·

UNIVERSIDADE DE COIMBRA

Ilda Patrícia Tavares da Silva Ribeiro

Head and Neck Squamous Cell Carcinoma: integrating genomic, epigenetic and transcriptomic data - from bench to clinical applications

Tese de doutoramento do Programa Interuniversitário de Doutoramento em Envelhecimento e Doenças Crónicas, orientada pela Senhora Professora Isabel Marques Carreira e pelo Senhor Professor Rui Manuel Reis e Coorientada pela Senhora Professora Joana Barbosa de Melo, o grau é conferido pelas Faculdade de Medicina da Universidade de Coimbra, Faculdade de Ciências Médicas da Universidade Nova de Lisboa e Escola de Ciências da Saúde da Universidade do Minho, e é apresentada à Faculdade de Medicina da Universidade de Coimbra.

Julho 2017

Universidade de Coimbra

Institutions and Funding

The presented research study involved the following institutions and funding:

Publications

The following original articles and abstracts have been published (or submitted) in peer-scientific journals, within the scope of the present work.

Book Chapters

Ilda P. Ribeiro, Joana B. Melo, Isabel M. Carreira (2017). High -Throughput Genomic Approaches for Head and Neck Cancer: Towards Molecular Biomarkers and Targeted Therapies, In: Head and Neck Cancer, Avid Science, Telangana, India, 2017. 2-31

Original articles (in chronological order)

Published

1. **Ilda P. Ribeiro**, Francisco Marques, Leonor Barroso, Jorge Miguéis, Francisco Caramelo, André Santos, Maria J. Julião, Joana B. Melo, Isabel M. Carreira. *Genetic and epigenetic characterization of the tumors in a patient with a tongue primary tumor, a recurrence and a pharyngoesophageal second primary tumor*. Molecular Cytogenetics, 2017. 10(13):2-9 doi: 10.1186/s13039-017-0310-z (Chapter 3.4.3)

2. **Ilda P. Ribeiro**, Francisco Caramelo, Francisco Marques, Ana Domingues, Margarida Mesquita, Leonor Barroso, Hugo Prazeres, Maria J. Julião, Isabel P. Baptista, Artur Ferreira, Joana B. Melo, Isabel M. Carreira. *WT1, MSH6, GATA5 and PAX5 as epigenetic oral squamous cell carcinoma biomarkers - a short report.* Cellular Oncology, 2016. 39(6):573-582 doi: 10.1007/s13402-016-0293-5 (Chapter 3.1.3)

3. **Ilda P. Ribeiro**, Leonor Barroso, Francisco Marques, Joana B. Melo, Isabel M. Carreira. *Early detection and personalized treatment in oral cancer: the impact of omics approaches.* Molecular Cytogenetics, 2016. 9(85):1-7 doi: 10.1186/s13039-016-0293-1

4. **Ilda P. Ribeiro**, Francisco Marques, Francisco Caramelo, João Pereira, Miguel Patrício, Hugo Prazeres, José Ferrão, Maria J. Julião, Miguel Castelo-Branco, Joana B. Melo, Isabel P. Baptista, Isabel M. Carreira. *Genetic gains and losses in oral squamous cell carcinoma: impact on clinical management*. Cellular Oncology, 2014. 37(1):29-39 doi:10.1007/s13402-013-0161-5 (Chapter 3.1.1)

5. **Ilda P. Ribeiro**, Francisco Marques, Francisco Caramelo, José Ferrão, Hugo Prazeres, Maria J. Julião, Widad Rifi, Suvi Savola, Joana B. Melo, Isabel P. Baptista, Isabel M. Carreira. *Genetic imbalances detected by Multiplex Ligation-dependent Probe Amplification in a cohort of patients with oral squamous cell carcinoma - the first step towards clinical personalized medicine*. Tumor Biology, 2014. 35 (5):4687-4695 doi: 10.1007/s13277-014-1614-9 (Chapter 3.1.2)

In Press

1. **Ilda P. Ribeiro**, Francisco Marques, Leonor Barroso, Joana Rodrigues, Francisco Caramelo, Joana B. Melo, Isabel M. Carreira. *Genomic profile of oral squamous cell carcinomas with an adjacent leukoplakia or with an erythroleukoplakia that evolved after the treatment of primary tumor: a report of two cases*. Molecular Medicine Reports. (Chapter 3.4.1)

2. **Ilda P. Ribeiro**, Joana Rodrigues, Alexandra Mascarenhas, Nadezda Kosyakova, Francisco Caramelo, Thomas Liehr, Joana B. Melo, Isabel M. Carreira. *Cytogenetic, genomic and epigenetic characterization of HSC-3 tongue cell line, with lymphnode metastasis*. Journal of Oral Sciences. (Chapter 3.5.1)

Submitted

1. **Ilda P. Ribeiro**, Francisco Caramelo, Luísa Esteves, Camila Oliveira, Francisco Marques, Leonor Barroso, Joana B. Melo, Isabel M. Carreira. *Genomic and epigenetic signatures associated with survival rate in oral squamous cell carcinoma patients*. (Chapter 3.1.4)

2. **Ilda P. Ribeiro**, Francisco Caramelo, Luísa Esteves, Joana Menoita, Francisco Marques, Leonor Barroso, Jorge Miguéis, Joana B. Melo, Isabel M. Carreira. *Genomic predictive model for recurrence and metastasis development in head and neck squamous cell carcinoma patients.* (Chapter 3.2.1)

3. **Ilda P. Ribeiro**, Isabel M. Carreira, Thomas Liehr, Joana B. Melo. *Chromosomal breaks and low copy repeat elements in head and neck squamous cell carcinoma.* (Chapter 3.2.2)

4. **Ilda P. Ribeiro**, Luísa Esteves, Francisco Caramelo, Francisco Marques, Leonor Barroso, Joana B. Melo, Isabel M. Carreira. *Expression of HIF3A, HOXA10 and LOX discriminates tumors of the tongue from the floor of the mouth.* (Chapter 3.3.1)

5. **Ilda P. Ribeiro**, Leonor Barroso, Francisco Marques, André Santos, Francisco Caramelo, Maria J. Julião, Joana B. Melo, Isabel M. Carreira. *Genomic and epigenetic characterization for the comparison of synchronous bilateral tongue squamous cell carcinomas - a case report.* (Chapter 3.4.2)

6. **Ilda P. Ribeiro**, Joana M. Rodrigues, Alexandra Mascarenhas, Vanessa Marques, Nadezda Kosyakova, Francisco Caramelo, Maria J. Julião, Thomas Liehr, Joana B. Melo, Isabel M. Carreira. *(Cyto)genomic and epigenetic characterization of BICR 10 cell line and three new established primary human head and neck squamous cell carcinoma cultures*. (Chapter 3.5.2)

Peer-Reviewed conference abstracts published on indexed journals

1. Isabel M. Carreira, **Ilda P. Ribeiro**, Susana I. Ferreira, Alexandra Mascarenhas, Joana Rodrigues, Vanessa Marques, Cláudia Pais, Francisco Marques, Joana B. Melo. *From the establishment of primary culture cell*

lines of oral squamous cell carcinoma to the identification of complex rearrangements. Molecular Cytogenetics, 2017. 10(Suppl1)2.20

2. **Ilda P. Ribeiro**, Francisco Caramelo, Francisco Marques, André Santos, Leonor Barroso, Margarida Mesquita, Teresa Lopes, Maria J. Julião, Artur Ferreira, Isabel P. Baptista, Joana B. Melo, Isabel M Carreira. *Genomic classification of oral squamous cell carcinoma: the role of copy number alterations.* European Journal of Human Genetics, Supplement. 2016; 24 (E-Supplement 1): 279

3. Isabel M. Carreira, **Ilda P. Ribeiro**, Francisco Marques, Francisco Caramelo, André Santos, Maria J. Julião, Isabel P. Baptista, Joana B. Melo. *Oral squamous cell carcinoma: can the genes predict a second primary tumor?* European Journal of Human Genetics, Supplement. 2016; 24 (E-Supplement 1): 295

4. **Ilda P. Ribeiro**, Francisco Marques, André Santos, Francisco Caramelo, Isabel P. Baptista, Joana B. Melo, Isabel M. Carreira. *Novel regions of gain and loss on chromosome 11 detected in the oral squamous cell carcinoma*. Revista Portuguesa de Pneumologia. 2016; 22 (Esp Cong 2):6

5. **Ilda P. Ribeiro**, Francisco Marques, André Santos, Joana Menoita, Isabel P. Baptista, Isabel M. Carreira, Joana B. Melo. *Non-random genomic breakpoints in oral cancer: which characteristic sequence motifs are behind the breaks?* Revista Portuguesa de Pneumologia. 2016; 22 (Esp Cong 2):13

6. **Ilda P. Ribeiro,** Francisco Caramelo, Francisco Marques, Ana Domingues, Alexandra Estevinho, Maria J. Julião, Artur Ferreira, Isabel P. Baptista, Joana B. Melo, Isabel M. Carreira. *From oral cancer genetic and epigenetic alterations to clinical outcome prediction.* European Journal of Human Genetics. 2015; (Supplement 1): 260

7. **Ilda P. Ribeiro,** Ana Cainço, Francisco Marques, Francisco Caramelo, Alexandra Estevinho, Maria J. Julião, Joana B. Melo, Artur Ferreira, Isabel P. Baptista, Isabel M. Carreira. *Decoding genetic and epigenetic signatures in oral squamous cell carcinoma.* Revista Portuguesa de Pneumologia. 2015. 21 (Esp Cong 1):4

8. **Ilda P. Ribeiro,** Francisco Marques, Francisco Caramelo, Susana I. Ferreira, Luís M. Pires, Joana Rodrigues, Maria J. Julião, Joana B. Melo, Artur Ferreira, Isabel P. Baptista, Isabel M. Carreira. G*enome-wide copy number analysis in oral squamous cell carcinoma.* Revista Portuguesa de Pneumologia. 2015. 21 (Esp Cong 1):9

9. **Ilda P. Ribeiro**, Francisco Marques, Susana I. Ferreira, Luís M. Pires, Francisco Caramelo, Maria J. Julião, Joana B. Melo, Isabel P. Baptista, Isabel M. Carreira. *The contribution of array-CGH to reveal genomic biomarkers for oral cancer*. Revista Portuguesa de Pneumologia. 2014. 20:2-3

10. **Ilda P. Ribeiro**, Francisco Marques, Francisco Caramelo, José Ferrão, Maria J. Julião, Joana B. Melo, Isabel P. Baptista, Isabel M. Carreira. *Genetic profile of oral cancer as a valuable ally for clinicians*. Revista Portuguesa de Pneumologia. 2014. 20:10-11

11. Francisco Marques, **Ilda P. Ribeiro**, Hugo Prazeres, Maria J. Julião, Isabel P. Baptista, Joana B. Melo, Isabel M. Carreira. HPV and head and neck cancer: guilty or not guilty? Revista Portuguesa de Pneumologia. 2014. 20:11

12. Francisco Marques, **Ilda P. Ribeiro**, Susana I. Ferreira, José Ferrão, Maria J. Julião, Isabel P. Baptista, Joana B. Melo, Isabel M. Carreira. *Head and neck cancer: a heterogeneous disease?* Revista Portuguesa de Pneumologia. 2014. 20:11-12

Conference abstract published in conference proceedings

1. **Ilda P. Ribeiro,** Leonor Barroso, Francisco Marques, André Santos, Francisco Caramelo, Maria J. Julião, Joana B. Melo, Isabel M. Carreira. *One patient and two synchronous tongue tumors: how genes help to discriminate bilateral carcinomas?* Livro de Resumos 20ª Reunião Anual da Sociedade Portuguesa de Genética Humana, Fundação Bissaya Barreto, Coimbra (2016). pp129

2. **Ilda P. Ribeiro,** André Santos, Francisco Caramelo, Nicole Pedro, Francisco Marques, Leonor Barroso, Maria J. Julião, Artur Ferreira, Joana B. Melo, Isabel M. Carreira. *Gene expression profiling to predict clinical outcome of oral cavity carcinomas.* Livro de Resumos 20ª Reunião Anual da Sociedade Portuguesa de Genética Humana, Fundação Bissaya Barreto, Coimbra (2016). pp62

3. **Ilda P. Ribeiro**, Joana B. Melo, Rui M. Reis, Isabel M. Carreira. *Genomic, transcriptomic and epigenetic data to identify biomarkers of prognosis in oral cancer*. Abstract Book, V Annual PhD OC meeting, Coimbra (2016). pp41

4. **Ilda P. Ribeiro**, Francisco Caramelo, Francisco Marques, Ana Domingues, Margarida Mesquita, Maria J. Julião, Isabel P. Baptista, Joana B. Melo, Artur Ferreira, Isabel M. Carreira. *Predictors of prognosis in oral squamous cell carcinoma: the role of DNA methylation.* Livro de Resumos 19ª Reunião Anual da Sociedade Portuguesa de Genética Humana, Ordem dos Médicos, Porto (2015). pp104

5. **Ilda P. Ribeiro**, Francisco Marques, Camila Oliveira, Maria J. Julião, Isabel P. Baptista, Joana B. Melo, Isabel M. Carreira. *Primary tongue tumor and second primary tumor in the floor of the mouth in the same patient - The discriminatory power of the genes.* Livro de Resumos 19ª Reunião Anual da Sociedade Portuguesa de Genética Humana, Ordem dos Médicos, Porto (2015). pp103

6. **Ilda P. Ribeiro**, Francisco Marques, Francisco Caramelo, Susana I. Ferreira, Luís M. Pires, Maria J. Julião, Joana B. Melo, Isabel P. Baptista, Isabel M. Carreira. *Molecular karyotype of oral squamous cell carcinoma.* Livro de Resumos da 18ª Reunião Anual da Sociedade Portuguesa de Genética Humana, Edifício Egas Moniz, Faculdade de Medicina da Universidade de Lisboa, Lisboa (2014). pp91

“It is the time you have wasted for your rose that makes your rose so important.”

Antoine de Saint-Exupéry, The Little Prince

“There are only two days in the year that nothing can be done. One is called **Yesterday** and the other is called **Tomorrow**. Today is the right day to Love, Believe, Do and mostly Live.”

Dalai Lama XIV

Acknowledgments /Agradecimentos

Terminada esta caminhada, depois de quatro anos de trabalho intenso, algum desespero e muitas conquistas, é hora de agradecer a todos aqueles que contribuíram para o seu sucesso. Embora pelo seu significado uma dissertação seja em nome individual e, por vezes um processo solitário, é também o resultado de muitos contributos preciosos que não posso deixar de realçar e, por isso, expressar publicamente os meus mais sinceros agradecimentos.

Em primeiro lugar tenho de expressar o meu profundo agradecimento à minha orientadora, a **PROFESSORA DOUTORA ISABEL MARQUES CARREIRA**, por me ter dado este magnifico presente que foi a possibilidade de colaborar neste projeto aquando do meu trabalho de mestrado. Obrigada por me ter impulsionado a crescer, expandir horizontes, ver para além do óbvio, evoluir e desenvolver espirito crítico e, assim, garantir as condições para seguir para doutoramento com este projeto que me entusiasmou e apaixonou desde o primeiro dia. Hoje sou apaixonada pelo "fazer ciência", "procurar respostas" e, acima de tudo realizada pelo percurso que escolhi e que me vem ajudando a trilhar através do seu exemplo de vida, profissionalismo e altruísmo pedagógico com que sempre partilhou os seus conhecimentos científicos. É também tempo agora para fazer uma confissão, talvez em jeito de desculpa, porque sempre me senti muito querida e porque a sua generosidade e paciência são imensas, durante estes anos abusei das mensagens e emails - "*estou-me a sentir ignorada*" disse várias vezes! Na realidade não o sentia e sabia que não era de todo verdade, mas por vezes abusamos e exageramos destes estados de espirito só para alcançarmos o céu - e eu queria que no final desta minha jornada, por tudo que investimos e demos de nós (juntas), o meu céu fosse perfeito!

Tenho igualmente de manifestar o meu agradecimento e reconhecimento à **PROFESSORA DOUTORA JOANA BARBOSA DE MELO**, pelo seu imprescindível contributo para este trabalho. São vários os agradecimentos que lhe podia e devia fazer: a partilha do seu vasto conhecimento científico, as horas de discussão que me proporcionou, os preciosos conselhos que me deu, o tempo que me dispensou quando percebeu que estava a precisar de apoio, todos os esforços que encetou para que este trabalho chegasse a bom-porto e fosse tudo aquilo que eu desejava, mas quero acima de tudo agradecer a sua amizade, generosidade e

preocupação que tanto me confortaram nestes anos, foi imensamente importante saber que estava ali para mim sempre que eu precisava!

Tenho também de expressar o meu sincero agradecimento ao **PROFESSOR DOUTOR FRANCISCO CARAMELO** pela sua ajuda imensurável na análise dos dados e algoritmos informáticos. Se o todo é maior que a soma das partes, é justo reconhecer que sem o seu contributo único este trabalho não teria a mesma relevância! Agradeço-lhe o seu enorme contributo científico e ensinamentos que me proporcionou, mas acima de tudo a sua disponibilidade e enorme esforço para solucionar com recurso a fórmulas e modelos matemáticos as minhas questões científicas! "Isso é possível" foi a expressão que mais me repetiu em reposta às minhas questões e que bom que era ouvir isso!

Tenho também de agradecer ao **PROFESSOR DOUTOR RUI REIS** pela coorientação deste trabalho e pela disponibilidade manifestada.

Não posso deixar de manifestar uma palavra de agradecimento à **Comissão Diretiva do Programa Interuniversitário de Doutoramento em Envelhecimento e Doenças Crónicas (PhDOC)**, por terem acreditado em mim e me terem confiado uma bolsa para a realização deste trabalho.

Ao **CIMAGO/ACIMAGO** pelo contributo financiamento que permitiu em parte a realização deste projeto.

Porque o nosso trabalho reflete o ambiente em que trabalhamos, não posso deixar de agradecer a toda a **EQUIPA DO LABORATÓRIO DE CITOGENÉTICA DA FACULDADE DE MEDICINA DA UNIVERSIDADE DE COIMBRA (LCG-FMUC)**, pelos ensinamentos, partilhas científicas, pelos momentos saborosos de convívio, mas principalmente por me receberem tão bem e me fazerem sentir parte da equipa desde o primeiro dia. Obrigada pela vossa amizade! Apesar de uma equipa ser o seu todo no coletivo e de ter recebido de **TODOS**, palavras de incentivo e ajuda, tenho forçosamente de destacar algumas pessoas pelos contributos científicos diretos neste trabalho. Obrigada à Susana Ferreira, ao Miguel Pires, à Alexandra Mascarenhas, à Cláudia Pais, à Patrícia Paiva e à Mariana Val.

Ao **DR. FRANCISCO MARQUES** do Serviço de Estomatologia e departamento de Dentária da Faculdade de Medicina da Universidade de Coimbra, às **DRA. LEONOR BARROSO, DRA. TERESA LOPES, DRA. MARGARIDA**

MESQUITA, DRA. ISABEL AMADO e ao **DR. ARTUR FERREIRA** do Serviço de Cirurgia Maxilo-Facial do Centro Hospitalar e Universitário de Coimbra, CHUC, EPE, ao **DR. JORGE MIGUÉIS** do Serviço de Otorrinolaringologia do Centro Hospitalar e Universitário de Coimbra, CHUC, EPE e à **PROFESSORA DOUTORA ISABEL POIARES BAPTISTA** da Faculdade de Medicina da Universidade de Coimbra, muito obrigada pela colheita de amostras, pelas informações clínicas e atualização das mesmas sempre que solicitado. Obrigada pelas discussões e ensinamentos do ponto de vista clínico e até mesmo do doente!

À **DRA. MARIA JOSÉ JULIÃO** do Serviço de Anatomia Patológica do Centro Hospitalar e Universitário de Coimbra, CHUC, EPE, obrigada pela disponibilidade e informações anátomo-patológicas dos doentes.

Ao **DOUTOR HUGO PRAZERES** do Departamento de Patologia Molecular do Instituto Português de Oncologia (IPO) Coimbra, obrigada pela ajuda na identificação e caracterização do vírus do papiloma humano.

A TODOS OS DOENTES, que com o seu consentimento permitiram a realização deste trabalho, o meu agradecimento.

A <u>TODOS</u> OS ALUNOS DE MESTRADO que nestes últimos anos passaram pelo Laboratório de Citogenética e Genómica da Faculdade de Medicina da Universidade de Coimbra (LCG-FMUC), obrigada pela vossa ajuda, por me permitirem transmitir os meus conhecimentos e também aprender convosco! Obrigada pelos momentos de boa disposição e pela amizade que ficou em tantos casos. Um agradecimento especial aos **MESTRES LUÍSA ESTEVES, JOANA MENOITA, ANDRÉ SANTOS, VANESSA MARQUES, CAMILA OLIVEIRA e ANA DOMINGUES** que muito contribuíram para este trabalho. Tenho ainda de evidenciar a Mestre **JOANA MATOS RODRIGUES** não só pela sua grande ajuda, apoio e contributo para este trabalho mas também pela amizade de todas as horas!

A TODOS OS MEUS AMIGOS, desculpem pelas minhas ausências, demora nas respostas e contínuas faltas de tempo. E obrigada porque faço parte do grupo dos privilegiados cujas verdadeiras amizades se adaptam e sobrevivem fortalecidas perante as exigências que se impõem nestas "vidas" tão diferentes e complicadas que decidimos viver!!!

AOS MEUS PAIS, obrigada pelo vosso apoio incondicional e gratuito, exemplo de força, perseverança e trabalho. Obrigada por me terem dado "as asas" para chegar aqui! Foram vocês que me fizeram continuar sempre, ultrapassando todos os obstáculos, e acreditando que com dedicação e trabalho tudo seria possível.

AO MEU IRMÃO, por ser o meu modelo a seguir, a força motriz para fazer sempre mais e melhor! Obrigada por elevares a fasquia. Seriam muitos os elogios e ainda mais os agradecimentos que te poderia e devia fazer, mas como és meu irmão vou-te simplesmente dizer OBRIGADA por estares presente na minha vida!

ÀS MINHAS SOBRINHAS, INÊS E SOFIA, obrigada por me ajudarem a equilibrar o meu tempo...porque os nossos momentos de brincadeira são reconfortantes!

Por último, mas nem por isso menos importante, **AO MICAEL**, obrigada pelo apoio, incentivo e milhões de sorrisos! Obrigada por me ensinares que é possível conciliar e viver em pleno as várias facetas da vida!

....Obrigada!

Index

Tables Index

Figures Index

Abbreviations

5-FU - 5-fluorouracil

A

ACER2 - Alkaline ceramidase 2

aCGH- Array comparative genomic hybridization

ADAM11 **-** ADAM metallopeptidase domain 11

ADM1 - Aberration detection module 1

ADM2 - Aberration detection module 2

AKT- Protein kinase B

AKT1 - V-akt murine thymoma viral oncogene homolog 1

ALDH2 **-** Aldehyde dehydrogenase 2 family (mitochondrial)

ANGPTL6 **-** Angiopoietin like 6

ANXA13 **-** Annexin A13

APC **-** APC, WNT signaling pathway regulator

APC2- APC2, WNT signaling pathway regulator

APEH **-** Acylaminoacyl-peptide hydrolase

APPL1 **-** Adaptor protein, phosphotyrosine interacting with PH domain and leucine zipper 1

ARAP1 **-** ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1

ARF4 **-** ADP ribosylation factor 4

ARHGEF18 **-** Rho/Rac guanine nucleotide exchange factor 18

ARHGEF3 **-** Rho guanine nucleotide exchange factor 3

ARID3A **-** AT-rich interaction domain 3A

ARIH2 **-** Ariadne RBR E3 ubiquitin protein ligase 2

ARL8B **-** ADP ribosylation factor like GTPase 8B

ASAP1 or ***DDEF1*** **-** ArfGAP with SH3 domain, ankyrin repeat and PH domain 1 or Development and differentiation enhancing factor 1

ATAD5 **-** ATPase family, AAA domain containing 5

ATM- ATM serine/threonine kinase

ATRIP **-** ATR interacting protein

AVEN - Apoptosis and caspase activation inhibitor

B

BAX **-** BCL2 associated X, apoptosis regulator

BBC3 **-** BCL2 binding component 3

BCL2 **-** BCL2, apoptosis regulator

BCL2L13 **-** BCL2 like 13

BCL6 **-** B-cell CLL/lymphoma 6

BIRC2 **-** Baculoviral IAP repeat containing 2

BIRC6 **-** Baculoviral IAP repeat containing 6

BMP5 **-** Bone morphogenetic protein 5

BNIPL **-** BCL2 interacting protein like

bp **-** Base pair

BRCA1 **-** BRCA1, DNA repair associated

BRCA2 **-** BRCA2, DNA repair associated

BRCAX **or** ***BRCA3*** **-** Breast cancer 3

BRD4 **-** Bromodomain containing 4

BRPF1 **-** Bromodomain and PHD finger containing 1

BTNL3 **-** Butyrophilin like 3

BUD23 **or** ***WBSCR22*** **-** BUD23, rRNA methyltransferase and ribosome maturation factor or Williams Beuren syndrome chromosome region 22

C

C21orf58 **-** Chromosome 21 open reading frame 58

CADM1 **-** Cell adhesion molecule 1

CARD8 **-** Caspase recruitment domain family member 8

CASP1**-** Caspase 1

CASP8 **-** Caspase 8

CASR **-** Calcium sensing receptor

CATPCA - Categorical principal components analysis

CBL **-** Cbl proto-oncogene

cCGH - Chromosomal comparative genomic hybridization

CCND1 **-** Cyclin D1

CCNL1 **-** Cyclin L1

CCNO **-** Cyclin O

CCR10 **-** C-C motif chemokine receptor 10

CCR4 **-** C-C motif chemokine receptor 4

CD302 **-** CD302 molecule

CD37 **-** CD37 molecule

CD44 **-** CD44 molecule (Indian blood group)

CDC25A **-** Cell division cycle 25A

CDH1 **-** Cadherin 1

CDH13 **-** Cadherin 13

CDH2 **-** Cadherin 2

CDIPT **-** CDP-diacylglycerol--inositol 3-phosphatidyltransferase

CDK11A **-** Cyclin dependent kinase 11A

CDK11B **-** Cyclin dependent kinase 11B

CDK2AP1- Cyclin dependent kinase 2 associated protein 1

CDK4 **-** Cyclin-dependent kinase 4

CDK6 **-** Cyclin dependent kinase 6

CDKN1B **-** Cyclin dependent kinase inhibitor 1B

CDKN2A **-** Cyclin dependent kinase inhibitor 2A

CDKN2B **-** Cyclin dependent kinase inhibitor 2B

CDKN2D **-** Cyclin dependent kinase inhibitor 2D

CELF2 **-** CUGBP, Elav-like family member 2

CFTR **-** Cystic fibrosis transmembrane conductance regulator

CGH - Comparative genomic hybridization

CHAF1B **-** Chromatin assembly factor 1 subunit B

CHCHD2 **-** Coiled-coil-helix-coiled-coil-helix domain containing 2

CHFR **-** Checkpoint with forkhead and ring finger domains

CHST13 -Carbohydrate (chondroitin 4) sulfotransferase 13

CI - Confidence interval

CIB1 **-** Calcium and integrin binding 1

CIDEC **-** Cell death inducing DFFA like effector c

CLDN3 **-** Claudin 3

CNA - Copy number alteration

CNV - Copy number variation

CO_2 - Carbon dioxide

COL11A2 - Collagen, type XI, alpha 2

COPS5 **-** COP9 signalosome subunit 5

COSMIC - Catalogue of somatic mutations in cancer

COX2 **-** Cyclooxygenase 2

CREB3L3 **-** cAMP responsive element binding protein 3 like 3

CREM **-** cAMP responsive element modulator

CRK **-** CRK proto-oncogene, adaptor protein

CSMD1 **-** CUB and Sushi multiple domains 1

CSPP1 **-** Centrosome and spindle pole associated protein 1

CT - Computed tomography

CTCF **-** CCCTC-binding factor

CTDSPL **or** ***RBSP3*** **-** CTD small phosphatase like or RB protein serine phosphatase from chromosome 3

CTNNB1 **-** Catenin beta 1

CTR - Chemoradiation therapy

CTTN **-** Cortactin

Cy3 - Cyanine 3

Cy5 - Cyanine 5

CYB5R3 or DIA1 **-** Cytochrome b5 reductase 3 or Diaphorase (NADH) (cytochrome b-5 reductase)

CYP21A2 **-** Cytochrome P450 family 21 subfamily A member 2

D

DAP3 **-** Death associated protein 3

DAPK **-** Death associated protein kinases

DAPK3 **-** Death associated protein kinase 3

DCC **-** DCC netrin 1 receptor

DDX19B **-** DEAD-box helicase 19B

DDX20 **-** DEAD-box helicase 20

DDX28 **-** DEAD-box helicase 28

DDX60L **-** DEAD-box helicase 60 like

del - Deletion

DEPDC1B **-** DEP domain containing 1B

der - Chromosome derivative of a translocation

DHX30 **-** DExH-box helicase 30

DLG1 **-** Discs large MAGUK scaffold protein 1

DLGAP2 **-** DLG associated protein 2

DMEM - Dulbecco's Modified Eagle's Medium

DNA - Deoxyribonucleic acid

DNASE1L3 **-** Deoxyribonuclease 1 like 3

dNTPs - Deoxynucleotide triphosphates

DOP-PCR- Degenerate oligonucleotide-primed Polymerase chain reaction

***DPP9* -** Dipeptidyl peptidase 9

***DUSP22* -** Dual specificity phosphatase 22

E

EBV - Epstein-Barr virus

ECACC - European Collection of Authenticated Cell Cultures

ECM - Extracellular matrix

***EFHB* -** EF-hand domain family member B

EGF - Epidermal growth factor

***EGFR* -** Epidermal growth factor receptor

eIF4E - Eukaryotic translation initiation factor 4E

***ELAVL1* -** ELAV like RNA binding protein 1

***ELMO3* -** Engulfment and cell motility 3

EMSY or ***C11orf30*** - EMSY, BRCA2 interacting transcriptional repressor

***ERBB4* -** Erb-b2 receptor tyrosine kinase 4

ERK1/2 - Extracellular signal-regulated kinase 1/2

***ESR1* -** Estrogen receptor 1

***EWSR1*-** EWS RNA binding protein 1

***EXT1* -** Exostosin glycosyltransferase 1

F

***FADD* -** Fas associated via death domain

FAK - Focal adhesion kinase

FAM - Fluorescein amidite

***FAM83A* -** Family with sequence similarity 83 member A

***FANCA* -** Fanconi anemia complementation group A

***FANCD2* -** Fanconi anemia complementation group D2

***FAT1* -** FAT atypical cadherin 1

FBS - Fetal bovine serum

***FBXL5* -** F-box and leucine rich repeat protein 5

***FDGF*-** Fibroblast-derived growth factor

FFPE - Formalin-fixed paraffin-embedded

FGF19 **-** Fibroblast growth factor 19

FGF3 **-** Fibroblast growth factor 3

FGF4 **-** Fibroblast growth factor 4

FHIT **-** Fragile histidine triad

FISH- Fluorescence *in situ* hybridization

FLCN - Folliculin

FLII - FLII, actin remodeling protein

FLNB **-** Filamin B

FLT4 **-**Fms related tyrosine kinase 4

FOLR3 **-** Folate receptor 3

FOXC1 **-** Forkhead box C1

FOXP1 **-** Forkhead box P1

FOXQ1 **-** Forkhead box Q1

FUT3 **-** Fucosyltransferase 3 (Lewis blood group)

FZR1 **-** Fizzy and cell division cycle 20 related 1

G

GAL3ST1 **-** Galactose-3-O-sulfotransferase 1

GALR1 **-** Galanin receptor 1

GATA4 - GATA binding protein 4

GATA5 **-** GATA binding protein 5

GFRAL **-** GDNF family receptor alpha like

GRINA **-** Glutamate ionotropic receptor NMDA type subunit associated protein 1

GSDMD **-** Gasdermin D

GSTP1 **-** Glutathione S-transferase pi 1

GSTT1- Glutathione S-transferase theta 1

GTF2H1 **-** General transcription factor IIH subunit 1

GTG - Giemsa - Trypsin - Giemsa

Gy – Gray

H

H2AFX **-** H2A histone family member X

HESX1 **-** HESX homeobox 1

HIC1 **-** HIC ZBTB transcriptional repressor 1

HIF **-** hypoxia-inducible factor

HIF3A - Hypoxia inducible factor 3, alpha subunit

HIGD1B **-** HIG1 hypoxia inducible domain family member 1B

HMGCLL1 **-** 3-hydroxymethyl-3-methylglutaryl-CoA lyase like 1

HN - Head and Neck

HNC - Head and neck cancer

HNSCC - Head and neck squamous cell carcinoma

HOXA10 -Homeobox A10

HPV - Human papilloma virus

HS3ST3B1 - Hparan sulfate-glucosamine 3-sulfotransferase 3B1

hsrs - Homogeneously staining regions

HUS1B **-** HUS1 checkpoint clamp component B

I

---

i(...) - Isochromosome

idic (...) - Isodicentric chromosome

IFI6 **-** Interferon alpha inducible protein 6

I-FISH - Interphase fluorescence *in situ* hybridization

IgG2 - Imunoglobulina G2

IL12B **-** Interleukin 12B

IL17RD **-** Interleukin 17 receptor D

IL2 **-** Interleukin 2

IL4 **-** Interleukin 4

ING2 **-** Inhibitor of growth family member 2

INPPL1- Inositol polyphosphate phosphatase like 1

IP6K2 **-** Inositol hexakisphosphate kinase 2

IQR - Inter- Quartile Range

IRAK2 **-** Interleukin 1 receptor associated kinase 2

ITGA9 **-** Integrin subunit alpha 9

J

JAK2 – Janus kinase 2

JARID2 **-** Jumonji and AT-rich interaction domain containing 2

JNK - c-Jun N-terminal kinase

K

KANK1 - KN motif and ankyrin repeat domains 1

KANSL1 - KAT8 regulatory NSL complex subunit 1

Kb - Kilobase

KCNK9 - Potassium two pore domain channel subfamily K member 9

KDR* or *VEGFR - Kinase insert domain receptor or vascular endothelial growth factor receptor 2

KIAA1456 - KIAA1456

KLF10 - Kruppel like factor 10

KLHDC8B - Kelch domain containing 8B

KLK3 - Kallikrein related peptidase 3

KLLN - Killin, p53-regulated DNA replication inhibitor

KMT2A - Lysine methyltransferase 2A

L

LAMA1 - Laminin subunit alpha 1

LINC01088 - Long intergenic non-protein coding RNA 1088

LINE - Long interspersed nuclear elements

LLGL1 - LLGL1, scribble cell polarity complex component

LLGL2 - LLGL2, scribble cell polarity complex component

LOC101927769 - Uncharacterized LOC101927769

LOH- Loss of heterozygosity

LOX - Lysyl oxidase

LPCAT1 - Lysophosphatidylcholine acyltransferase 1

LRP - Loco-regional progression

LRP12 - LDL receptor related protein 12

LRTM1 - Leucine rich repeats and transmembrane domains 1

LSCC - Laryngeal squamous cell carcinoma

LTR - Long Terminal Repeats

M

mAb - Monoclonal antibody

MAML3 - Mastermind like transcriptional coactivator 3

MANEAL -Mannosidase, endo-alpha-like

MAPKAPK5 **-** Mitogen-activated protein kinase-activated protein kinase 5

MAPKs - Mitogen-activated protein kinase

mar - Marker chromosome

MARCH7 **-** Membrane associated ring-CH-type finger 7

MATLAB - Matrix Laboratory

Mb - Megabase

MCCC1 **-** Methylcrotonoyl-CoA carboxylase 1

MCL1 **-** MCL1, BCL2 family apoptosis regulator

MCM4 **-** Minichromosome maintenance complex component 4

MDC1 **-** Mediator of DNA damage checkpoint 1

MDM2 **-** MDM2 proto-oncogene

MET **-** MET proto-oncogene, receptor tyrosine kinase

mFISH - Multicolor Fluorescence In Situ Hybridization

MGMT **-** O-6-methylguanine-DNA methyltransferase

MLH1 **-** MutL homolog 1

MLH3 **-** MutL homolog 3

MLPA - Multiplex Ligation-dependent Probe Amplification

MME **-** Membrane metalloendopeptidase

MRE11A- MRE11 homolog, double strand break repair nuclease

MRI - Magnetic resonance imaging

MRI1 **-** Methylthioribose-1-phosphate isomerase 1

mRNA - Messenger RNA

MSH6 **-** MutS homolog 6

MS-MLPA - Methylation-Specific Multiplex Ligation-dependent Probe Amplification

mTOR - Mammalian target of rapamycin

mTORC1 – mTOR complex1

mTORC2 – mTOR complex2

MTSS1 **-** MTSS1, I-BAR domain containing

MTUS1 **-** Microtubule associated scaffold protein 1

MVP **-** Major vault protein

MXD4 **-** MAX dimerization protein 4

MYC **-** MYC proto-oncogene, bHLH transcription factor

N

NA - Not available

***NAE1* -** NEDD8 activating enzyme E1 subunit 1

***NCKIPSD* -** NCK interacting protein with SH3 domain

***NEFL* -** Neurofilament light

***NFKB1* -** Nuclear factor kappa B subunit 1

NKD2 -Naked cuticle homolog 2

***NOP14* -** NOP14 nucleolar protein

***NPEPPS* -** Aminopeptidase puromycin sensitive

***NQO1* -** NAD(P)H quinone dehydrogenase 1

***NQO2* -** N-ribosyldihydronicotinamide:quinone reductase 2

***NSMCE2* -** NSE2/MMS21 homolog, SMC5-SMC6 complex SUMO ligase

O

OPMLs - Oral potentially malignant lesions

OR - Odds ratio

***ORAOV1* or *TAOS1* -** Oral cancer overexpressed 1

OSCC - Oral squamous cell carcinoma

P

---

p - Short arm of a chromosome

***P2RY2* -** Purinergic receptor P2Y2

***PAH* -** Phenylalanine hydroxylase

***PAK2* -** p21 (RAC1) activated kinase 2

***PALLD* -** Palladin, cytoskeletal associated protein

***PARD6A* -** Par-6 family cell polarity regulator alpha

***PAX5* -** Paired box 5

***PAX6* -** Paired box 6

***PAX9* -** Paired box 9

PC - Primary cell cultures

PCA - Principal component analysis

PCR- Polymerase chain reaction

PD-1 - Programmed cell death protein 1

PDF **-** Peptide deformylase (mitochondrial)

PEG10 **-** Paternally expressed 10

PET - Positron emission tomography

PFKFB4 **-** 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4

PGPEP1 **-** pyroglutamyl-peptidase I

PI3K - Phosphoinositide 3-kinase

PIK3CA **-** Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha

PIK3CB **-** phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta

PINX1 PIN2 **or** ***TERF1*** **-** interacting telomerase inhibitor 1

PKT - Protein tyrosine kinase

PLA2G1B **-** Phospholipase A2 group IB

PMP22 **-** Peripheral myelin protein 22

PMP22 - Peripheral myelin protein 22

POLD3 **-** DNA polymerase delta 3, accessory subunit

PP2D1 or ***C3orf48*** - Protein phosphatase 2C like domain containing 1

PPAN-P2RY11 **-** PPAN-P2RY11 readthrough

PPFIA1 **-** PTPRF interacting protein alpha 1

PPME1 **-** Protein phosphatase methylesterase 1

PPP4R2 **-** Protein phosphatase 4 regulatory subunit 2

PPP6C **-** Protein phosphatase 6 catalytic subunit

pRb - Retinoblastoma protein

PRDX1 **-** Peroxiredoxin 1

PRDX2 **-** Peroxiredoxin 2

PRIM2 **-** Primase (DNA) subunit 2

PRKCA **-** Protein kinase C alpha

PRKDC **-** Protein kinase, DNA-activated, catalytic polypeptide

PRPF19 - Pre-mRNA processing factor 19

PRUNE **-** Prune exopolyphosphatase 1

PTCH1 **-** Patched 1

PTEN **-** Phosphatase and tensin homolog

PTK2 **-** Protein tyrosine kinase 2

PTP4A3 **-** Protein tyrosine phosphatase type IVA, member 3

PTPRM **-** Protein tyrosine phosphatase, receptor type M

PTRH2 **-** Peptidyl-tRNA hydrolase 2

PUF60 **-** Poly(U) binding splicing factor 60

PUM1 **-** Pumilio RNA binding family member 1

PVT1 **-** Pvt1 oncogene (non-protein coding)

PYCARD **-** PYD and CARD domain containing

Q

q - Long arm of a chromosome

Q-MSP - Quantitative Methylation-Specific PCR

qRT-PCR - Quantitative reverse transcription PCR

QT - Chemotherapy

R

R0 - Negative resection margin

R1 - Positive resection margin

RAB25 **-** RAB25, member RAS oncogene family

RAB5A **-** RAB5A, member RAS oncogene family

RAD17 **-** RAD17 checkpoint clamp loader component

RAP1B **-** RAP1B, member of RAS oncogene family

RARB **-** Retinoic acid receptor beta

RASD1 - Ras related dexamethasone induced 1

RASL10A **-** RAS like family 10 member A

RASSF1 **-** Ras association domain family member 1

RB1 **-** RB transcriptional corepressor 1

RBM5 - RNA binding motif protein 5

RECQL4 - RecQ like helicase 4

RERE - Arginine-glutamic acid dipeptide repeats

RFX1 - Regulatory factor X1

RIN - RNA integrity number

RNA - Ribonucleic acid

RNF139 **-** Ring finger protein 139

RNF168 **-** Ring finger protein 168

RNF169 **-** Ring finger protein 169

RPM - Revolutions per minute

RPS15 **-** Ribosomal protein S15

RPS8 **-** Ribosomal protein S8

RRAGA - Ras related GTP binding A

RT - Radiotherapy

RTK - Receptor tyrosine kinase

RT-PCR – Reverse transcription PCR

S

S100PBP **-** S100P binding protein

S1PR2 **-** Sphingosine-1-phosphate receptor 2

S6K1 **-** Ribosomal protein S6 kinase beta-1

SAFB **-** Scaffold attachment factor B

SAP30BP **-** SAP30 binding protein

SCAI **-** Suppressor of cancer cell invasion

SCCs - Squamous cell carcinomas

SCRIB **-** Scribbled planar cell polarity protein

SCRT1 **-** Scratch family transcriptional repressor 1

SDHA - Succinate dehydrogenase complex flavoprotein subunit A

SEC13 **-** SEC13 homolog, nuclear pore and COPII coat complex component

SELENBP1 **-** Selenium binding protein 1

SEMA3F **-** Semaphorin 3F

SEPT14 **-** Septin 14

SERPINH1 **-** Serpin family H member 1

SESN2 **-** Sestrin 2

SF3B3 **-** Splicing factor 3b subunit 3

SGIP1 **-** SH3 domain GRB2 like endophilin interacting protein 1

SHARPIN **-** SHANK associated RH domain interactor

SHISA5 **-** Shisa family member 5

SINE - Short interspersed nuclear elements

SIRT6 **-** Sirtuin 6

SKY - Spectral karyotyping

SLC9A3 - Solute carrier family 9 member A3

SMAD4 **-** SMAD family member 4

SNP - Single Nucleotides Polymorphisms

SOX2 **-** SRY-box 2

SOX7 - SRY-box 7

SPIDR **or** ***KIAA0146*** **-** Scaffolding protein involved in DNA repair

SPIN1- Spindlin 1

SPSS – Statistical Package for the Social Sciences

SPT - Second primary tumor

STAG2 **-** Stromal antigen 2

STAT3 **-** Signal transducer and activator of transcription 3

STAT5 **-** Signal transducer and activator of transcription 5A

STK11 **-** Serine/threonine kinase 11

STK19 **-** Serine/threonine kinase 19

STR - Short Tandem Repeats

SUMO2 **-** Small ubiquitin-like modifier 2

SVM - Support vector machine

SYT7 - Synaptotagmin 7

T

TADA3 - Transcriptional adaptor 3

TAGLN **-** Transgelin

TAOK2 **-** TAO kinase 2

TCGA **-** The Cancer Genome Atlas

TFPI2 **-** Tissue factor pathway inhibitor 2

THBS1 **-** Thrombospondin 1

TIMP **-** TIMP metallopeptidase inhibitor 1

TIMP3 **-** TIMP metallopeptidase inhibitor 3

TKI **-** Tyrosine kinase inhibitor

TMEM161A **-** Transmembrane protein 161A

TMEM16A or ANO1 **-** Transmembrane protein 16A or Anoctamin 1

TNFRSF1A **-** TNF receptor superfamily member 1A

TNIP2 **-** TNFAIP3 interacting protein 2

TNM **-** Tumor/node/metastasis classification

TNXB **-** Tenascin XB

TP53- Tumor protein p53

TP63 **-** Tumor protein p63

TP73 **-** Tumor protein p73

TRADD **-** TNFRSF1A associated via death domain

TREX1 **-** Three prime repair exonuclease 1

TRIAP1 **-** TP53 regulated inhibitor of apoptosis 1

TRIM37 **-** Tripartite motif containing 37

TRIM71 **-** Tripartite motif containing 71

TTYH2 **-** Tweety family member 2

TUSC3 **-** Tumor suppressor candidate 3

TFPI2 **-** Tissue factor pathway inhibitor 2

U

UBA7 **-** Ubiquitin like modifier activating enzyme 7

UBE2K **-** Ubiquitin conjugating enzyme E2 K

UCSC - University of California, Santa Cruz genome browser

UGDH **-** UDP-glucose 6-dehydrogenase

UGT2B15 **-** UDP glucuronosyltransferase family 2 member B15

UGT2B28 **-** UDP glucuronosyltransferase family 2 member B28

UIMC1 **-** Ubiquitin interaction motif containing 1

V

VEGF **-** Vascular endothelial growth factor

VHL **-** Von Hippel-Lindau tumor supressor

W

WISP1 **-** WNT1 inducible signaling pathway protein 1

WNT11 **-** Wnt family member 11

WNT5A **-** Wnt family member 5A

WT1 **-** Wilms tumor 1

WWP2 **-** WW domain containing E3 ubiquitin protein ligase 2

X

XRCC2 **-** X-ray repair cross complementing 2

XRCC6 **-** X-ray repair cross complementing 6

XRRA1 **-** X-ray radiation resistance associated 1

Y

YAC - Yeast artificial chromosomes

Z

ZC3H8 **-** Zinc finger CCCH-type containing 8

ZIC1 **-** Zic family member 1

ZIC4 **-** Zic family member 4

ZNF385C **-** Zinc finger protein 385C

ZNRF3 **-** Zinc and ring finger 3

ZRANB1 **-** Zinc finger RANBP2-type containing 1

Abstract

Head and neck squamous cell carcinoma (HNSCC) is an emergent health problem worldwide. These tumors present heterogeneity at phenotypic, aetiological, biological and clinical level. In developed countries, smoking and alcohol are implicated in the increase of HNSCC cases, and human papillomavirus is an important risk factor, especially in the rise of oropharyngeal tumors without smoke and alcohol habits. A significant percentage of HNSCC patients develop loco-regional and distant recurrences. Even with progresses in surgery, radiation and chemotherapy, approximately half of all patients die of the disease. Risk stratification for HNSCC is essential in order to decrease mortality and improve quality of life of the patients. The great HNSCC heterogeneity makes difficult to understand the molecular carcinogenesis process as well as to develop early detection and therapeutic strategies for these tumors. Nowadays, the majority of genome-wide molecular profiling studies of HNSCC are limited to single approaches, which hampers the identification of accurate and robust biomarkers of early diagnosis and prognosis. Indeed, there is a lack of proven biomarkers for predicting clinical outcomes and response to treatment. The present work aimed to perform a molecular characterization of HNSCC in order to predict recurrence/metastasis development and signaling pathways associated to targeted therapy and resistance to conventional drugs through the identification of different molecular groups with apparently different survival profiles using genomic, epigenetic and transcriptomic approaches. We analyzed the same HNSCC patients through different molecular technologies, being the identified biomarkers and molecular signatures validated with TCGA (The Cancer Genome Atlas) data. First, we performed a direct genetic and epigenetic characterization of HNSCC patients, using specific Multiplex Ligation-dependent Probe Amplification (MLPA) and Methylation-Specific MLPA (MS-MLPA) probe panels. We reported different genetic signatures related to tumor stage and anatomic site as well as tobacco use. Additionally, specific genomic and epigenetic signatures associated to patients' risk of recurrence/metastasis development after treatment of primary tumor and also to survival were identified. The genetic analysis of non-tumor samples (from surgical margins) revealed some imbalances similar to those identified in the tumor samples, which reinforce the importance of molecularly analyze the high-risk patients even before the visible morphological changes and also the suspicious lesions in order to early diagnose these

tumors and their recurrences. Secondly, we moved forward to a high-throughput genomic and transcriptomic approaches and we identified molecular signatures with capability to predict the recurrent/metastatic disease development and clinical outcome. In these studies, using either direct probe panels or genome-wide approaches we identified the most common chromosomal regions with imbalances and altered genes. As expected, whole-genome techniques revealed new chromosomal regions and genes that seem to have a role in HNSCC development and behavior. Overall, through these comprehensive genomic, epigenetic and transcriptomic characterization we identified biomarkers and molecular signatures of prognosis and survival, which open the door for personalized medicine in HNSCC patients. Finally, we applied these genomic and epigenetic technologies to perform a molecular characterization of four paradigmatic HNSCC patients in order to prove the benefit of these molecular knowledgement to the clinical management of the HNSCC patients. Several chromosomal regions and genes related to radiation and/or chemotherapy resistance and to patients' prognosis and survival were identified, which could help and guide the type or intensify of treatment modalities. Moreover, molecular characterization of commercial HNSCC cell lines and primary cell cultures established from these patients was conducted, which revealed the ploidy and the complex structural chromosomal rearrangements of HNSCC tumors. This comprehensive characterization enables cell models for further studies both in radiation and pharmacogenomics fields, as well as to understand the molecular mechanisms of HNSCC development and progression. With this work we performed a robust molecular characterization of HNSCC, using different omic approaches in the same tumor samples, which allowed the identification of new prognostic biomarkers and molecular signatures with potential to be translated to clinical practice.

Keywords: Biomarkers; Molecular diagnosis and prognosis; Genetic and epigenetic signatures; Gene expression; Methylation; Copy number alterations; Recurrence/metastasis development; Oral cavity tumors; Head and neck squamous cell carcinoma; Omics integration; Non-invasive screening

Resumo

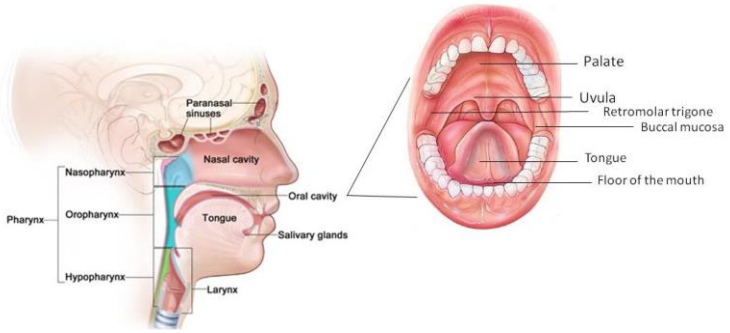
O carcinoma epidermóide da cabeça e pescoço (CECP) é um problema emergente de saúde em todo o mundo. Estes tumores são heterogéneos a nível fenótipico, etiológico, biológico e clínico. Nos países desenvolvidos, o tabaco e o álcool estão implicados no aumento do número de casos de CECP e o papiloma vírus humano é um fator de risco importante para o aumento dos tumores da orofaringe não relacionados com hábitos tabágicos e de álcool. Uma percentagem significativa de doentes com CECP desenvolve recidivas loco-regionais e à distância. Mesmo com os progressos na cirurgia, radioterapia e quimioterapia, cerca de metade de todos os doentes morre devido ao CECP. A estratificação do risco de CECP é essencial de forma a contribuir para a diminuição da mortalidade e melhoria da qualidade de vida destes doentes. A heterogeneidade do CECP dificulta por um lado a compreensão dos processos moleculares da carcinogénese e por outro lado o desenvolvimento de estratégias de deteção precoce e de terapêutica. Atualmente, a maioria dos estudos moleculares de grande escala são restritos, o que dificulta a identificação robusta e precisa de biomarcadores de diagnóstico e prognóstico. De facto, há falta de biomarcadores para predizer o desenlace clínico e resposta ao tratamento. O presente trabalho teve como objetivo caraterizar molecularmente o CECP de forma a prever o desenvolvimento de recidivas/metástases e a identificação de vias de sinalização associadas a terapias alvo e resistência às terapias convencionais, através da identificação de diferentes grupos moleculares com diferentes sobrevivências, usando abordagens de genómica, epigenética e transcriptómica. Neste estudo, analisámos os mesmos doentes com CECP usando diferentes tecnologias moleculares, tendo validado os biomarcadores e assinaturas moleculares identificados usando dados do portal TCGA (The Cancer Genome Atlas). Em primeiro lugar, realizámos uma caraterização genética e epigenética do CECP direcionada, utilizando painéis de sondas específicos de *Multiplex Ligation-dependent Probe Amplification* (MLPA) e *Methylation-Specific* MLPA (MS-MLPA). Identificaram-se diferentes assinaturas genéticas relacionadas com o estadio do tumor e as localizações anatómicas, bem como com o consumo de tabaco. Adicionalmente, uma assinatura genética e epigenética associada ao risco dos doentes desenvolverem recidivas/metástases após o tratamento do tumor primário e também associada à sobrevivência, foi identificada. A análise genética das amostras não tumorais (provenientes das margens cirúrgicas) revelou alguns desequilíbrios similares aos identificados nas amostras tumorais, o que reforça a importância de analisar

molecularmente os doentes de elevado risco mesmo antes de qualquer alteração morfológica visível e também das lesões suspeitas, de forma a diagnosticar precocemente estes tumores e as suas recidivas. Na segunda parte do estudo utilizámos abordagens genómicas e transcriptómicas de larga escala e, identificámos assinaturas moleculares capazes de prever o desenvolvimento de recidivas/metástases e evolução clínica dos doentes. Estes estudos, usando quer painéis de sondas direcionados quer abordagens de todo o genoma, permitiram identificar as regiões cromossómicas e genes mais comummente alterados. As técnicas de análise de todo o genoma revelaram novas regiões cromossómicas e genes que parecem desempenhar um papel no desenvolvimento e evolução clínica do CECP. No geral, através desta caraterização genómica, epigenética e trasncriptómica, identificámos biomarcadores e assinaturas moleculares de prognóstico e sobrevivência, o que abre novas portas para a medicina personalizada no CECP. Finalmente, utilizámos estas tecnologias de genómica e epigenética para caraterizar quatro doentes paradigmáticos com CECP de forma a provar o benefício deste conhecimento molecular na conduta clínica. Várias regiões cromossómicas e genes relacionados com a resistência à radiação, quimioterapia, prognóstico e sobrevivência foram identificados, o que poderia ajudar na escolha do tipo e intensidade das modalidades de tratamento. Adicionalmente, foi realizada a caraterização molecular de linhas comerciais de CECP e de culturas primárias estabelecidas a partir destes doentes de CECP, o que revelou a ploidia e rearranjos estruturais complexos destes tumores, garantindo modelos celulares para futuros estudos no campo da radiação e farmacogenómica e ainda para uma melhor compreensão dos mecanismos moleculares de desenvolvimento e progressão do CECP. Este trabalho permitiu, de uma forma robusta, caracterizar molecularmente o CECP, usando diferentes abordagens ómicas nas mesmas amostras tumorais, ajudando assim a identificar novos biomarcadores de prognóstico e assinaturas moleculares com potencial translação à clínica.

Palavras-chave: Biomarcadores; Diagnóstico e prognóstico molecular; Assinaturas genéticas e epigenéticas; Expressão genética; Metilação; Alterações do número de cópias; Desenvolvimento de recidivas/metástases; Tumores da cavidade oral; Carcinoma epidermóide da cabeça e pescoço; Integração ómica; Rastreio não-invasivo

1. General Introduction

Head and neck cancer (HNC) is a general term that describes malignancies originating from a variety of sites in the upper aerodigestive tract, being the oral and nasal cavities, pharynx, and larynx the most affect sites (Figure 1) [Ragin et al. 2007].



Chapter 1 - Figure 1. The most common anatomic sites of HNC.

The most common (~90%) histologic type of HNC is squamous cell carcinoma (HNSCC) [Leemans et al. 2011]. This carcinoma is frequently aggressive in its biologic behavior, so patients exhibit malignant phenotypes with invasion of surrounding tissue and distant metastasis.

At the time of the diagnosis, the majority (60–70%) of the HNSCC patients present loco-regionally advanced disease, but only 10% present metastatic disease [Argiris et al. 2008]. Patients presenting early stage (stage I and II) HNSCC remain frequently disease-free after treatment, while patients with more advanced disease stage (stage III and IV) often relapse, either loco-regionally and/or at distant sites [Colevas 2006; Vermorken 2005]. Patients who develop an inoperable recurrence or metastasis have a poor prognosis with a median survival of 6-8 months [Goerner et al. 2010]. Approximately 33-50% of patients in western countries relapse [Molin and Fayette 2011]. Within two years after first diagnosis, almost half of advanced stage HNSCC patients develop loco-regional relapses regardless of the treatment modalities [Cooper et al. 2004] and even 10-30% of the patients that develop local recurrent disease had histopathologically tumor-free surgical margins after resection of the primary tumor [Tabor et al. 2001]. Despite advances in therapy and diagnostic approaches, the long-term survival of HNSCC patients remains poor, with only 40‑50% of patients surviving for 5 years [Leemans et al. 2011]. The frequent late diagnosis, the development of metastasis, locoregional recurrences and second primary tumors are referred as the major reasons for the lack of significant improvements in the survival rate of HNSCC in the last decades.

HNC patients need to deal not only with a life-threatening diagnosis but also with the serious disfiguring and functional consequences of the treatment, being HNC for that a psychologically traumatic cancer.

Overall awareness of HNC is low compared with other cancers and includes a lack of knowledge of the term “head and neck cancer”, its common symptoms and risk factors, which is unfortunate since the risk factors could be avoided and HNC prognosis is good when early detected [Luryi et al. 2014].

Epidemiology, Etiology and Risk Factors

Worldwide, HNSCC is the sixth most common neoplasm and represents the third cause of cancer death [Galbiatti et al. 2013], with 900 000 new cases and 300 000 deaths annually [van Monsjou et al. 2013]. It is mainly a disease of the adult and aging population, almost 98% of the diagnoses occur in patients older than 40 years and 50% in patients older than 60 years [Mehanna et al. 2010] and is more common in males (2:1) [Leemans et al. 2011; Lim et al. 2016]. However, this disparity in the male/female ratio has become less marked in the last years probably due to the more similar exposure of both genders to the HNSCC risk factors. Likewise, the incidence of oropharyngeal cancer has been on the rise, particularly among individuals under 45 years of age [Pai and Westra 2009]. The incidence of this malignancy is higher in developed countries. India presents the highest rate of HNSCC, accounting for approximately 25–30% of all cancer cases [Ferlay et al. 2013], mainly due to the lifestyle [Nagpal and Das 2003].

The primary risk factors related to HNSCC development include tobacco use, alcohol consumption, human papilloma virus (HPV) infection (in oropharyngeal cancer) and Epstein-Barr virus (EBV) infection (in nasopharyngeal cancer).

Tobacco smoking is a well-established independent risk factor for HNSCC, correlating with the intensity and duration of the smoking habit [Hashibe et al. 2007]. This risk increases in patients which smoked for more than 20 years and more than 20 daily cigarettes smoked [Petti 2009]. The cigarette has approximately 4 700 substances, and at least 50 of these are carcinogenic, including nitrosamines and polycyclic hydrocarbons, which have genotoxic effects and therefore can change the molecular profile of the individuals, causing mutations that lead to malignant transformation of normal cells [Choi and Myers 2008; Galbiatti et al. 2013]. For example, HNSCC patients that were smokers presented more frequently mutations in the *TP53* gene [Pai and Westra 2009]. Smoking cessation reduces

but does not eliminate the risk of cancer development [Kumar et al. 2008]. Patients who keep smoking habits during radiotherapy are more likely to develop osteoradionecrosis and to require hospitalization, causing the continuation of the smoking habit during treatment to be associated with an adverse effect on local control and survival [Shaw and Beasley 2016]. Betel quid - a chewed preparation composed of areca nut, lime, and sometimes tobacco, all wrapped in betel leaf is an important risk factor especially in Southeast Asia where there is a high incidence of oral cavity cancers [Trivedy et al. 2002].

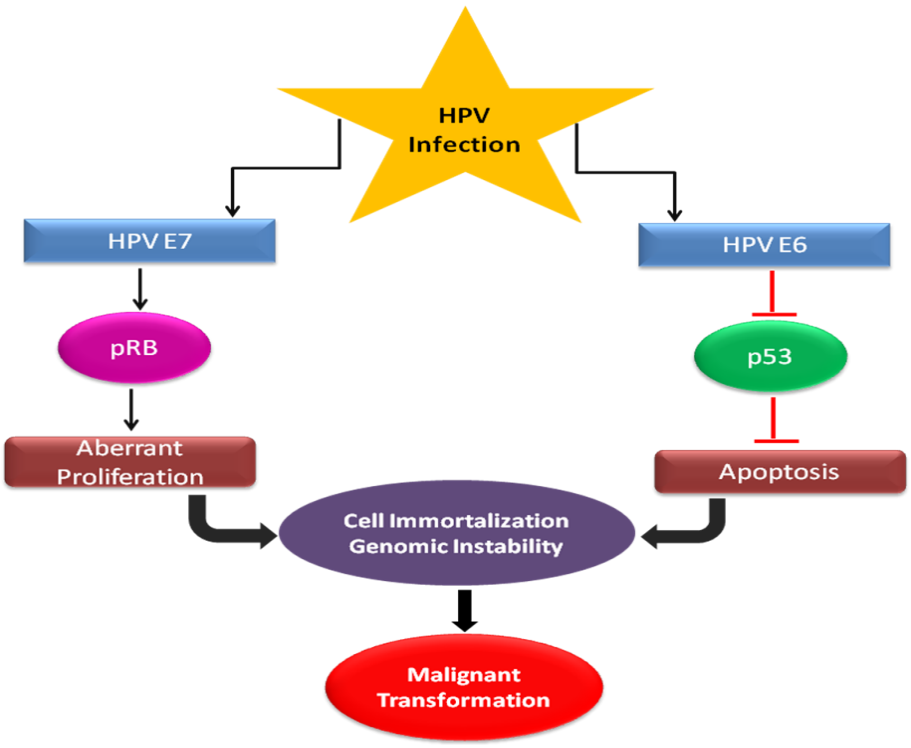
Alcohol does not seem to be carcinogenic *per se*; however, acts as a solvent to enhance mucosal exposure to carcinogens, increasing their cellular uptake. The acetaldehyde, a metabolite of alcohol, can form DNA adducts, that interfere with DNA synthesis and repair [Poschl and Seitz 2004]. Patients that keep drinking heavily after treatment for HNSCC have a significantly worse quality of life [Potash et al. 2010] with a negative impact on survival [Fortin et al. 2009]. Only after more than 20 years of quitting alcohol the risk of developing HNC reaches the values of non-drinkers [Marron et al. 2010].

Alcohol and tobacco have a synergistic effect, so their simultaneous consumption increases the HNSCC risk in about 40-50 fold [Marur and Forastiere 2008; Rodriguez et al. 2004]. Since not all smokers and alcohol users develop HNSCC, the individual variation in the genetic susceptibility to the malignant transformation of the cells is also of paramount importance [Liang et al. 2012]. Prevalence of smoking and alcohol consumption has declined in the last years but the overall incidence of HNSCC has risen, which seems to be due to the increase in HPV infection [Walden and Aygun 2013].

Human papiloma virus, a group of small circular, double-stranded DNA virus, is in the origin of several proliferative epithelial lesions. There are several types of HPVs that infect the human mucosa. These types are classified as high-risk and low-risk HPV, being almost 90% of HPV-related HNC linked to the presence of HPV type 16, a high-risk HPV type [Tornesello et al. 2014]. The high-risk HPV types are considered risk factors in about 25% of HNSCC, independent of other known risk factors, such as alcohol and tobacco [Hashibe et al. 2007]. Overall, HPV positive HNSCC patients have better outcomes, with tumors more responsive to treatment, a better survival and consequently a lower risk of dying from disease [Tornesello et al. 2014]. So, positive HPV status is an important prognosticator of response to therapy and survival.

The mode of transmission of HPV in HNC has not been fully determined, but has been linked to patterns of sexual activity, mainly a history of multiple sexual partners, early age of first intercourse, and frequent oral sexual practices [D'Souza et al. 2007].

HPV mediates carcinogenesis mainly through its E6 and E7 viral oncogenes. E6 oncoprotein targets TP53 for proteasomal degradation in an ubiquitin dependent way and E7 oncoprotein binds to and inhibits the pRb, which promotes the aberrant cell proliferation (Figure 2) [Leemans et al. 2011].



Chapter 1 - Figure 2. Schematic representation of malignant transformation in HNSCC mediated by the HPV oncoproteins E6 and E7.

Some inherited disorders that predispose to HNSCC are described, namely Fanconi's anemia, hereditary non-polyposis colorectal cancer, Li-Fraumeni syndrome and ataxia telangiectasia [Argiris et al. 2008].

The diet can also be associated with decreased risk for HNSCC. Diet with high consumption of fruit and vegetable and low intake of red meat, was associated with HNSCC reduced risk [Chuang et al. 2012].

Deficient oral hygiene, dental plaque formation and chronic irritation to the lining of the mouth are also risk factors for HNSCC development. The formation of polymicrobial supragingival plaque has a relevant mutagenic interaction with saliva which might trigger the development of oral cavity carcinomas, and periodontal diseases resulting from poor oral hygiene can lead to infections with consequent release of inflammatory mediators such as cytokines and the reactions against inflammation can also promote cancer development [Galbiatti et al. 2013]. The loss of teeth leads to the alterations of oral flora, with the reduction of nitrites and nitrates and the production of

acetaldehyde, which promotes the formation of DNA adducts [Bloching et al. 2007; Choi and Myers 2008].

The family history might also play an important factor in HNSCC development due to the inheritable genetic factors related to HNSCC risk, such as genetic polymorphisms involved in the carcinogens metabolism and DNA repair [Galbiatti et al. 2013]. Occupational activity also seems to be associated with HNSCC. Manual occupational activities, low income, low occupational-social class, low educational attainment and unemployment correlate with increased risk for HNSCC development [Conway et al. 2010], due to frequent exposure to sunlight and carcinogenic substances.

Signs and symptoms

HNSCC development is a multi-step process involving the transition from potentially malignant lesions to metastatic tumors. The 5-year survival of HNSCC patients is directly linked to stage at diagnosis, being in this way early detection the cornerstone to improve the survival and quality of life of the survivors.

Precursor lesions and earliest stages of HNSCC are often very subtle and asymptomatic, while later stage symptoms, depending on the primary cancer site, often include nonhealing sores or ulcers in the mouth, bleeding, loosening of teeth, difficulty wearing dentures, dysphagia, dysarthria, odynophagia, hoarseness, nasal bleeding or blockage, pain in the ear and growth of a neck mass [Marur and Forastiere 2008; Neville and Day 2002; Sanderson and Ironside 2002]. Enlargement of cervical lymph nodes, as the first symptom, is frequent in the tongue base, supraglottis and nasopharynx tumors [Sanderson and Ironside 2002]. The invasive oral squamous cell carcinoma is sometimes preceded by clinically identifiable potentially malignant lesions of the oral mucosa, which could be either white or red patches, named leukoplakia and erythroplakia that with malignant transformation evolve to a non-healing ulcer [Neville and Day 2002].

The head and neck (HN) area is highly sensitive to pain due to its rich innervation and the confinement of many anatomical structures to a small space, and thus the persistent pain could be due to a direct tumor effect or nerve invasion.

It is important to refer that the general public has a scant knowledge about the signs, symptoms and risk factors of HNSCC, and since this carcinoma is preventable and curable through avoidance of risk factors and treatment at early stages, increased public awareness should benefit both prevention and survival [Luryi et al. 2014].

Screening and diagnosis

Numerous efforts have been dedicated to the development of new strategies for detection, diagnosis and treatment of HNSCC. Screening for this neoplasm would be crucial to early diagnosis and consequently treatment with less aggressive modalities, which would improve the quality-of-life and the overall survival of the patients.

Medical screening is a strategy that allows detecting a disease or a high probability of developing a disease in individuals without any signs or symptoms [Awan 2014]. HNC screening is based on thorough history and physical examination, namely visual and tactile examination of the oral cavity and neck, as well as indirect mirror or direct fiber optic examination of nasal cavity, oropharynx, larynx and hypopharynx [Luryi et al. 2014]. Visual and tactile examination is subjective and dependent of the clinicians' training and skills. Biopsy of suspicious lesions followed by histopathological analyses is still the core method to identify and diagnose malignant disorders, presenting the major limitations of being invasive, expensive and lengthy, as well as the inability to perform several analyses in short periods of time during the course of treatment and in the subsequent years of follow up [Awan 2014].

Some methods that might help in the early diagnosis of these tumors include brush biopsy, toluidine blue staining, auto fluorescence and spectroscopy. The identification of putative biomarkers in body fluids using genomic, transcriptomic, proteomic and metabolomic approaches is also being intensively explored, but without implementation in clinical practice until now [Guerra et al. 2016].

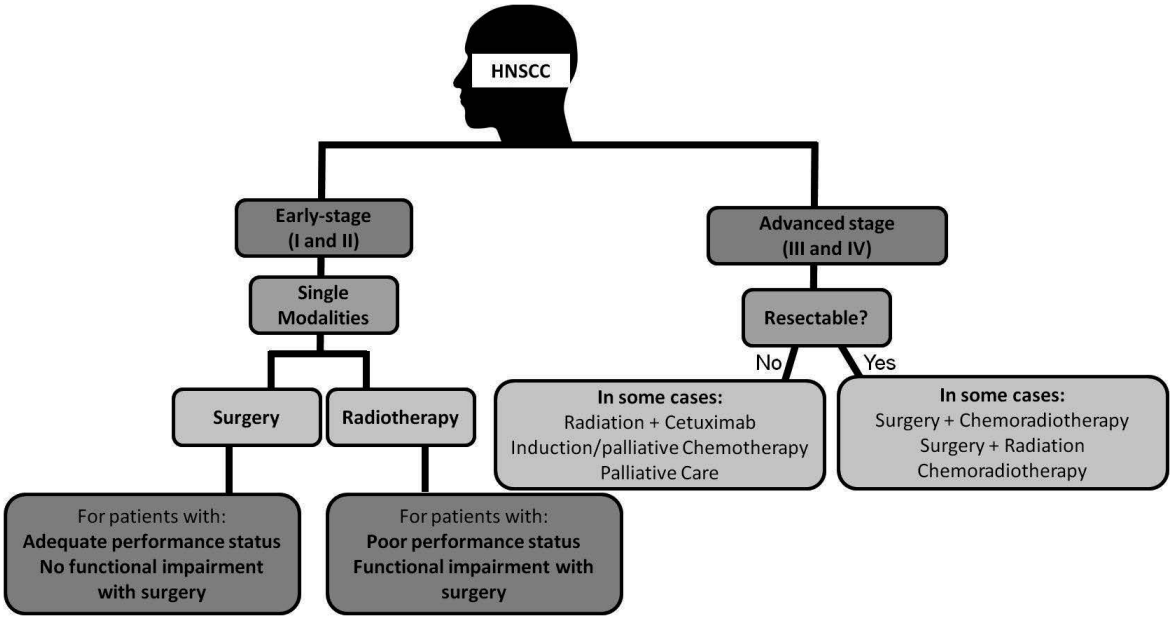
There are several imaging techniques, such as nasopharyngolaryngoscopy, computed tomography (CT), magnetic resonance imaging (MRI) and positron emission tomography (PET), habitually used to observe the extent of the primary tumor, invasion, regional lymph node status, and distant metastatic disease, guiding in the HNSCC staging [Neville and Day 2002].

Treatment

Currently, HNSCC is still classified only according to the T (size of tumor) N (involvement of loco-regional lymph nodes) and M (evidence of distant metastasis) system with different criteria for different primary sites. The tumors are staged I-IV with stages III-IV representing a more advanced disease. The tumor stage is vital in

clinical practice to guide the therapeutic decision making and to compare outcomes of therapy; however, this classification system does not take into account the variability in clinical outcome, lacking for instance biological and molecular markers, which leads to the use of the same treatment approach for different tumors [Prince et al. 2010].

Treatment for HNC is complex and usually debilitating. There are three traditional treatment modalities for HNSCC: surgery, chemotherapy, and radiotherapy (RT), which are characteristically aggressive and cause side effects that hamper the patient quality of life (Figure 3).



Chapter 1 - Figure 3. Most common treatment algorithm for HNSCC patients; however, additional options and combinations could be followed by clinicians.

During the last years, the timing and combinatorial approaches of these treatment options have changed, with inclusion of induction chemotherapy and concurrent chemoradiation therapy (CTR).

Induction chemotherapy is given before surgery or radiation, and CRT is given during radiotherapy.

Surgical techniques have suffered advances namely with the development of microvascular free tissue transfer and the endoscopic and robotic techniques [Prince et al. 2010]. In conventional RT, the primary tumor and gross adenopathy are irradiated with 70 Gy, delivered in single fractions of 1.8 to 2.0 Gy [Marur and Forastiere 2008]. Nowadays, new techniques to better locoregional control consist in the alteration of dose and frequency of irradiation, such as hyperfractionated RT that consists of giving to the patients treatments more often but at lower doses (1.2 Gy twice daily; a total of 81.6 Gy over the

course of 7 weeks), accelerated fractionation RT that consists of giving to the patients more irradiation in a shorter period of time (1.6 Gy twice daily; 67.2 Gy over the course of 6 weeks) and accelerated fractionation with a concomitant boost (1.8 Gy daily and 1.5 Gy daily as a boost for the last 12 days only; 72 Gy over the course of 6 weeks) [Marur and Forastiere 2008]. The chemotherapeutic drugs more extensively used in HNSCC are methotrexate, cisplatin, 5-fluorouracil (5-FU) and taxanes [Molin and Fayette 2011].

Early diagnosed HNSCC (stage I or II) has a good prognosis and have been successfully treated with single modalities, surgery or radiation [Cognetti et al. 2008], with a 90% cure rate [Schutt et al. 2012].

The standard treatment modalities for patients with locally advanced and regional HNSCC (stage III and IV) are primary surgery followed by chemoradiotherapy. Induction chemotherapy is also an option that can be considered, since a 5% increase in survival was reported when performed with cisplatin and 5-fluorouracil [Pignon et al. 2000]. Additionally, a reduction in metastases and an increase in local control of the primary tumor were also observed, which is particularly relevant, not only because at 2 years, almost 20% of patients with locally advanced disease, initially treated with concurrent chemoradiotherapy present distant metastasis, but also because the reduction of tumor size can turn an unresectable lesion into a tumor eligible for surgery [Haddad and Shin 2008]. All treatments present side-effects, and neutropenia, febrile neutropenia, and diarrhea are linked to induction chemotherapy, causing this treatment to be most useful in younger patients with good performance status, large primary tumors or bulky lymph node involvement, and very symptomatic disease [Haddad and Shin 2008].

Postoperative chemotherapy should be offered to patients that present positive margins, extracapsular invasion and the involvement of many lymph nodes [Schutt et al. 2012].

Frequently, patients with a loco-regional recurrence or metastatic disease are only eligible for palliative treatment, which includes supportive care, single-agent chemotherapy, combination chemotherapy or targeted therapies either alone or in combination with cytotoxic agents [P. M. Specenier and Vermorken 2008].

The current treatment modalities present tremendous side effects that limit their use in some patients, as example the disfiguring and the loss of important head and neck functions as result of surgery and radiation, or internal organ damage and dysfunction

caused by chemotherapy, specially the platinium salts that have been linked to neuropathy, nephrotoxicity, hearing loss, nausea, and vomiting [Tsao et al. 2006].

At the moment, cetuximab (Erbitux®), a chimeric monoclonal antibody of the immunoglobulin G1 class that binds with high affinity to the extracellular domain of the human EGFR (Epidermal Growth Factor Receptor), is the only targeted agent approved for the treatment of HNSCC, namely in combination with RT for the treatment of locally advanced disease and in combination with platinum based chemotherapy for the treatment of recurrent and/or metastatic disease.

EGFR is a transmembrane tyrosine kinase receptor overexpressed in almost 90% of HNSCC [Ang et al. 2002]. Thus, the high expression of EGFR ligands in HNSCC induces dimerization of EGFR, autophosphorylation of its intracellular kinase domain and, consequently the activation of several oncogenic pathways related to tumor cell growth, angiogenesis, and invasion [Kalyankrishna and Grandis 2006].

The affinity of Cetuximab for EGFR is approximately 5-10 fold higher than that of the endogenous ligands leading to the inhibition of the receptor function by blocking of ligand binding, prevention of receptor dimerization, and induction of the internalization and degradation of EGFR [P. Specenier and Vermorken 2013]. Cetuximab together with RT in the treatment of locally advanced HNSCC presented a benefit of 10% in the overall survival of the patients without increasing toxicity [Bonner et al. 2006].

There are other targeted agents under evaluation, namely panitumumab (a fully humanized anti-EGFR IgG2 mAb), erlotinib and gefitinib (tyrosine kinase inhibitors, TKIs, of EGFR), lapatinib (a dual TKI of EGFR and human epidermal growth factor receptor 2), sorafenib (a multitargeted TKI) and bevacizumab (a humanized anti-vascular endothelial growth factor monoclonal antibody), either in monotherapy and/or combined therapies for locally advanced or in recurrent/metastatic disease [Frampton 2010].
Cancer immunotherapy is based in the principle that tumors can be recognized as foreign rather than as self and, therefore, can be effectively attacked by an activated immune system. So, tumor progression results from the acquisition of traits that permit cancer cells evading immune surveillance and, consequently an effective immune response [Ferris 2015]. Immunotherapy, most significantly the immune checkpoint inhibitors targeting cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4) and programmed cell death protein 1 (PD-1) holds great potential to fight HNSCC [Ferris 2015].
In this sense, the therapeutic promise for HNSCC patients reside in immunomodulatory therapies that are able to overcome immune suppressive signals, such as cancer vaccines

using tumor peptide antigens, viral, bacterial, and DNA-based vector and tumor antigen-specific monoclonal antibodies [Ferris 2015].

Cancer treatment requires a fragile balance between complete eradication of tumor and preservation of the anatomical form and function of the organs. Regarding HN there are unique challenges in the treatment of these patients since this region supports several vital functions such as respiration, articulation, mastication, and deglutition. Quality of life is also an important factor in clinical decision-making for treatment planning since HN structures are the basis for human social interactions and individual identity of the person [Prince et al. 2010].

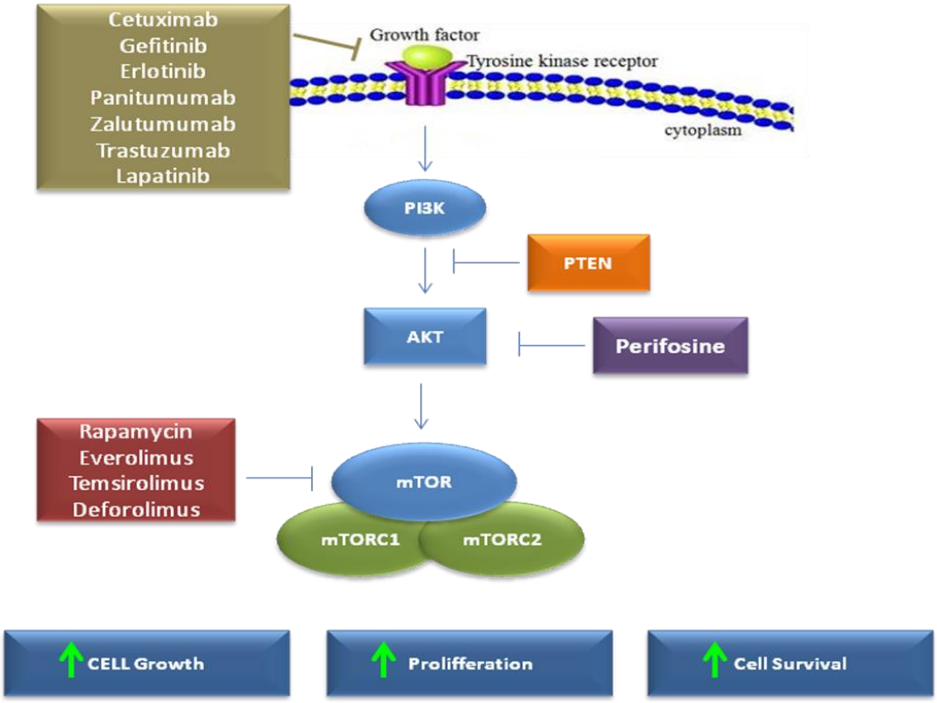
The choice of treatment modalities should take into account the performance status, co-morbidities prior treatments, symptoms and also patient preference. It is important to reinforce that HNC has a huge psychological traumatic impact in the patients, due to the negative effects of the tumor and also the consequences of the treatment in the swallowing, speaking, tasting, and smelling, as well as at esthetic level, which create difficulties with the social interactions leading to depression, anxiety, isolation and shame, being more than one third of HNC patients confronted with psychological problems [Zabora et al. 2001].

Molecular alterations in HNSCC

HNSCC is a genetically and biologically heterogeneous disease that results from the accumulation of numerous genetic and epigenetic alterations in several signaling pathways, which hampers not only the efforts to precisely predict the prognosis, but also the identification of the altered genes that are behind the origin of tumor development and could become targets for new therapies [Leemans et al. 2011].

It is crucial to understand and identify the full spectrum of molecular alterations of HNSCC in order to identify new therapy targets. Molecular alterations linked to different pathways involved in proliferation, squamous epithelial differentiation, cell survival, invasion and metastasis were reported in HNC [Rothenberg and Ellisen 2012]. EGFR-PI3K-AKT-mTOR signaling cascade seems to be a vital player in the tumorigenesis of HNSCC, where several components of this pathway such as EGFR, PI3K and mTOR have been shown to be highly activated in HNSCC patients, due to genetic and epigenetic alterations [Freudlsperger et al. 2011]. This pathway is attractive for molecular target therapy (Figure 4).

Nowadays, targeting EGFR was introduced in the HNSCC therapy, with cetuximab being used in clinical practice in patients with locally advanced tumors and recurrent or metastatic disease. However, numerous preclinical and clinical studies are still evaluating the role of different inhibitors of this pathway to treat HNSCC, being this a hard task, as frequently clinical trials did not confirm the expected results, probably due to the complexity of this pathway and the crosstalk and redundancy with other pathways [Freudlsperger et al. 2011].



Chapter 1 - Figure 4. Overview of EFGR-PI3K-AKT-mTOR signaling pathway with some targeted agents under clinical trials for the treatment of HNC.

- **Epidermal Growth Factor Receptor (EGFR) signaling**

EGFR signaling has been related to carcinogenesis, tumor progression and response to therapy in HNSCC. This was the first receptor tyrosine kinase (RTK) to be discovered. The receptors of the EGFR family all are composed of an extracellular ligand binding domain, a single transmembrane domain and a cytoplasmic domain containing a conserved protein tyrosine kinase (PTK) core, flanked by regulatory sequences [Schlessinger 2002]. The activation of receptor tyrosine kinases occurs when ligands bind to EGFR receptors, leading to a dimerization that is followed by intermolecular autophosphorylation of key tyrosine residues [Stadler et al. 2008]. Phosphorylated receptor has the potential to trigger several oncogenic intracellular

signaling pathways and, consequently hallmarks of cancer [Citri and Yarden 2006]. Several downstream effectors of EGFR are activated in HNSCC, namely ERK1/-2, akt, STAT3, STAT5 and COX2.

In HNSCC, EGFR is overexpressed in almost 90% of tumors, especially in more advanced-stage and also in poorly differentiated carcinomas, being this overexpression associated with decreased patient survival rates and with resistance to some therapeutic modalities [Kalyankrishna and Grandis 2006].

Nowadays there are numerous ongoing trials in HNSCC patients investigating potential target agents, both tyrosine kinase inhibitors and antibodies against several receptors.

The combination of RT with EGFR-specific antibody cetuximab showed an improvement in locoregional control and progression-free survival over RT alone, and also some increase in survival in the palliative setting in locoregionally advanced and metastatic tumors [Bonner et al. 2006; Vermorken et al. 2008].

- **Phosphatidylinositol-3-kinase/protein kinase B pathway (PI3-K/Akt)**

The phosphatidylinositol-3-kinase/protein kinase B (PI3-K/Akt) signal transduction pathway has been shown to regulate several cellular processes related to carcinogenesis. This pathway has been shown to be activated in up to 50–80% of HNSCCs and is one of the main downstream signaling pathways activated by EGF/tyrosine Kinase receptor [Lothaire et al. 2006]. This pathway could also be activated through direct mutation or amplification of *PIK3CA*, amplification of *AKT1*, activation of the RAS oncogenes, and/or decreased expression of the tumor-suppressor protein PTEN, an inhibitor of the PI3-K/Akt pathway [Bussink et al. 2008]. *PTEN* is a tumor suppressor gene that presents homozygous deletions or inactivating mutations in HNSCC, correlating the lack of PTEN expression with aggressive tumors and poor overall survival [Di Cristofano and Pandolfi 2000]. Activated AKT promotes the increased expression of numerous proteins related to apoptosis inhibitors, cell cycle inhibitors and transcription factors, such as Bcl-2, Bcl-x and NF-kB, which leads to cell proliferation and survival [Tafe 2017]. The potential clinical benefit of PI3-K/Akt pathway as therapeutic target in HNSCC holds significant promise.

- **Mammalian Target of Rapamycin (mTOR)**

Mammalian target of rapamycin (mTOR), an atypical serine/threonine kinase, is a major downstream effector of Akt and regulates cell growth through controlling growth factor and nutrient signaling [Shamji et al. 2003]. There are two different mTOR complexes: 1) mTOR complex 1 (mTORC1), that is regulated by growth factors, nutrients, energy status, oxygen and cellular stress that promote protein synthesis, cell proliferation and survival, ribosome biogenesis, angiogenesis, migration, invasion, and metastasis; and 2) mTOR complex 2 (mTORC2), that functions in actin remodeling, cell-cycle progression, and cell survival [Gao et al. 2012]. This signaling pathway is activated by several growth factors, such as VEGF, FDGF, EGF, IGF1, hormones, nutrients and oxygen [Dancey 2006]. This pathway plays a central role in HNSCC and, consequently the expression levels of mTOR and downstream targets eIF4E, 4EBP1, S6K1, and S6 are promising diagnostic and prognostic biomarkers for HNSCC [Gao et al. 2012]. Rapamycin, a mTOR inhibitor and its derivatives showed antitumor activity in HNSCC both *in vitro* and *in vivo* [Freudlsperger et al. 2011].

Biomarkers

Nowadays, there is a growing amount of molecular markers with potential clinical applicability at prognostic and diagnostic level in HNC field. The understanding of these biological markers and the corresponding disrupted signaling pathways is still beginning due to the great heterogeneity of these tumors, hampering the establishment of a truly personalized treatment for these patients. By tailoring treatment to the specific molecular profile of each tumor, the treatment response and consequently the survival could be improved. Thus, understanding the molecular mechanisms of HNSCC is pivotal to identify biomarkers, which could be detected in tumor tissue, saliva or in the serum and that could 1) lead to early cancer detection in diagnostic screenings; 2) be used as prognosticators to guide treatment modality decisions and 3) monitor patients during the course of treatment and in the follow-up period [Schutt et al. 2012]. Nowadays, several potential biomarkers have been pointed out; however, none has been translated to clinical practice yet. Several examples of putative biomarkers for HNSCC are listed in table 1.

Chapter 1 - Table 1. Several biomarkers for HNSCC described in the literature.

Chromosomal region	Genes	Alteration	Clinical association	Technique	Sample size	Sample Type	Reference
1p33-34, 2q31, 7q21, 7q22-31, 11q12, 12q11, 14q13		Amplification	Chemoradiosensitivity	aCGH	52	FFPE	[van den Broek et al. 2007]
1p32, 3q24, 7p11.1, 7p11.2-12, 8p11.1, 8p11.1-12, 12q15, 13q21, 15q21, 18p11.3, 18q11		Amplification	Chemoradioresistance	aCGH	52	FFPE	[van den Broek et al. 2007]
1q21-q22, 7q11.2, 11q13,		Gain	Metastasis	aCGH	54	Fresh-frozen	[Bockmuhl et al. 2002]
1q34		Gain	LRP-free survival	aCGH	117	FFPE	[Bauer et al. 2008]
2q22-q25, 5p11-pter, 7q11-q22		Loss	Chemoradiosensitivity	aCGH	52	FFPE	[van den Broek et al. 2007]
3p		Loss	Early event	microsatellite PCR analysis	43	Fresh-Frozen	[Hogg et al. 2002]
3p22, 3p24.1, 3p26.1			Early Event	aCGH	20	FFPE	[C. Garnis et al. 2003]
3p11-pter, 4p11-pter, 18p11-pter		Loss	Chemoradioresistance	aCGH	52	FFPE	[van den Broek et al. 2007]
3p22.3	***MLH1, ITGA9, RBSP3***	Loss/promoter methylation	Early stages of disease	PCR-based methylation-sensitive restriction analysis (MSRA)	148	Fresh-Frozen	[Ghosh et al. 2010]
3p21.31	***RASSF1A***	Methylation	Poor prognosis	Bisulphite DNA sequencing	24	Fresh-Frozen	[Hogg et al. 2002]

Chapter 1 - Table 1. Several biomarkers for HNSCC described in the literature (continued).

Chromosomal region	Genes	Alteration	Clinical association	Technique	Sample size	Sample Type	Reference
3p21.31	***SEMA3F***	Loss	Metastasis and decreased survival	Imunohistochemistry	4	Cell Lines	[Doci et al. 2015]
3p14.3		LOH	Survival	PCR	163	FFPE + Fresh-frozen	[Coon et al. 2004]
3p14.2	***FHIT***	Loss	Early event	aCGH	20	FFPE	[C. Garnis et al. 2003]
3q		Gain	Poor prognosis; Early marker for tumor invasion	CGH; aCGH	113; 54	Fresh-frozen	[Bockmuhl et al. 2000] [Bockmuhl et al. 2002]
3q11-q13, 3q21-q26.1, 6p11-pter, 6q22-q27, Xq11-qter		Gain	Chemoradioresistance	aCGH	52	FFPE	[van den Broek et al. 2007]
3q25-q27		Gain	Reduced overall survival	CGH	45	Fresh-frozen	[Ashman et al. 2003]
3q25.31	***CCNL1***	Amplification	Loco-regional metastases	FISH	280	Fresh-frozen	[Sticht et al. 2005]
3q26.32	***PIK3CA***	Amplification	Poor prognosis	qPCR and aCGH	115	Fresh-frozen	[Suda et al. 2012]
3q26.33	***SOX2***	Amplification	Metastasis, Worse outcome, Resistance to cisplatin	FISH, Imunohistochemistry	496	Fresh-frozen	[Schrock et al. 2014]
4p		Loss	Metastasis	CGH, (DOP)-PCR	35	Fresh-frozen	[Noutomi et al. 2006]

Chapter 1 - Table 1. Several biomarkers for HNSCC described in the literature (continued).

Chromosomal region	Genes	Alteration	Clinical association	Technique	Sample size	Sample Type	Reference
4q12 **5q35.3** **7q21.3**	***KDR*** ***FLT4*** ***TFPI2***	Methylation	Early Stage	Illumina GoldenGate Methylation Cancer Panel	40	Fresh-frozen	[Y. F. Li et al. 2015]
4q31.1	***MAML3***		Poor Survival	Affymetrix Genome-wide Human SNP array 6.0	75	Fresh-frozen	[X. Chen et al. 2015b]
4q32.3	***PALLD, DDX60L***						
4q35.1	***ING2***	LOH	Advanced stage	Microsatellite analysis	80	Fresh-frozen	[Borkosky et al. 2009]
5q11-q12, 6q23-q27, 8p21-p23, 10q11-q22, 15q13-q26, 17q, 18q21-q23, 22		Gain	Chemoradiosensitivity	aCGH	52	FFPE	[van den Broek et al. 2007]
7p		Gain	Nodal metastases, poor prognosis	CGH	97	Fresh-frozen	[Pathare et al. 2011]
7p11.2	***EGFR***	Amplification	Poor prognosis	FISH, Immunohistochemistry	134	FFPE	[Sheu et al. 2009]
7q31	***MET***	Overexpression	Metastasis	Immunohistochemistry	34	Fresh-frozen	[Galeazzi et al. 1997]
8p		Loss	Nodal metastases	CGH	97	Fresh-frozen	[Pathare et al. 2011]
8p23		Loss	Poor prognosis	Microsatellite polymorphism analysis	70	Fresh-frozen	[Bockmuhl et al. 2001]
8p21- p22		Loss	Advanced tumor stages and poor survival	CGH	13	Fresh-frozen	[Bockmuhl et al. 2000]

Chapter 1 - Table 1. Several biomarkers for HNSCC described in the literature (continued).

Chromosomal region	Genes	Alteration	Clinical association	Technique	Sample size	Sample Type	Reference
8p21.2	***NEFL***	LOH	Increased mortality risk	PCR	163	FFPE + Fresh-frozen	[Coon et al. 2004]
8q22-23		Gain	Early stages of disease	CGH, (DOP)-PCR	35	Fresh-frozen	[Noutomi et al. 2006]
8q24-21	***MYC***	Amplification	Advanced primary tumors/ late event in the tumorigenesis	PCR	58	Fresh-frozen	[Rodrigo et al. 1996]
9p		Gain	Unfavorable outcome	CGH	97	Fresh-frozen	[Pathare et al. 2011]
9p21.3	***CDKN2A***	Loss/mutation	Poor Prognosis	Immunohistochemical	148	FFPE	[Bova et al. 1999]
	CDKN2A, RARB, MGMT	Methylation	Early Event	MSP	32	Fresh-frozen	[Maruya et al. 2004]
9q21.33	***DAPK***	Methylation	Early Event	Q-MSP	40	Fresh-frozen	[Arantes et al. 2015]
10p12, 10q, 11p14, 11q14-qter, 14q		Loss	Metastasis	aCGH	54	Fresh-frozen	[Bockmuhl et al. 2002]
11q13	***CCND1, TAOS1, PPFIA, CTTN***	Amplification	Reduced overall survival	aCGH	20	Cell Lines	[Freier et al. 2010]

Chapter 1 - Table 1. Several biomarkers for HNSCC described in the literature (continued).

Chromosomal region	Genes	Alteration	Clinical association	Technique	Sample size	Sample Type	Reference
11q13	***CCND1***	Amplification	Presence of occult cervical lymph node metastasis	FISH	45	Fresh-frozen	[Myo et al. 2005]
	ANO 1	Amplification	Poor overall survival, metastases	FISH, Immunohistochemistry	365	FFPE	[Ruiz et al. 2012]
11q22.1-q22.2		Gain	Loco-regional recurrence	I-FISH	60	Fresh-frozen	[Ambatipudi et al. 2011]
12q24.33	***CHFR***	Methylation	Late stage disease	MS-MLPA	28	Fresh-frozen	[K. Chen et al. 2007]
13q13.1; 13q14.2; 13q21.2-q22.1; 13q31.1	***BRCA2; RB1; and BRCAX***	LOH	Poor patient outcome	PCR	55	Fresh-frozen	[Sabbir et al. 2006]
16q24.3	***FANCA***	Gain	LRP-free survival	aCGH	117	FFPE	[Bauer et al. 2008]
17p13.3; 11q23-q25		Loss	Shorter survival	I-FISH	60	Fresh-frozen	[Ambatipudi et al. 2011]
17p13.1	***TP53***	LOH	Nonresponse to neoadjuvant chemotherapy		106		[Cabelguenne et al. 2000]
18q		LOH	Poor patient survival	PCR	27	FFPE + Fresh-frozen	[Pearlstein et al. 1998]
18q21.2	***DCC***	Methylation	Early Event	Q-MSP	40	saliva	[Arantes et al. 2015]
18q22			LRP-free survival	aCGH	117	FFPE	[Bauer et al. 2008]
22q		Loss	Poor clinical outcome	CGH	45	Fresh-frozen	[Ashman et al. 2003]

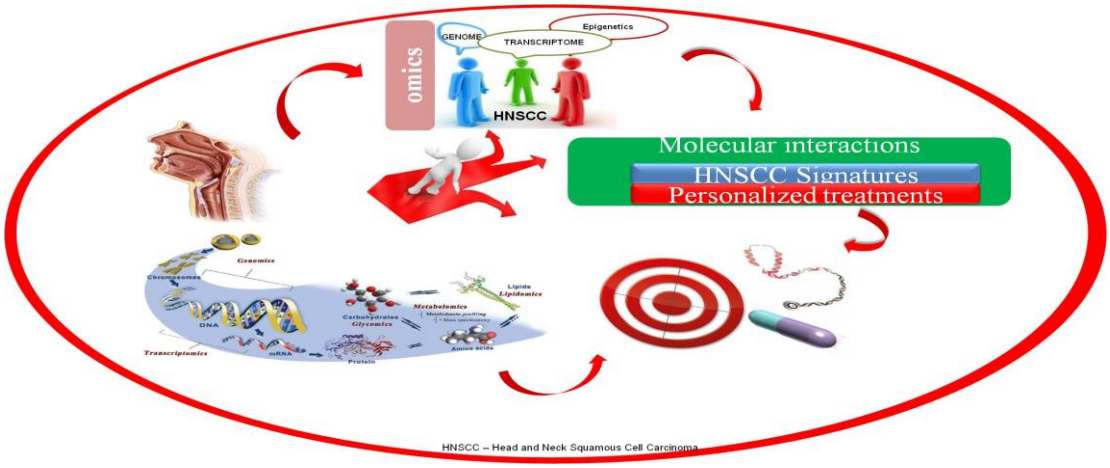
Chapter 1 - Table 1. Several biomarkers for HNSCC described in the literature (continued).

Chromosomal region	Genes	Alteration	Clinical association	Technique	Sample size	Sample Type	Reference
22q12.3	***TIMP***	Methylation	Early Event	Q-MSP	40	saliva	[Arantes et al. 2015]
22q13.2 - 13.31	***DIA1***	Loss	Worse prognosis, decreased survival	qRT-PCR	40	Fresh-frozen	[Reis et al. 2002]

aCGH - array Comparative genomic hybridization; **CGH -** Comparative genomic hybridization; **DOP-PCR -** Degenerate oligonucleotide-primed Polymerase chain reaction; **FFPE -** Formalin-fixed paraffin-embedded; **FISH -** Fluorescence *in situ* hybridization; **I-FISH -** Interphase fluorescence *in situ* hybridization **LRP -** Loco-regional progression; **LOH -** Loss of heterozygosity; **MS-MLPA -** Methylation-Specific Multiplex Ligation dependent Probe Amplification; **PCR -** Polymerase chain reaction; **Q-MSP -** Quantitative Methylation-Specific PCR; **qRT-PCR -** Quantitative reverse transcription PCR

2. Aims

Bearing in mind the clinical and molecular heterogeneity of HNSCC and their multifactorial nature, it seems crucial to combine several molecular biomarkers in order to early diagnose these tumors and the relapses. This study has focus on omics profiles to predict local relapses/metastasis, signaling pathways associated to targeted therapy and resistance to conventional drugs through the identification of different molecular groups with apparently different clinical outcomes using genomics, transcriptomics and epigenetics (Figure 1).



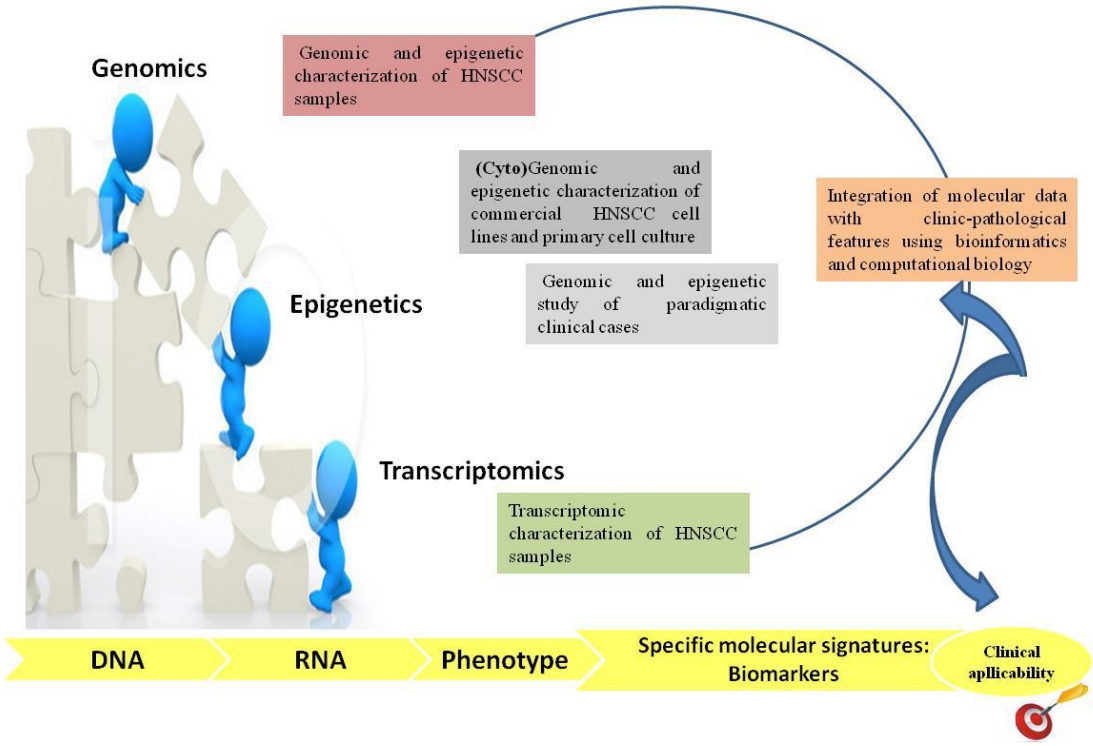
Chapter 2 - Figure 1. Omics integration in this study for HNSCC characterization.

The main research question was: "Different omics profiles are associated with different clinical outcomes in HNSCC?" and consequently: "Can these different profiles be associated with biomarkers that can guide early diagnosis, prognosis and therapeutic management of these patients?" To address these questions the specific aims of this work were:

- Identify specific molecular patterns for each stage and anatomic site of HNSCC and for the tumors associated with different risk factors (tobacco and alcohol or HPV);
- Identify molecular signatures that could have a role in the prediction of metastatic/local recurrent disease, which would have an essential impact to choose or intensify the treatment modalities;
- Identify possible molecular patterns related to radiation and/or chemotherapy resistance;

- Characterize HNSCC recurring to an integrative omic approach and correlate with clinical outcome;
- Develop an easy non-invasive approach with costs compatible with routine diagnostic to screen high-risk populations and patients follow-up.

This work is an example of translation research in terms of strategy design with data integration, which can contribute to a significant step forward in the body of knowledge of HNSCC, with practical implications in the management of these patients (Figure 2).



Chapter 2 - Figure 2. Design of the study with the different phases evaluated.

3. Results

3.1 Genetic and Epigenetic Studies

3.1.1 Genetic gains and losses in oral squamous cell carcinoma: impact on clinical management

Ilda P. Ribeiro, Francisco Marques, Francisco Caramelo, João Pereira, Miguel Patrício, Hugo Prazeres, José Ferrão, Maria J. Julião, Miguel Castelo-Branco, Joana B. Melo, Isabel P. Baptista, Isabel M. Carreira

Cellular Oncology
Volume 37, Issue 1, Pages 29-39, (2014)
doi:10.1007/s13402-013-0161-5

Abstract

Purpose The identification of genetic markers associated with oral cancer is considered essential to improve the diagnosis, prognosis, early tumor and relapse detection and, ultimately, to delineate individualized therapeutic approaches. Here, we aimed at identifying such markers. ***Methods*** Multiplex Ligation-dependent Probe Amplification (MLPA) analyses encompassing 133 cancer-related genes were performed on a panel of primary oral tumor samples and its corresponding resection margins (macroscopically tumor-free tissue) allowing, in both types of tissue, the detection of a wide arrange of copy number imbalances on various human chromosomes. ***Results*** We found that in tumor tissue, from the 133 cancer-related genes included in this study, those that most frequently exhibited copy number gains were located on chromosomal arms 3q, 6p, 8q, 11q, 16p, 16q, 17p, 17q and 19q, whereas those most frequently exhibiting copy number losses were located on chromosomal arms 2q, 3p, 4q, 5q, 8p, 9p, 11q and 18q. Several imbalances were highlighted, i.e., losses of *ERBB4*, *CTNNB1*, *NFKB1*, *IL2*, *IL12B*, *TUSC3*, *CDKN2A*, *CASP1*, and gains of *MME*, *BCL6*, *VEGF*, *PTK2*, *PTP4A3*, *RNF139*, *CCND1*, *FGF3*, *CTTN*, *MVP*, *CDH1*, *BRCA1*, *CDKN2D*, *BAX*, as well as exon 4 of *TP53*. Comparisons between tumor and matched macroscopically tumor-free tissues allowed us to build a logistic regression model to predict the tissue type (benign versus malignant). In this model, the *TUSC3* gene showed statistical significance, indicating that loss of this gene may serve as a good indicator of malignancy. ***Conclusions*** Our results point towards relevance of the above mentioned cancer-related genes as putative genetic markers for oral cancer. For practical clinical purposes, these genetic markers should be validated in additional studies.

Keywords: Oral squamous cell carcinoma; Genetic profile; Chromosomal imbalances; Copy number losses and gains

Introduction

Oral cavity tumors constitute a subgroup of head and neck tumors that rank sixth in prevalence, with an annual incidence of almost 600.000 cases worldwide [Jemal et al. 2009]. The most common histological subtype is oral squamous cell carcinoma (OSCC). Tumors of the oral cavity exhibit a complex etiology involving multiple environmental, toxic, and viral factors. In addition to tobacco and alcohol consumption, human papilloma virus (HPV) infection is a well-known risk factor for OSCC [Ambatipudi et al. 2011]. In spite of advances that have been made in diagnostic technologies and treatment modalities, these tumors are still diagnosed at relatively late stages and, consequently, no major improvements in survival rates have been made. Patients with a positive OSCC diagnosis undergoing primary treatment show recurrence rates ranging from 25-45% [Kowalski 2002; Liu et al. 2007; Nathanson et al. 1989; Rate et al. 1991; W. Wang et al. 2011; Wong et al. 2003].

Early detection is considered the gold key for decreasing morbidity and mortality rates, as well as for reducing healthcare costs. OSCC development results from the accumulation of both genetic and epigenetic changes, and studies have been reported aimed at identifying aberrantly expressed genes that can be used in the classification, diagnosis and prognosis of OSCC, including the prediction of treatment outcome [B. J. Braakhuis et al. 2005; Weiss et al. 2013]. Although copy number imbalances have been reported to occur in almost all chromosomes, it appears that some chromosomal regions are recurrently affected in these tumors [Guervos et al. 2007]. Overall, however, oral cancer displays a vast genetic and biologic heterogeneity, and the most challenging task is to establish the clinical relevance of each molecular subgroup associated with specific histopathological features. Therefore, establishing correlations between molecular data and disease phenotypes appears to be crucial to (i) confirm the histological type and the stage of the tumor and (ii) predict more accurately the patient's outcome. Due to the currently limited clinical and pathological capability of identifying patients at high-risk of treatment failure, better biomarkers for prognosis are urgently needed. Ultimately, it will be imperative to take into account the genetic profile of each individual patient in order to be able to delineate personalized therapeutic strategies. In the present study, we have established the genetic profiles of 35 OSCC samples and correlated the results obtained with its corresponding clinicopathological characteristics. The putative relevance of the most frequent genetic changes encountered, including their applicability in routine clinical practice are discussed.

Materials and Methods

Tumors and control samples

The present study was conducted on 35 primary oral tumor samples with corresponding resection margins (macroscopically tumor-free tissue) from 28 patients. These samples were obtained between 2010 and 2012 from the Maxillofacial Surgery and Stomatology Unit of the Coimbra Hospital and University Centre, CHUC, EPE, Portugal. All patients were submitted to surgery and the histopathologic diagnoses of the mirror sections of the samples were performed by two different pathologists. Hematoxylin and eosin staining was used to evaluate the tumor content in each specimen. In our cohort, all samples contained at least 50% tumor cells. Diagnosis and staging were performed according to the American Joint Committee on Cancer TNM staging system [Wittekind et al. 2003] for OSCCs. All patients provided informed consent in accordance with the regulations in the Declaration of Helsinki. The study was approved by Ethics committee of the Faculty of Medicine of the University of Coimbra. Detailed characteristics of our OSCC cohort are listed in table 1. As controls, gingival tissues from healthy donors subjected to "wisdom teeth" removal were included. DNAs from patient and control samples were extracted using a High Pure PCR Template Preparation Kit (Roche GmbH, Mannheim, Germany) according to the manufacturer's instructions, and quantified using a Nanodrop 1000 Spectrophotometer (Thermo Scientific, USA).

Multiplex Ligation-dependent Probe Amplification (MLPA)

MLPA was performed using four tumor-specific MLPA probe panels. Overall, these four panels (P005, P006, P007 and P014; MRC-Holland, Amsterdam, The Netherlands) included 154 probes targeting 133 different genes located on all human autosomes (supplementary Table 1). The P014 panel was exclusively designed for chromosome 8, allowing a more comprehensive study of this chromosome as compared to the other ones. Details of the probe sequences, gene loci and chromosomal locations can be found at: www.mrc-holland.com. All MLPA reactions were performed according to the protocol described by Schouten et al [Schouten et al. 2002]. Briefly, DNA samples (5 µl) were heated at 98 ºC for 10 min. After the addition of the probe mix, samples were heated for 1 min at 95 ºC and then incubated for 16h at 60 ºC. Ligation of the annealed oligonucleotide probes was performed for 15 min at 54 ºC in buffer containing Ligase-

65. After inactivating the ligase by heating at 98 ºC for 5min, multiplex PCR was carried out using FAM-labeled primers, dNTPs and SALSA polymerase. PCRs were performed for 35 cycles of 30 sec at 95 ºC, 30 sec at 60 ºC and 1 min at 72 ºC. All the reactions were carried out in a thermal cycler equipped with a heat lid (ABI 2720, Applied Biosystems, Foster City, CA, USA). Finally, the PCR products were heat denatured and analyzed using a Gene Scan ABI PRISM 3130 capillary electrophoresis system (Applied Biosystems, Foster City, CA, USA). Three normal controls and a negative control (without DNA) were included in each MLPA assay. The results are displayed as ratios between references and experimental samples. For each MLPA probe we determined specific cut-off values for gain and loss, using the values limiting the 95% confidence interval (CI) as determined on non-cancer samples. A numerical gain was scored when the ratio was higher than 1.2 and a numerical loss was defined when the ratio was lower than 0.8.

HPV typing

All tumor tissue samples were analyzed for HPV infection as described by Nobre et al [Nobre et al. 2010]. Briefly, PCR was performed using established general consensus and degenerate primer sets, i.e., GP5+/GP6+ and MY09/MY11, which were designed to amplify a fragment of the L1 gene of mucosatropic HPVs. For genotyping we performed Sanger sequencing of intra-primer segments within the amplified fragments in order to determine the specific types of HPV present in the samples. In addition to DNA sequencing, we complemented our analyses by DNA microarray hybridizations, using HPV CLART2 arrays (Genomica), to address cases that showed infection with multiple HPV genotypes.

Statistical analysis

The statistical analysis was carried out using the statistical software package IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp. The significance level adopted was $p = 0.05$. From detailed descriptive analyses and chi-square tests corrected for multiple comparisons (Bonferroni correction), it was possible to reduce the number of genes with statistical meaning for distinguishing tumor tissue from macroscopically tumor-free tissue, selecting only those that were significantly imbalanced between the two groups. We then performed a logistic regression (forward:

conditional) model using these genes (thirteen in total), in order to assess their usefulness as predictors of the tissue type.

Chapter 3.1.1 - Table 1. Patient and tumor characteristics

	Patients (n = 35)
Mean Age, years (range)	**61.5** (37-84)
Sex	
Male	30
Female	5
Smoking (cigarettes/day)	
≥20	21
< 20	3
None	11
Alcohol	
Yes	4
None	4
Not recorded	27
Stage	
I and II	14
III and IV	21
Site	
Tongue	13
Floor of the mouth	12
Buccal mucosa	4
Retromolar trigone	4
Gingival	2
Pathological margin status	
Positive	2
Negative	33
Treatment	
Surgery + RT	9
QT + Surgery	3
Surgery + RT + QT	7
Surgery only	16
Clinical outcome	
Alive	22
Death from the disease	12
Dead from the other cause	1

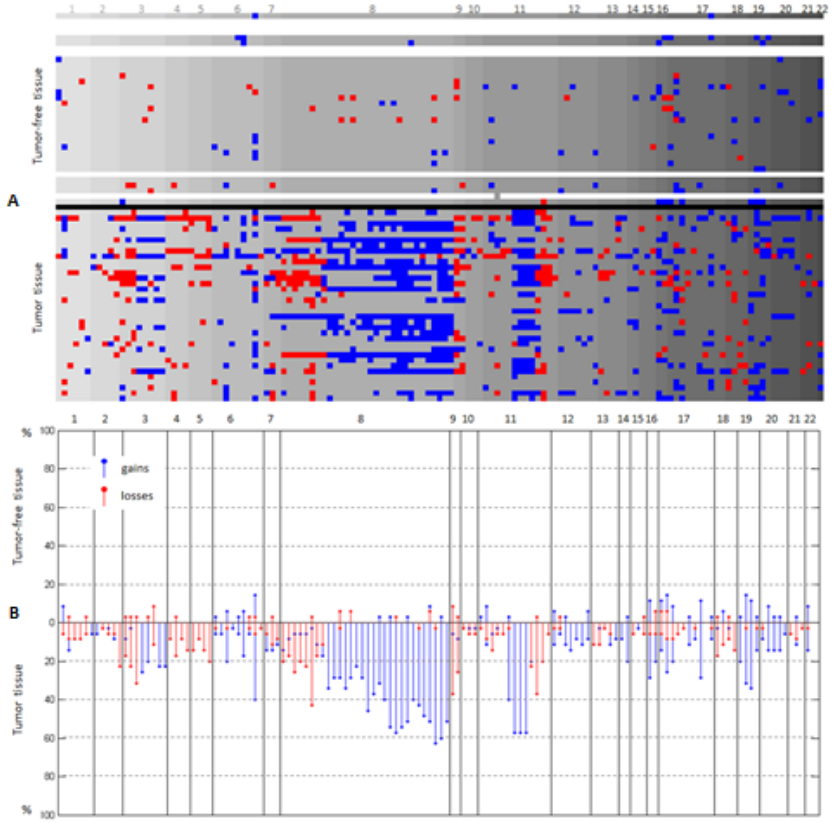
RT - Radiation therapy; **QT -** Chemotherapy

Results

Genetic profiles of tumor tissues and macroscopically tumor-free tissues

All oral tumor samples included in this study were analyzed by MLPA in order to establish their genetic profiles. The sex chromosomes were excluded from the analyses since the control and tumor samples were not gender-matched. We observed genetic alterations in all 35 tumor samples analyzed and, in addition, in 23 of the 28

macroscopically tumor-free tissue samples recovered from surgical margins (Figure 1A). The numbers of alterations detected in both tissues were, however, very different (Figure 1A, B). As expected, we detected more copy number imbalances in the tumor tissues than in the macroscopically tumor-free tissues. Besides this, the distribution of the imbalances in terms of losses and gains by chromosome was very consistent for some chromosomes, i.e., 3p and 8p showed mostly losses, whereas 3q and 8q frequently showed gains; chromosomes 4 and 5 only showed losses for the genes analyzed. Additionally, on chromosomes 19 and 20 we encountered more frequently gains than losses for the genes analyzed. In the tumor tissues of these 35 patients, we found that from the 133 genes analyzed, those with the most frequent gains were localized on chromosomal arms 3q, 6p, 8q, 11q, 16p, 16q, 17p, 17q and 19q, whereas those with the most frequent losses were localized on chromosomal arms 2q, 3p, 4q, 5q, 8p, 9p, 11q and 18q (Figure 1B). In the macroscopically tumor-free tissues, the most frequent gains were localized on chromosomal arms 6p, 16p, 17p, 17q and 19q, whereas the most frequent losses were localized on chromosomal arms 3p and 9p (Figure 1B). In contrast, we found that the genes tested on chromosomes 10 and 15 did not exhibit imbalances (gains or losses) in more than two patients.



Chapter 3.1.1 - Figure 1. Genetic imbalances (gains and losses) in 35 OSCC patients detected by using four MLPA probe mixes. Losses of genetic material are represented in red, gains in

blue. **A)** The results above the black line correspond to macroscopically tumor-free tissue and the results below this line correspond to tumor tissue. Each line represents one patient and each pixel in this line corresponds to one gene. Gray represents genes without alteration, and each shade of gray shows the localization of the genes on each specific chromosome. From left to right, the genes are ordered by chromosome, from the short arm to the long arm. Each white pixel on macroscopically tumor-free tissue means that no information, since we did not have macroscopically tumor-free tissue for all patients analyzed. **B)** Picture showing the percentage of imbalances by chromosome in tumor tissue and in macroscopically tumor-free tissue for each gene analyzed, excluding the imbalances detected in HPV-positive patients. Each arrow represents one gene.

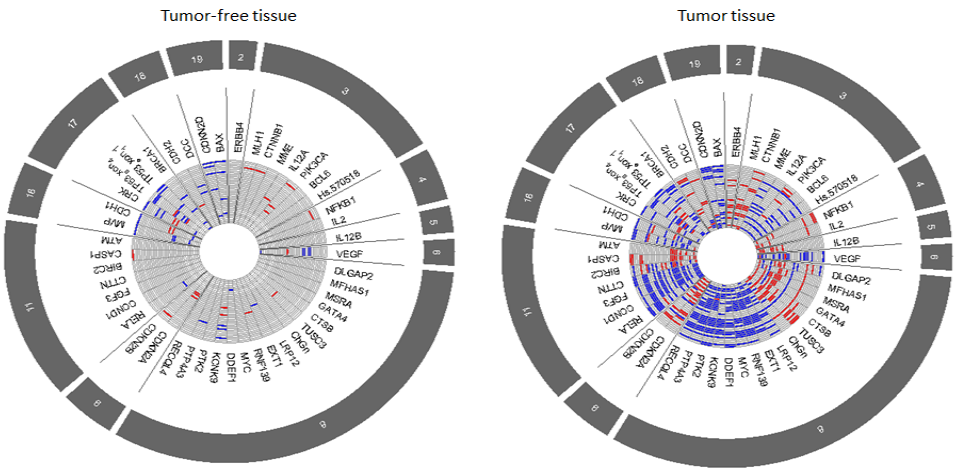
Overall, we found that the number of samples that showed losses of genetic material was lower than those showing gains. Chromosome 8 showed most gains and losses in the largest number of patients (Figure 1B). It is important note here that in this study we used one probe panel specific for chromosome 8 with probes for 30 genes mapped on this chromosome, which allowed a more comprehensive assessment of this chromosome compared to the other ones. We also found that no single gene was altered in all patients. Figure 2 illustrates the most commonly altered genes on the 12 aforementioned chromosomes in tumor tissues, as well as in macroscopically tumor-free tissues. Thus, from the 133 genes analyzed (Figure 2), the following were the ones showing frequent losses in tumor samples: *ERBB4* (2q33.3-q34), *CTNNB1* (3p21), *NFKB1* (4q24), *IL2* (4q26-q27), *IL12B* (5q31.1-q33.1), *TUSC3* (8p22), *CDKN2A* (9p21) and *CASP1* (11q23). It is also important to note that we detected homozygous deletions for both the *CDKN2A* and *CDKN2B* genes in one patient. Additionally, the following genes: *MME* (3q25.2), *BCL6* (3q27), Hs. 570518 (3q28), *VEGF* (6p21.1), *PTK2* (8q24.3), *PTP4A3* (8q24.3), *RNF139* (8q24), *CCND1*, *FGF3*, *CTTN* (11q13), *MVP* (16p11.2), *CDH1* (16q22.1), exon 4 of *TP53* (17p13.1), *BRCA1* (17q21), *CDKN2D* (19p13) and *BAX* (19q13.3-q13.4) frequently showed gains in the tumor samples. In the macroscopically tumor-free tissues the genes that showed the most frequent losses were: *PIK3CA* (3q26.3) and *CDKN2A* (9p21), whereas *VEGF* (6p21.1), *KCNK9* (8q24.3), *MVP* (16p11.2), exon 4 of *TP53* (17p13.1), *BRCA1* (17q21), *CRK* (17p13.3), *CDKN2D* (19p13) and *BAX* (19q13.3-q13.4), showed the most frequent gains.

Genetic imbalances predicting clinicopathological features

Next, we generated a logistic regression model to predict the type of tissue (tumor versus macroscopically tumor-free) using the *DLGAP2, TUSC3, EXT1, RNF139,*

MYC, *DDEF1*, *PTK2*, *PTP4A3* and *RECQL4* genes on chromosome 8, and the *CCND1*, *FGF3*, *CTTN* and *BIRC2* genes on chromosome 11 as predictors. The final model (-2LL = 24.004; Cox and Snell R^2 = 0.630; Nagelkerke R^2 = 0.843) included the *TUSC3*, *PTK2* and *CCDN1* genes, but only the *TUSC3* gene showed statistical significance (p = 0.041). The accuracy of this model was 93.7%, compared to 55.6% if the prediction had been random. In this model, the probability of being tumor tissue was 27-fold higher when the *TUSC3* gene showed loss versus this gene being normal (OR = 27.000 with CI 95% [2.091; 348.661]). If we only take the *TUSC3* gene into account to make this regression model, the accuracy drops to 68.3% (-2LL = 69.852; Cox and Snell R^2 = 0.233; Nagelkerke R^2 = 0.312). Nonetheless, in this simplified model the value of this gene remains statistically significant (p = 0.017), with an OR = 21.316 for losses with CI 95% [2.590; 175.398].

With respect to other clinicopathological features, including tumor stage, development of metastasis and tobacco consumption, we found similar patterns of copy number losses and gains across the genome in both stage I + II and stage III + IV, as well as in the presence or absence of metastases, and in the smokers and non-smokers groups (Figure 3). No single gene showed statistical significance, and we were unable to genetically differentiate patients belonging to stages I or II from those belonging to stages III or IV. The same lack of statistical significance was found for the presence or absence of metastasis, and for the consumption or non-consumption of tobacco. We did, however, observe differences in losses of genetic material between smokers and non-smokers, i.e., only the smokers showed losses at 3p (*MLH1*) and 11q (*ATM*).



Chapter 3.1.1 - Figure 2. Radial heatmap of the genes frequently altered in 12 most commonly affected chromosomes for tumor tissue and for macroscopically tumor-free tissue based on the

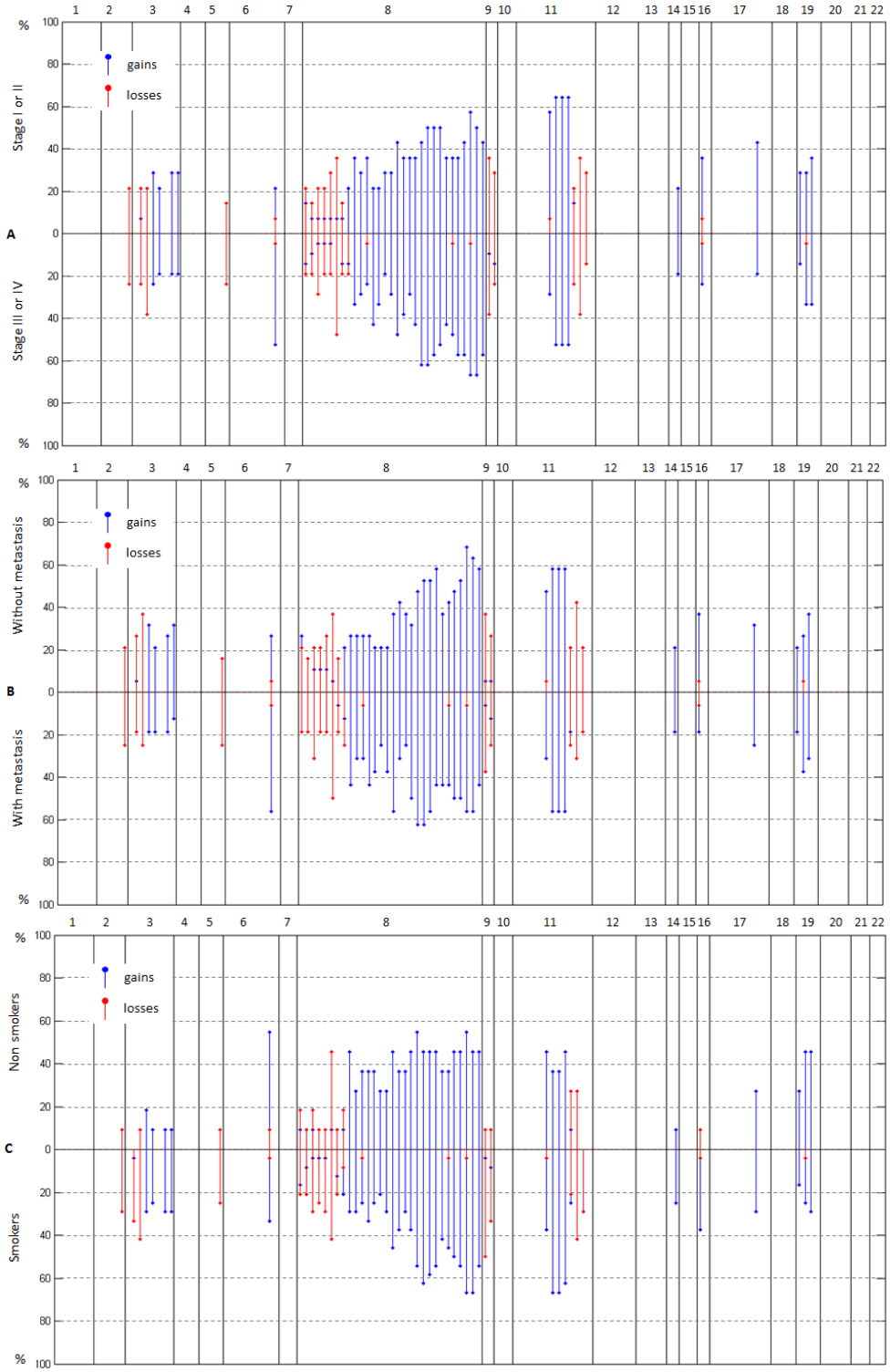
use of four MLPA probe sets. Each line represents one patient. Red lines represent losses of genetic material and blue lines represent gains.

HPV infections in OSCC samples

Among the 35 tumor samples analyzed, two were found to be HPV-positive (data not shown). One of these samples showed a HPV type belonging to a low-risk class, i.e., HPV type 42, whereas the other sample showed a HPV type belonging to a high-risk class, i.e., HPV type 31. The low-risk HPV patient did not exhibit any other risk factors, such as tobacco use or alcohol consumption.

Discussion

Cancer is considered to be a disease of the genome and to result from the sequential acquisition of DNA alterations by somatic cells. Besides being useful for early diagnostics, these alterations may also serve as specific targets for therapy. As such, they provide a window of hope and promise. Since the identity of the most relevant oral cancer-related genes is still unknown, we set out to identify at least some of them through genomic profiling.



Chapter 3.1.1 - Figure 3. Percentages of imbalances by chromosome in 35 OSCC tumors. Each arrow represents one gene. Losses of genetic material are represented in red, gains in blue. **A)** Genetic profiles of tumors in stage I or II, and in stage III or IV. **B)** Genetic profiles of tumors that developed metastases, and tumor without metastases. **C)** Genetic profiles of tumors in smokers and non-smokers.

Genomic profiling of oral cancer

Until now, no single gene alteration has been identified that is exclusive for oral cancer. After analyzing in detail the most commonly altered chromosomes in our cohort, we found that we could highlight specific genes that frequently show copy number gains or losses. Thus, on chromosome 2 we observed frequent losses of the *ERBB4* gene (2q33.3-q34) only in tumor tissue. In the past, in breast cancer decreased ERBB4 protein expression has been correlated with increased recurrence rates [Barnes et al. 2005]. This gene should, therefore be taken into account for OSCC in order to redefine its predictive value for recurrence risk and, consequently, for choosing additional treatment options. On chromosome 3, we found that most frequent losses and gains occurred at 3p and 3q, respectively. In a wide range of patient samples we observed copy number gains at 3q, highlighting the genes *MME* (3q25.2), *BLC6* (3q27), Hs.570518 (3q28) and *IL12A* (3q25.33) as being putatively related to oral carcinogenesis. Indeed, previously Freier et al [Freier et al. 2010] detected frequent DNA copy number gains at 3q, and emphasized the possibility of several candidate proto-oncogenes being located within the region 3q25-qter. Intriguingly, *PIK3CA* (3q26.3) showed both gains and losses at the same frequencies in tumor tissues and, additionally, in a macroscopically tumor-free tissues this gene was the only one in this chromosome that showed alterations in a substantial number of patients (8.5%). In another study on head and neck cancer, the aforementioned gene had already been pointed out as a strong candidate oncogene [Singh et al. 2001b]. Losses observed on 3p highlighted several putative tumor suppressor genes, such as *CTNNB1* (3p21), *MLH1* (3p21.3) and *VHL* (3p25.3). Previously, losses of *CTNNB1* (3p21) have significantly been associated with precursor lesions showing progression towards laryngeal carcinoma [Marcos et al. 2010]. It will be crucial to further elucidate the role of these genes in oral tumor initiation and to assess their role in tumor progression during clinical follow-up of the patients. Previously chromosome 3p has been described as being frequently involved in loss of heterozygosity (LOH) in OSCC. So, other tumor suppressor genes located within this region, such as *FHIT*, may also be relevant for the development of this carcinoma [Uzawa et al. 2001]. On chromosome 4 we only observed copy number losses, highlighting the genes *NFKB1* (4q24) and *IL2* (4q26-q27). Loss of the *NFKB1* gene was also observed in the macroscopically tumor-free tissue of one patient. Deletions associated with this chromosome were initially considered to be relatively rare as compared to other chromosomal aberrations. However, Pershouse et al revealed that 92% of head and neck tumors showed deletions

involving chromosome 4 with the highest frequency of loss in band q25 [Pershouse et al. 1997; X. L. Wang et al. 1999]. Until now, no tumor suppressor gene within this region has firmly been established as being important for oral tumor development. On chromosome 5, inactivation of genes located at 5q21-22 is thought to be common and, thus, they may be associated with the initiation or progression of OSCC [Mao et al. 1998]. Besides that, we might speculate about the putative importance of the *IL12B* (5q31.1-q33.1), *IL4* (5q31.1) and *RAD17* (5q13) genes, which frequently showed losses in our cohort. On chromosome 6 we observed a preferential gain of the *VEGF* (6p21.1) gene, both in tumor tissue and in macroscopically tumor-free tissue. Previously, the expression of *VEGF* and its receptors has been found to increase during tumor growth [Brekken et al. 2000], thus turning them into potential targets for cancer therapy. In our cohort, copy number alterations on chromosome 8 were observed in the largest number of patient samples. We showed that in two samples all the genes analyzed on 8q exhibited copy number gains, and that three samples showed losses of all genes analyzed on 8p. These results are in line with other studies and are compatible with the formation of isochromosomes 8q [da Silva Veiga et al. 2003]. We find, however, that isochromosomes do not represent the main rearrangements that occur in this chromosome, since our data, as well as data reported by others, strongly suggest gains and losses that do not match with such scenario. We found that gains in 8q were more frequent than losses in 8p. Similar to our results, Lin et al identified gain of 8q as the most prevalent chromosomal anomaly in these tumors [S. C. Lin et al. 2002]. Band 8q24, which harbors the *MYC* gene, is considered to be most important. Garnis et al raised the possibility of additional oncogenes near the *MYC* gene to be relevant for the progression of oral cancer [C. Garnis et al. 2004a]. The same authors [Cathie Garnis et al. 2004b] observed amplifications in band 8q22 and postulated that the *LRP12* gene, which maps in this region, could be implicated in oral carcinogenesis. Our study showed that the distal region of chromosome 8 is most frequently altered, i.e., the *PTK2* (8q24.3) gene followed by the *PTP4A3* (8q24.3) gene. The *MYC* and *LRP12* genes exhibited less frequent gains (Figure 2). Gains in *PTK2* copy numbers have also been reported by others [Cha et al. 2011; Susanne M. Gollin 2001; Sparano et al. 2006], suggesting its relevance for oral carcinogenesis. Thus, in our cohort it seems that two events are prevalent: amplification of the 8q24 region and isochromosome 8q formation with a concomitant loss of 8p. These events were not always seen in the same tumors. Regarding the losses at 8p, our study highlights the gene *TUSC3* (8p22) as being putatively important to oral carcinogenesis. This gene has already been suggested as an

important target of genomic rearrangements in epithelial cancers [Cooke et al. 2008]. On chromosome 9 we observed preferential copy number losses at 9p21, where the genes *CDKN2A* and *CDKN2B* are located. The *CDKN2A* and *CDKN2B* genes are positioned in tandem, spanning a region of approximately 80 kb, with *CDKN2B* located 25 kb centromeric to *CDKN2A* [Worsham et al. 2003]. The *CDKN2A* gene is considered to be the major tumor suppressor gene that is targeted in a wide variety of human cancers [Rocco and Sidransky 2001]. In our study, we identified more frequently losses of *CDKN2A* than of *CDKN2B*, both in tumor tissues and in macroscopically tumor-free tissues, supporting the idea that *CDKN2A* has a more important role in oral carcinogenesis than *CDKN2B.* This genomic imbalance has been pointed out as the most common of all genetic changes occurring in the early progression of oral tumors [van der Riet et al. 1994]. Concerning the second chromosome most commonly altered in terms of gain of genetic material, i.e., chromosome 11, we found that band 11q13 was most frequently increased (57%). Similarly, others have reported amplifications of band 11q13 in about 30-45% of head and neck cancers [C. Jin et al. 2006; Parikh et al. 2007]. Initially, the *CCND1* gene was considered to be the most important 11q13 target gene [Freier et al. 2006; Mineta et al. 2000; Namazie et al. 2002; Yu et al. 2005]. More recently, however, other candidate genes have been considered as possible cofactors of *CCND1,* i.e., *FGF3* and *CTTN* [Freier et al. 2006]. In our study we observed co-amplification of these three genes in 20 patient samples, which corroborates the findings described in the literature and provides strong evidence for the importance of these genes in oral carcinogenesis. However, copy number gain of other genes located in this region, such as the *TMEM16A* gene, may also contribute to oral carcinogenesis [Duvvuri et al. 2012]. We also identified losses in 11q, mostly the *CASP1* (11q23) gene. Additionally, two other genes appear to be interesting, i.e., *ATM* (11q22-q23) and *BIRC2* (11q22). The last one showed copy number losses and gains in seven and eight patients, respectively. As opposed to the gains in the 11q13 region, where some genes were already identified and correlated to clinical outcome, such information is as yet scarce on the losses observed in 11q. Parikh et al proposed that haploinsufficiency or copy number loss of the *ATM* gene may contribute to defects in DNA damage responses and reduced sensitivity to ionizing radiation, which consequently may lead to tumor progression [Parikh et al. 2007]. Based on our data, we can raise the hypothesis that not only the *ATM* gene may be important but also, and perhaps even more so, the *CASP1* gene, since its encoded cysteine-aspartic acid protease 1 is known to play a central role in apoptosis. Thus, loss of this gene may contribute to a decrease in apoptotic signaling.

Interestingly, this may open options for pro-apoptotic drugs as a valid choice for treatment [Nilsson 2013]. On chromosome 16 we identified mostly gains of the *MVP* (16p11.2) and *CDH1* (16q22.1) genes. MVP/vaults have been associated with chemoresistance in primary tumors and various tumor cell lines. Therefore, *MVP* is frequently considered as a negative prognostic factor for response to chemotherapy, as well as disease-free survival and/or overall survival [Mossink et al. 2003]. The *CDH1* gene, which encodes E-cadherin, is one of the most important genes regulating cell-cell adhesion in epithelial tissues [Pecina-Slaus et al. 2005]. This gene is considered to be an invasion-suppressor gene and loss of function of its encoded protein has been correlated with increased invasiveness and metastatic potential of tumors [Pecina-Slaus et al. 2005]. On chromosome 17 we detected frequent copy number gains of the *BRCA1* (17q21), *TP53* (17p13.1) and *CRK* (17p13.3) genes. Previously, Hardisson et al detected chromosome 17 anomalies using fluorescence in situ hybridization (FISH) analysis in pharynx and larynx carcinomas [Hardisson et al. 2004]. As *TP53* is well-known for its role as guardian of the genome, loss was expected for this region. Additionally, besides deletions, *TP53* gene mutations and protein inactivation were also reported [Scully et al. 2000]. On chromosome 18 we observed gene losses on its q-arm. Some of these genes may specifically be important for oral carcinogenesis, i.e., *CDH2* (18q11.2), *BCL2* (18q21.3) and *DCC* (18q21.3). In head and neck cancer loss of 18q is commonly observed, and the putative importance of the *DCC* gene for these tumors has already been highlighted [Susanne M. Gollin 2001]. On chromosome 19 we identified copy number gains in all genes analyzed. The *BAX* (19q13.3-q13.4) and *CDKN2D* (19p13) genes showed gains both in tumor tissues and in macroscopically tumor-free tissues. Amplifications of the 19q13 region in oral and esophageal carcinomas have already been described [Carneiro et al. 2008; Martin et al. 2008]. Unexpectedly, we found that some of the genes that have previously been described as tumor suppressor genes were amplified in our study. This observation could be explained by the fact that massive DNA rearrangements occurred randomly in some of the chromosomes. Some of these rearranged regions may harbor both tumor suppressor genes and dominantly acting oncogenes which, ultimately, may have led to a gain of these regions. A simpler explanation could be that these tumor suppressor genes are non-functional (e.g. due to hypermethylation) and, thus may have been amplified as a passenger event in samples showing genetic instability. Clearly, further studies are required in order to correctly interpret these rearrangements.

Importance of evaluating surgical margins

It is worth noting that surgical margins in two of the patients included in our study were histologically evaluated as tumor positive. Surgical margins represent macroscopically tumor-free tissues. We detected genetic imbalances in both histologically positive and negative margins. Delineating the exact area of excision is a great challenge for physicians, since the presence of tumor cells within or close to a surgical margin may be indicative for a risk of relapse which, in turn, affects additional treatment options [B. J. M. Braakhuis et al. 2010]. Unfortunately, relapses also occur in patients with histologically tumor-free margins after surgery. Therefore, our understanding of the transition from normal mucosa to potentially malignant oral mucosa, and from that to tumor needs to be explored in more detail. Several questions need to be addressed, such as how many genetic events need to occur in a cell that clonally expands and spreads to normal epithelium in order to create a field of transformed epithelium? Is there a common early genetic event that drives the development of multiple tumors or recurrences? If at the onset of carcinogenesis the cells do carry genetic alterations but at the histological level seem to be normal, what happens genetically, in quantitative and qualitative terms, to make the distinction between a histologically normal and a histologically malignant appearance? The molecular understanding of this transition is crucial for the development and implementation of biomarkers in routine diagnostics. The analysis of macroscopically tumor-free tissues from surgical margins made us aware of the fact that these surgical margins frequently show genetic imbalances similar to those encountered in the tumor from the same patient. Such findings may be indicative for an increased risk of relapses. In light of that, a rigorous follow-up of these patients seems mandatory, as well as the development of a non-invasive way to perform this surveillance.

Prediction models correlating genetic profiles with clinical features

Despite the fact that genomic profiling can be extremely useful for distinguishing different tumor sub-types, most probably because of the relatively small number of patients enrolled, we could not make a distinction between oral tumors with different clinicopathologic features on basis on their genomic profiles. We found that the genomic profiles of tumors in stages I or II and tumors in stages III or IV were very similar. The same was found for patients that exhibited metastases and those who did

not, as well as for smokers and non-smokers. We did find, however, that smokers exhibited losses at 3p (*MLH1*) and 11q (*ATM*) (Figure 3C), suggesting the occurrence of specific imbalances that could be characteristic for smokers. Such data might be important for classifying OSCC patients in different subgroups with different prognostics as well as different therapeutic responses. We anticipate that with a larger cohort it may be possible to molecularly identify such OSCC subgroups.

Califano et al were the first to describe a progression model for OSCC [Califano et al. 1996]. In this model, losses at chromosomal regions 3p, 9p and 17p were considered early events in the carcinogenic process. In the present study, comparisons between tumor and matched macroscopically tumor-free tissues allowed us to build a logistic regression model to predict the two types of tissue. By applying this model, the *TUSC3* gene turned out to be the only one that reached statistical significance, which may be indicative for its relevance in the development of oral tumors. Thus, the *TUSC3* gene may play a role in the transition from normal oral mucosa to potentially malignant oral mucosa. Precise and adequate prediction models are essential to determine the patients eligibility for clinical trials and to predict the disease outcome, as well as to select individualized therapies for each case.

Role of HPV typing

In this study, we only identified two HPV-positive patient samples. This result is not surprising taking into account that in the group of head and neck cancers HPV infection has been reported most particularly in association with oropharynx carcinoma, where its incidence may in some cases reach up to 60% [Bahl et al. 2013; Gillison 2004]. It has been reported that HPV-positive and HPV-negative tumors may exhibit distinct clinicopathological and molecular features [Leemans et al. 2011]. In our study it was impossible to assess whether HPV-positive cases represent a distinct group with a genetic profile different from HPV-negative cases (Figure 1A). It is relevant to note here that the mean age of our cohort was 61.5 years, whereas the incidence of oral tumors has increased mostly among younger people. Perhaps in these cases HPV vaccination could be an option.

In spite of continuous technological progress, our understanding of oral tumors is still limited. It is, therefore, peremptory to identify genes that may serve as good candidates for further studies, in order to validate them as biomarkers and to translate their application into routine clinical practice. Our current results not only reinforce

previous reports, but also revealed novel imbalances in chromosomes 2, 3, 4, 5, 6, 8, 9, 11, 16, 17, 18 and 19, with a putative impact in terms of clinical management. Selection of the most frequently altered genes may be instrumental for the development of biomarkers distinguishing between different susceptibilities for relapses and, possibly, different chemotherapeutic agents. In order to distinguish tumor tissue from tumor-free tissue, the *TUSC3* gene may serve as a bona fide biomarker. In the future the logistic regression model presented here could help clinicians to optimize the clinical management of OSCC patients by improving the estimation of the risk of relapses, the survival rates and, ultimately, the prognosis.

Acknowledgements

The authors are grateful to Dr. Artur Ferreira, Director of the Maxillofacial Surgery Unit from Coimbra Hospital and University Centre, for his contribution in the collection of the samples. This work was supported in part by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra.

Chapter 3.1.1 - Supplementary Table 1. All genes analyzed in this study using four MLPA probemixes (P005, P006, P007 and P014).

Pos. chr.	Genes	Pos. chr.	Genes	Pos. chr.	Genes	Pos. chr.	Genes	Pos. chr.	Genes	Pos. chr.	Genes
Chromosome 1		**Chromosome 5**		8p12	*FGFR1*	**Chromosome 10**		**Chromosome 13**		**Chromosome 18**	
1p36.22	*TNFRSF1B*	5q13.2	*RAD17*	8q11.21	*PRKDC*	10p14	*UPF2*	13q13.1	*BRCA2*	18q11.2	*CDH2*
1p34.2	*CTPS*		*IL13*	8q12.1	*MOS*	10p11.21	*CREM*	13q14.2	*RB1*	18q21.2	*DCC*
1p22.1	*BCAR3*	5q31.1	*IL4*	8q12.2	*CHD7*	10q23.31	*PTEN*	13q32.1	*ABCC4*	18q21.32	*PMPAIP1*
1p21.3	*F3*	5q31.1-q33.1	*IL12B*	8q13.1	*MYBL1*	**Chromosome 11**			*ING1*	18q21.33	*BCL2*
1p13.2	*NRAS*	**Chromosome 6**		8q13.3	*NCOA2*	11p15.5	*HRAS*	13q34	*ARHGEF7*	**Chromosome 19**	
1q32.1	*IL10*		*HLAF*	8q21.13	*TPD52*		*EHF*	**Chromosome 14**		19p13.2	*DNMT1*
Chromosome 2			*MDC1*	8q21.2	*E2F5*		*HIPK3*	14q12	*TINF2*	19q13.2	*CDKN2D*
2p21	*MSH2*		*IER3*	8q22.1	*RAD54B*	11p13	*LMO2*	14q13.2	*NFKBIA*		*BAX*
2q13	*IL1A*	6p21.3	*LTA*	8q22.3	*LRP12*		*CD44*	**Chromosome 15**		19q13.33	*KLK3*
2q24.2	*TANK*		*TNF*	8q23.1	*EIF3E*	11q13.1	*RELA*	15q21.1	*B2M*	**Chromosome 20**	
2q33.1	*CFLAR*		*BAK1*	8q24.11	*EIF3H*	11q13.2	*CCND1*	15q25.1	*MESDC1*	20p11.21	*THBD*
2q33.3-q34	*ERBB4*	6p21.2	*CDKN1A*		*EXT1*	11q13.3	*FGF3*	15q26.3	*IGF1R*	20q11.21	*BCL2L1*
Chromosome 3		6p21.1	*VEGF*	8q24.13	*RNF139*	11q13	*CTTN*	**Chromosome 16**		20q13.12	*NCOA3*
3p25.3	*VHL*	6q23.3	*MYB*		*MYC*	11q22.2	*BIRC2*	16p11.2	*MVP*	20q13.13	*PTPN1*
3p21.3	*MLH1*	**Chromosome 7**		8q24.21	*DDEF1*	11q23	*CASP1*	16q22.1	*CDH1*	20q13.31	*AURKA*
3p21	*CTNNB1*	7q21.12	*ABCB1*	8q24.22	*KCNQ3*		*ATM*	**Chromosome 17**		**Chromosome 21**	
3q25.2	*MME*	7q31.2	*MET*		*SLA*	11q23.1	*IL18*	17p13.3	*CRK*	21q11.2	*HSPA13*
3q25.33	*IL12A*	7q34-q35	*CASP2*	8q24.23	*KNDRBS3*	**Chromosome 12**		17p13.1	*TP53 - exon 4*	21q22.13	*SIM2*
3q26.33	*PIK3CA*	**Chromosome 8**			*KCNK9*	12p13.3	*CCND2*	17p13.1	*TP53 - exon 11*	21q22.3	*TFF1*
3q27	*BCL6*	8p23.3	*DLGAP2*		*PTK2*		*TNFRSF1A*	17p12	*PMP22*	**Chromosome 22**	
3q28	*Hs.570518*		*MFHAS1*	8q24.3	*PTP4A3*	12p13.31	*CD27*		*LHX1*	22q11.23	*MIF*
Chromosome 4			*MSRA*		*RECQL4*	12p12.1	*LRMP*	17q12	*ERBB2*		
4q22.1	*ABCG2*	8p23.1	*GATA4*	**Chromosome 9**			*IFNG*	17q21.2	*TOP2A*		
4q24	*NFKB1*		*CTSB*		*CDKN2A*	12q15	*MDM2*	17q21.31	*BRCA1*		
4q25	*CASP6*	8p22	*TUSC3*	9p21	*CDKN2B*	12q24.31	*BCL7A*	17q24.1	*AXIN2*		
4q26-q27	*IL2*	8p21.3	*ChGn*					17q25.3	*TIMP2*		

3.1.2 Genetic imbalances detected by Multiplex Ligation-dependent Probe Amplification in a cohort of patients with oral squamous cell carcinoma - the first step towards clinical personalized medicine

Ilda P. Ribeiro, Francisco Marques, Francisco Caramelo, José Ferrão, Hugo Prazeres, Maria J. Julião, Widad Rifi, Suvi Savola, Joana B. Melo, Isabel P. Baptista, Isabel M. Carreira

Tumor Biology
Volume 35, Issue 5, Pages 4687-95, (2014)
doi: 10.1007/s13277-014-1614-9

Abstract

Objectives Oral tumors are a growing health problem worldwide, thus, it is mandatory to establish genetic markers in order to improve diagnosis, early detection of tumors, control relapses and ultimately, to delineate individualized therapies. This study was the first to evaluate and discuss the clinical applicability of a MLPA probe panel directed to head and neck cancer. ***Materials and Methods*** 30 primary oral squamous cell tumors were analyzed using the P428 MLPA probe panel. ***Results*** We detected genetic imbalances in 26 patients and we observed a consistent pattern of distribution of genetic alterations in terms of losses and gains for some chromosomes, particularly for chromosomes 3, 8 and 11. Regarding the latter, some specific genes were highlighted due to frequent losses of genetic material - *RARB*, *FHIT*, *CSMD1*, *GATA4*, *MTUS1* – and others due to gains - *MCCC1*, *MYC*, *WISP1*, *PTK2*, *CCND1*, *FGF4*, *FADD* and *CTTN*. We also verified that the gains of *MYC* and *WISP1* genes seem to suggest higher propensity of tumors localized in the floor of the mouth. ***Conclusions*** This study proved the value of this MLPA probe panel for a first-tier analysis of oral tumors. The probemix was developed to include target regions that have been already shown to be of diagnostic/prognostic relevance for oral tumours. Furthermore, this study emphasized several of those specific genetic targets, suggesting its importance to oral tumors development, to predict patients' outcomes and also to guide the development of novel molecular therapies.

Keywords: Genetic imbalances, MLPA technique, Gains and losses, Oral squamous cell carcinoma

Introduction

Oral tumors belong to head and neck cancers, of which more than 90% are squamous cell carcinoma [Forastiere et al. 2001]. The overall incidence of oral cancer seems to be rising worldwide: for 2013, 36 000 new cases and 6 850 deaths are estimated, including lip, oral cavity, and pharynx cancer [Mitka 2013]. This shows that oral squamous cell carcinoma (OSCC) is an aggressive neoplasm frequently diagnosed at an advanced stage and associated with high mortality and morbidity. Its high incidence in young people has been correlated with tobacco, alcohol consumption and, mostly, with human papilloma virus (HPV) infection [Llewellyn et al. 2001]. At the time of diagnosis, more than half of oral cancer patients have local or distant metastases, which mean a 5-year survival rate of only 33.2% [Mitka 2013]. Albeit the technological progress and the improvements in therapeutic modalities, the survival rates of oral cancer have remained the same in the recent decades, which could indicate that the markers of prognosis, such as tumor size, lymph node involvement and tumor stage, are not enough to predict clinical outcome, choose the best treatment and least of all to improve the early detection. Genetic classification of oral tumors identifies subgroups with a different prognosis [Wittekindt et al. 2012]. In this respect, as oral cancer is a heterogeneous disease, attempts have been made to identify the genetic imbalances associated with the histological progression of oral cancer, but it remains imperative to establish biomarkers for oral tumors, in order to choose the most appropriate clinical therapy according to each patient's genetic profile of tumor. Nowadays, in order to facilitate the early detection and the management of oral tumors, potentially useful diagnostic tools at clinical and molecular levels have been developed [Shah et al. 2011]. Multiplex Ligation-dependent Probe Amplification (MLPA) is a molecular technique developed in 2002 by Schouten et al that allows amplification and semi-quantitative detection of up to 50 sequences of DNA in the same reaction with a single pair of primers [Schouten et al. 2002]. This technique has been shown to be applicable not only for the diagnostic of inherited or congenital diseases, but also for tumor diagnostic and prediction of cancer progression [Homig-Holzel and Savola 2012]. This study is the first one to perform an evaluation of the genetic profile of oral cavity carcinomas through copy number variations using a MLPA probe panel specific for head and neck cancer, in a cohort of 30 patients. We also discuss the routine clinical application of this MLPA probe panel, highlighting the

most frequent imbalances detected in these tumors and, additionally, correlating these genetic alterations with clinical characteristics.

Material and Methods

Patients and DNA Isolation

The present study was conducted on 30 primary OSCC samples. These samples were obtained between 2010 and 2012 from the Maxillofacial Surgery and Stomatology Unit, of the Coimbra Hospital and University Centre, CHUC, EPE, Portugal. All patients were submitted to surgery and the histopathologic diagnosis of the mirror sections of these fragments were performed by two different pathologists. Hematoxylin and eosin staining was used to evaluated tumor content in each specimen. In our set, all samples contained at least 50% tumor cells. Diagnosis and staging were performed in accordance with the American Joint Committee on Cancer's TNM staging system for OSCCs [Wittekind et al. 2003]. The patients gave their informed consent in accordance with the regulations in the Declaration of Helsinki. The study was approved by Ethics committee of the Faculty of Medicine of the University of Coimbra. The detailed characterization of our cohort with OSCC diagnosis is illustrated in table 1. Controls were 15 gingival tissue from healthy subjects submitted to the "wisdom teeth" removal. DNA from patients and controls were extracted from fresh frozen tissue using a High Pure PCR Template Preparation Kit (Roche GmbH, Mannheim, Germany), according to the manufacturer's instructions and quantified by UV spectrophotometric analysis using Nanodrop 1000 Spectrophotometer (Thermo Scientific, USA).

MLPA probemix development and MLPA analysis

MLPA analysis was performed using a MLPA probe panel specific for head and neck squamous cell carcinoma (HNSCC). This MLPA assay was designed and optimized at MRC-Holland (Amsterdam, the Netherlands). For selection of probes included in this assay an extensive literature search was performed to pinpoint genes, which are suggested to be of diagnostic and/or prognostic importance in HNSCC. The probes included in this assay were selected from the MRC-Holland MLPA probe database, and if there were no existing MLPA probes available, new probes were designed. All the MLPA probes included in this assay were tested under various conditions by changing hybridization temperature as well as salt, probe and polymerase

concentration. Moreover, the variability of these MLPA probes was tested on healthy individuals (n = 24) and only probes that were shown to have standard deviation under 0.10 were included this assay. After the optimization and quality testing phase, this P428-B1-lot1111 MLPA panel) included 41 probes targeting 36 different genes, located on chromosomes 3, 4, 5, 7, 8, 11, 13 and 18 (Figure 1). Additionally, 11 references probes were included, which detect 10 different autosomal chromosomal regions considered relatively quiet in HNSCC. This MLPA assay contains also nine quality control probes to assess DNA denaturation and DNA quantity in the MLPA reaction. All MLPA reactions were performed according to the standard MLPA reaction protocol described by Schouten et al [Schouten et al. 2002]. Briefly, 100 ng of DNA samples were heated at 98 ºC for 10 min. After the addition of the probemix, samples were heated for 1 min at 95 ºC and then incubated for 16h at 60ºC. Ligation of the annealed oligonucleotide probes was performed for 15min at 54ºC in buffer containing Ligase-65 enzyme. After inactivating the ligase enzyme by heating at 98ºC for 5min, multiplex PCR was carried out using FAM- labeled primers, dNTPs and SALSA polymerase. PCR was performed for 35 cycles of 30s at 95ºC, 30s at 60ºC and 1min at 72ºC. All the reactions were carried out in a thermal cycler equipped with a heat lid (ABI 2720, Applied Biosystems, Foster City, CA, USA). PCR products were heat denatured and analyzed on a Gene Scan ABI PRISM 3130 capillary electrophoresis system (Applied Biosystems, Foster City, CA, USA). Three controls and a negative control (without DNA) were always included in each MLPA experiment. The sample results are displayed as a ratio between the reference and experimental samples. Data-binning of raw data and comparative analysis was performed using Coffalyser.NET software [Coffa and Berg 2011]. For each MLPA probe we determined specific cutoff values for gain and loss, using the values limiting the 95% confidence interval as determined on non-cancer subjects. A numerical gain was scored when the values exceeded the 1.2 and a numerical loss was defined when the values were lower than 0.8.

HPV Typing

All tumor tissue samples were analyzed for HPV infection as described by Nobre et al [Nobre et al. 2010]. Briefly, we used PCR methods employing established general consensus and degenerate primer sets, namely GP5+/GP6+ and MY09/MY11, which were designed to amplify a fragment of the L1 gene of mucosatropic HPVs. For

genotyping we performed the sequencing of intra-primer DNA in the amplified fragments in order to determine the specific types of HPV present in the samples. In addition to DNA sequencing, we complemented our analysis by using DNA array hybridization, using HPV CLART2 arrays (Genomica), as a way to address cases that shown infection with multiple HPV genotypes.

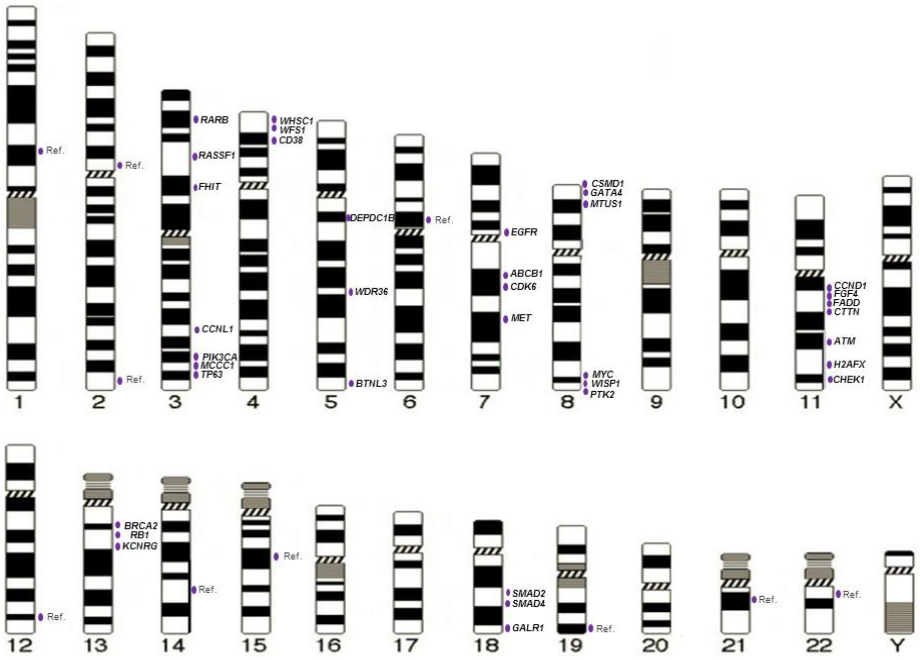
Chapter 3.1.2 -Table 1. Patient and tumor characteristics

	Patients (n = 30)
Median Age, years (range)	**63.1** (37-84)
Sex	
Male	26
Female	4
Smoking (cigarettes/day)	
≥20	17
< 20	3
None	10
Stage	
I and II	9
III and IV	21
Site	
Tongue	13
Floor of the mouth	12
buccal mucosa	4
Palate	1
Treatment	
Surgery + RT	7
QT + Surgery	2
Surgery + RT + QT	6
Surgery only	15
Clinical outcome	
Death from the disease	10
Dead of other cause	1
Alive	19

RT - Radiation Therapy; **QT -** Chemotherapy

Statistical analysis

The statistical analysis was carried out with IBM SPSS v.20 software package, assuming a significance level of 0.05. Descriptive analysis of losses and gains frequencies in genes, in each mouth location and tumor staging, was performed. In addition, we created convenient charts for better understanding the relationships between genes, locations and tumor staging. In order to verify the existence of an association, with statistical meaning, between gains/losses in genes and location, a chi-square test was made. The same test was also applied to the association between gains/losses in genes and tumor staging.



Chapter 3.1.2 - Figure 1. Chromosomal distribution of the genes studied using P428 MLPA probe panel as well as references probes (Ref).

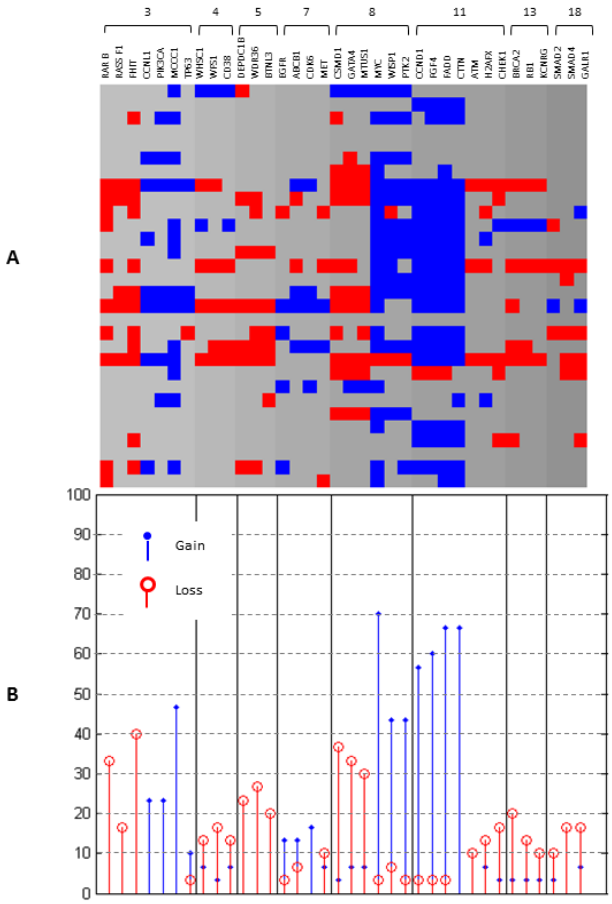
Results

Genetic profile of oral tumor tissue

This study showed the existence of genetic imbalances in 26 of the 30 oral tumor specimens analyzed (Figure 2A). Four tumor samples (13.3%, 4 of 30 specimens) did not harbor copy number aberrations in the target regions of the MLPA probemix. These patient samples were further analyzed with an oligonucleotide microarray 180 K (Agilent Technologies, Santa Clara, CA, USA) and we did not detect copy number alterations in the regions targeted by MLPA assay (data not shown). Simultaneously, we did not observe any genetic imbalance for the 36 genes analyzed with this MLPA probe panel in all the 15 gingival tissues from healthy subjects submitted to the "wisdom teeth" removal (data not shown), which correspond to the control group.

Regarding the genetic profile of tumor samples, overall, we verified that the number of tumors that carried gains of genetic material was higher than the number of tumors that carried losses (Figure 2B). It was also possible to observe a pattern of imbalances in some chromosomes; therefore the distribution of genetic alterations in terms of losses and gains was very consistent for some chromosomes (Figure 2A). In this sense, 3p and 8p showed mostly losses, while 3q and 8q frequently displayed gains. Chromosomes 4, 5, 13 and 18 presented mainly losses for the analyzed genes. On the

other hand, chromosome 7 showed more frequent gains than losses for the analyzed genes. Regarding chromosome 11, in band q13 we verified almost exclusively gains and in q-distal we observed predominantly losses. Regarding all these imbalances (Figure 2A, B), we verified that the most frequently altered chromosomes were 3, 8 and 11. Moreover, chromosome 8 showed the highest frequency of gains (70%) in the largest number of patients. The second most frequently altered chromosome in terms of gains was chromosome 11 (66.7%), followed by 3q (46.7%). Regarding the losses of genetic material, the most commonly altered chromosome was 3p (40%), followed by 8p (36.7%).



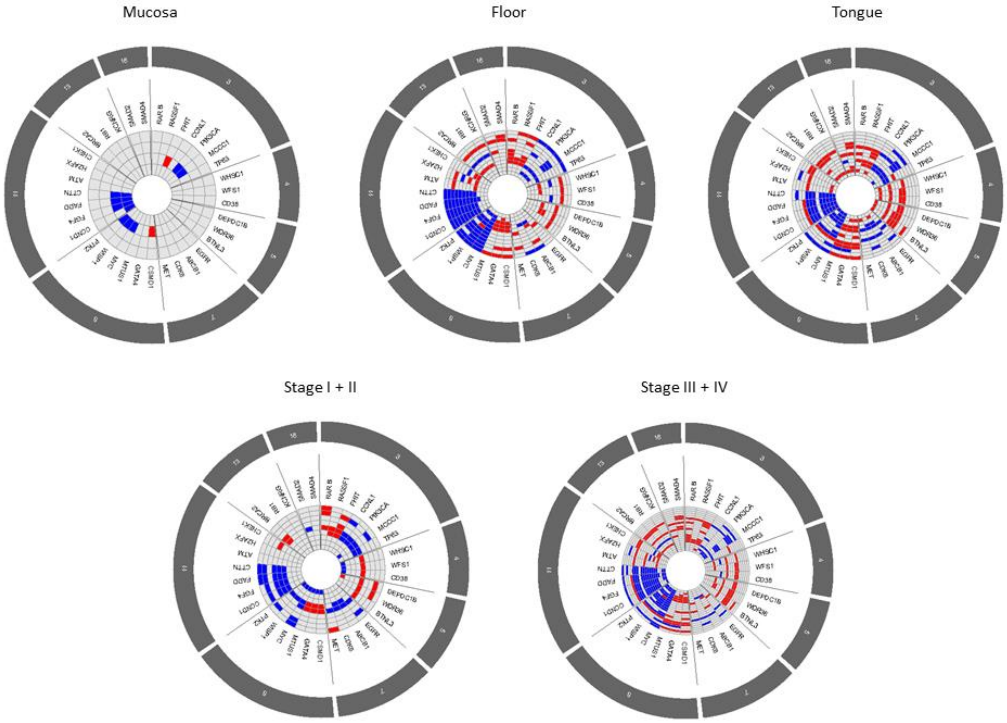
Chapter 3.1.2 - Figure 2. Genetic imbalances (gains and losses) in the 30 OSCC patients detected using P428 MLPA probe panel. Losses of genetic material are represented by red, and blue represents gains. **A)** Each row represents one patient and each pixel in this row corresponds to one gene analyzed. The gray color represents genes without alteration, and each shade of gray shows the localization of the genes in each specific chromosome. The genes analyzed, from the left to the right, are ordered by chromosome from the short arm to the long arm; **B)** The picture shows the percentage of the imbalances in all genes analyzed. Each line represents one gene analyzed.

No single gene was altered in all patients. On the other hand, all genes analyzed presented gain and/or loss of genetic material in at least three patients. It is important to

notice that the following genes only presented exclusively loss of genetic material in these oral tumor samples: *RARB* (3p24.2), *RASSF1* (3p21.31), *FHIT* (3p14.2), *DEPDC1B* (5q12.1), *WDE36* (5q22.1), *BTNL3* (5q35.3), *ATM* (11q22.3) and *SMAD4* (18q21.2). Additionally, the following genes were the ones that presented exclusively gain of genetic material: *CCNL1* (3q25.32), *PIK3CA* (3q26.33), *MCCC1* (3q26.33), *CDK6* (7q21.12) and *CTTN* (11q13). The remaining 23 genes analyzed in this study presented both gain and loss of genetic material (Figure 2A, B); however, for some of those genes it was possible observe that one specific event is much more prevalent than the other, e.g. gain or loss. In this sense, after analyzing in detail every gene in the most commonly altered chromosomes, we verified that from the 36 genes studied it was possible to highlight some of those that frequently showed gains or losses of genetic material. Thus, taking only into account the genes altered in ≥ 30% of the oral tumor samples (Figure 2A, B), the following genes were the ones to present the most frequent losses: *RARB* (3p24.2), *FHIT* (3p14.2), *CSMD1* (8p23.1), *GATA4* (8p23.1) and *MTUS1* (8p22); and these were the ones that presented gains: *MCCC1* (3q26.33), *MYC* (8q24.21), *WISP1* (8q24.22), *PTK2* (8q24.3), *CCND1* (11q13.3), *FGF4* (11q13.3), *FADD* (11q13.3) and *CTTN* (11q13.3).

Association between clinicopathological features and genetic imbalances

Overall, regarding the different anatomic localizations and stages of tumors (Figure 3) it was not possible to reach statistical significance in order to discriminate the patients according their genetic profile and clinicopathologic characteristics. Taking into account only three anatomic localizations (buccal mucosa, floor of the mouth and tongue) - and excluding the palate due to the fact that only one tumor was located there - and performing correction for multiple comparisons, we did not observe any gene with statistical significance to perform this separation; however, if the correction of multiple comparisons (Bonferroni correction) was not applied, two genes from chromosome 8 showed statistical significance. In this sense, gain in these two genes, *MYC* and *WISP1* could suggest that the tumor is predominantly localized in the floor of the mouth. Regarding HPV infection, we only detected two HPV-positive patients from the 30 patients' analyzed (data not shown). One showed a HPV type belonging to high-risk classes, HPV type 31, and another belonging to low-risk classes, HPV type 42. The latter HPV patient did not have any other risk factors, such as tobacco or alcohol consumption.



Chapter 3.1.2 - Figure 3. Radial heatmap of the genetic imbalances for tumor samples in the three anatomic localizations (buccal mucosa, floor of the mouth and tongue) and in stage I or II, and in stage III or IV of the tumors. Each line represents one patient. Red lines represent losses of genetic material and blue lines gains.

Discussion

Genetic profile

Oral cavity tumors develop through multistep genetic pathways, involving typically losses of tumor suppressor genes and gains of oncogenes. Presently, the probe panel P428 is the only one exclusively developed MLPA assay for head and neck cancer. This is, therefore, the first study using this MLPA probe panel in patients with OSCC diagnosis. In our patient cohort the most frequently altered chromosomes were 3, 8 and 11. We observed losses at 3p and gains at 3q, a pattern that had already been previously described for HNSCC [Bockmuhl et al. 1997]. Imbalances at 3p are associated to the early HNSCC [Redon et al. 2001]. In our cohort the most common losses in 3p included *FHIT* (40%) and *RARB* (33.3%) tumor suppressor genes. *FHIT* is frequently deleted in epithelial cancer cell lines [Ohta et al. 1996]. In 1996, Virgilio et al described that 55% of HNSCC cell lines express aberrant *FHIT* transcripts and that one or both *FHIT* alleles are deleted in many of these cell lines, which suggests that inactivation of *FHIT* is important for development and progression of HNSCC [Virgilio et al. 1996]. Also Saldivar et al have shown that loss of expression of Fhit protein causes DNA damage and genome

instability and otherwise expression of Fhit reduces this DNA damage and contribute to protection of genome stability, thus, loss of Fhit provides a selective advantage in sporadic cancers [Saldivar et al. 2012; Saldivar et al. 2013]. Concerning *RARB*, in 1995, Lotan et al verified that 60% of oral potential malignant lesions did not express this gene, in comparison to all the samples of normal tissue where this gene were expressed [Lotan et al. 1995]. This way, loss of RARB expression could enhance carcinogenesis through loss of response to retinoids [Zou et al. 2001]. Hence, this receptor could eventually be a valuable intermediate marker in trials of retinoids for the prevention of oral carcinogenesis. This point needs to be more carefully explored, since retinoids have proven activity in treating early potential malignant lesions [O'Shaughnessy et al. 2002]. Additionally, we detected in a wide range of copy number gains at 3q, namely at *MCCC1* (46.7%). In our cohort, concerning 3q, this gene was the one that presented the highest frequency of gains, besides aberrations in other genes, such as *PIK3CA* have already been described as extremely important for oral tumors development [Murugan et al. 2008].

We observed frequent losses in 8p, which is in line with literature describing deletions at 8p in oral tumors [Hermsen et al. 2001]. In our cohort, most frequently deleted genes in 8p were *CSMD1* (36.7%), *GATA4* (33.3%) and *MTUS1* (30%). *CSMD1* is considered to be a strong candidate at 8p23 [Sun et al. 2001] and loss of this gene is also present in other epithelial cancers, namely in breast and lung cancer [Ma et al. 2009]. *GATA4* has been implicated in colorectal and gastric development, where it is epigenetically silenced and expected to contribute to the tumor progression [Akiyama et al. 2003]; however, its function in human cancers is not yet fully understood. A tumor suppressor role for *GATA4* has been described, due to the methylation and loss of expression of this gene observed in lung, colorectal, and gastric cancer [Akiyama et al. 2003; Hellebrekers et al. 2009; Zheng and Blobel 2010]. *MTUS1* gene was initially identified as a candidate tumor suppressor gene in pancreatic cancer; moreover the ectopic expression of *MTUS1* gene products has been shown to inhibit the cell proliferation [Seibold et al. 2003]. In 2006, Zhou et al hypothesized that the reduction of *MTUS1* expression could be associated with advanced oral tongue squamous cell carcinoma [Zhou et al. 2006]. In 2007, Ye et al suggested that *MTUS1* gene is a potential tumor suppressor gene also for HNSCC and a promising candidate for further functional analysis [Ye et al. 2007]. Regarding the long arm of chromosome 8, we observed more frequently gains of the genetic material in our cohort, which is in line with the literature [Squire et al. 2002]. The altered genes mapped in 8q in the largest number of patients were *MYC* (70%), *WISP1* (43.3%) and *PTK2* (43.3%). Nowadays, *MYC* gene is

considered the most significant oncogene mapped on 8q and has been associated to the pathogenesis of several human cancers [da Silva Veiga et al. 2003]. Amplification and overexpression of this gene has been observed in 10-40% of human oral tumors [Saranath et al. 1989]. In the literature, tongue carcinomas showed the highest incidence of gains (75%) in *MYC* gene among oral tumors, resulting in significantly poor survival rates [Vora et al. 2003a]. In our cohort the highest incidence of gains in *MYC* gene was observed in the floor of the mouth. Regarding the role of *WISP1* gene, it has been observed that this gene could enhance or inhibit tumor growth. In colon carcinoma cell lines and in colon tumors significant increase of genomic copies and mRNA of *WISP1* were observed, in comparison to normal mucosa [Babic et al. 1998]. In melanoma cells WISP1 expression was inversely correlated with proliferation, metastasis and growth, as well as with metastasis of lung cancer [Hashimoto et al. 1996],[Soon et al. 2003]. Further studies will be needed to determine if *WISP1* has, in fact, prognostic value in predicting oral cancer metastasis. Amplification in *PTK2* (alias, *FAK*) gene has already been reported in HNSCC [Canel et al. 2006]. Similarly, FAK overexpression has been shown in tumor biopsy samples from a wide variety of tumors [Canel et al. 2006]. This overexpression has been associated with the invasive potential of the tumor [Owens et al. 1995].

In our patient cohort gains at 11q13 were detected and included *CCND1* (56.7%), *FGF4* (60%), *FADD* (66.7%) and *CTTN* (66.7%). Gains at band 11q13 are frequent in human cancer, and this event is also extremely common in HNSCC [Schuuring 1995]. *CCND1* and *CTTN* genes were the first to be identified [Schuuring 1995] and reported as candidates for driving 11q13.3 amplification, due to the fact that they are both recurrently co-amplified and this amplification has been correlated with its over-expression [Gibcus et al. 2007]. Amplification of 11q13 occurs in 30-60% of HNSCC, which also includes *FGF4* [Lese et al. 1995]. In 2007, Gibcus et al identified *FADD* gene as a potential driver gene in the 11q13 amplicon for head and neck cancer [Gibcus et al. 2007]. Similarly to this study, also in our cohort the *FADD* gene presented gains in the highest number of patients, when compared to the other genes analyzed in this chromosome. This gene plays a significant role in cell cycle regulation, which suggests its importance in the response to cytotoxic drugs. Thus, Gibcus et al hypothesized that HNSCC patients with 11q13.3 amplification and concomitant FADD overexpression could benefit from the administration of Taxol-based chemoradiotherapy over radiotherapy alone [Gibcus et al. 2007]. Similarly, loss of distal 11q contributes to chromosomal instability

and consequently to tumor progression as well as to resistance to therapy, namely reduced sensitivity to ionizing radiation [Parikh et al. 2007].

Correlation between genetic imbalances and clinicopathological features

Taking into account the genetic differences among the anatomic localizations, we have in our cohort a relatively low number of patients distributed by the four anatomic localizations in the oral cavity, which hinder this interpretation. In 2002, Huang et al found evidence indicating that subsets of HNSCC have different genetic patterns, allowing to perform the distinction by site of disease within the upper aerodigestive tract [Huang et al. 2002]. In line with that, we believe that some gains and losses could be tumor-site specific (Figure 3); however, in this first pilot study due to the few patients enrolled in each anatomic site, it was not possible to reach statistical significance in order to molecularly identify different subgroups of the oral cavity carcinoma. Besides that, in chromosome 8, gains of *MYC* and *WISP1* genes seem to suggest higher propensity of tumors localized in the floor of the mouth. This finding could be extremely important from a clinical point of view, because it helps to subdivide the patients according to these clinicopathological characteristics, which could signify different prognosis and, ultimately, different therapies. Concerning the genetic profile between tumors in stages I or II and in stages III or IV (Figure 3), the lack of statistical significance is probably due to the small number of patients enrolled in each stage. In this sense, further studies evaluating a larger cohort seem to be crucial in order to separate these tumors according the genetic profile. Regarding HPV status, we only identified two HPV-positive patients. Although it has been described that HPV-positive and HPV-negative tumors exhibited distinct clinicopathological and molecular entities,[Leemans et al. 2011] we cannot clarify whether HPV-positive patients have indeed a relatively favorable prognosis due to the presence of specific genetic imbalances (Figure 2A). Nowadays, however, some clinicopathologic variables have been validated to classify and prognosticate oral tumor patients, the predictive value of which is, in general, very low [Gibcus et al. 2007; Nix et al. 2004]. We believe that with a higher number of patients it might be possible to molecularly identify different subgroups of OSCC with different clinical relevance.

Detection of genetic imbalances - a step forward towards clinical personalized medicine

There is a very urgent need to establish genetic markers in order to improve diagnosis, help in the redefinition and reclassification of histology, stage of tumors and predict therapeutic outcome of oral tumor patients. Our results with this MLPA probe panel gave a step forward in this direction. MLPA is a cost-efficient technique with little hands-on time and compatible with routine diagnostic. Other extremely important advantage in tumor analysis is the capability of this technique to identify imbalances even in samples with wild-type tissue contamination,[Homig-Holzel and Savola 2012; Jeuken et al. 2006] as well as to provide information about the intra-tumoral mosaicism.

We can conclude that our study showed the value of MLPA technique for the analysis of oral tumor tissue in order to detect imbalances in specific genomic regions. Moreover, our study emphasized several specific genetic targets suggesting its putative importance to improve diagnostic, predict patients' outcome and also guide the development of novel molecular therapies, which must be further assessed and validated in a larger cohort of patients. This MLPA probe panel specific for HNSCC detected genetic imbalances in 86.7% of samples of our patient cohort, encompassing some genes that seem to be associated to early stage of oral tumors, other that are likely to be associated to development of metastasis and other possibly associated to therapeutic response, which could in the future represent the first step to help in the introduction of a clinical personalized medicine. Besides these promising results, the need to establish more specific biomarkers for these oral squamous tumors remains evident in order to implement a simple way of screening a wide range of imbalances with clinical utility not only in terms of diagnosis and prognosis but also to follow up high risk populations.

Conflict of interest

W. Rifi and S. Savola are employed by MRC-Holland, manufacturer of commercially available MLPA probemixes. All other authors declare no competing financial interests.

Acknowledgements

The authors are grateful to Dr. Artur Ferreira, Director of Maxillofacial Surgery Unit from Coimbra Hospital and University Centre, for the contribution in the collection

of the samples. This work was supported in part by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology, Faculty of Medicine, University of Coimbra).

3.1.3 *WT1, MSH6, GATA5* and *PAX5* as epigenetic oral squamous cell carcinoma biomarkers - a short report

Ilda P. Ribeiro, Francisco Caramelo, Francisco Marques, Ana Domingues, Margarida Mesquita, Leonor Barroso, Hugo Prazeres, Maria J. Julião, Isabel P. Baptista, Artur Ferreira, Joana B. Melo, Isabel M. Carreira

Cellular Oncology
Volume 39, Issue 6, Pages 573-582, (2016)
doi: 10.1007/s13402-016-0293-5

Abstract

Purpose Oral squamous cell carcinoma (OSCC) is a frequently occurring aggressive malignancy with a heterogeneous clinical behavior. Based on the paucity of specific early diagnostic and prognostic biomarkers, which hampers the appropriate treatment and, ultimately the development of novel targeted therapies, we aimed at identifying such biomarkers through a genetic and epigenetic analysis of these tumors. ***Methods*** 93 primary OSCCs were subjected to DNA copy number alteration (CNA) and methylation status using methylation-specific multiplex ligation-dependent probe amplification (MS-MPLA) analyses. The genetic and epigenetic OSCC profiles obtained were associated with the patients' clinic-pathological features. ***Results*** We found that *WT1* gene promoter methylation is a predictor of a better prognosis and that *MSH6* and *GATA5* gene promoter methylation serve as predictors of a worse prognosis. *GATA5* gene promoter methylation was found to be significantly associated with a shorter survival rate. In addition, we found that *PAX5* gene promoter methylation was significantly associated with tongue tumors. None of these genes exhibited CNAs. To the best of our knowledge, this is the first study that highlights this specific set of genes as epigenetic diagnostic and prognostic biomarkers in OSCC. ***Conclusions*** Our data highlight the importance of epigenetically assessing OSCCs to identify key genes that may serve as diagnostic and prognostic biomarkers and, potentially, as candidate therapeutic targets.

Keywords: DNA methylation; Copy number gains and losses; Predictors of prognosis; Oral squamous cell carcinoma

Introduction

Oral squamous cell carcinoma (OSCC) encompasses a range of tumors that belong to the group of head and neck cancers. OSCC elicits serious health problems due to its high mortality rate and its severe impact on the quality of life [Arantes et al. 2015]. The incidence of OSCC is increasing among younger individuals [Markopoulos 2012] and the risk factors include life style, i.e., tobacco and alcohol use, and infections with high risk human papillomaviruses (HPV) [Mohankumar et al. 2014; Salazar et al. 2014]. Although slight improvements have been made over the past decade, the 5-year survival rate of OSCC patients is still poor (~50-60%) [Nakaoka et al. 2014; Pulte and Brenner 2010]. In spite of the easy access of the oral cavity for clinical examination, OSCCs are often diagnosed in advanced stages due to the absence of visible changes in the mucosa at early stages. This late diagnosis leads to a poor prognosis and, consequently, one of the worst overall survival rates among human cancer patients [Tokumaru et al. 2004]. The molecular and clinical heterogeneity of OSCCs, which results from the accumulation of multiple genetic and epigenetic alterations, further hampers its diagnosis, prognosis and treatment [Leemans et al. 2011]. In routine clinical practice, TNM classification is used to predict the prognosis and to select the most appropriate treatment modalities, but this approach frequently fails to improve the disease outcome [Bezabeh et al. 2005]. Since OSCC patients often develop loco-regional recurrences, distant metastases and second primary tumors, the identification of diagnostic and prognostic biomarkers is of paramount importance for the detection of, next to primary tumors, relapses and metastasis and, consequently, for therapeutic decision making. Here, we link tumor-specific (epi)genetic alterations identified through methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) in 93 OSCC samples to its clinic-pathological features. By doing so, we found statistically significant associations of *GATA5* and *MSH6* gene promoter methylation with a worse prognosis and *WT1* gene promoter methylation with a better prognosis. In addition, we found that *PAX5* gene promoter methylation is associated with a localization at the tongue.

Materials and methods

Study population

The study protocol was approved by the Committee on Ethics in Research of the Faculty of Medicine of the University of Coimbra. All patients provided their written consent to participate in the study after being informed about the research purposes, following the regulations in the Declaration of Helsinki.

The study cohort includes tissue specimens from 93 OSCC patients who underwent tumor resection. The patients were recruited between October 2010 and March 2015 from the Maxillofacial Surgery and Stomatology Units of the Coimbra Hospital and University Centre, CHUC, EPE, Portugal. Diagnosis and staging were performed in accordance with the American Joint Committee on Cancer TNM staging system. The participants in this study answered a survey regarding lifestyle and risk factors for upper aerodigestive tract malignancies, including alcohol and tobacco consumption. Patients were followed-up through hospital revisits during routine clinical appointments. The final date of follow-up was April 30, 2015. The follow-up periods ranged from 1 to 54 months. Details of our study cohort are listed in table 1. For the control group, 16 gingival tissues from healthy subjects (6 males and 10 females), with ages ranging from 18 to 81 years and subjected to "wisdom teeth" removal, were used.
The tissue samples were snap-frozen in liquid nitrogen within 30 min after resection and stored at -80ºC until use.

DNA extraction, HPV typing and MS-MLPA

DNA from fresh frozen tissues of patients and controls were extracted using a High Pure PCR Template Preparation Kit (Roche GmbH, Mannheim, Germany), according to the manufacturer's instructions. The DNAs were quantified by UV spectrophotometric analysis using a Nanodrop 1000 Spectrophotometer (Thermo Scientific, USA). All tumor tissue samples were analyzed for HPV infection as described by Nobre et al [Nobre et al. 2010].

MS-MLPA analyses were performed using MS-MLPA probe sets ME001 and ME002 (MRC-Holland, Amsterdam, The Netherlands), which can simultaneously detect copy number alterations (CNAs) in 54 tumor suppressor genes and aberrant methylation patterns in a subset of 35 of these genes (Supplementary Table 1). All MS-MLPA reactions were performed according to a standard protocol described by Nygren

et al [Nygren et al. 2005], with minor modifications. Briefly, 100 ng of each DNA samples was denatured and, after the addition of the probemix, the probes were allowed to hybridize for 15h at 60°C. Subsequently, the samples were divided into two groups, i.e., half of the samples was directly ligated and in the other half the ligation was combined with *HhaI* digestion. Multiplex PCR was carried out for 35 cycles of 30s at 95°C, 30s at 60°C, and 1min at 72°C. All the reactions were carried out in a thermal cycler equipped with a heat lid (ABI 2720, Applied Biosystems, Foster City, CA, USA). The PCR products were heat-denatured and analyzed on a Gene Scan ABI PRISM 3130 capillary electrophoresis system (Applied Biosystems, Foster City, CA, USA). Three controls selected from 16 previously analyzed samples without CNAs and methylation values below 20%, as well as a negative control (without DNA), were always included in each MLPA assay. Binning of the raw data and comparative analyses were performed using Coffalyser.NET software. For each MLPA probe we determined the specific cutoff values for gain and loss, using 95% confidence intervals as determined on tissues from non-cancer subjects. A copy number gain was scored when a value exceeded 1.2 and a copy number loss was scored when a value was lower than 0.8. These cut-off values were based in our previous work [Ribeiro et al. 2014a]. With respect to methylation, we considered a gene promoter as methylated when the methylation dosage ratio was ≥ 0.20, which means that at least 20% of the DNA was methylated, as described in [Chmelarova et al. 2013].

Statistical analyses

The statistical analyses comprised three parts: first, a variable reduction step was carried out followed by a multivariate analysis and, finally, a survival analysis. For the reduction step the percentage of patient samples showing methylation at each gene was calculated. Similarly, the percentages of patient samples exhibiting copy number loss or gain at each gene were calculated. Only genes with a patient percentage > 20% were considered to be relevant for the remaining analyses. After selecting the most frequently methylated genes, a principal component analysis (PCA) was carried out with these genes and other clinic-pathological variables (i.e., pT and pN classification) aiming at reducing the number of variables and exploring the probable existence of relations between them by visual inspection of the PCA chart. Since nominal variables (e.g. gender and metastasis) were considered in the analyses, a nonlinear PCA was carried out resorting to the CATPCA algorithm available in the IBM® SPSS® statistical

software. A more tuned examination on these relations was considered by employing logistic regression (enter method) using as independent variable the methylation status.

The gene promotor methylation status was also studied with respect to its possible impact on patient survival. Therefore, a survival analysis was carried out. In order to assess differences between factors in the survival analyses, Log Rank and Breslow tests were applied. The statistical analyses were performed using IBM® SPSS® version 20 and MATLAB Release 2014, the MathWorks. The significance level adopted for statistical meaning was 5% ($\alpha = 0.05$).

Chapter 3.1.3 - Table 1. Clinic-pathologic characteristics of study population.

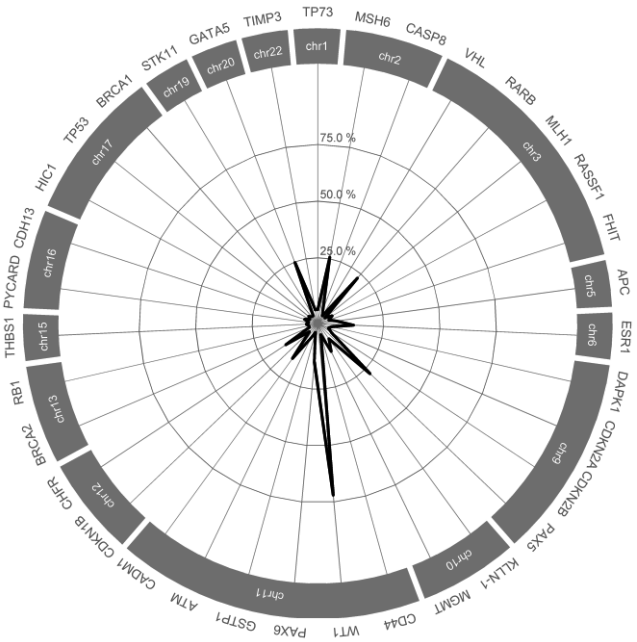
Patients (n = 93)			
	n (%)		n (%)
Gender			
Male	79 (85)	Nodal status	
Female	14 (15)	**N0**	41(44.1)
Age at diagnosis (Years)		**N1+N2**	31(33.3)
<60	48 (51.6)	**NA**	21 (22.6)
≥60	45 (48.4)	Metastasis	
Anatomic Subsite		**M0**	12 (12.9)
Tongue	40 (43)	**M1**	1 (1.1)
Floor of the mouth	26 (28)	**NA**	80 (86)
Retromolar Trigone	8 (8.6)	Margins	
Jugal Mucosa	7 (7.5)	**R0**	47 (50.5)
Alveolar ridge	6 (6.4)	**R1**	24 (25.8)
Palate	5 (5.4)	**NA**	22 (23.7)
Tonsil	1(1.1)	Differentiation	
Tobacco (nº cigarettes/day)		**Well**	59 (63.5)
≥20	38 (41)	**Moderate**	15 (16.1)
<20	7 (7.5)	**NA**	19 (20.4)
None	19 (20.4)	Treatment	
NA	29 (31.1)	**surgery alone**	33(35.5)
Alcohol		**Surgery + RT**	41(44)
Yes	20 (21.5)	**Surgery + RT + QT**	17 (18.3)
No	9 (9.7)	**Surgery + QT**	1 (1.1)
NA	64 (68.8)	**NA**	1 (1.1)
TNM stage		Vital status	
Tumor size		**Relapses/Metastasis in follow-up**	31 (33.3)
T1 + T2	56 (60.2)	**Dead - OSCC**	25 (26.8)
T3+T4	29 (31.2)	**Dead-non-OSCC**	1 (1.1)
NA	8 (8.6)	**Alive without disease**	36 (38.8)

NA- Not Available; **RT -** Radiotherapy; **QT -** Chemotherapy

Results and discussion

From the 35 genes analyzed, we found that the most frequently methylated genes were *WT1* (69/93), followed by *MSH6* and *PAX5* (29/93), *GATA5* (25/93), *RARB* (21/93), *CADM1* (14/93), *PAX6* (13/93), *CHFR* (12/93) and *ESR1* (11/93) (Figure 1). Based on our finding that the methylation levels for all genes in all 16 control samples

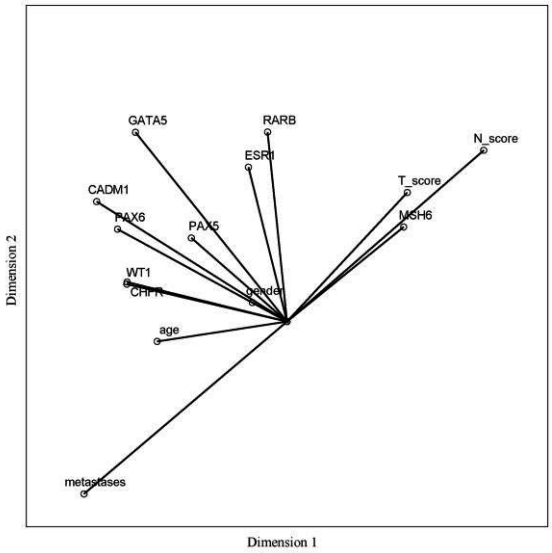
tested were < 20% (data not shown), this value was used as a biological cut-off for the presence or absence of methylation. For the nine genes mentioned above, the methylation levels varied from 20 to 100% in the different patient samples included in the cohort. In contrast, we found that the *VHL*, *MLH1*, *GSTP1*, *BRCA2*, *RB1*, *PYCARD*, *HIC1*, *BRCA1* and *STK11* genes did not exhibit methylation in any of the patient samples tested.



Chapter 3.1.3 - Figure 1. Radar chart with all methylated genes in our 93 OSCC cohort, highlighting the nine genes most frequently methylated: *WT1, MSH6, PAX5, GATA5, RARB, CADM1, PAX6, CHFR* and *ESR1*. The genes are ordered by chromosome, from the short arm to the long arm. The scale represents the percentage of subjects presenting methylated genes.

Subsequent multivariate analyses of the nine methylated genes and the various clinic-pathological parameters of the patients revealed that the variables tested contributed differently to two dimensions (Cronbach's Alpha Coefficient = 0.862) (Figure 2). The first dimension (Cronbach's Alpha Coefficient = 0.695) encompasses the pN and pT classification scores, the *MSH6* methylation status and, at the opposite end of the spectrum, the presence of metastases/relapses diagnosed during follow-up. The second dimension (Cronbach's Alpha Coefficient = 0.584) encompasses the majority of the nine genes, as well as the age and gender of the patients. Using a logistic regression model we observed a statistically significant association between the selected genes and the pT classification (G^2 (3) = 15.045; p = 0.002), in which *WT1* gene promoter methylation (p = 0.002; OR_{adj} = 6.949) and *GATA5* gene promoter methylation (p = 0.010, OR_{adj} = 0.194) showed distinct predictive effects, i.e., *WT1* gene promoter

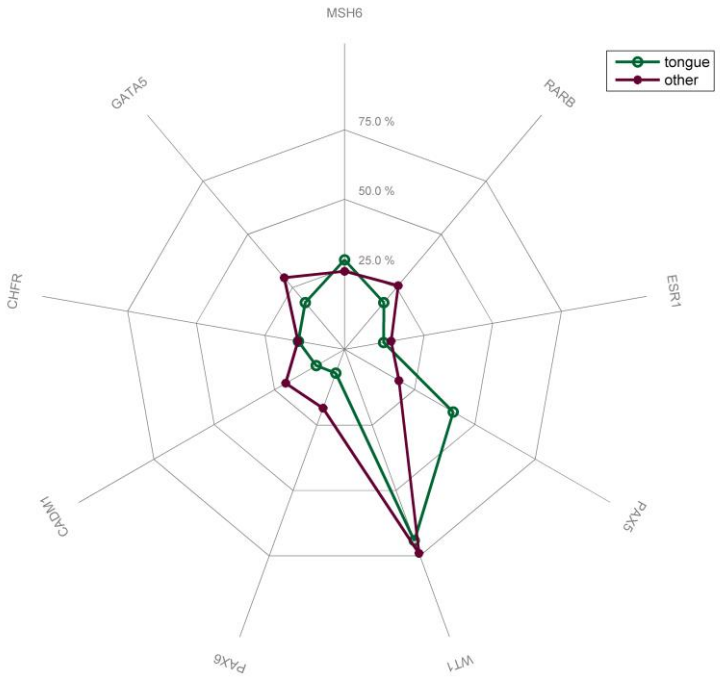
methylation was associated with T1 and T2 and *GATA5* gene promoter methylation with T3 and T4 tumor stages. With respect to pN status we found a statistically significant (G^2 (1) = 6.794; p = 0.009) predictive value for the *MSH6* gene (p = 0.011, OR_{adj} = 0.259), i.e., *MSH6* gene promoter methylated samples exhibited an approximately 4x higher probability to be N1 or N2 versus the unmethylated samples. Additionally, we found that *MSH6* gene promoter methylation was statistically associated (p = 0.008, OR_{adj} = 4.16) with the development of metastases and/or relapses during or after treatment.



Chapter 3.1.3 - Figure 2. PCA using the nine genes most frequently methylated in our 93 OSCC cohort and the following variables of the patients: gender, age, pN and pT classification of tumors and the development of metastases/relapses during follow-up. The first dimension is plotted on the X axis and captures 69.5% of the variance. This first dimension shows an association between the methylation status of *MSH6*, pN and pT tumor classification and the development of metastases/relapses during follow-up. The second dimension is plotted on the Y axis and captures 58.4% of the variance. This second dimension encompasses the majority of the genes, as well as the age and gender of the patients.

Our cohort comprises tumors derived from several sites within the oral cavity, of which those of the tongue are the most frequent. Considering the primary tumor locations, we decided to divide the cohort into two groups: one with tumors derived from the tongue (40 patients) and one with tumors derived from all remaining locations (53 patients). By doing so, we found that there was a statistically significant association (χ^2 (1) = 7.015; p = 0.008) between a location in the tongue and the methylation status of the *PAX5* gene (Figure 3), i.e., when this gene was methylated the probability of the

tumor being derived in the tongue was ~3 times higher than when this gene was unmethylated [OR = 3.36 (IC95% (1.34; 8.33)].



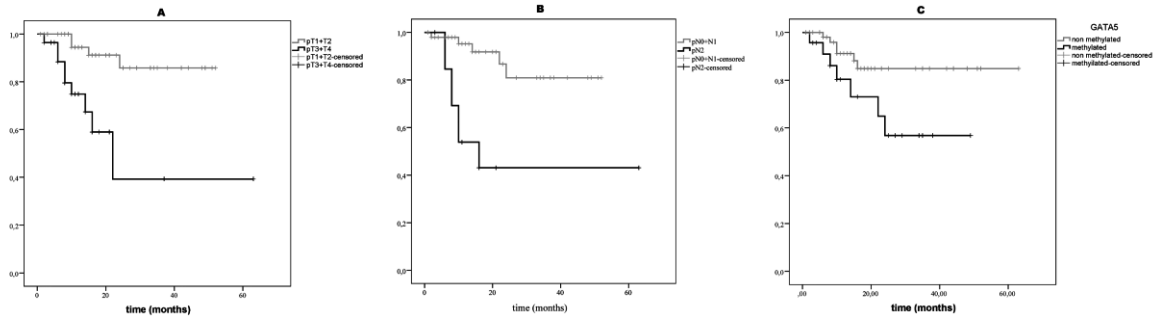
Chapter 3.1.3 - Figure 3. Radar chart with the nine genes most frequently methylated in our 93 OSCC cohort, separated in two groups according the anatomical sites of the primary tumors. The green line represents the tongue site and the brown line represents all the other anatomic sites. Methylation of the *PAX5* gene was associated with tongue tumors. The scale represents the percentage of subjects presenting methylated genes.

Of the 93 patients, 33 were only treated by surgery and 60 were treated by surgery plus radiotherapy or chemoradiotherapy. The follow-up periods of the patients ranged from 1 to 54 months. During follow-up 25/93 patients died. We found statistically significant survival differences with respect to tumor size (Log Rank: $\chi^2(1) = 11.187$: $p = 0.001$; Breslow: $\chi^2(1) = 10.143$; $p = 0.001$), i.e., patients with T1 or T2 tumors exhibited a better mean survival time, 46.96 months (CI95% [42.3; 51.6]) compared to those with T3 or T4 tumors, 33.21 months (CI95% [17.4; 49.0]) (Figure 4A).

We also observed significant differences between survival and nodal status (Log Rank: χ^2 (1) = 13.654: $p < 0.001$; Breslow: χ^2 (1) = 13.662; $p < 0.001$), i.e., 45.55 months for N0 or N1 (CI95% [40.3; 50.8]) and 40.46 months for N2 (CI95% [27.6; 53.3]) (Figure 4B).

3. Results
3.1. Genetic and Epigenetic Studies

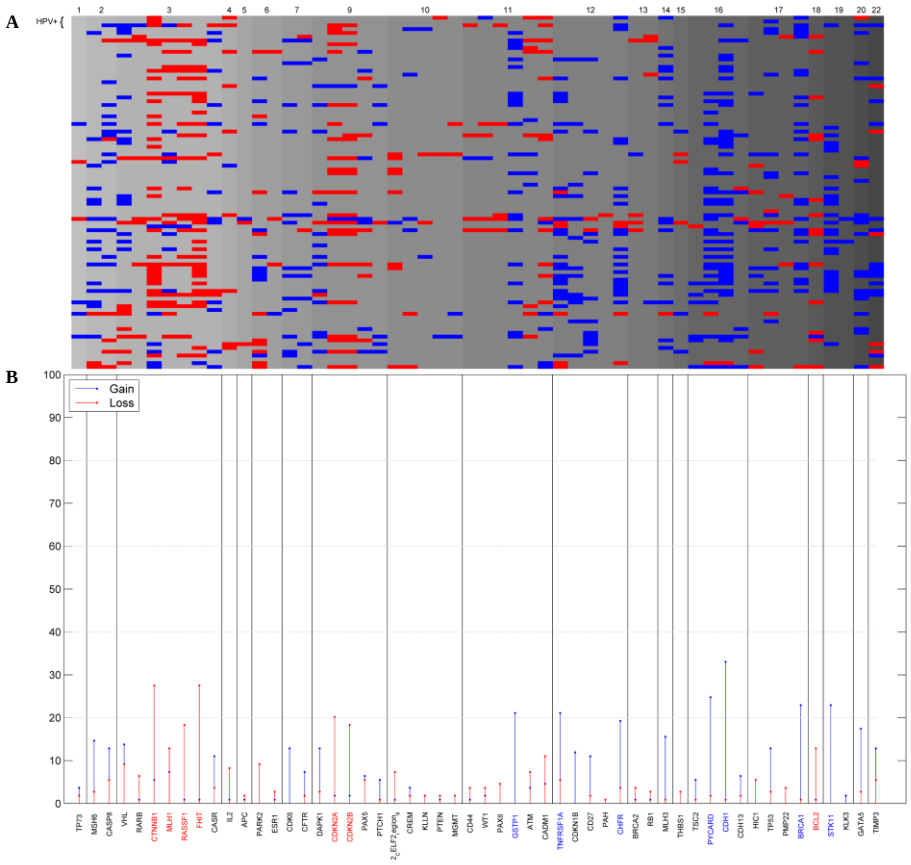
We observed only marginal statistically significant differences between survival (Log Rank: $\chi^2(1) = 3.886$; $p = 0.049$; Breslow: $\chi^2(1) = 3.384$; $p = 0.066$) and the *GATA5* methylation status, i.e., patients with a methylated *GATA5* gene promoter exhibited the worst survival rates [33.92 months (CI95% [25.1; 42.7])], whereas patients with an unmethylated *GATA5* gene promoter exhibited the best survival rates [55.21 months (CI95% [49.4; 61.0])] (Figure 4C).



Chapter 3.1.3 - Figure 4. Kaplan-Meier curves showing overall survival in the 93 OSCC patients according to: **A)** pT1+T2 and pT3+T4 tumor classification; small size tumors (pT1+T2) present a better survival; **B)** pN0 + pN1 and pN2 classification; tumors with pN2 presented worse survival; **C)** *GATA5* methylation status; tumors with unmethylated *GATA5* have a better survival. Censured data means subjects that left the study.

For the *MSH6* and *WT1* genes we did not find any statistically significant associations with survival (Log Rank: $\chi^2(1) = 0.052$: $p = 0.820$; Breslow: $\chi^2(1) = 0.051$; $p = 0.821$ and Log Rank: $\chi^2(1) = 2.194$: $p = 0.139$; Breslow: $\chi^2(1) = 1.950$; $p = 0.163$, respectively). Survival parameters by the presence or absence of compromised surgical margins did not show any statistically significant differences (Log Rank: $\chi^2(1) = 0.507$: $p = 0.477$; Breslow: $\chi^2(1) = 1.184$; $p = 0.277$).

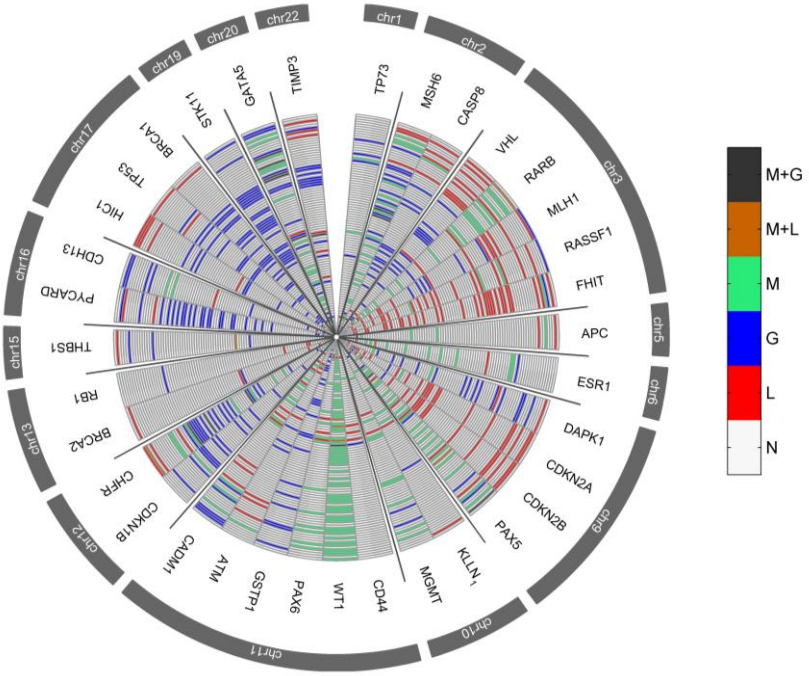
Genetic imbalances were observed for all 54 genes analyzed (Figure 5).



Chapter 3.1.3 - Figure 5. Genetic imbalances (gains and losses) in the 93 OSCC patient samples detected using two MS-MLPA probe mixes encompassing 54 genes. Losses of genetic material are represented in red, gains in blue. **A)** Each line represents one patient and each pixel in this line corresponds to one gene. Gray represents genes without alteration, and each shade of gray shows the localization of the genes on each specific chromosome. From left to right, the genes are ordered by chromosome, from the short arm to the long arm. **B)** Picture showing the percentage of imbalances by chromosome in tumor tissue for each gene analyzed. Each line represents one gene and the size of the line is proportional to the percentage of patients that present this alteration. The most frequent gains were detected in genes mapping to chromosomes 16, 17, 19, 11 and 12, being the genes highlighted in blue. The most frequent losses were detected in genes mapping to chromosomes 3, 9 and 18, being the genes highlighted in red.

We found that the same gene may exhibit either copy gain or copy number loss in different patient samples. The most frequent copy number gains were observed in chromosome 16 (*CDH1* and *PYCARD* genes), followed by the chromosomes 17 and 19 (*BRCA1* and *STK11* genes, respectively), chromosome 11 (*GSTP1* gene) and chromosome 12 (*TNFRSF1A* and *CHFR* genes). The most frequent copy number losses were observed in chromosome 3 (*FHIT*, *CTNNB1*, *RASSF1* and *MLH1* genes), followed by chromosome 9 (*CDKN2A* and *CDKN2B* genes) and chromosome 18 (*BCL2* gene). We also found genes that exclusively exhibited copy number gain: *CDK6* (7q21.2), *GSTP1* (11q13.2), *CDKN1B* (12p13.1), *STK11* (19p13.3) and *KLK3* (19q13.33). In

contrast, we found that the *PAX6* (11p13), *PAH* (12q23.2) and *PMP22* (17p12) exhibited only copy number losses. Genes that were less frequently altered in our cohort were *APC* (5q22.2), *ESR1* (6q25.1), *KLLN* (10q23.31), *PTEN* (10q23.31), *MGMT* (10q26.3), *PAH* (12q23.2), *RB1* (13q14.2), *PMP22* (17p12) and *KLK3* (19q13.33), with copy number gains and/or losses in at most four of the patient samples. We found that none of the genes evaluated exhibited both copy number loss and methylation (Figure 6).



Chapter 3.1.3 - Figure 6. Radar chart with all analyzed genes for methylation and the respective copy number alterations in our 93 OSCC cohort. Legend: M+G - methylation and copy number gains; M+L - methylation and copy number losses; M - methylation; G - gains; L - losses; N - Normal.

In contrast, we observed copy number gain and methylation in the *MSH6*, *CASP8*, *PAX5*, *WT1*, *CHFR*, *GATA5* and *TIMP3* genes. With respect to HPV infection, we identified three HPV-positive patients, two of whom were reported by us previously [Ribeiro et al. 2014a]. Two of the HPV infections were found to belong to the high-risk class, HPV type 31 and 33, whereas the third belonged to the low-risk class, HPV type 42.

Since gene methylation is considered to be an early event in the development of solid tumors, including OSCC [Maruya et al. 2004], this phenomenon may be relevant for its early diagnosis. In addition, it may serve as a therapeutic target based on its reversibility [S. Sharma et al. 2010a]. Currently, there are several human trials ongoing aimed at reversing

altered epigenetic patterns in both hematologic and solid tumors [Ho et al. 2013], including head and neck cancers [Worsham et al. 2014]. So far, however, epigenetic therapies have yielded relatively low efficacies in solid tumors [Federico and Bagella 2011], possibly due to their heterogeneity, their poor penetration and/or the development of tumor resistance [Ho et al. 2013]. We highlight *WT1* methylation as a possible early prognostic biomarker for OSCC, since we found a statistically significant association between methylation of this gene and early tumor stages (i.e., T1 and T2). In line with our results, Oji et al [Oji et al. 2003] suggested that *WT1* could serve as a prognostic factor in these tumors, as they found that a high *WT1*expression significantly correlated with a poor tumor differentiation and, consequently, an advanced tumor stage. Likewise, it has been found that *WT1* mRNA expression levels increase with the progression of chronic myeloid leukemia and myelodysplastic syndrome [Sugiyama 2010]. Although *WT1* may act as either a tumor suppressor or as an oncogene [Yang et al. 2007] its oncogenic role seems to be evident in OSCC. The *WT1* mRNA and protein expression levels remain to be evaluated in these tumors in order to establish its use in clinical diagnostic practice, as well as to establish its therapeutic applicability through e.g. WT1 peptide vaccination, which has already shown promising results in phase I and II clinical trials for at least some solid tumors [Qi et al. 2015].

MSH6 gene promoter methylation turned out to be a marker for a poor prognosis in our cohort, being statistically associated with the presence of metastases and relapses during/after treatment. Since *MSH6* is a DNA mismatch repair gene, its methylation (i.e., silencing) may enhance genomic instability, leading to a more aggressive phenotype and, consequently, a worse prognosis. Interestingly, Yip et al [Yip et al. 2009] reported that *in vitro* inactivation of the wildtype MSH6 protein in glioblastoma cells resulted in increased temozolomide (alkylating agent) resistance and that *in vitro* reconstitution of MSH6 protein expression restored the temozolomide sensitivity in MSH6-null glioblastoma cells. These results underscore the need to further explore the role of this DNA mismatch repair gene in oral carcinogenesis, including the effect of its methylation on the therapeutic response of OSCC patients.

We also observed a statistically significant association between *GATA5* gene promoter methylation and advanced OSCC, turning methylated *GATA5* into a biomarker for a worse prognosis. Likewise, Peters et al [Peters et al. 2014a; Peters et al. 2014b] reported that methylation and reduced *GATA5* mRNA expression levels were associated with a poor clinical outcome in renal cell carcinoma patients, highlighting its role as

molecular marker for aggressiveness and disease progression. Importantly, it has been found that treatment of colorectal cancer cells, in which the expression of *GATA4* and *GATA5* was silenced, with a DNA methyltransferase inhibitor could lead to reactivation of these genes. As a consequence, their proliferation and migration were found to be impaired, substantiating the importance of *GATA4* and *GATA5* expression loss in cancer development [Zheng and Blobel 2010]. Thus, *GATA4* and *GATA5* gene reactivation may serve as a therapeutic approach for OSCC. With respect to the *WT1* and *MSH6* genes, we did not find any statistically significant associations between promoter methylation and survival rates, probably due to the short follow-up periods for the last patients enrolled in the study.

Regarding the different primary tumor locations within the oral cavity, we found that when the *PAX5* gene promotor is methylated the tongue is the most likely site of origin. Previously, the *PAX5* gene has been found to be methylated and down-regulated in head and neck cancers [Guerrero-Preston et al. 2014]. Our current results indicate that the epigenetic heterogeneity of OSCCs may explain their different clinical outcomes. Since different epigenetic patterns in tumors from different anatomic locations may explain differences in therapeutic response, this should be taken into consideration for the selection of therapeutic modalities. The molecular and clinical heterogeneity of OSCCs has been reported before [Pai and Westra 2009; Worsham et al. 2012]. The identification of specific biomarkers for subgroups of these tumors opens up new possibilities for their clinical management. It is imperative to note here that patients often receive inappropriate treatment due to difficulties in the detection of metastases [Robbins et al. 2002]. *MSH6* gene promoter methylation may serve as such. In relation to copy number alterations, we [Ribeiro et al. 2014a] and others [Cancer Genome Atlas 2015] found that several chromosomes and genes are frequently altered in OSCCs. These alterations did, however, not coincide with the methylation patterns observed in our current cohort. The prevalence of HPV in OSCC has been reported to be between 5.9-21.3% [Chinn and Myers 2015]. In our cohort this prevalence was lower, which could be explained by the frequent presence of traditional risk factors and also the high number of relatively older patients included. It is important to reinforce that the role of HPV in OSCC remains ambiguous and controversial due to the relatively small number of positive patients in comparison to those in oropharyngeal cancer [Chinn and Myers 2015].

In conclusion, we identified gene promoter methylation signatures (*WT1*, *MSH6*, *GATA5* and *PAX5)* that are strongly correlated to, and can have a predictive value for

the clinical outcome of OSCC patients. The establishment of methylation signature scores offers a novel strategy to stratify patients into those with a better and those with a worse prognosis and, eventually, those with a risk of treatment failure. Extended follow-up and larger multicenter validation studies are required to confirm the potential prognostic power of these biomarkers before its application in routine clinical practice.

Acknowledgments

Ribeiro I.P. is a recipient of a PhD fellowship (SFRH/BD/52290/2013) from the Portuguese Foundation for Science and Technology. This work was in part supported by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra) and by the Portuguese Foundation for Science and Technology (grant: UID/NEU/04539/2013).

Conflict of interest

The authors have no conflict of interest to disclose.

Chapter 3.1.3 - Supplementary Table 1. Summary of gene function and chromosomal localization of the 54 genes in study

Symbol	Name	Function	Chromosomal localization
TP73	Tumor protein p73	Apoptosis related gene	1p36.32
MSH6	mutS homolog 6	DNA mismatch repair	2p16.3
CASP8	Caspase 8, apoptosis-related cysteine peptidase	Apoptosis related gene	2q33.1
VHL	Von Hippel-Lindau tumor suppressor, E3 ubiquitin protein ligase	Ubiquitination and proteasomal degradation	3p25.3
RARB	Retinoic acid receptor, beta	Transcription regulation, cell growth and differentiation	3p24.2
MLH1	mutL homolog 1	DNA mismatch repair	3p22.2
CTNNB1	Catenin (cadherin-associated protein), beta 1, 88kDa	Component of Wnt signaling pathway, cell growth and adhesion regulation	3p22.1
RASSF1	Ras association (RalGDS/AF-6) domain family member 1	Cell cycle control	3p21.31
FHIT	Fragile histidine triad	Apoptosis, cell proliferation, survival and nucleotide metabolism	3p14.2
CASR	Calcium-sensing receptor	Cellular calcium homeostasis	3q21.1
IL2	Interleukin 2	Immune response	4q27
APC	Adenomatous polyposis coli	Antagonist of the Wnt signaling pathway, cell migration and adhesion, transcriptional activation, and apoptosis	5q22.2
ESR1	Estrogen receptor 1	Transcription regulation, cellular proliferation and differentiation	6q25.1
PARK2	Parkin RBR E3 ubiquitin protein ligase	Ubiquitination	6q26
CDK6	Cyclin-dependent kinase 6	Differentiation and Cell cycle control	7q21.2
CFTR	Cystic fibrosis transmembrane conductance regulator (ATP-binding cassette sub-family C, member 7)	Transport of chloride ions	7q31.2
CDKN2A	Cyclin-dependent kinase inhibitor 2A	Cell cycle control, apoptosis regulation	9p21.3
CDKN2B	Cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)	Cell cycle control	9p21.3
PAX5	Paired box 5	Regulator in early development	9p13.2
DAPK1	Death-associated protein kinase 1	Cell survival, apoptosis, and autophagy.	9q21.33
PTCH1	Patched 1	Receptor for sonic hedgehog	9q22.33
CELF2	CUGBP, Elav-like family member 2	RNA processing	10p14
CREM	cAMP responsive element modulator	Component of cAMP-mediated signal transduction	10p11.21
KLLN	killin, p53-regulated DNA replication inhibitor	Cell cycle control	10q23.31
PTEN	Phosphatase and tensin homolog	Cell cycle regulation	10q23.31
MGMT	O-6-methylguanine-DNA methyltransferase	DNA repair	10q26.3
CD44	CD44 molecule (Indian blood group)	Cell-cell and cell-matrix interactions	11p13
WT1	Wilms tumor 1	Trasncription factor binding	11p13
PAX6	Paired box 6	Transcription regulation	11p13
GSTP1	Glutathione S-transferase pi 1	Apoptosis regulation	11q13.2
ATM	ATM serine/threonine kinase	Cell cycle control	11q22.3
CADM1	Cell adhesion molecule 1	Cell adhesion	11q23.3
CD27	CD27 molecule	Putative role in apoptosis	12p13.31
TNFRSF1A	Tumor necrosis factor receptor superfamily, member 1A	Apoptosis regulation	12p13.31
CDKN1B	Cyclin-dependent kinase inhibitor 1B (p27, Kip1)	Cell cycle control	12p13.1
PAH	Phenylalanine hydroxylase	Phenylalanine catabolism	12q23.2
CHFR	Checkpoint with forkhead and ring finger domains, E3 ubiquitin protein ligase	Cell cycle control	12q24.33
BRCA2	Breast cancer 2, early onset	DNA repair, Cell cycle control	13q13.1
RB1	Retinoblastoma 1	Cell cycle control	13q14.2
MLH3	mutL homolog 3	DNA mismatch repair	14q24.3
THBS1	Thrombospondin 1	Cell-cell and cell-matrix interactions	15q14
TSC2	Tuberous sclerosis 2	Cell cycle control	16p13.3
PYCARD	PYD and CARD domain containing	Apoptosis regulation	16p11.2
CDH1	Cadherin 1, type 1	Cell adhesion	16q22.1
CDH13	Cadherin 13	Cell adhesion	16q23.3
HIC1	Hypermethylated in cancer 1	Transcription regulation	17p13.3
TP53	Tumor protein p53	Cell cycle control and apoptosis	17p13.1
PMP22	Peripheral myelin protein 22	Growth regulation	17p12
BRCA1	Breast cancer 1, early onset	DNA repair, Cell cycle control	17q21.31
BCL2	B-cell CLL/lymphoma 2	Apoptosis regulation	18q21.33
STK11	Serine/threonine kinase 11	Cell metabolism, cell polarity, apoptosis and DNA damage response	19p13.3
KLK3	kallikrein-related peptidase 3	Angiogenesis regulation	19q13.33
GATA5	GATA binding protein 5	Transcription factor binding	20q13.33
TIMP3	TIMP metallopeptidase inhibitor 3	Matrix metalloproteinases inhibition	22q12.3

3.1.4 Genomic and epigenetic signatures associated with survival rate in oral squamous cell carcinoma patients

Ilda P. Ribeiro, Francisco Caramelo, Luísa Esteves, Camila Oliveira, Francisco Marques, Leonor Barroso, Joana B. Melo, Isabel M. Carreira

Submitted

Abstract

Objectives Although oral squamous cell carcinoma (OSCC) presents great mortality and morbidity worldwide, the mechanisms behind its clinical behavior remain unclear. Biomarkers are needed to forecast patients' survival and, among those patients undergoing curative therapy, which are more likely to develop tumor recurrence/metastasis. Demonstrating clinical relevance of these biomarkers could be crucial both for surveillance and in helping to establish adjuvant therapy strategies. We aimed to identify genomic and epigenetic biomarkers of OSCC prognosis as well as to explore a noninvasive strategy to perform its detection. ***Materials and Methods*** OSCC tumor and non-tumor tissue samples and cells scrapped from the tumor surface were genomic and epigenetically evaluated by Methylation-Specific Multiplex Ligation-dependent Probe Amplification technique. ***Results*** Copy number alterations in *ATM, CASR, TP73, CADM1, RARB, CDH13, PAX5, RB1* genes and *GATA5, PAX6, CADM1* and *CHFR* promoter methylation were shown to be associated with worse OSCC patients survival. Copy number alterations in *BRCA1, CDKN2A, CHFR, GATA5, PYCARD, STK11, TP53, VHL* genes and *GATA5, CADM1, KLLN, MSH6, PAX5, WT1* promoter methylation were shown to be associated with development of metastasis/relapses during or after OSCC patients treatment. We also found a good agreement in the status of *CDKN2A* promoter methylation evaluated noninvasively or in the tumor tissue. ***Conclusion*** Genomic and epigenetic signatures were validated in a larger and geographically separate cohort, from TCGA database, which reinforce their clinical applicability. Noninvasive methodologies for detection of these signatures require further studies before translation to clinical practice.

Keywords: Copy number alterations; Methylation; Biomarkers, OSCC survival, Recurrence, TCGA data

Introduction

Oral cancer is the most common neoplasm of head and neck region, being squamous cell carcinoma (OSCC) the most frequent histological type. Although the oral cavity region is accessible to visual examination, oral cancer is still diagnosed at advanced stage when signs and symptoms are already present [van der Waal et al. 2011]. On contrary, in early stages oral malignant lesions are usually benign in appearance and, as they are often asymptomatic, patients tend to depreciate the small signs [Cowpe et al. 1988]. When oral cancer is detected at early stages, 5-year survival rate is around 90%, whereas in later stages the survival rate decreases to about 30% [Omar 2015]. As so, clinical staging at the time of diagnosis is of utmost importance and can be used as a predictor of recurrence and mortality in oral cancer patients. On the other hand, the treatment of advanced tumors is frequently very mutilating, leading to visible deformations which, in turn, results in social stigmatization, speech disabilities and nutrition problems [Campbell et al. 2004; Holloway et al. 2005; Meyer et al. 2004]. Additionally, the behavior of OSCC is difficult to predict using solely conventional clinical and histopathological parameters [Oliveira and Ribeiro-Silva 2011]. Accurate prognostic biomarkers should have immediate applicability on the clinical set, allowing to select patients for more effective tailored treatment strategies and conceivably for shorter monitoring programs. Tissue biopsy with histological assessment remains the gold standard in oral diagnosis, which needs a trained health-care provider, and is considered invasive, painful, expensive and time consuming [Nair et al. 2012]. Nowadays, several advancements have been made in the development of potentially useful diagnostic tools at the clinical and molecular level for early detection of oral cancer and its relapses, such as exfoliative cytology (consisting in cells collection from mucosal surfaces by scrapping or brushing) that presents as advantages the fact that it is painless, non-invasive, easy to perform and can be repeated several times to diagnose and follow-up the patients [Bremmer et al. 2005; Verma et al. 2015]. Even so, non-invasive tools have not yet proved their value in clinical routine since they were unable to reduce the problem of late diagnosis [Messadi 2013].

In this study we identified a specific genomic and epigenetic profile associated with OSCC survival and risk of relapse/metastasis development and, consequently, with patients prognosis. This genomic signature was validated using TCGA data, and may contribute for accurately stratifying patients for personalized clinical management. Additionally, we developed a noninvasive approach of identifying this genomic and

epigenetic signature by evaluating in tumor samples and in cells scrapped from the tumor surface of the same OSCC patients through Methylation-Specific Multiplex Ligation-dependent Probe Amplification (MS-MLPA) technique. This noninvasive approach showed good agreement for *CDKN2A* promoter methylation.

Materials and methods

Study population

The study protocol was approved by the Committee on Ethics in Research of the Faculty of Medicine of the University of Coimbra. All patients provided their written consent to participate in the study after being informed about the research purposes, following the regulations in the Declaration of Helsinki.

The study cohort includes tumor tissue, tissue from surgery resection margin (macroscopically tumor-free tissue) and exfoliated cells, scrapped from the tumor surface, of 49 OSCC patients who underwent tumor resection. The surface of tumor, before surgery, was gentle scrapped with a blunt instrument to obtain exfoliated neoplastic cells. The material obtained was then taken in a falcon tube with phosphate buffered saline buffer and stored at 4ºC until DNA extraction, which was performed within one week after collection. The tissue samples were snap-frozen in liquid nitrogen within 30 min after resection and stored at -80ºC until use. The patients were recruited between October 2010 and January 2016 from the Maxillofacial Surgery Unit of the Coimbra Hospital and University Centre, CHUC, EPE, Portugal. Diagnosis and staging were performed in accordance with the American Joint Committee on Cancer TNM staging system. The participants in this study answered a survey regarding lifestyle and risk factors for upper aerodigestive tract malignancies, including alcohol and tobacco consumption. Patients were followed-up through hospital revisits during routine clinical appointments. The follow-up periods ranged from 1 to 63 months. Details of our study cohort are listed in table 1. For the control group of tissue samples, gingival tissues from 16 healthy donors (6 males and 10 females, with ages ranging from 18 to 81 years) subjected to wisdom teeth removal were used. For the control group of tumor exfoliated cell samples, cells acquired by scrapping the oral surface of 14 healthy donors (4 males and 10 females, with ages ranging from 23 to 60 years) were used.

Chapter 3.1.4 - Table 1. Clinical-pathologic characteristics of study population.

Patients (n = 49)			
	n (%)		n (%)
Gender		Age at diagnosis (Years)	
Male	39 (80)	**<60**	24 (49)
Female	10 (20)	**≥60**	25 (51)
Anatomic Subsite		Invasion peri(neural)	
Tongue	26 (53)	**Yes**	26 (53)
Floor of the mouth	12 (25)	**No**	20 (41)
Retromolar Trigone	4 (8)	NA	3 (6)
Jugal Mucosa	2 (4)	Differentiation	
Palate	2 (4)	**Well**	38 (78)
Alveolar ridge	3 (6)	**Moderate**	9 (18)
Tobacco		**Poor**	1 (2)
Yes	31 (63)	**NA**	1 (2)
No	15 (31)	Margins	
NA	3 (6)	**R0**	29 (59)
Alcohol		**R1**	16 (33)
Yes	29 (59)	**NA**	4 (8)
No	15 (31)	HPV	
NA	5 (10)	**Positive**	1 (2)
TNM stage		**Negative**	48 (98)
I	9 (18)	Vital status	
II	14 (29)	**Relapses/Metastasis in follow-up**	16 (33)
III	7 (14)	**Dead - OSCC**	14 (29)
IV	19 (39)	**Dead-non-OSCC**	1 (2)
Treatment			
surgery alone	13 (27)		
Surgery + RT	26 (53)		
Surgery + RT + QT	6 (12)		
NA	4 (8)		

NA- Not Available; **RT -** Radiotherapy; **QT -** Chemotherapy

DNA extraction, HPV typing and MS-MLPA

DNA from fresh frozen tissues of patients and controls were extracted using a High Pure PCR Template Preparation Kit (Roche GmbH, Mannheim, Germany), according to the manufacturer's instructions. The DNAs were quantified by UV spectrophotometric analysis using a Nanodrop 1000 Spectrophotometer (Thermo Scientific, USA). All tumor tissue samples were analyzed for HPV infection as previous described [Ribeiro et al. 2014b; Ribeiro et al. 2014a].

MS-MLPA analyses were performed using MS-MLPA probe set ME002 (MRC-Holland, Amsterdam, The Netherlands), which can simultaneously detect copy number alterations (CNAs) in 38 tumor suppressor genes and aberrant methylation patterns in a subset of 25 of these genes. All MS-MLPA reactions were performed according our previous work [Ribeiro et al. 2016]. Three controls selected from the previously analyzed control group of 16 from tissue samples and 11 from exfoliated cells, without CNAs and

methylation values below 20%, as well as a negative control (without DNA), were always included in each MS-MLPA assay. Binning of the raw data and comparative analyses were performed using Coffalyser.NET software. For each probe we determined the specific cutoff values for gain and loss, using 95% confidence intervals as determined on non-cancer subjects. A copy number gain was scored when a value exceeded 1.2 and a copy number loss was scored when a value was lower than 0.8. We considered a gene promoter as methylated when the methylation dosage ratio was ≥ 0.20, which means that at least 20% of the DNA was methylated. These cut-off values were based in our previous works [Pinto-Leite et al. 2014; Ribeiro et al. 2014b; Ribeiro I.P 2016].

Validation cohort from TCGA data portal

Copy number data obtained by SNP array and patients' clinical data of 314 OSCC were downloaded from the TCGA Data Portal, available at https://tcga-data.nci.nih.gov/tcga/, (accessed on the 23rd October, 2015). The available copy number data was Level 3 data. The clinical-pathologic features of the validation cohort are listed in table 2.

Statistical analysis

Data analysis was essentially divided in three parts: one focusing the relationship between genetic/epigenetic data and survival information, other attempting to obtain a genetic and epigenetic biomarker for relapses or metastases and, finally, a part regarding agreement between tissue samples and scrapped cells of the tumor surface.

Copy number and methylation data were obtained for tumor and non-tumor cells, comparing the same genes between the two tissues, thus aiming to enlighten which genes are really contributing to tumoral features. This comparison was performed resorting to the difference between the entropy [Hausser and Strimmer 2009] computed for each gene, and genes that showed larger differences were considered to be significant to explain dissimilarities between tumor and non-tumor cells. Afterwards a factorial analysis, with two factors, was performed aiming at achieving a smaller group of genes and the relationship of them to each factor and to each observation. Additionally, a cluster analysis (two step method) was carried out and two different groups of observations were found. Finally, survival data was compared between these two groups using Kaplan-Meyer analysis and the Log Rank test.

In order to achieve a form of estimating the risk of relapses or metastases a classification algorithm was implemented. Variable importance plots based on Gini index [Menze et al. 2009] were employed to select the most informative genes and a Random Forest method was implemented with a k-fold cross validation technique.

Agreement between genetic data coming from tissues samples and scrapped cells was computed using Kappa statistics for each gene.

Statistical analysis was performed using R statistics platform (R version 3.3.2) and Matlab (R2016b, Matworks™). The statistical level of significance adopted was 0.05.

Chapter 3.1.4 - Table 2. Clinical-pathologic characteristics of validation cohort from TCGA.

Patients (n = 314)			
	n (%)		n (%)
Gender		Age at diagnosis (Years)	
Male	209 (67)	**<60**	133 (42.4)
Female	105 (33)	**≥60**	180 (57.3)
Anatomic Subsite		Invasion peri(neural)	
Tongue	131 (42)	**Yes**	134 (42.7)
Oral cavity	73 (23)	**No**	113 (36)
Floor of the mouth	63 (20)	NA	63 (20.1)
Buccal Mucosa	22 (7)	Margins	
Alveolar Ridge	18 (6)	**R0**	229 (72.9)
Hard Palate	7 (2)	**R1**	37 (11.8)
Tobacco		**Close**	34 (10.8)
Yes	215 (68.5)	**NA**	14 (4.5)
No	90 (28.6)	HPV	
NA	9 (2.9)	**Positive**	32 (10.2)
Alcohol		**Negative**	281 (89.5)
Yes	203 (64,6)	**NA**	1 (0.3)
No	104 (33,1)	Country	
NA	7 (2,3)	**United States**	212 (67)
TNM stage		**Canada**	37 (11.8)
I	12 (3.8)	**Brazil**	10 (3.2)
II	76 (24.2)	**Poland**	9 (2.9)
III	63 (20)	**Other**	9 (2.9)
IV	155 (49.4)	**NA**	37 (11.8)
NA	8 (2.5)	Vital status	
Treatment		**Relapses/Metastasis in follow-up**	189 (60.2)
QT	82 (26.1)		
Immunotherapy	2 (0.6)		
Targeted Molecular Therapy	1 (0.3)		
NA	229 (72.9)		

NA- Not Available; **RT -** Radiotherapy; **QT -** Chemotherapy

Results

Genomic signature associated with OSCC survival

We evaluated whether the genetic differences between the tumor and non-tumor tissue samples could be used as survival predictors using entropy differences between the results of CNAs for the genes analyzed in these two kinds of samples. Firstly, we reduced data by selecting only the genes presenting a difference higher than 50%. The selected genes were: *ATM, BRCA2, CADM1*, *CASR*, *CD44*, *CDH13, CDK6, CREM, ESR1, KLLN, MGMT, MLH3, PAX5, PAX6, PTEN, RARB, RB1, THBS, TP73* and *WT1* (Figure 1A).

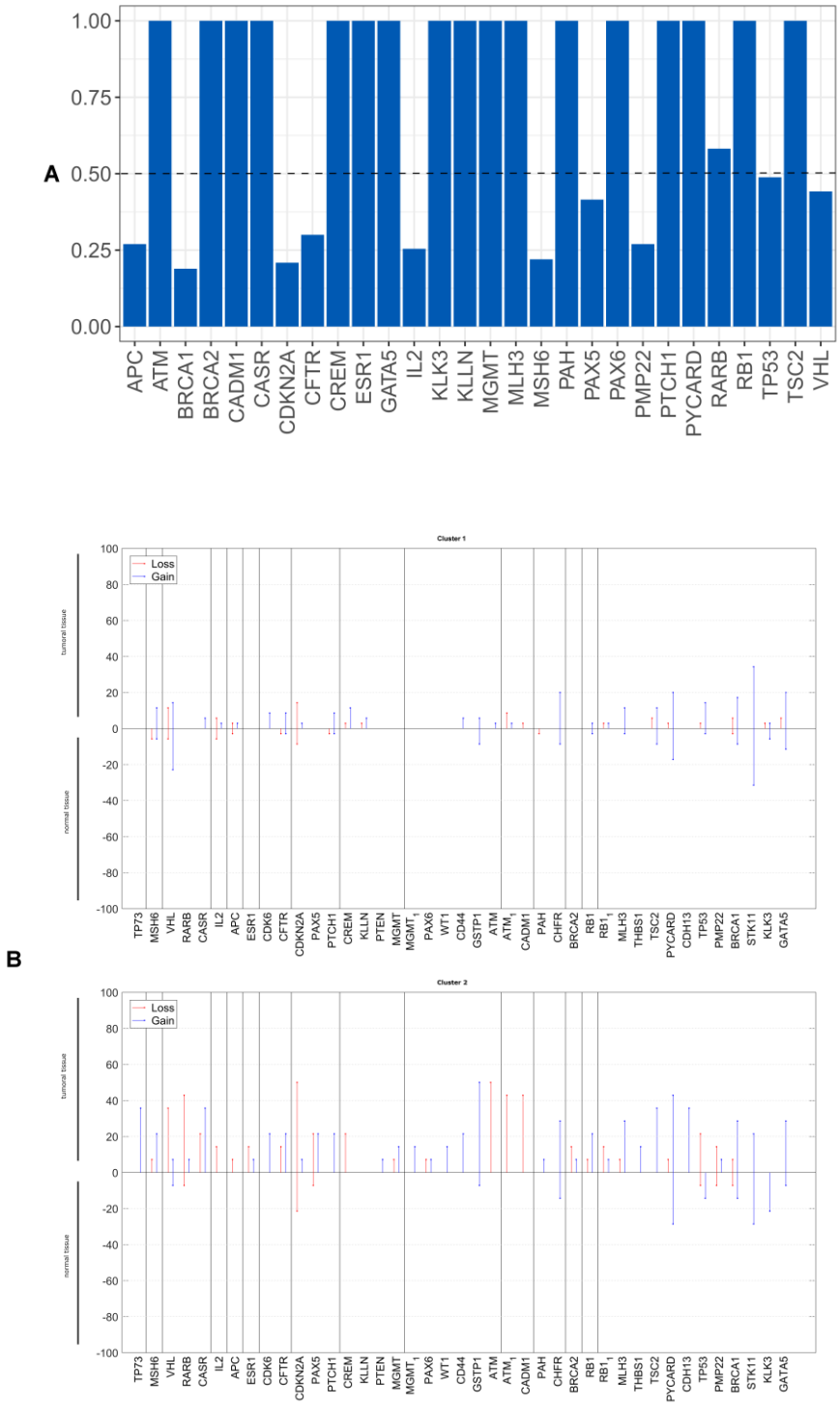
Applying factorial and clusters analysis, we reached a division of our cohort into two clusters with a Cronbach alpha of 0.781 (internal consistency) and a silhouette Coefficient of 0.7 (cluster consistency). The genes that most contribute for this division were *ATM*, *CASR*, *TP73*, *CADM1*, *RARB*, *CDH13*, *PAX5* and *RB1*.

Cluster 1 presented a proportion of 71% (35 patients) and cluster 2 of 29% (14 patients).

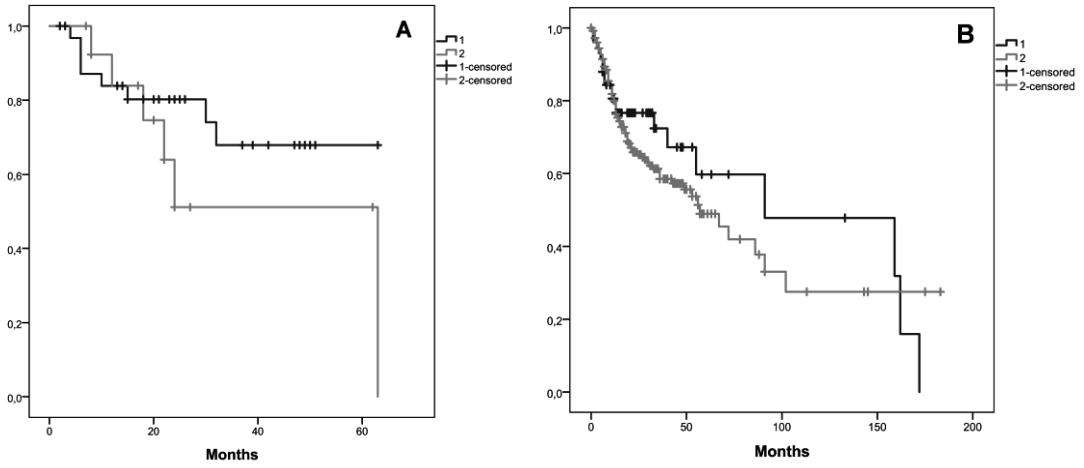
The genomic signature of these two clusters is different, cluster 1 did not presented CNA in the selected genes and cluster 2 presented gains and losses in these genes (Figure 1B).

Kaplan-Meier curves for the two clusters, considering cluster 1 with all patients without CNAs for the selected genes and cluster 2 with patients that have at least one gene with CNA were performed. We observed that the average survival for cluster 1 was 48.2 months (CI95% [39.5; 56.9]) whereas for cluster 2 was 40.9 months (CI95% [25.2; 56.6]) (Figure 2A). We did not reach statistical significance (Log-Rank: $p = 0.243$).

A validation of these results using TCGA data was performed, including 58 patients in cluster 1 and 254 in cluster 2. In this validation cohort, the average survival for cluster 1 was 98.6 months and for cluster 2 was 79.9 months (Figure 2B). In this validation cohort it was possible to observe 18.7 months of survival difference between the two clusters, which is higher than those observed in our cohort (7.3 months). Despite, this clinical relevant difference, we did not find statistical significance (Log-Rank: $p = 0.377$).



Chapter 3.1.4 - Figure 1 A) Differences of entropy observed among CNA results for the genes analyzed in the OSCC tumor and non-tumor tissue samples. **B)** Copy number gains and losses detected in tumor and non-tumor tissue samples of cluster 1 and cluster 2. Loss is represented by red and gain by blue.



Chapter 3.1.4 - Figure 2. Kaplan-Meier curves for the two clusters identified, **A)** in our cohort, **B)** in the validation cohort from the TCGA database. Cluster 1 is represented by 1 and cluster 2 by 2.

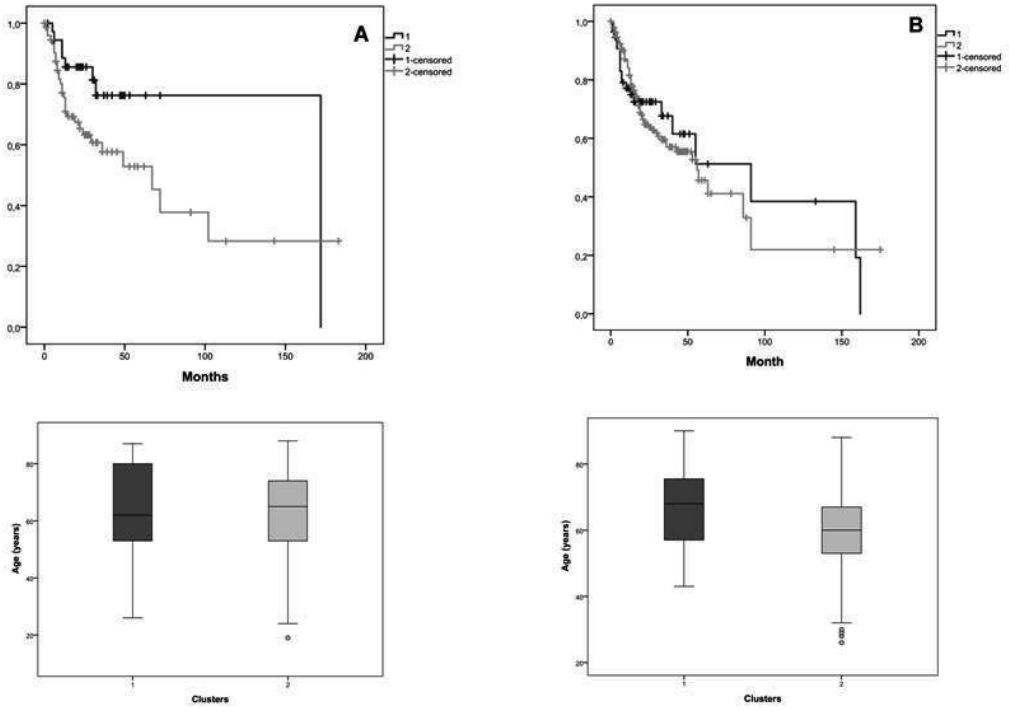
Considering the observed clinical relevant difference in survival of both clusters and the absence of statistical significance, we decided to joint both databases (our database from Portuguese patients and database from TCGA) in order to try maximize the observed differences and enlarge the number of patients. In this joint database the number of patients in cluster 1 was 93 and in cluster 2 was 268. The average survival for cluster 1 was 100 months and for cluster 2 was 77.8 months (data not shown). This joint database showed a difference of 22.2 months in the survival of both clusters. Despite this meaningful survival difference, statistical significance was not reached (Log-Rank: p = 0.167).

We also researched the effect of tumor stage in the survival, so we considered in this joint database, tumor stage I + II and tumor stage III + IV. Considering tumor stage I + II, Kaplan-Meier curves for the two clusters, cluster 1 with 37 patients and cluster 2 with 74 patients, showed that the average survival for cluster 1 was 135.4 months and for cluster 2 was 80.0 months (Figure 3A). Patients with tumor stage I+II with and without CNAs (cluster 1 and 2) in the selected genes, exhibited a marginal statistical significance difference (Log-Rank: p = 0.049) of 55.4 months in the survival of both clusters.

Considering tumor stage III + IV, Kaplan-Meier curves for the two clusters, cluster 1 with 55 patients and cluster 2 with 187 patients, showed that the average survival for cluster 1 was 84.9 months and for cluster 2 was 71.0 months (Figure 3B).

We verified a 13.9 months of difference in survival between these clusters; however, we did not reach statistical significance (Log-Rank: p = 0.809).

The distribution of patients' age is very similar in both clusters, which do not constitute a bias in the observed statistical significance survival differences (Figure 3).



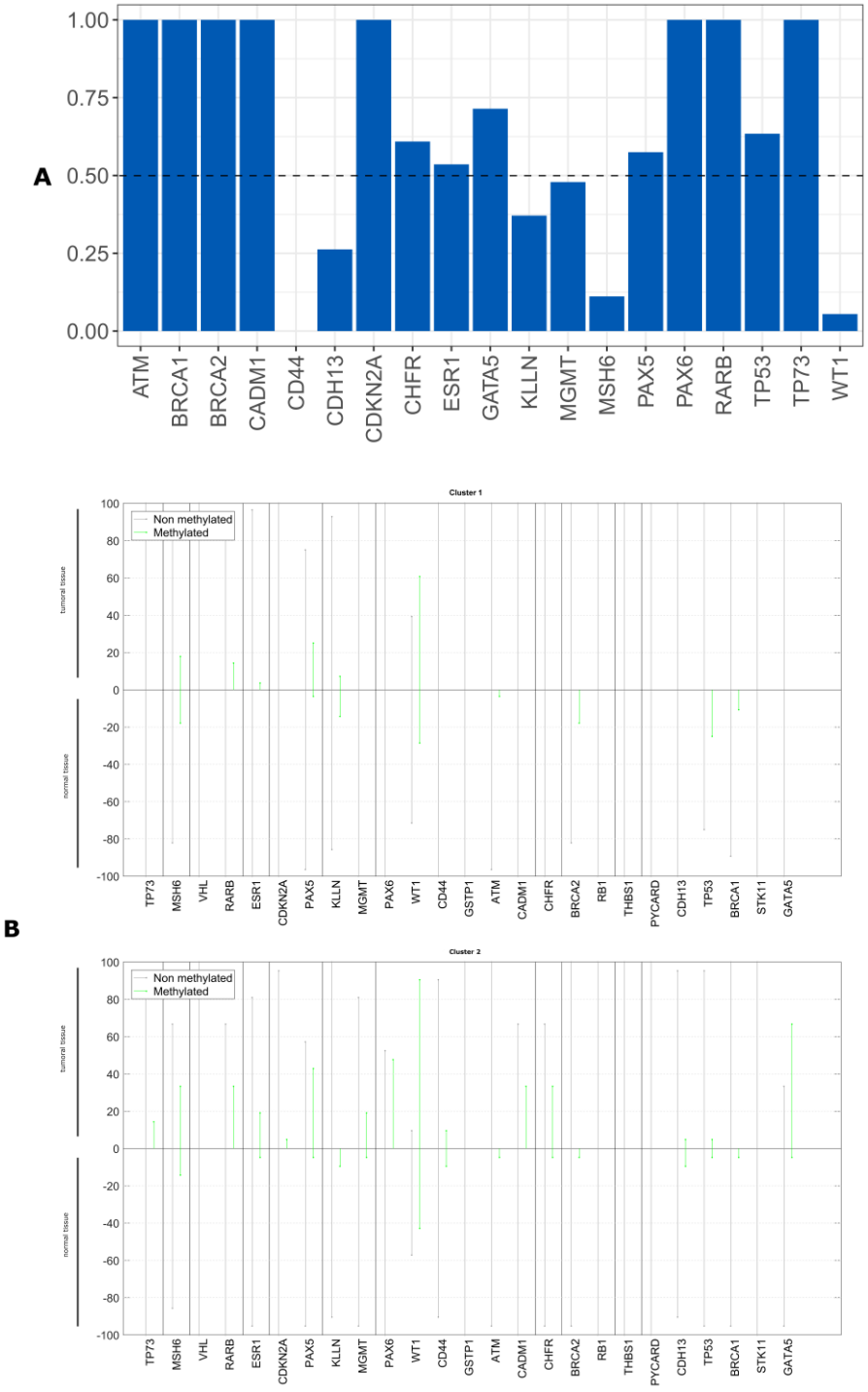
Chapter 3.1.4 - Figure 3. Kaplan-Meier curves for the two clusters identified in the joint database (our and validation cohorts) and the distribution of patients' age in both clusters, **A)** for tumor stage I + II, **B)** for tumor stage III + IV. Cluster 1 is represented by 1 and cluster 2 by 2.

Epigenetic signature associated with OSCC survival

We evaluated whether the epigenetic differences between the tumor and non-tumor tissue samples could be used as survival predictors using entropy differences between the results of gene promoter methylation in these two kinds of samples. We reduced the data, selecting only the genes with a difference higher than 50%. The selected genes were: *ATM, BRCA1, BRCA2, CADM1, CDKN2A, CHFR, ESR1, GATA5, MGMT, PAX5, PAX6, TP53* and *TP73* (Figure 4A).

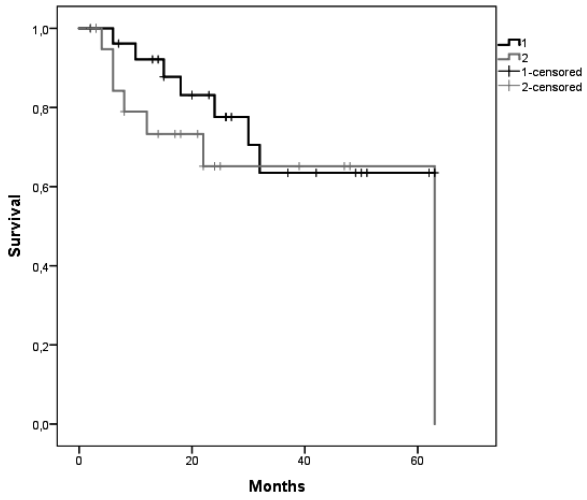
Applying factorial and clusters analysis with these selected genes, we reached a division of our cohort into two clusters with a Cronbach alpha of 0.718 and a silhouette Coefficient of 0.7. The cluster 1 presented a proportion of 57% with 28 patients and the cluster 2 of 43% with 21 patients.

The epigenetic profile of these two clusters is different. We observed that cluster 1 did not present any gene promoter methylation and cluster 2 presented methylation in these genes (Figure 4B).



Chapter 3.1.4 - Figure 4. A) Differences of entropy observed among methylation status for the analyzed genes in the OSCC tumor and non-tumor tissue samples, **B)** Methylation profile detected in tumor and non-tumor tissue samples of our cohort for the cluster 1 and the cluster 2.

Kaplan-Meier curves for the two clusters were performed, being 47.8 months the average survival of cluster 1 (CI95% [38.5; 57.1]) and 44.8 months of cluster 2 (CI95% [31.8; 57.7]) (Figure 5). We did not reach statistical significance (Log-Rank: p = 0.330).



Chapter 3.1.4 - Figure 5. Kaplan-Meier curves for the two clusters identified using gene promoter methylation results of our cohort. Cluster 1 is represented by 1 and cluster 2 by 2.

Validation using TCGA data was not done because methylation data is not comparable with our own data.

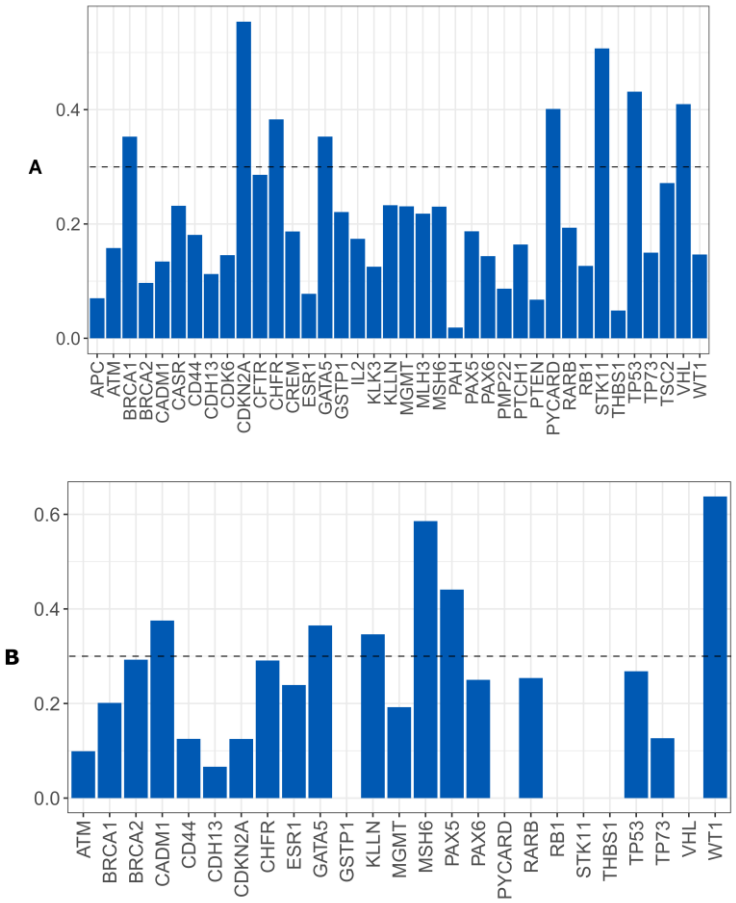
Genomic and epigenetic signatures associated with OSCC relapse/metastasis risk

Random Forest model using CNA and gene promoter methylation results to search biomarkers with capability of predicting the relapse/metastases development risk in our OSCC cohort was performed. For CNA, we verified that a set of eight genes has the power to discriminate between patients that developed relapse/metastases during clinical follow up from those without recurrence. The genomic signature with potential to perform this discrimination comprises the following genes: *BRCA1, CDKN2A, CHFR, GATA5, PYCARD, STK11, TP53* and *VHL* (Figure 6A). We did not reach statistical significance.

A validation of genomic results using TCGA data was performed, but we did not find statistical significance (data not shown).

For gene promoter methylation, we verified that a set of six genes has the power to discriminate between patients that developed relapse/metastases during clinical follow up from those without recurrence.

The epigenetic signature with potential to perform this discrimination comprises the following genes: *CADM1*, *GATA5*, *KLLN*, *MSH6*, *PAX5* and *WT1* (Figure 6B). We did not reach a statistical significance.



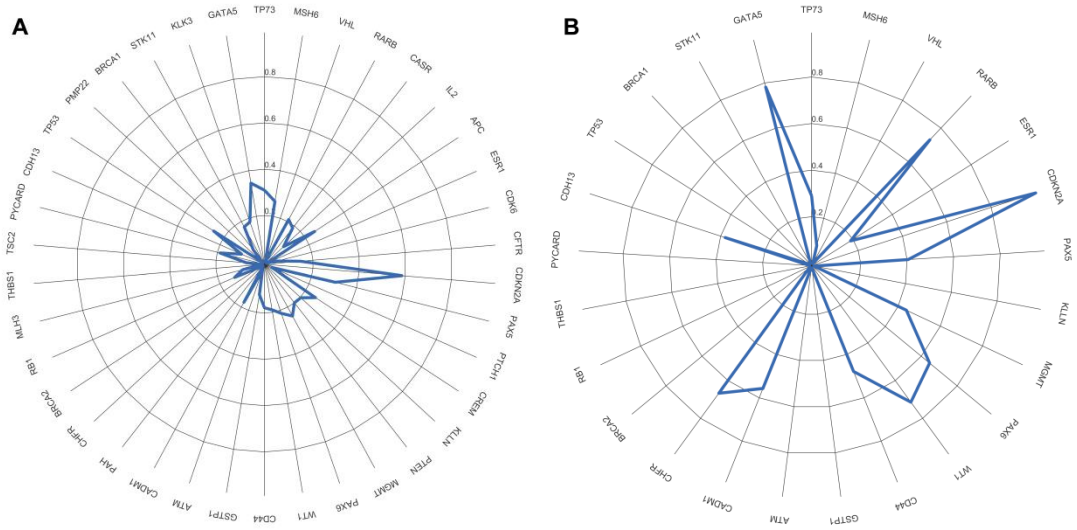
Chapter 3.1.4 - Figure 6. Importance level of the genes analyzed to discriminate between the patients that develop or not metastases/relapses, **A)** using CNA data, **B)** using gene promoter methylation genes.

Agreement between results obtained from tumor tissue samples and scrapped cells of the tumor surface

Evaluation of *CDKN2A* gene promoter methylation status in both tumor tissue and scrapped cells exhibited a total agreement, *Kappa* = 1 (Figure 7A). From all analyzed genes, in our cohort only the promoter methylation status of *CDKN2A* gene can be assessed using a non-invasive approach with the same result as in the tumor tissue.

For the overall copy number alterations (0.1568929 CI95% [0.0764; 0.2368]) and methylation status (0.005028447 CI95% [-0.0872; 0.1150]) of the analyzed genes

in the tumor tissue and in the scrapped cells (Figure 7 A, B) we observed a reduced agreement.



Chapter 3.1.4 - Figure 7. Agreement measured by Kappa value between tumor tissue and scrapped cells. **A)** using CNAs, **B)** using methylation status.

Discussion

Using a clinically well characterized OSCC cohort, followed during 1 to 63 months, we identified a specific genomic and epigenetic signature associated with survival and risk of metastasis/relapse development during/after treatment. Nowadays we are witnessing an explosion in the knowledge of biological markers related to pathogenesis and progression of OSCC, which might emerge as a possibility to complement the selection of patients to more aggressive treatment modalities; however, single markers have proved insufficient predictive power, so, no biomarker was yet translated to clinical practice [G. R. Thomas et al. 2005]. Copy number alterations in *ATM, CASR, TP73, CADM1, RARB, CDH13, PAX5, RB1* genes and *GATA5, PAX6, CADM1* and *CHFR* promoter methylation were shown to be associated with worse patient survival in our OSCC patients. Copy number alterations in *BRCA1, CDKN2A, CHFR, GATA5, PYCARD, STK11, TP53, VHL* genes and *GATA5, CADM1, KLLN, MSH6, PAX5* and *WT1* promoter methylation were shown to be associated with development of metastasis/relapses during or after treatment of our OSCC patients. These genomic signatures were validated in a geographically separated cohort (from TCGA database), reinforcing their potential for clinical application, since they were associated with

differences in patient survival around 20 months and more than 50 months if we considered only tumor stage I + II. The presence of these set of genes with copy number alterations could guide the clinical management of OSCC patients, indicating which patients have a better prognosis and also which patients must be monitored in closer follow up or should have an intensification of therapy due to the major risk of metastasis/relapse.

Focusing on the individual gene members of the genomic and epigenetic signatures related to survival and metastasis/relapse risk, we see that these genes are highly relevant to OSCC disease. Chief among them is *TP53* gene, in which aberrations are the most frequent molecular events in human cancers. This gene encodes the p53 protein, which is involved in many key events in the cell like regulation of cell cycle and glucose metabolism in cancer cells, DNA-repair, apoptosis, and senescence, which are induced by various stress signals, including DNA damage and inflammation [Bensaad and Vousden 2007; Hussain and Harris 2006]. Another gene is *CDKN2A*, in which inactivation was found in 57% of HPV-negative head and neck cancer [Cancer Genome Atlas 2015]. This gene regulates cell cycle progression by blocking the activity of *CCND1* and its associated kinases, CDK6 and CDK4, which phosphorylate and inactivate the tumor suppressor *RB1* [Beck et al. 2015; Burke et al. 2012]. Patients that harbor simultaneous amplified *CCND1* and deleted *CDKN2A* presented worse prognosis [Beck et al. 2015]. *BRCA1* overexpression in leukoplakia followed by subsequent underexpression in tongue squamous cell carcinoma was also already described [Vora et al. 2003b]. *CADM1* downregulation through epigenetic silencing or loss of heterozygosity is related to tumor cell invasion and metastatic potential [van den Berg et al. 2011; van Kempen et al. 2014]. *ATM* plays a role in maintaining genome integrity, being its aberrant promoter methylation linked to lost or greatly diminished expression of several tumor suppressors, namely *BRCA1* [Esteller et al. 2000], *CDKN2A* [Costello et al. 1996] and *VHL* [Herman et al. 1994]. *ATM* promoter hypermethylation showed a significant correlation with decreased overall OSCC patient survival [Ai et al. 2004]. We previously reported the association of *MSH6* and *GATA5* promoter methylation with OSCC worse prognosis, being the later gene also significantly associated with shorter survival rate [Ribeiro et al. 2016]. *PAX5* promoter methylation was considered useful to identify patients at high risk of locoregional recurrence, after surgical treatment [Hayashi et al. 2015].

Even with some limitations, such as, the different platform used to obtain the genomic results (MS-MLPA in our study cohort and SNP-microarray in TCGA cohort)

and also the different clinical follow-up frame time of the two cohorts, our results were in general validated in this larger and geographical distinct cohort (from TCGA database). This validation proves the clinical relevance of the identified signatures for OSCC patients. The reduced follow-up time of our patients (1 - 63 months) and also the junction of two databases with different patients' follow-up time could explain the absence of survival statistical significance results; nevertheless, the survival difference between the two identified clusters is clear, representing a huge difference in the life time of these patients. Considering the genomic signature to predict the risk of relapses/ metastases, the absence of statistical significance in both cohorts could be due to the reduced number of patients that developed metastases and relapses during the time of follow-up.

Screening oral cancer implies searching oral potentially malignant and cancerous/recurrence lesions, typically before symptoms occur in people belonging to groups of risk. In 2005 Bremmer et al showed that the Multiplex Ligation-dependent Probe Amplification (MLPA) technique is suitable for detection of genetic alterations in noninvasive samples, opening the door to improve the early diagnostic of oral potential malignant lesions [Bremmer et al. 2005]. Four years later, the same group [Bremmer et al. 2009] compared the results from exfoliated cells and oral lesions biopsies of the same patients showing a high sensitive rate of this noninvasive technique. However, up to now, this technique was not validated, being still a promise for clinical practice without evidence of a truly practical application. We showed in our cohort that using MS-MLPA technique we could evaluate the *CDKN2A* gene promoter methylation status with same result either in scrapped cells from surface of tumor or in tumor tissue. The major limitation of these non-invasive samples could be the admixture with normal cells which hampers the detection of genetic and epigenetic alterations of tumor cells. Further studies regarding the collection of these non-invasive samples and also the technologies used to perform the molecular evaluations are needed.

Conclusions

Our results suggest the involvement of a specific set of genes that together has the ability to predict the patients' prognosis. The genomic and epigenetic signatures proposed include genes that individually were already related to OSCC and together proved to be associated to survival and risk of metastasis/relapse development. These

findings are of particular clinical relevance since accurate prognostic biomarkers are still urgently needed in the clinical management of OSCC patients. Our genomic signatures were validated in a larger and geographically separate cohort, from the TCGA database, which reinforce their strength to future biomarker-driven clinical trials. We obtained the same result from *CDKN2A* gene promoter methylation status either in scrapped cells from the surface of tumor or in tumor tissue. Non-invasive methodologies to perform molecular screening in the high risk populations are promising; however, further studies are needed before a translation to clinical practice, namely regarding sample collection in this hard to reach population.

Acknowledgments

Ribeiro I.P. is a recipient of a PhD fellowship (SFRH/BD/52290/2013) from the Portuguese Foundation for Science and Technology. This work was in part supported by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra).

Conflict of interest

The authors have no conflict of interest to disclose.

3.2 Genomic Studies

3.2.1 Genomic predictive model for recurrence and metastasis development in head and neck squamous cell carcinoma patients

Ilda P. Ribeiro, Francisco Caramelo, Luísa Esteves, Joana Menoita, Francisco Marques, Leonor Barroso, Jorge Miguéis, Joana B. Melo, Isabel M. Carreira.

Submitted

Abstract

The head and neck squamous cell carcinoma (HNSCC) population consists mainly of high-risk for recurrence and locally advanced stage patients. Increased knowledge of the HNSCC genomic profile can improve early diagnosis and treatment outcomes. The development of models to identify consistent genomic patterns that distinguish HNSCC patients that will recur and/or develop metastasis after treatment is of utmost importance to decrease mortality and improve survival rates. In this study, we used array comparative genomic hybridization data from HNSCC patients to implement a robust model to predict HNSCC recurrence/metastasis. This predictive model showed a good accuracy (> 80%) and was validated in an independent population from TCGA data portal. This predictive genomic model comprises chromosomal regions from 5p, 6p, 8p, 9p, 11q, 12q, 15q and 17p, where several upstream and downstream members of signaling pathways that lead to an increase in cell proliferation and invasion are mapped. The introduction of genomic predictive models in clinical practice might contribute to a more individualized clinical management of the HNSCC patients, reducing recurrences and improving patients' quality of life. The power of this genomic model to predict the recurrence and metastases development should be evaluated in other HNSCC populations.

Keywords: Copy number alteration; Head and neck carcinoma; Recurrence and metastasis; Candidate genes; Predictive model

Introduction

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common type of cancer worldwide [Ferlay et al. 2010]. The overall five-year survival rate remains of approximately 50% even with treatment advances [Thariat et al. 2015]. HNSCC patient outcomes are strongly linked to tumor stage. Patients with early stage tumors (I and II) have 60-95% possibility of successful treatment; however, when diagnosed two thirds of patients already exhibit disease in advanced stage (III and IV) [Worsham 2011]. Tumor recurrence and metastasis lead to a poor prognosis and quality of life, being the recurrence rate in HNSCC patients of about 50% during the first two years after the diagnosis of the primary tumor [Argiris et al. 2008]. Patients with failure after first-line therapy have a median overall survival of less than one year [Vermorken et al. 2008]. Some clinical-pathological parameters have been pointed out to prognosis, recurrence, and survival, namely tumor primary site, nodal involvement, tumor thickness, and the status of the surgical margins [Jerjes et al. 2010]. However, in the actual HNSCC clinical practice, treatment modalities and prognosis are still based only in the TNM staging system classification, which leads to a homogeneous treatment for different HNSCC tumors. Genomic factors also play an important role in the aetiology of these tumors, being the malignant transformation of the cells characterized by a progressive and sequential acquisition of genomic abnormalities, which provide a selective growth advantage to cancer cells. Accurate and reliable methods to predict which HNSCC patients are most likely to recur or to develop distant metastases would significantly enhance the choice of personalized treatment modalities and consequently improve survival of patients. This stratification of patients has been difficult to obtain due to the numerous anatomic sites, the unpredictable clinical behavior and heterogeneous molecular features of these tumors [Worsham et al. 2012]. In this study, we used whole genome copy number alterations (CNAs) to predict recurrence/metastasis development in HNSCC patients. Our predictive model presented an accuracy of more than 80% and it was validated in a TCGA cohort, representing a step further in the identification of clinically significant biomarkers with predictive value for HNSCC management.

Material and methods

Study population

The study protocol was approved by the Committee on Ethics in Research of the Faculty of Medicine of the University of Coimbra. All patients provided their written consent to participate in the study after being informed about the research purposes, following the regulations in the Declaration of Helsinki.

The study cohort includes tissue specimens from 104 HNSCC patients who underwent treatment with curative intent. The patients were recruited between October 2010 and August 2015 from the Maxillofacial Surgery and the Department of Otorhinolaryngology - Head and Neck Surgery, of the Coimbra Hospital and University Centre, CHUC, EPE, Portugal. Diagnosis and staging were performed in accordance with the American Joint Committee on Cancer TNM staging system. The participants in this study answered a survey regarding lifestyle and risk factors for upper aerodigestive tract malignancies, including alcohol and tobacco consumption. Patients were followed-up through hospital revisits during routine clinical appointments. The final date of follow-up was February 29, 2016. The follow-up periods ranged from 6 to 64 months. Details of our study cohort are listed in table 1.

For control, gingival tissues from healthy donors subjected to wisdom teeth removal were used.

The tissue samples were snap-frozen in liquid nitrogen within 30 min after resection and stored at -80ºC until use.

DNA extraction and array-CGH analysis

DNA from fresh frozen tissues of patients and controls were extracted using a High Pure PCR Template Preparation Kit (Roche GmbH, Mannheim, Germany), according to the manufacturer's instructions. The DNAs were quantified by UV spectrophotometric analysis using a Nanodrop 1000 Spectrophotometer (Thermo Scientific, USA).

High-resolution whole genome analyses were performed using Agilent SurePrint G3 Human Genome microarray 180 K (Agilent Technologies, Santa Clara, CA, USA), according Pinto-Leite et al 2014 [Pinto-Leite et al. 2014]. DNA of tumor samples was labelled with Cy5 by random primer labelling. DNA from controls was labelled with Cy3.

Results were analysed using Agilent Genomic Workbench v6.5 software with the following settings: ADM1 as aberration algorithm, threshold of 6.0, moving average 2 Mb. The results are according to Human Genome build 19 and include imbalances with at least three consecutive probes with abnormal log_2 ratios.

Chapter 3.2.1 - Table 1. Clinic-pathologic characteristics of study population - our cohort.

Patients (n = 104)			
	n (%)		n (%)
Gender		Age at diagnosis (Years)	
Male	88 (84,6)	**<60**	52 (50)
Female	16 (15,4)	**≥60**	52 (50)
Anatomic Subsite		Invasion peri(neural)	
Tongue	44 (42,3)	**Yes**	47(45,2)
Floor of the mouth	28 (26,9)	**No**	40(38,5)
Retromolar Trigone	8 (7,7)	NA	17(16,3)
Jugal Mucosa	6 (5,8)	Differentiation	
Palate	4 (3,8)	**Well**	76 (73,1)
Alveolar ridge	6 (5,8)	**Moderate**	22 (21,2)
Tonsil	2 (1,9)	**Poor**	1(1,0)
Hypopharynx	2 (1,9)	**NA**	5 (4,8)
Larynx	2 (1,9)	Margins	
Supraglottis	1 (1,0)	**R0**	59(56,7)
Epiglottis	1 (1,0)	**R1**	28(26,9)
Tobacco		**NA**	17(16,3)
Yes	76 (73,1)	HPV	
No	28 (26,9)	**Positive**	3(2,9)
Alcohol		**Negative**	101(97,1)
Yes	70 (67,3)	Treatment	
No	31 (29,8)	**surgery alone**	33 (31,7)
NA	3 (2,9)	**Surgery + RT**	42 (40,4)
TNM stage		**Surgery + RT + QT**	13 (12,5)
I	18 (17,3)	**Surgery + QT**	1 (1,0)
II	27 (26,0)	**RT + QT**	12 (11,5)
III	20 (19,2)	**RT alone**	3 (2,9)
IV	39 (37,5)	Vital status	
		Relapses/Metastasis in follow-up	40 (38,5)
		Dead - HNSCC	33 (31,7)
		Dead-non - HNSCC	6 (5,8)
		Alive	65 (62,5)

NA- Not Available; **RT -** Radiotherapy; **QT –** Chemotherapy

Validation cohort from TCGA data portal

Copy number data, obtained by SNP array using Affymetrix Genome-Wide Human SNP Array 6.0 were downloaded from The Cancer Genome Atlas (TCGA) Data Portal along with the patients' clinical data of 95 HNSCC, available at https://tcga-data.nci.nih.gov/tcga/, on the 23rd October, 2015. The available copy number data was Level 3 data.

Tissue samples were collected by TCGA with appropriate informed consent from newly diagnosed HNSCC patients at the time of their surgical resection. Human Genome Version 19 samples without germline CNVs were used.

Only copy number information for tumor samples that came from the anatomical locations of the tumors contained in our cohort was selected. In addition, only patients that had available recurrence/metastasis status information were considered. The clinical-pathologic features of the validation cohort are listed in table 2.

Chapter 3.2.1 - Table 2. Clinic-pathologic characteristics of study population - TCGA cohort.

Patients (n = 95)			
	n (%)		n (%)
Gender		Age at diagnosis (Years)	
Male	67 (70.5)	**<60**	40 (42.1)
Female	28 (29.5)	**≥60**	55 (57.9)
Anatomic Subsite		Invasion peri(neural)	
Tongue	44 (46.3)	**Yes**	49 (51.6)
Floor of the mouth	13 (13.7)	**No**	34 (35.8)
Buccal Mucosa	6 (6.3)	NA	12 (12.6)
Palate	5 (5.3)	Margins	
Alveolar ridge	7 (7.4)	**R0**	74 (77.9)
Oral Cavity	20 (21.1)	**R1**	7 (7.4)
Tobacco		**NA**	14 (14.7)
Yes	70 (73.7)	Treatment	
No	24 (25.3)	**NA**	72 (75.8)
NA	1 (1.1)	**Surgery + RT + QT**	1 (1.1)
Alcohol		**Surgery + QT**	5 (5.3)
Yes	67 (70.5)	**RT + QT**	14 (14.7)
No	27 (28.4)	**RT alone**	3 (3.2)
NA	1 (1.1)	Vital status	
TNM stage		**Relapses/Metastasis in follow-up**	27 (28.4)
I	6 (6.3)	**Dead - HNSCC**	17 (17.9)
II	18 (18.9)	**Dead-non - HNSCC**	4 (4.2)
III	18 (18.9)	**Alive**	69 (72.6)
IV	53 (55.8)	**NA**	5 (5.3)

NA- Not Available; **RT -** Radiotherapy; **QT -** Chemotherapy

Statistical analysis

Data Preparation

Chromosomes were binned by mean size of alterations in each chromosome, both in our cohort and the TCGA patients. Then, the alterations present in the patients were distributed by those bins, reducing significantly the number of regions to analyse as well as generating a more structured dataset that can be compared across cohorts.

Statistical Classification

Both our cohort and the cohort of patients from TCGA Data Portal were analysed. The latter was used as a means of external validation.

A three-class support vector machine (SVM) algorithm for statistical classification was applied to these data. The three studied classes were: patients with recurrence/metastasis, patients without recurrence/metastasis and patients of unidentifiable class.

The most important regions for the distinction between classes were selected by Gini's coefficient given by a Variable Importance Plot in a bootstrapping scheme applied to a balanced set regarding the number of cases in each class. Since the number of observations in both data sets was limited, few genomic regions were used as well: the six most important variables, for a minimum number of 32 cases, were selected.

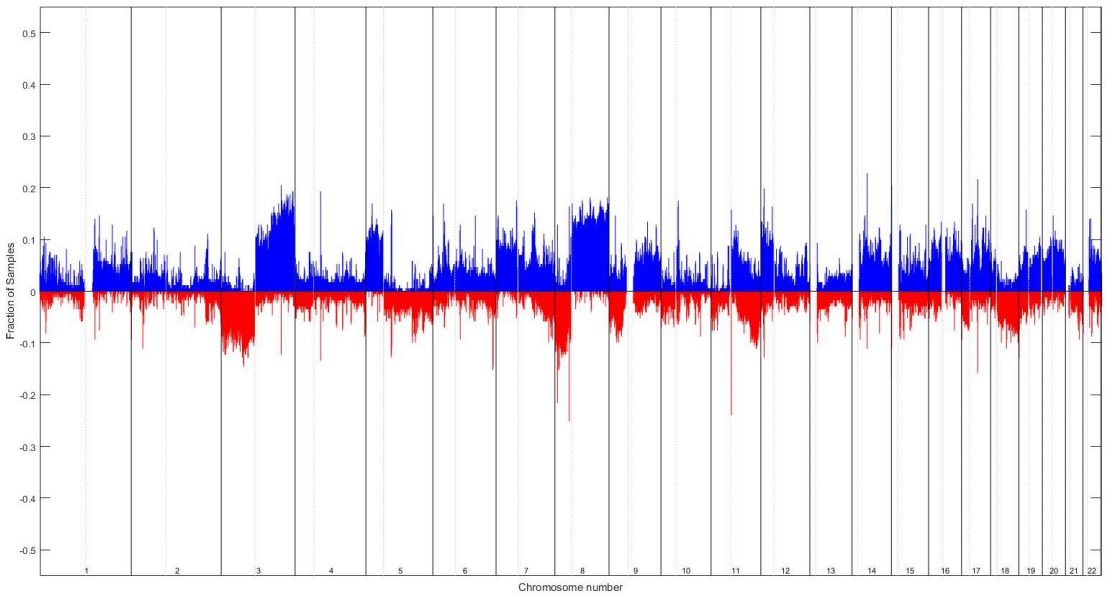
The final multiclass classifier is obtained by the combination of three binary classifiers that distinguishes between: i) having or not recurrence/metastasis; ii) not having recurrence/metastasis and being unidentifiable; iii) having recurrence/metastasis and being unidentifiable. The three binary classifiers are applied to data and their responses are combined using a voting strategy, thus obtaining the final classification.

Model performance is reported across 5000 iterations, executed twice. The algorithm's performance was evaluated by the accuracy considering balanced sets. All analyses were performed using R (version 3.4.0) and Matlab (R2016b).

Results

CNAs detection in HNSCC cohort

The genomic characterization of HNSCC through whole genome array-CGH revealed several copy number gains and losses in all chromosomes (Figure 1), being the chromosomes 3, 5, 7, 8, 9, 11, 12, 14, 15, 16, 17, 18, 19, 20 and 22 the most frequently altered. The most frequent copy number gains were observed at chromosomes 3q, 5p, 7p, 7q, 8q, 11q, 12p, 14, 15, 16, 17, 19, 20, 22 and the most frequent copy number losses were observed at chromosomes 3p, 8p, 9p, 11qter and 18 (Figure 1).



Chapter 3.2.1 - Figure 1. Profile of chromosomal imbalances detected in HNSCC patients using array-CGH technique. Blue represents copy number gains and red copy number losses. The fraction of samples means the fraction of patients that exhibited the imbalance.

Development of a genomic predictive model for HNSCC recurrence and metastasis

The identified chromosomal alteration profile of our HNSCC patients together with their follow up clinical data (6 - 64 months of clinical follow up) were used to build a predictive model for HNSCC recurrence/metastasis development. This model includes three phases: i) identification of patients with vs. without recurrence/metastasis developed after the diagnosis and treatment of primary tumor, during follow up ii) distinction between patients without recurrence and unidentifiable, iii) distinction between patients with recurrence and unidentifiable.

This three-phase model presented an accuracy of 83.6 % CI95% [66.7; 94.4] %, correct prediction of patients without recurrence/metastasis of 92.3% CI95% [66.7; 100] %, correct prediction of patients with recurrence/metastasis of 87.0% CI95% [50.0; 100] % and correct prediction of patients unidentified of 71.4% CI95% [33.0; 100] %.

This model was developed in three phases since we observed that in the first phase of the classification some patients were frequently and systematically misclassified. Interestingly, these misclassified patients presented an overall genomic profile similar to the patients that developed recurrence/metastasis (data not shown). After this observation, we decided to develop two new phases in this predictive model, where we considered the misclassified patients in the first phase of the predictive model as a new category labeled as unidentifiable patients.

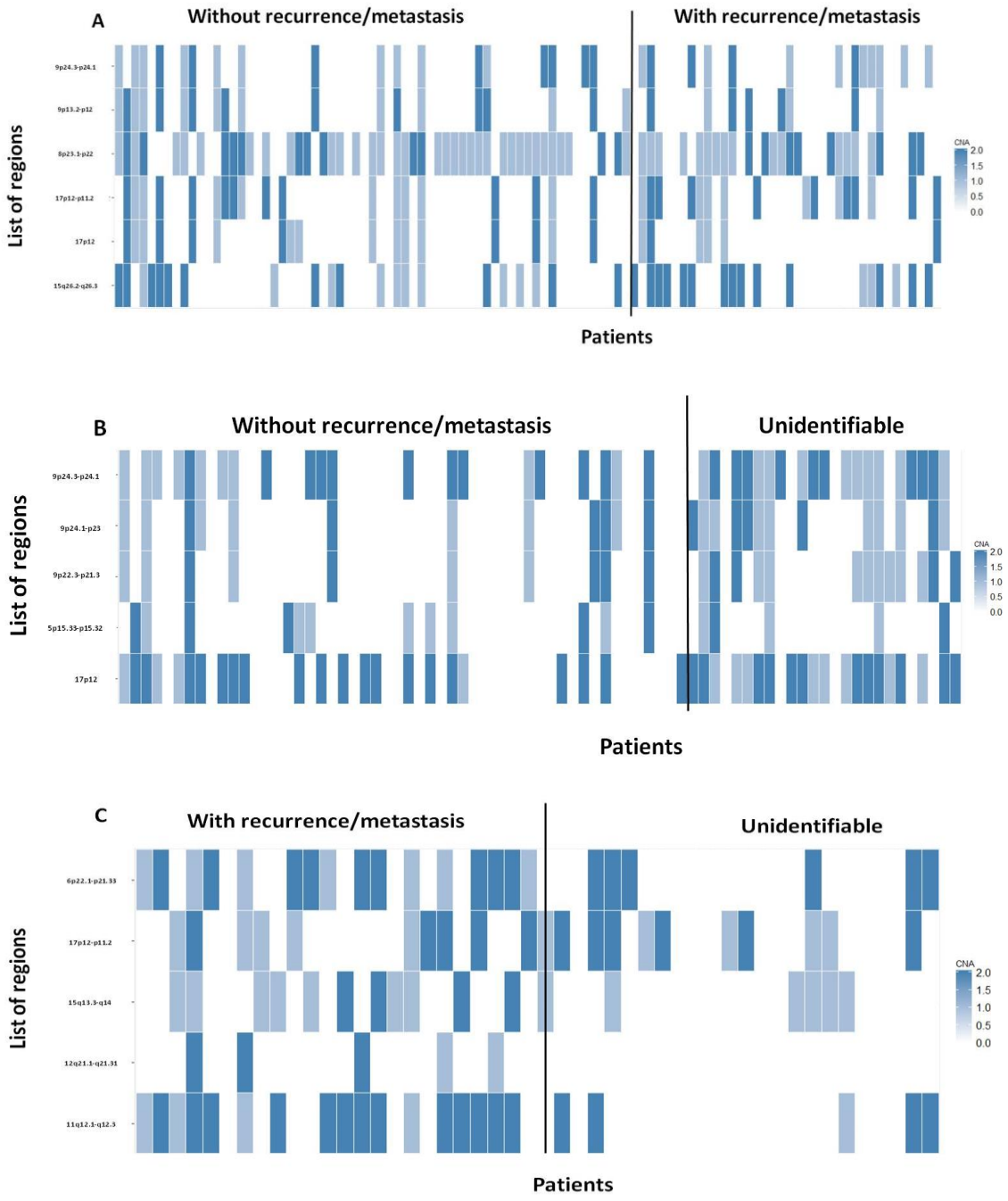
This predictive model, in the first phase: identification of patients with vs. without recurrence/metastasis used six chromosomal regions: 8p23.1-p22, 9p13.2-p12, 9p24.3-p24.1, 15q26.2-q26.3, 17p12 and 17p12-p11.2 (Figure 2A).

The second (patients without recurrence vs. unidentifiable) and third (patients with recurrence vs. unidentifiable) phases of this predictive model considered the following specific chromosomal regions, 8p23.1-p22, 9p13.2-p12, 9p24.3-p24.1, 15q26.2-q26.3, 17p12, 17p12-p11.2 and 6p22.1-p21.33, 11q12.2-q12.3, 12q21.2-q21.31, 15q13.3-q14 and 17p12-p11.2, respectively (Figures 2B, C).

The bands 9p24.3-p24.1 and 17p12 are important for the discrimination of patients that developed or not recurrence/metastasis and for patients without recurrence/metastasis and the unidentifiable ones.

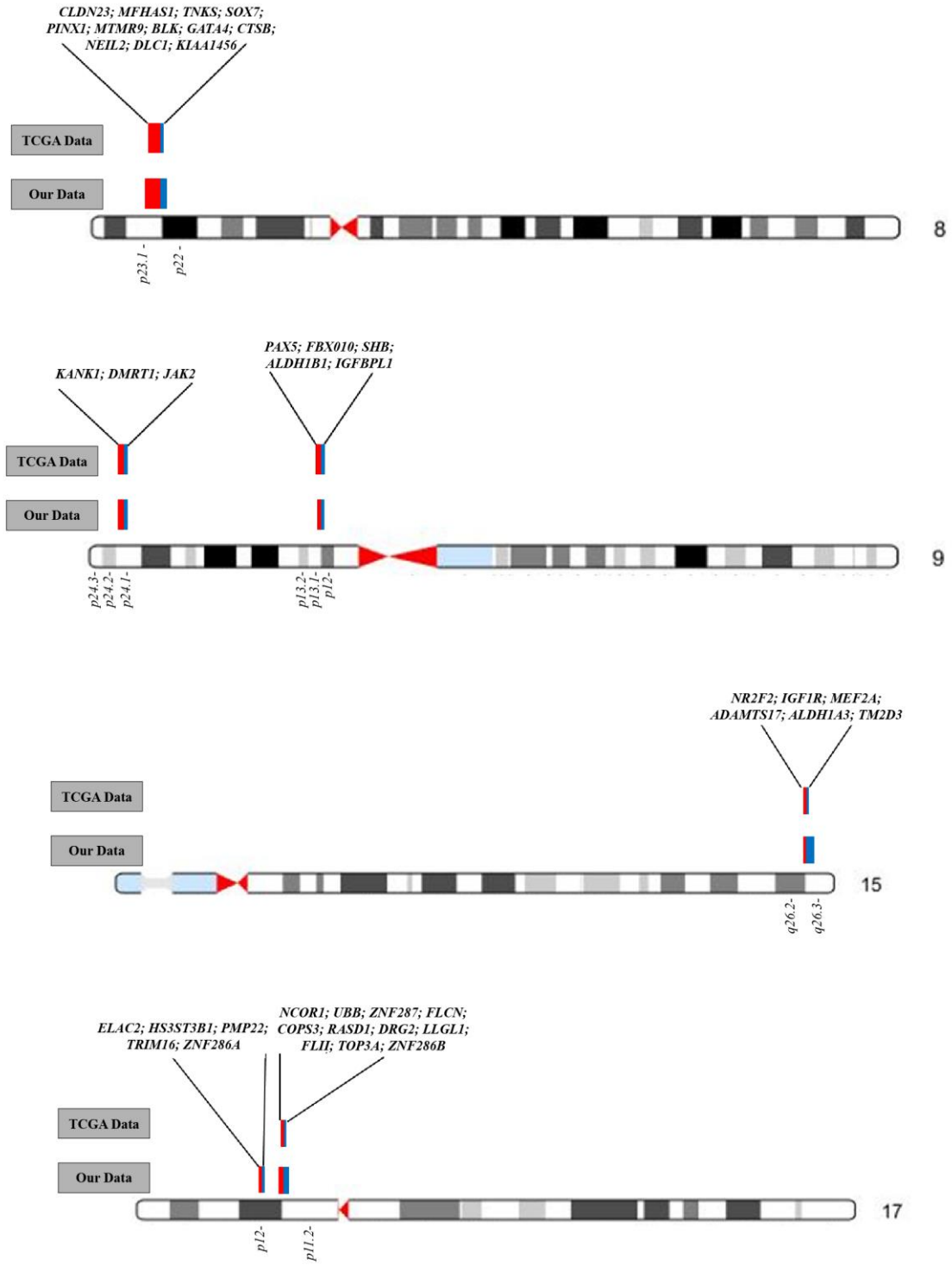
The band 17p12-p11.2 is important for the discrimination of patients that developed or not recurrence/metastasis as well as of patients that develop recurrence/metastases from those that are unidentifiable.

In the specific chromosomal regions of chromosomes 5p, 6p, 8p, 9p, 11q, 12q, 15q and 17p that were used in this three-phase predictive model are mapped important genes for the carcinogenesis process. Through the analysis of the genes mapped in these chromosomal regions using UCSC genome browser (https://genome.ucsc.edu/) and GeneCards - Human gene database (http://www.genecards.org/) we identified some potential candidate genes connected with signaling pathways that control processes associated with tumorigenesis (Figure 3 and supplementary tables 1, 2 and 3).

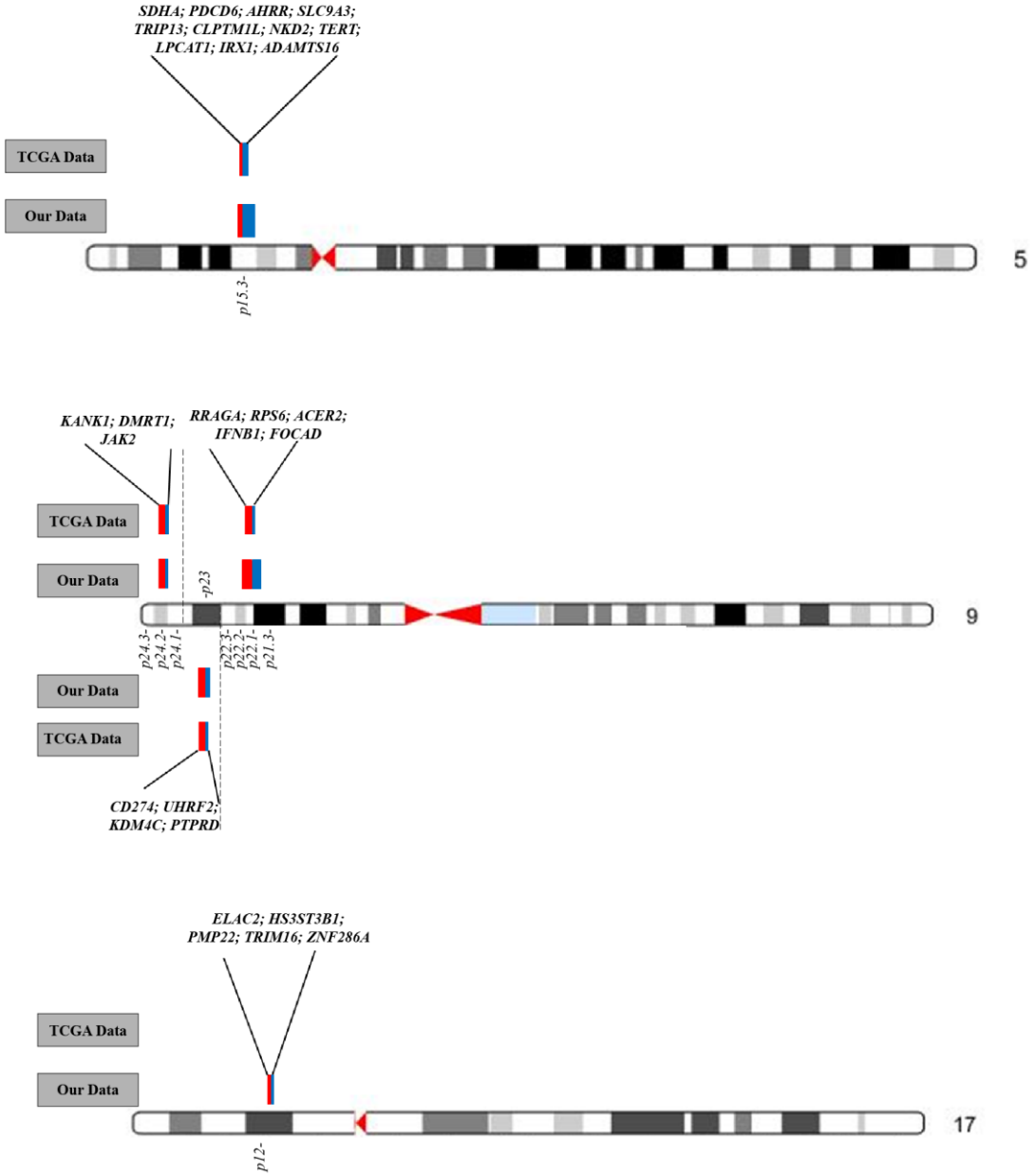


Chapter 3.2.1 - Figure 2. Heatmap with copy number alteration profile in the chromosomal regions used by the predictive genomic model, **A)** in patients with vs. without recurrence/metastasis - first phase of the predictive model; **B)** in patients without recurrence and those unidentifiable - second phase of the predictive model; **C)** in patients with recurrence and those unidentifiable - third phase of the predictive model.

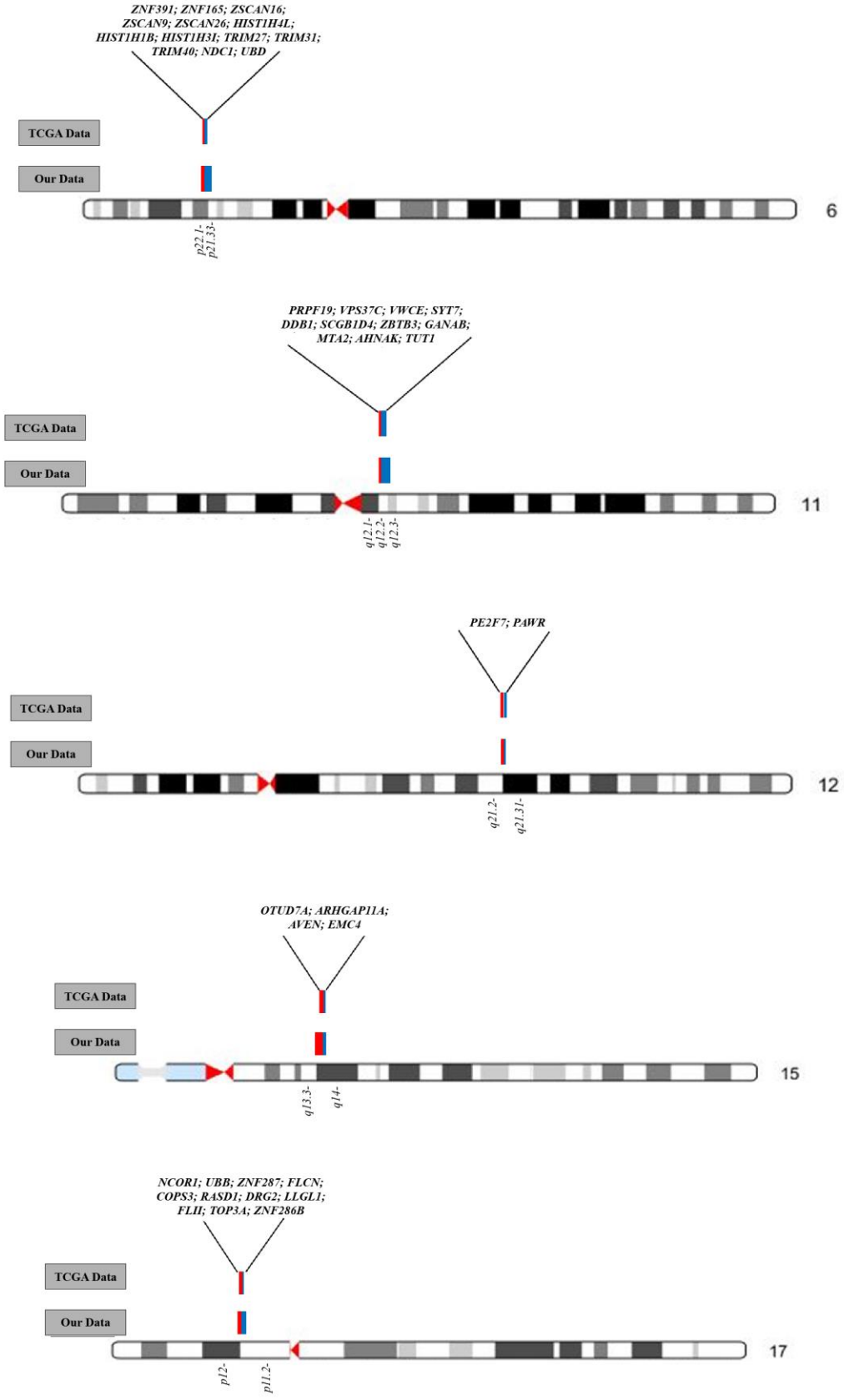
A - Patients with vs. without recurrence/metastasis - first phase of the predictive model



B - Patients without recurrence and those unidentifiable - second phase of the predictive model



C - Patients with recurrence and those unidentifiable - third phase of the predictive model



Chapter 3.2.1 - Figure 3. Ideogram with chromosomal regions used by predictive genomic model and the highlighted candidate genes in these regions. **A)** in patients with vs. without recurrence/metastasis - first phase of the predictive model; **B)** in patients without recurrence and those unidentifiable - second phase of the predictive model; **C)** in patients with recurrence and those unidentifiable - third phase of the predictive model. Blue represents the proportion of copy number gains and red represents the copy number losses identified in these specific chromosomal regions both in our HNSCC patients and in TCGA database.

Validation of the predictive model for recurrence and metastasis development using a TCGA cohort

The usefulness of the developed predictive model based on CNAs was validated using TCGA data from 95 HNSCC, which presented full information regarding their vital status.

Our three-phase model, when applied to TCGA cohort presented an accuracy of 59.4 % CI95% [38.9; 77.8] %, correct prediction of patients without recurrence of 74.7% CI95% [33.3; 100] %, correct prediction of patients with recurrence of 55.1% CI95% [16.7; 100] % and correct prediction of patients unidentified of 48.5% CI95% [0; 100] %.

Discussion

The development and progression of HNSCC is significantly correlated with the accumulation of genomic alterations, allowing cells to escape homeostatic controls that suppress inappropriate proliferation, which result not only in increased proliferation but also in, metabolic changes, genetic instability, induction of angiogenesis, resistance to cell death and increased migratory capacity [Sever and Brugge 2015]. Identification of genome-wide high resolution DNA copy number changes through array-CGH has been applied to a wide range of tumors including HNSCC [Cha et al. 2011; Y. J. Chen et al. 2004; Sparano et al. 2006]. Genomic instability is a common characteristic of cancer cells, with aneuploidy and large-scale DNA rearrangements being frequently observed [Ribeiro et al. 2014b; Ribeiro et al. 2014a; Ribeiro et al. 2016]. However, the relevant chromosomal variations and genes that play a central role in the HNSCC development and progression as well as in the recurrence and metastasis development are not still fully elucidated. In this study we observed several copy number gains and losses in all chromosomes, with chromosomes 3, 5, 7, 8, 9, 11, 12, 14, 15, 16, 17, 18, 19, 20 and 22 being the most frequently altered in our HNSCC cohort (Figure 1). These results revealed the great genomic complexity that underlies

HNSCC. Additionally, these tumors exhibit great heterogeneity in their clinical behavior that cannot currently be predicted using only the available set of clinical markers; therefore, the development of a prognosis predictive model is a novel and promising strategy to increase the HNSCC survival rate and improve the quality of life of the patients. We genomically characterized HNSCC through array-CGH technology, highlighting specific chromosomal alterations. The identified genomic signature was used to build a predictive statistical model of recurrence and metastasis development (Figure 2). This predictive multivariate model presented average accuracy higher than 80% and was validated in a TCGA cohort. This model comprises several upstream and downstream members of signaling pathways that lead to an increase in the cell proliferation and invasion (Figure 3). Components of PI3K/Akt, mTOR, Wnt, Hedgehog, Hippo, Notch, MAP/ERK, were identified as affected in our cohort since several downstream nuclear targets of these signaling pathways are deregulated in the chromosomal regions used by the developed predictive model, such as, *NKD2* (5p15.33), *SOX7* (8p23.1), *RRAGA* (9p22.1) *KANK1* (9p24.3), *JAK2* (9p24.1), *LLGL1* (17p12.2) and *FLCN* (17p11.2). Likewise, genes related to regulation of telomerase, cytoskeletal, metabolism and DNA repair were also frequently altered in tumors and used in this predictive model, namely *SDHA* and *LPCAT1* (5p15.33), *MDC1* (6p21.33), *PINX1* (8p23), *ACER2* (9p22.1), *PRPF19* (6p21.33), *AVEN* (15q14) and *FLII* (17p11.2). As consequence of tumor progression, neoplastic cells become more migratory and develop the capability to invade surrounding tissues, which is accompanied by alterations in adhesion, cell polarity, cytoskeletal dynamics and morphology [Sever and Brugge 2015]. Cell growth is coordinated with metabolic processes involved in the synthesis of macromolecules, thus, cancer cells display metabolic plasticity, altering their metabolic profile during tumorigenesis and metastasis. Altogether, these results suggest a specific set of chromosomal regions and genes that seem to have an important role in the development and prediction of HNSCC recurrence/metastasis. This model could also help in the design of targeted therapies; however, cancer cells seem to develop resistance to inhibition of a particular signaling pathway by expressing alternate protein isoforms or up-regulating compensatory pathways; thus, cancer therapeutic strategies should involve targeting simultaneously multiple deregulated signaling pathways. The observed genomic instability of HNSCC reflects a failure of checkpoint signaling and/or DNA repair mechanisms, denoting the clear need for further research in this area to establish a precise link between these

highlighted specific candidate genes and the HNSCC recurrence/metastasis and, consequently with patients' prognosis.

It is important to stress some limitations of this study, namely the fact that our cohort presented a relatively reduced clinical follow-up time (range from 6 to 64 months), so, some patients with a genomic profile similar to those with recurrence/metastasis could be incorrectly classified in the first phase of the predictive model only because the patients were not followed up enough time to be diagnosed with recurrence/metastasis. This scenario would justify the great majority of unidentified patients. Another limitation is the fact that patients with different anatomic tumors in the head and neck region were analyzed as a homogenous entity, but these tumors were indeed already described as clinical and molecular different entities. In the future, this predictive model should be tested in larger cohorts of different populations of the different head and neck anatomic subsites. Further studies and larger follow-up times should be performed to better characterize the unidentified patients. The validation of this predictive model in the TCGA cohort presented overall reduced quality comparatively to our cohort, which could be due to the fact that we are testing a model specifically developed to array-CGH data in results obtained with a different platform, SNP- microarray. Larger cohorts would also permit the use of more genes, as well as the fine tuning of the set of genes that are more important to the prediction and thus more involved in the development of metastasis/recurrence. The clinical application of this genomic predictive model is promising since it is possible to identify newly diagnosed HNSCC patients with risk of development of recurrence/metastasis and, in this sense, monitor them closely, avoiding or performing early detection of the recurrences and even provide more aggressive and personalized treatment in order to reduce the morbidity and mortality associated with this disease. The complexity of the cancer signaling pathways presents a significant challenge to the development of targeted therapies due to the redundancy of the pathways that control cell proliferation and survival, the crosstalk between pathways, and the feedback inhibition mechanisms that cause pathway reactivation; however, we highlighted in this study several chromosomal regions and genes that could be good candidates for targeted therapy studies.

Since HNSCC has a poor overall prognosis with a high tendency to recur at the primary site and to involve the cervical lymph nodes, this predictive model for recurrence and metastasis development may pave the way to a more practical and individualized patient management and targeted drug design.

Acknowledgments

Ribeiro I.P. is a recipient of a PhD fellowship (SFRH/BD/52290/2013) from the Portuguese Foundation for Science and Technology. This work was in part supported by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra).

Conflict of interest

The authors have no conflict of interest to disclose.

Chapter 3.2.1 - Supplementary Table 1. List of candidate genes identified in the chromosomal regions used by the first phase of predictive model: identification of patients with vs. without recurrence/metastasis.

	Candidate genes	Biological function	Related pathways (GeneCards)
8p23.1-p22	*CLDN23*	Signal transduction	Blood-Brain Barrier and Immune Cell Transmigration: VCAM-1/CD106 Signaling Pathways and Tight junction
	MFHAS1	Potential oncogene	-
	TNKS	Wnt signaling pathway, telomere length and vesicle trafficking	Nicotinate and nicotinamide metabolism and HIV Life Cycle
	SOX7	Tumorigenesis	Wnt Signaling and ERK Signaling
	PINX1	Tumor suppressor	Regulation of Telomerase
	MTMR9	Cell proliferation	-
	BLK	Cell proliferation and differentiation	CXCR4-mediated signaling events and B cell receptor signaling pathwa
	GATA4	Transcriptional regulation	DREAM Repression and Dynorphin Expression and Cardiac conduction
	CTSB	Associated with esophageal adenocarcinoma	Immune System and Toll-Like receptor Signaling Pathways
	NEIL2	Base excision repair of DNA	Recognition and association of DNA glycosylase with site containing an affected pyrimidine and Telomere C-strand (Lagging Strand) Synthesis
	DLC1	Tumor suppressor	G-protein signaling_Regulation of CDC42 activity and Regulation of RhoA activity
	KIAA1456	Tumor suppressor	tRNA processing and Gene Expression
9p13.2-p12	*PAX5*	Differentiation	Transcriptional misregulation in cancer
	FBX010		-
	SHB	Angiogenesis; Apoptosis	EPH-Ephrin signaling and Development VEGF signaling via VEGFR2 - generic cascades
	ALDH1B1	Alcohol metabolism	Metabolism and Cytochrome P450 - arranged by substrate type
	IGFBPL1	Putative tumor suppressor	-
9p24.3p24.1	*KANK1*	Putative tumor suppressor	PI3K/Akt signaling
	DMRT1	Putative tumor suppressor	-
	JAK2	Cell growth, differrentiation and histone modifications	Interferon gamma signaling and RET signaling
15q26.2-q26.3	*NR2F2*	Transcriptional regulation	Oct4 in Mammalian ESC Pluripotency and Regulation of lipid metabolism by Peroxisome proliferator-activated receptor alpha (PPARalpha)
	IGF1R	Anti-apoptotic agent by enhancing cell survival	IL-2 Pathway and Development Ligand-independent activation of ESR1 and ESR2
	MEF2A	Cell growth control and apoptosis	Activated TLR4 signalling and Immune System
	ADAMTS17	Peptidase activity and metalloendopeptidase activity	O-glycosylation of TSR domain-containing proteins and HIV Life Cycle
	ALDH1A3	Aldehyde dehydrogenase (NAD) activity	Drug metabolism - cytochrome P450 and Tyrosine metabolism.
	TM2D3	Regulation of cell death or proliferation signal cascades	-
17p12	*ELAC2*	Mediated growth arrest	tRNA processing and rRNA processing in the mitochondrion
	HS3ST3B1	Epithelial-mesenchymal transition	Cytochrome P450 - arranged by substrate type and Metabolism
	PMP22	Growth regulation	a6b1 and a6b4 Integrin signaling and Neural Crest Differentiation
	TRIM16	Cell growth, differentiation and pathogenesis	-
	ZNF286A	Transcriptional regulation	Gene Expression
17p12-p11.2	*NCOR1*	Chromatin remodeling and repression of transcription	Regulation of nuclear SMAD2/3 signaling and BMAL1-CLOCK,NPAS2 activates circadian gene expression
	UBB	Involved in the maintenance of chromatin structure, the regulation of gene expression, and the stress response.	Interferon gamma signaling and Activated TLR4 signalling
	ZNF287	Transcriptional regulation	Gene Expression
	FLCN	Putative tumor suppressor	ErbB signaling pathway and mTOR signaling pathway
	COPS3	Signal transduction	Vesicle-mediated transport and Transcription-Coupled Nucleotide Excision Repair (TC-NER)
	RASD1	Cell morphology, growth and cell-extracellular matrix interactions	Neuroscience and MAP Kinase Signaling.
	DRG2	Cell growth and differentiation	-
	LLGL1	Cytoekeletal network	Tight junction and Hippo signaling pathway
	FLII	Regulation of cytoekeletal rearrangements	Cytoskeletal Signaling
	TOP3A	Controls and alters the topologic states of DNA during transcription	Meiosis and Cell Cycle Checkpoints
	ZNF286B	Transcriptional regulation	-

Chapter 3.2.1 - Supplementary Table 2. List of candidate genes identified in the chromosomal regions used by the second phase of predictive model: identification of patients without recurrence/metastasis. vs. unidentifiable

	Candidate genes	Biological function	Related pathways (GeneCards)
5p15.33-p15.32	*SDHA*	Putative tumor suppressor	Metabolism and Citrate cycle (TCA cycle)
	PDCD6	May inhibit KDR/VEGFR2-dependent angiogenesis; the function involves inhibition of VEGF-induced phosphoprylation of the Akt signaling pa	-
	AHRR	Cell growth and differentiation	Regulation of lipid metabolism by Peroxisome proliferator-activated receptor alpha (PPARalpha) and Metabolism
	SLC9A3	Signal transduction	Transport of glucose and other sugars, bile salts and organic acids, metal ions and amine compounds and Protein digestion and absorption.
	TRIP13	Putative role in early-stage non-small cell lung cancer	-
	CLPTM1L	Enhances cisplatin-mediated apoptosis	-
	NKD2	Negative regulators of Wnt receptor signaling	Wnt Signaling Pathway and Pluripotency and Wnt / Hedgehog / Notch
	TERT	Cellular senescence and oncogenesis	HTLV-I infection and Chromosome Maintenance
	LPCAT1	Progression of oral squamous cell, prostate, breast, and other human cancers	Metabolism and Acyl chain remodelling of PE
	IRX1	Tumor suppressor in gastric and head and neck cancers	-
	ADAMTS16	Metalloendopeptidase activity	O-glycosylation of TSR domain-containing proteins and HIV Life Cycle
9p22.3-p21.3	*RRAGA*	Inhibitor of TNF-alpha functions, affecting cell death	mTOR signalling and TP53 Regulates Metabolic Genes
	RPS6	Cell growth and proliferation	Activation of the mRNA upon binding of the cap-binding complex and eIFs, and subsequent binding to 43S and Metabolism
	ACER2	Cell proliferation and survival	Metabolism and sphingosine and sphingosine-1-phosphate metabolism
	IFNB1	Cell differentiation and anti-tumor defenses	Interferon gamma signaling and Regulation of nuclear SMAD2/3 signaling
	FOCAD	Potential tumor suppressor in gliomas	-
9p24.3p24.1	*KANK1*	Putative tumor suppressor	PI3K/Akt signaling
	DMRT1	Putative tumor suppressor	-
	JAK2	Cell growth, differrentiation and histone modifications	Interferon gamma signaling and RET signaling
9p24.1-p23	*CD274*	Prognostic value in colon cancer and renal cell carcinoma	Immune System and IgA-Producing B Cells in the Intestine
	UHRF2	Cell cycle regulation	
	KDM4C	Regulation of gene expression and chromosome segregation	Activated PKN1 stimulates transcription of AR (androgen receptor) regulated genes KLK2 and KLK3 and Chromatin organization
	PTPRD	Cell growth, differentiation, mitotic cycle, and oncogenic transformation	Transmission across Chemical Synapses and Protein-protein interactions at synapses.
17p12	*ELAC2*	Mediated growth arrest	tRNA processing and rRNA processing in the mitochondrion
	HS3ST3B1	Epithelial-mesenchymal transition	Cytochrome P450 - arranged by substrate type and Metabolism
	PMP22	Growth regulation	a6b1 and a6b4 Integrin signaling and Neural Crest Differentiation
	TRIM16	Cell growth, differentiation and pathogenesis	-
	ZNF286A	Transcriptional regulation	Gene Expression

Chapter 3.2.1 - Supplementary Table 3. List of candidate genes identified in the chromosomal regions used by the third phase of predictive model: identification of patients with recurrence /metastasis. vs. unidentifiable.

	Candidate genes	Biological function	Related pathways (GeneCards)
6p22.1-p21.33	*ZNF391; ZNF165; ZSCAN16; ZSCAN9; ZSCAN26*	Transcriptional regulation	-
	HIST1H4L	Transcription regulation, DNA repair, DNA replication and chromosomal stability	Mitotic Prophase and Activated PKN1 stimulates transcription of AR (androgen receptor) regulated genes KLK2 and KLK3
	HIST1H1B	Regulator of individual gene transcription through chromatin remodeling, nucleosome spacing and DNA methylation	Apoptosis induced DNA fragmentation and Regulation of activated PAK-2p34 by proteasome mediated degradation
	HIST1H3I	Transcription regulation, DNA repair, DNA replication and chromosomal stability	Mitotic Prophase and Immune System
	TRIM27	Transcriptional repressor activity; Apoptosis	-
	TRIM31	Negative regulator of cell growth	Interferon gamma signaling and Immune System
	TRIM40	Negative regulator against inflammation and carcinogenesis	-
	MDC1	Cell cycle checkpoints in response to DNA damage	DNA Double Strand Break Response and ATM Pathway
	UBD	Apoptosis	PEDF Induced Signaling and Beta-Adrenergic Signaling
11q12.1-q12.3	*PRPF19*	Cell survival and DNA repair	mRNA Splicing - Major Pathway and Transcription-Coupled Nucleotide Excision Repair (TC-NER)
	VPS37C	Cell growth and differentiation	Vesicle-mediated transport and HIV Life Cycle
	VWCE	Target for chemoprevention of hapatocellular carcinoma	-
	SYT7	Associated with prostate cancer.	Protein-protein interactions at synapses and Transmission across Chemical Synapses
	DDB1	Damaged DNA binding	Toll-like receptor signaling pathway and Nucleotide excision repair
	SCGB1D4	Involved in the regulation of chemotactic cell migration and invasion	-
	ZBTB3	Transcriptional regulation	-
	GANAB	Associated to lung tumor	Transport to the Golgi and subsequent modification and Calnexin/calreticulin cycle
	MTA2	Transcriptional regulation (It is closely related to another member of this family, a protein that has been correlated with the metastatic potential of certain carcinomas)	Activated PKN1 stimulates transcription of AR (androgen receptor) regulated genes KLK2 and KLK3 and RNA Polymerase I Promoter Escape
	AHNAK	Tumor metastasis	-
	TUT1	Regulation of gene expression and cell proliferation	-
12q21.2-q21.31	*E2F7*	Regulation of cell cycle progression	TP53 Regulates Transcription of Cell Cycle Genes and Gene Expression
	PAWR	Tumor suppressor	-
15q13.3-q14	*OTUD7A*	Putative tumor suppressor	Ovarian tumor domain proteases and Metabolism of proteins
	ARHGAP11A	Cell-cycle arrest and apoptosis	Signaling by GPCR and p75 NTR receptor-mediated signalling
	AVEN	Apoptosis	Apoptosis and Autophagy
	EMC4	Anti-apoptotic activity	-
17p12-p11.2	*NCOR1*	Chromatin remodeling and repression of transcription	Regulation of nuclear SMAD2/3 signaling and BMAL1-CLOCK,NPAS2 activates circadian gene expression
	UBB	Involved in the maintenance of chromatin structure, the regulation of gene expression, and the stress response.	Interferon gamma signaling and Activated TLR4 signalling
	ZNF287	Transcriptional regulation	Gene Expression
	FLCN	Putative tumor suppressor	ErbB signaling pathway and mTOR signaling pathway
	COPS3	Signal transduction	Vesicle-mediated transport and Transcription-Coupled Nucleotide Excision Repair (TC-NER)
	RASD1	Cell morphology, growth and cell-extracellular matrix interactions	Neuroscience and MAP Kinase Signaling.
	DRG2	Cell growth and differentiation	-
	LLGL1	Cytoekeletal network	Tight junction and Hippo signaling pathway
	FLII	Regulation of cytoekeletal rearrangements	Cytoskeletal Signaling
	TOP3A	Controls and alters the topologic states of DNA during transcription	Meiosis and Cell Cycle Checkpoints
	ZNF286B	Transcriptional regulation	-

3.2.2 Chromosomal breaks and low copy repeat elements in head and neck squamous cell carcinoma

Ilda P. Ribeiro, Isabel M. Carreira, Thomas Liehr, Joana B. Melo

Submitted

Abstract

Head and neck squamous cell carcinoma (HNSCC) present complex chromosomal rearrangements, however, the molecular mechanisms behind HNSCC-development remain elusive. These patients are often diagnosed in advanced stage, which is among others a consequence of the lack of useful diagnostic biomarkers. The identification of recurrent chromosomal breakpoints could help to understand the molecular mechanisms behind HNSCC development, behavior, and consequently provide to the establishment of diagnostic biomarkers. Array comparative genomic hybridization was performed in 104 HNSCC patients and the chromosomal breaks being involved in gene amplification or loss were analyzed. It turned out that low copy repeat DNA sequences are localized at or in close proximity to breakpoint sites identified. Frequent non-random break events clustered in chromosomes 12p, 8p, 3q, 14q, 6p, 4q, Xq and 8q. Chromosomes 6, 14, 3, 8 and X exhibited a higher susceptibility to breaks than other chromosomes. The probability to find repeat DNA elements showed a maximum 100 kb downstream and 100 kb to 10 Mb upstream of breakpoint, respectively. LINEs, SINEs and LTRs were the most frequent repeat elements identified. Thus, specific non-random genomic breaks involve low repeat chromosomal regions in HNSCC genomic instability.

Keywords: Breakpoint, DNA repeat elements; Non-random breaks; Head and neck cancer

Introduction

Head and neck cancer is a heterogeneous epithelial tumor group of the upper aerodigestive tract, of which ~90% are squamous cell carcinomas [Pai and Westra 2009]. These tumors arise from the oral cavity, oropharynx, hypopharynx and larynx and are grouped together as head and neck squamous cell carcinoma (HNSCC). This carcinoma goes together with an accumulation of several complex chromosomal rearrangements. As frequent chromosomal alterations in these tumors, deletions, translocations, isochromosomes, and marker chromosomes were observed, whereas duplications, insertions, inversions, ring-, dicentric-, and endoreduplicated chromosomes have been reported less frequently [Susanne M. Gollin 2001]. These chromosomal abnormalities are continually acquired over the time leading to malignant transformation and consequently to genomic instability, a hallmark of cancer [Negrini et al. 2010]. The molecular basis behind this genomic instability is not yet understood. Some studies suggest that replication stress through common fragile sites is a major cause of genomic instability [Mrasek et al. 2010]. In HNSCC, despite diverse aberrations identified, several non-random chromosomal breakpoints are observed, being chromosomal rearrangements involving the centromeric or juxtacentromeric bands [Susanne M. Gollin 2001]. The origin of these common, repeated aberrations in HNSCC patients could be influenced by the local genomic architecture, which includes the presence of repeat sequences within or flanking the breakpoints. Repeat elements could be linked to malignant transformation process and also in the promotion of increased mutation and recombination rates, since their activity is known to be higher in tumor cells compared to normal cells [Kozeretska et al. 2011]. Interspersed repeats represent the most common type of low copy repeats, covering almost 44% of the human genome [L. Chen et al. 2014]. The great majority of transposable elements are retrotransposons, which are mobile elements with the ability to integrate into different regions. Long interspersed nuclear elements (LINEs), short interspersed nuclear elements (SINEs), and Long terminal repeats (LTRs) represent the three major categories of mammalian retrotransposons [Cardoso et al. 2016] and reach about 20%, 13% and 8.3% of the human genome, respectively [Kozeretska et al. 2011; Lander et al. 2001]. L1 is the most common LINE and Alu is the most common SINE element, encompassing approximately 17% and 11% of human genomic DNA, respectively [Kozeretska et al. 2011; Lander et al. 2001]. A full-length LINE element has approximately 8 kb, while SINE elements are shorter with 100 - 300 bp, as are the LTR transposons with 5 kb [Treangen and Salzberg

[2011]. Within the human genome L1 transposon activity persisted over time being associated with genomic instability and tumor development. Thus L1 transposon activity is a merely non-random event that precedes clonal expansion [Lee et al. 2012]. It was pointed out as an important contributor to destabilization of the cancer genome by integration/mobilization in or near oncogenes or tumor suppressor genes [Iskow et al. 2010]. The presence of repeat elements within breakpoint region can also play a role in non-random break events in cancer.

In this study we investigated if there are detectable features of DNA sequence-structure, which could be involved in HNSCC breakpoint development, i.e. if some sequence motifs are co-localized with these breakpoints. Interestingly, chromosomal regions in the peri-centromeric and telomeric regions were more frequently involved in breakpoints than to be expected. Also the DNA repeat elements LINEs, SINEs and LTRs could be shown to be mostly absent in breakpoint regions themselves.

Material and Methods

Study population

The study protocol was approved by the Committee on Ethics in Research of the Faculty of Medicine of the University of Coimbra. All patients provided their written consent to participate in the study after being informed about the research purposes, following the regulations in the Declaration of Helsinki.

The cohort in analysis includes tumor tissue specimens from 104 HNSCC patients who underwent tumor resection. Patients were recruited between October, 2010 and April, 2015 from the Maxillofacial Surgery and Stomatology Units, of the Coimbra Hospital and University Centre, CHUC, EPE, Portugal. Diagnosis and staging were performed in accordance with the American Joint Committee on Cancer's TNM staging system. The detailed characterization of cohort in study is illustrated in table 1. For the control group, 16 gingival tissues from healthy subjects (6 males and 10 females), with age range from 18 to 81 years old, submitted to "wisdom tooth" (molar 8) removal were used.

Tissue samples were snap-frozen in liquid nitrogen within 30 min after resection and stored at -80ºC until use.

DNA extraction

DNA from tumor and control samples was extracted using High Pure PCR Template Preparation Kit (Roche GmbH, Mannheim, Germany), according to the manufacture's recommendations. DNA concentration and purity were measured using a NanoDrop1000 Spectrophotometer (Thermo Scientific, Waltham, USA).

Chapter 3.2.2 - Table 1. Clinic-pathologic characteristics of study population.

Patients (n = 104)			
	n		n
Gender		Age at diagnosis (Years)	
Male	88	**<60**	52
Female	16	**≥60**	52
Anatomic Subsite		TNM stage	
Tongue	44	**I**	18
Floor of the mouth	28	**II**	27
Retromolar Trigone	8	**III**	20
Jugal Mucosa	6	**IV**	39
Palate	4	Differentiation	
Alveolar ridge	3	**Well**	76
Gingival ridge	3	**Moderate**	22
Tonsil	2	**Poor**	1
Hypopharynx	2	**NA**	5
Larynx	2	Margins	
Supraglottis	1	**R0**	59
Epiglottis	1	**R1**	28
Tobacco		**NA**	17
Yes	76	HPV	
No	28	**Positive**	3
Alcohol		**Negative**	101
Yes	70	Invasion peri(neural)	
No	31	**Yes**	47
NA	3	**No**	40
		NA	17

NA - Not Available.

Array CGH

High-resolution whole genome analyses were performed using Agilent SurePrint G3 Human Genome microarray 180 K (Agilent Technologies, Santa Clara, CA, USA), according to [Pinto-Leite et al. 2014]. DNA of head and neck tumor samples was labelled with Cy5 by random primer labelling. DNA from control was labelled with Cy3. Results were analyzed using Agilent Genomic Workbench v6.5 software with the following settings: ADM1 as aberration algorithm, threshold of 6.0, moving average 2 Mb. The results are according to Human Genome GRCh37/hg 19 and include imbalances with at least three consecutive probes with abnormal $\log_2$ ratios.

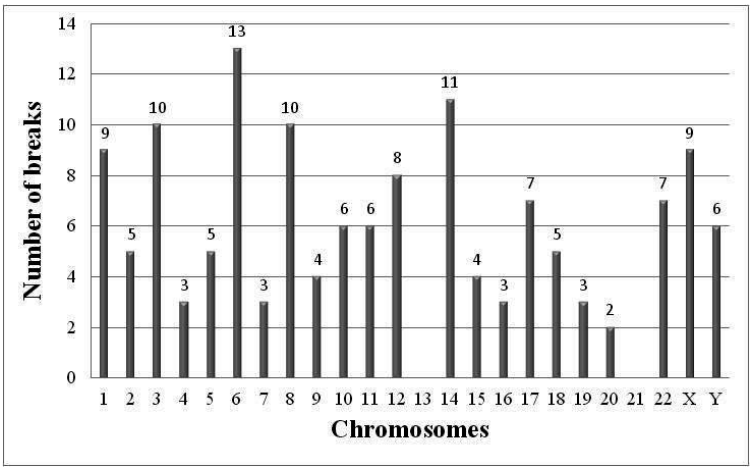
We identified the common breakpoint sites of the copy number gains and losses presented in more than 10% of our cohort.

In Silico Analysis of breakpoints identified

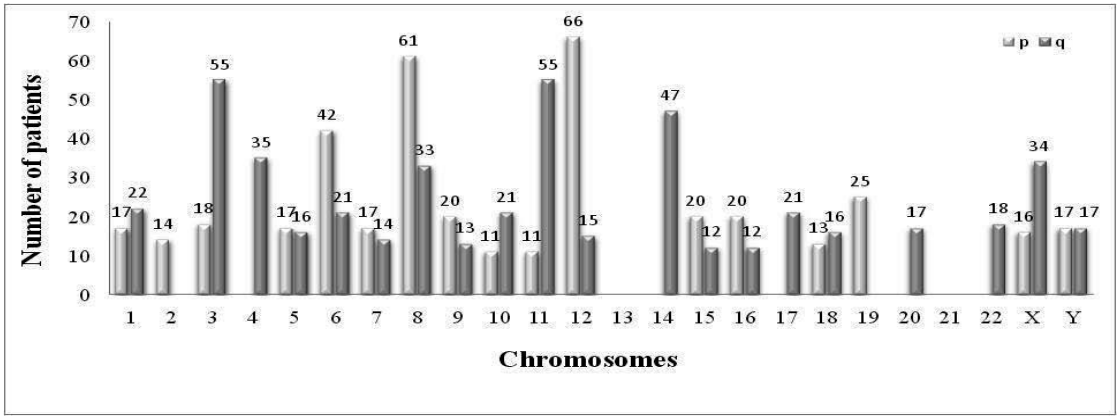
The genomic sequences being present in or surrounding the identified breakpoint sites were retrieved from UCSC genome browser (https://genome.ucsc.edu/; GRCh37/hg 19). The Repeat Masker track and joined fragments of interrupted repeats of the UCSC genome browser was used to determine repeat DNA elements in the breakpoints site, 100bp, 300bp, 500bp, 100 Kb, 500 Kb, 1 Mb and 10 Mb upstream and downstream flanking each breakpoint.

Results

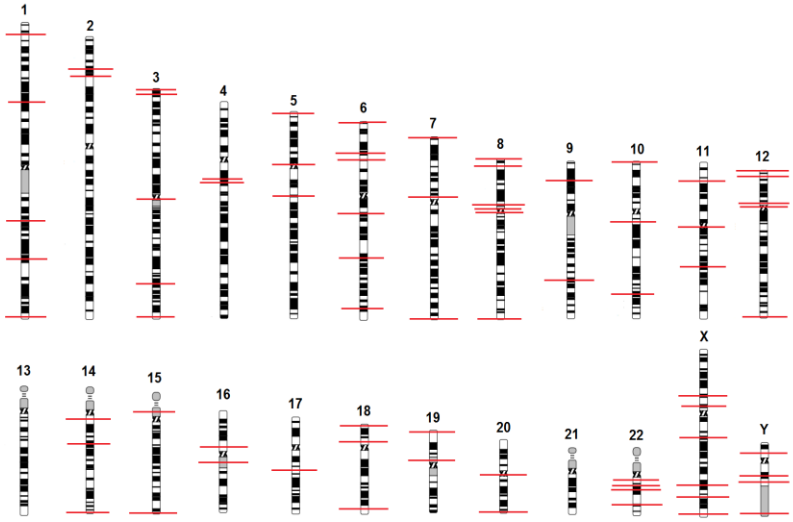
In our HNSCC cohort we observed breaks in almost each and every chromosome. However, chromosomes 6, 14, 3, 8 and X presented higher number of breaks *per* chromosome (Figure 1). Most frequently observable were breaks in chromosomal arms 12p, 8p, 3q, 11q, 14q, 6p, 4q, Xq, while they were underrepresented in 8q, 2q, 4p, 13, 17p, 19q, 20p and 21 (less than 10% of our cohort) (Figure 2). Regarding the distribution along chromosome-axis, we observed several peri-centromeric and telomeric breakpoints (Figure 3).



Chapter 3.2.2 - Figure 1. Average number of breaks observed per chromosome in 104 HNSCC cases studied (cut-off of 10%).

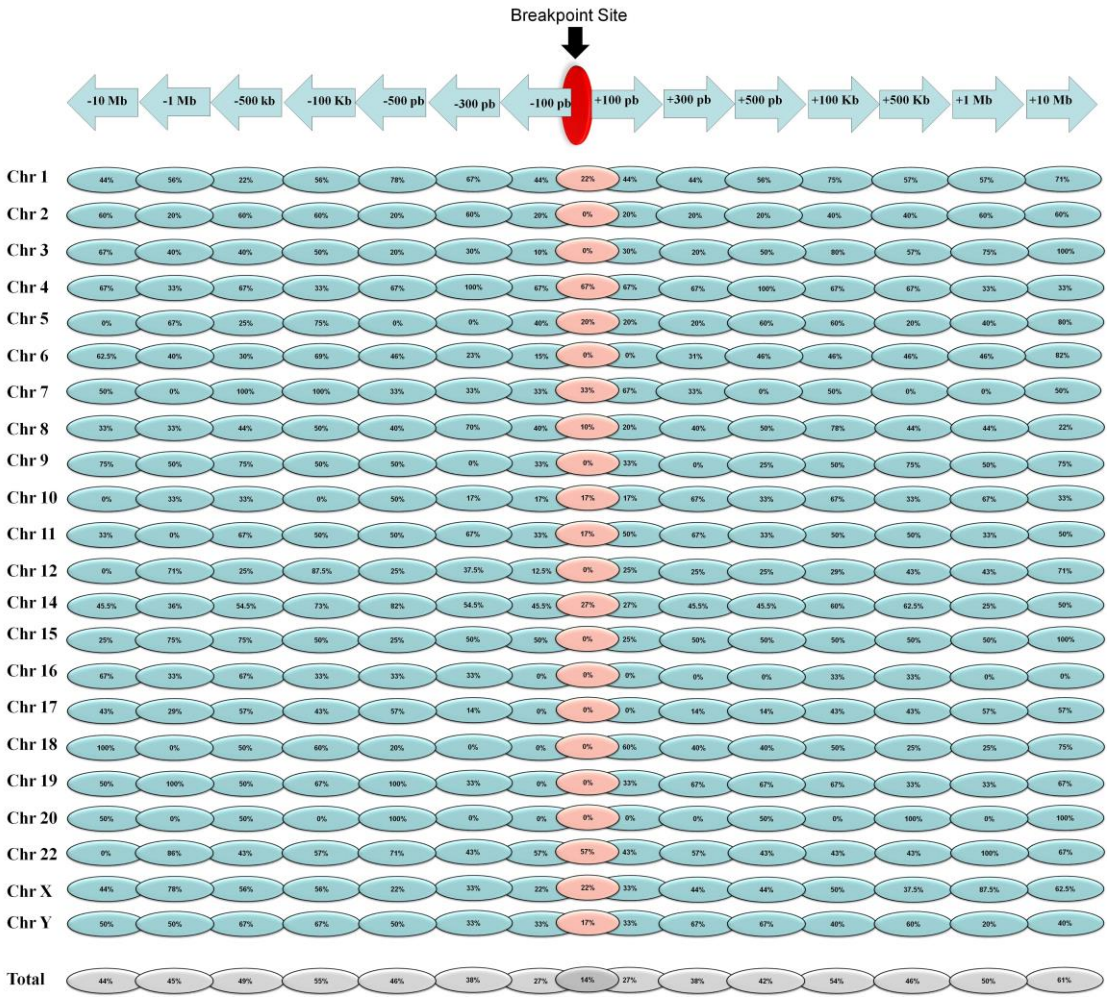


Chapter 3.2.2 - Figure 2. Representation of the most frequent breakpoint per chromosome arm and the number of patients presenting this breakpoint. It was considered for each chromosome arm only one single break, the one most common to a higher number of patients.



Chapter 3.2.2 - Figure 3. Chromosomal distribution of the identified breakpoints in our cohort. Each line represents an identified breakpoint. Several breaks in the same chromosomal band are represented only by one line.

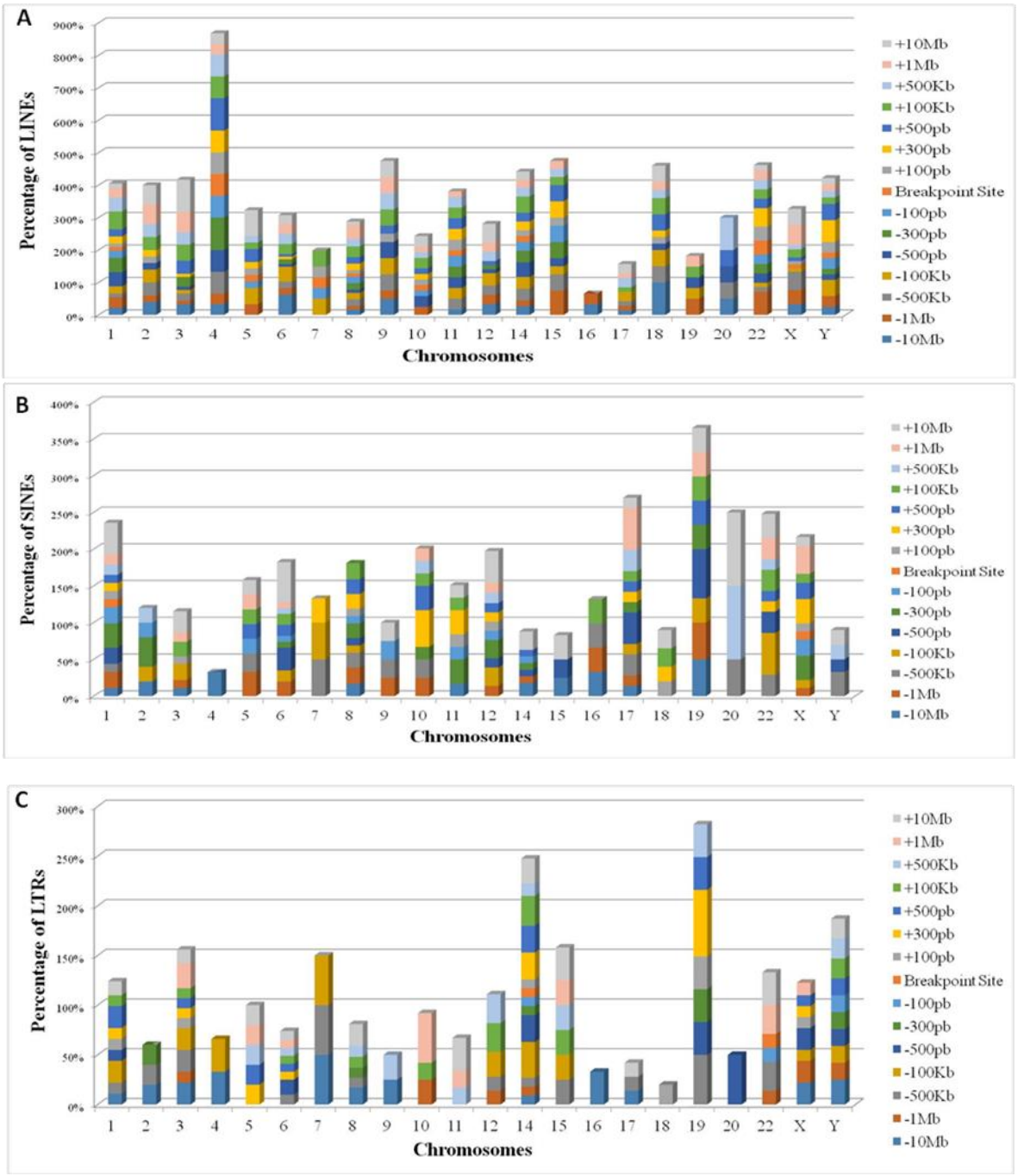
Also we determined the presence of low-repetitive DNA elements along the breakpoints site, 100bp, 300bp, 500bp, 100 Kb, 500 Kb, 1 Mb and 10 Mb upstream and downstream flanking each breakpoint identified. The percentage of characteristic repeat element is lower in the exact breakpoint regions identified in our HNSCC cohort (Figure 4). Moreover, we observed a higher percentage of these sequences around +/- 100 Kb and +10 Mb of the breakpoint site (Figure 4). The great majority of the low copy repeat elements found in the breakpoint regions and in the regions ups- and downstream each breakpoint identified in our cohort were LINEs, SINEs and LTRs (Table 2).



Chapter 3.2.2 - Figure 4. Percentage of low copy repeat elements in the breakpoint site, 100bp, 300bp, 500bp, 100 Kb, 500 Kb, 1 Mb and 10 Mb upstream and downstream flanking each breakpoint identified in our HNSCC cohort.

Analyzing the distribution of these three most common repeat elements in all chromosomes, we highlighted chromosome 4 that includes the highest number of LINEs, especially at 300bp downstream and 500bp upstream the breakpoint site and chromosome 19 that presented the highest percentage of SINEs and LTRs, especially at 500bp downstream and 300bp upstream the breakpoint site, respectively (Figure 5).

Considering the breakpoints being shared by several patients, we observed that LINE, especially the LINE-1 family is the most common repeat element present (Table 3). Additionally, for all chromosomes the region with the presence of more repeat elements is located at 10 Mb upstream the breakpoint site (Tables 2 and 3).



Chapter 3.2.2 - Figure 5. Percentage of repeat elements observed in each chromosome at the breakpoint site, 100bp, 300bp, 500bp, 100 Kb, 500 Kb, 1 Mb and 10 Mb upstream and downstream flanking each breakpoint identified in our HNSCC cohort. **A)** LINEs, **B)** SINEs, **C)** LTRs.

Chapter 3.2.2 - Table 2. Classes of repeat elements found in the breakpoint sites identified in our cohort and in each of the analyzed regions upstream and downstream flanking each breakpoint identified.

Repeat element	-10Mb	-1Mb	-500Kb	-100Kb	-500bp	-300bp	-100bp	Breakpoint Site	+100bp	+300bp	+500bp	+100Kb	+500Kb	+1Mb	+10Mb
LINE	28	32	29	38	31	26	19	15	21	27	33	45	37	38	39
SINE	12	16	20	16	15	18	12	2	7	17	15	17	10	17	26
LTR	15	8	14	16	10	5	3	2	6	10	11	12	10	12	13
DNA Repeat Elements	1	2	4	5	7	4	3	1	1	2	3	3	5	4	6
Simple Repeat	0	2	3	2	2	3	2	0	1	2	0	1	3	0	1
Low Complexity	0	1	2	0	1	0	1	0	2	0	1	1	3	1	1
Satellite	0	0	4	6	6	5	0	0	1	0	0	3	2	0	0
Other	0	0	0	1	0	0	0	0	1	1	1	0	0	0	0
Unknown	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0

Chapter 3.2.2 - Table 3. Classes and/or family of repeat elements found in the most common breakpoint identified in each chromosome for our cohort, analyzed in and around the breakpoint site.

Chr.	Number of patients	-10Mb	-1Mb	-500Kb	-100Kb	-500bp	-300bp	-100bp	Breakpoint Site	+100bp	+300bp	+500bp	+100Kb	+500Kb	+1Mb	+10Mb
1q31.3	22	No	No	No	LINE	LINE	LINE	LINE	No	No	No	No	LINE	LINE, L2c	LINE L1PA7, L1PA4	LTR
2p23.1	14	LINE, L2c	No	No	LINE	No	No	No	No	No	No	No	LINE	SINE, AluY, L1MEg	No	LINE
3q26.1	55	LINE, L1MDa	No	No	LTR	No	No	No	No	No	No	No	LINE	Simple, L1ME1, L1MA5	No	SINE, L1MD1
4q13.2	35	LTR, L1MC2	No	LINE, L1M3	No	No	LINE, L1M1	L1M1	L1M1	L1M1	L1M1	LINE, L1M1	LINE	LINE	No	No
5p15.33	17	-	-	-	-	No	No	No	No	No	No	LINE	LINE; L1MEF	No	No	LINE, L1M3f
6p25.3	42	-	-	-	LINE, L1PA7	SINE, L1PA16	LINE, L1PA16	No	No	No	DNA	DNA	No	SINE	DNA	SINE, L1MC1
7p22.3	17	-	-	-	-	No	No	LINE	No	No	No	No	No	No	No	No
8p11.22	61	SINE, MER34-Int	No	SINE	No	No	No	No	No	No	No	No	LTR	LTR, L1MEc, LTR	LINE, L1MEe	No
9p22.2	20	LINE	No	LINE, L1M4c	LINE, L1ME1	LINE, L2c	No	No	No	LINE	No	No	LINE, L1ME1	LINE, L1MA9 MER5B	No	LINE, L1ME4a
10q11.22	21	No	No	DNA	No	LINE, L1M5	No	L1M5	L1M5	L1M5	No	SINE, L1M5	LINE, L1M4C	MIR	LTR, HERVL-int L1M5	LINE
11q11	55	SINE, MIR3	No	Satelite, ALR/Alpha	LINE	No	No	No	No	No	No	No	No	L1MEc	LINE, L1M1	LTR, LTR49-Int
12p13.31	66		L2a	LINE, L1MEf	LINE, L1MB7	SINE, L2a	SINE, L2a	No	No	No	No	No	No	No	LINE, L1MA8	SINE, L1MC4a
14q21.1	47	No	No	LINE, L1MC3	No	LINE, L1ME1	LINE, L1ME1	LINE, L1ME1	L1ME1	L1ME1	LTR, L1ME1	LTR, L1ME1	No	Low Complexity	LINE, L1PA3	SINE, PRIMA4-Int, Ricksha-c
14q21.1	47	SINE, L2c	No	No	No	LINE	No	No	No	No	No	No	No	LINE, L1M4b, L1MA3	No	No

Chapter 3.2.2 – Table 3. Classes and/or family of repeat elements found in the most common breakpoint identified in each chromosome for our cohort, analyzed in and around the breakpoint site (continued).

15q11.2	20	No	L1ME2z	No	LTR, ERVL-B4-int, LTR39	SINE, L1PA3	LINE, L1PA3	LINE, L1PA3	No	LINE, L1PREC2	LINE, L1PREC2	LINE, L1PREC2	LTR, ERVL-B4-int, LTR39-int	No	No	LTR
16p11.2	20	LINE, L1MEc	SINE, L3, AluSx, AluJb	MIRb	No	No	No	No	No	No	No	No	SINE, MER33	Satelite	No	No
17q21.31	21	No	SINE	SINE, MER20	No	No	SINE	No	No	No	No	No	No	LINE, L3	SINE, L1M4, AluJb	SINE, MIR
18q23	16	LINE	No	No	No	No	No	No	No	LTR	No	No	-	-	-	-
19p12	25	No	LINE, L1MB7	LTR, LTR7	LINE, L1PREC2	LTR, AluSx, LTR2	LTR, LTR2, AluSx	No	No	LTR, HERVH-int	LTR, HERVH-int	LTR, HERVH-int	No	No	SINE, L1PA15	No
22q11.21	18	No	LINE, L2b	No	No	No	No	No	No	No	LINE	No	No	No	SINE	No
Xq21.1	34	LINE, L1MA2, L1PREC2	LINE, L1MA3	No	No	DNA Repeat elements, MER103C	No	No	No	SINE	SINE	SINE	Low Complexity	No	SINE	SINE, Charlie18a
Yp11.2	17	No	LTR, HERVL-int, HERVK22-int	No	LINE, L1M1	Satellite, ALR/Alpha	No	No	No	LINE	LINE	LINE	No	No	No	SINE, L1M4c
Yq12	17	No	No	Satelite	LINE, L1PA3	No	No	No	No	No	No	No	-	- 107	-	-

Discussion

Malignant transformation is usually considered as an accumulation of genomic alterations, which could disrupt tumor suppressor genes or activate proto-oncogenes, promoting genomic instability [Hanahan and Weinberg 2011]. Tumor genomes display several DNA breakpoints, being some of them promoters of tumor development through the disruption of functional elements, which tend to recur in tumors of the same type, suggesting common progression mechanisms [Tolosi et al. 2013]. The breakpoints that appear in random locations of the genome are proposed to have none or reduced effect on tumor progression [Tolosi et al. 2013]. We observed in our HNSCC cohort numerous non-random breakpoints, i.e. several structural rearrangements involving the same chromosomal regions were consistently present in our HNSCC patients. Since the most common rearrangement events were often clustered within multiple focal regions in the chromosomes 12, 8, 3 and 11 (Figure 2), we can hypothesize that this damage so highly localized in these clinical heterogeneous HNSCC patients can represent a non-random event. The importance of these chromosomes in HNSCC tumorigenesis has been well documented [Ribeiro et al. 2014b; Ribeiro et al. 2014a]. Interestingly, in our cohort the chromosome that presented breaks in a higher number of patients (chromosome 12) is different from the chromosome that exhibited higher number of breaks (chromosome 6), and chromosomes 3 and 8 have presented both breaks in a higher number of patients and exhibited higher number of breaks (Figures 1, 2), supporting their importance for HNSCC carcinogenesis. Genomic instability dependent on chromosomal breakage events seem to be non-random in HNSCC, targeting some chromosomes clearly more than others.

In our cohort, chromosome 12 showed to have the highest number of HNSCC patients with breaks, especially in 12p13.31, this fact could be related to its unique structural features, namely the abundant segmental duplications in this chromosome that represents 2.66% out 5.37% for the entire genome, with particular activity in the peri-centromeric region of the p arm, and the telomeres [Scherer et al. 2006]. In relation to noncoding RNAs, LINEs and SINEs distribution, CpG island distribution, and overall GC content, this chromosome is considered typical; however, chromosome 12 is rich in disease-associated loci and several genes linked to cancer are mapped [Scherer et al. 2006].
Chromosome 6 is the one that showed the highest density of breaks in different sites, which might be related with its repeat content that represented almost 44%, where

LINEs, SINEs, LTRs and DNA transposons represent almost 21%, 11%, 8% and 3% of the chromosome sequence, respectively [Mungall et al. 2003]. Additionally, several tandemly arranged, adjoining Alu repeat fragments in 6pter are indicative of a complex series of duplications, where several disease phenotypes involving chromosome rearrangements at 6pter, which include neoplasms, were also identified [Mungall et al. 2003].

Reconstructions of ancestral karyotypes proposed that evolutionary breakpoints are clustered in specific chromosomal regions that are prone to breakage and reorganize, showing the dynamic nature of mammalian genome organization [Hinsch and Hannenhalli 2006; Murphy et al. 2005; Peng et al. 2006; Ruiz-Herrera and Robinson 2007]. Similarly, DNA breakage is suggested as an initiating event of tumor genomic instability, having regions in the genome especially susceptible to breakage: the "fragile sites". Fragile sites can be caused by environmental and chemical agents, such as alcohol and cigarette smoke [Richards 2001], which are typical risk factors shared by the majority of our HNSCC patients. Moreover, repeat elements are implicated in evolutionary breakpoints and also in the genomic instability that cause some human genomic disorders [Ruiz-Herrera and Robinson 2008]. We showed in our HNSCC cohort that in less than 14% of the breakpoints site low copy repeat elements are located. However, we observed in the vicinity of the breakpoint region an increased frequency of repeat elements (Figures 4, 5). As expected according to the overall distribution in the human genome, also in our cohort the repeat elements more frequently found were LINEs, SINEs and LTRs (Tables 2 and 3), being L1 the most common LINE element. The most interesting finding from this study is that the frequency of repeat elements increased as we increase upstream and downstream distance from each breakpoint site, reaching the maximum of repeat elements at -/+ 100 Kb and + 10 Mb of the breakpoint site (Figure 4). It seems that the repeat elements may not be the exclusive explanation for the recurrent non-random breakpoints observed. We verified that HNSCC displayed numerous rearrangements on several chromosomes simultaneously, while showing few specific chromosomes without or with scarce aberrations. This pattern of instability on these chromosomes is not incompatible of chromothripsis, as described by Stephens [Stephens et al. 2011]. We observed in our cohort that frequently the same patient presented simultaneous breaks in 13-16 chromosomes (data not shown). Using this breakpoint analysis approach, we identified the variability of non-random breaks and consequently the genomic instability in HNSCC samples.

Identifying relevant breakpoint hotspots and characterizing their influence on tumor development and behavior could be the missing element to understand the

molecular mechanisms behind tumor progression. In our cohort of HNSCC patients, we identified frequent non-random chromosome breakpoints underlying rearrangements clustered in chromosomes 12p, 8p, 3q, 14q, 6p, 4q, Xq and 8q. Moreover, chromosomes 6, 14, 3, 8 and X exhibited a higher susceptibility to break in different sites, presenting a higher number of breaks. These results strongly suggest the importance of recurrent breakage and genomic instability in HNSCC. Evaluating the repeat DNA elements in and surrounding the breakpoints site, we verified that the percentage of these elements increased with distance of breakpoint site, presenting the maximum downstream 100 Kb, upstream 100 Kb and 10 Mb of the each breakpoint site. The great majority of the repeat sequences found were LINEs, SINEs and LTRs. Overall our results demonstrate selective, non-random genomic breaks involving several chromosomal regions. The presence of low copy repeat elements in or close to the breakpoint site may contribute to the genomic instability, but may not be the only explanation for the common rearrangement events observed in specific chromosome regions of HNSCC patients.

3.3 Transcriptomic Studies

3.3.1. Expression of *HIF3A*, *HOXA10* and *LOX* discriminates tumors of the tongue from the floor of the mouth

Ilda P. Ribeiro, Luísa Esteves, Francisco Caramelo, Francisco Marques, Leonor Barroso, Joana B. Melo, Isabel M. Carreira

Submitted

Abstract

Purpose Oral squamous cell carcinoma (OSCC) is a frequent neoplasm with an aggressive and unpredictable biological behavior and consequently unfavorable prognosis. Genome-wide mRNA expression profiling has been used for the classification of solid tumors into relatively homogeneous groups, predicting their clinical outcome. However, in clinical practice there is no biomarker to stratify these tumors, in order to predict the survival and help the management of the patients. In this study we analyzed gene expression patterns of OSCC in order to understand the molecular basis behind the variability in clinical outcomes. ***Material & Methods*** Gene expression microarray analysis in 48 OSCC patients was performed. Data analysis using BRB-array tools was conducted to identify gene expression signatures with ability to discriminate tongue from floor of the mouth tumors. ***Results*** From the 3028 differentially expressed genes in our cohort, a statistically significant gene expression signature of 10 genes able to discriminate tongue from floor of the mouth tumors was identified. *HIF3A, HOXA10* and *LOX* were differentially expressed in tongue and floor of the mouth tumors, which seems to be related to a high aggressive behavior and worse prognosis in the floor of the mouth tumors. ***Conclusion*** We identified a set of genes that seems to play a role in the behavior and prognostic prediction of the tongue and floor of the mouth tumors. This transcriptomic signature should be evaluated and validated in a larger and independent cohort, opening the door for personalized treatments of the different anatomic sites tumors.

Keywords: Gene expression signatures; Transcriptomic profiling; Outcome prediction; Oral cavity carcinomas; Prognostic biomarkers

Introduction

Head and Neck Squamous Cell Carcinoma (HNSCC) is the sixth most common cancer worldwide and includes tumors from the upper aerodigestive tract. These tumors from numerous anatomic locations exhibit different behaviors, namely drug response and survival that are not predictable by histopathology of the primary tumor [Belbin et al. 2008; Severino et al. 2008]. Oral squamous cell carcinoma (OSCC) and oropharyngeal squamous cell carcinoma are the most common tumors of HNSCC, being associated with high mortality and morbidity [Bose et al. 2013]. Despite improvements in nonselective treatments and the development and approval of targeted therapies, approximately 50% of these patients will experience local/regional recurrence or distant metastases and die in the following 5-years after diagnosis [Marta et al. 2015; Pulte and Brenner 2010]. Presently, the clinical decisions regarding therapy modalities are usually based upon clinical staging, which relies on nodal status and tumor size [Chung et al. 2004]. The development and progression of OSCC is a multistep process that comprises genomic instability and several molecular alterations. The understanding of the carcinogenesis process and the molecular mechanisms behind the different tumor behavior will help in the clinical management of this aggressive neoplasm. Nowadays, there is no biomarker used clinically to stratify OSCC in order to predict and improve the survival outcomes of these patients [Mendez et al. 2009]. Gene expression analysis seems to be useful for the classification of HNSCC into relatively homogeneous subtypes with a potential role in the prediction of clinical outcome [Belbin et al. 2002; Belbin et al. 2008; Chung et al. 2004; Roepman et al. 2006; Severino et al. 2008]; however, its clinical applicability needs to be further explored and confirmatory studies are required.

We performed genome-wide expression analysis in 48 OSCC patients and identified a gene expression signature of 10 genes that discriminate tumors of tongue from those of the floor of the mouth. Our findings highlight the molecular heterogeneity in the different anatomic sites of primary OSCCs, showing that the tumor outcome and survival could be influenced by the tumor anatomic site and consequently its different gene expression signature could be used to predict the behavior of disease and to select the best therapeutic strategy.

Material and methods

Study population

The study protocol was approved by the Committee on Ethics in Research of the Faculty of Medicine of the University of Coimbra. All patients provided their written consent to participate in the study after being informed about the research purposes, following the regulations in the Declaration of Helsinki.

The study cohort includes tissue specimens from 48 OSCC male patients (34 tongue and 14 floor of the month tumors). The mean age of the patients was 61 years old, range from 37 to 84. Patients were recruited between January 2011 and August 2015 from the Maxillofacial Surgery Unit of the Coimbra Hospital and University Centre, CHUC, EPE, Portugal. Diagnosis and staging were performed in accordance with the American Joint Committee on Cancer TNM staging system. The participants in this study answered a survey regarding lifestyle and risk factors for upper aerodigestive tract malignancies, including alcohol and tobacco consumption. Patients were followed-up through hospital revisits during routine clinical appointments. The final date of follow-up was February 29, 2016. The follow-up periods ranged from 6 to 61 months. Details of our study cohort are listed in table 1.

After resection, the tumor tissue samples were stored in RNAlater (Sigma Aldrich, USA) at -20ºC until use.

RNA extraction and gene expression microarray technology

Total RNA was purified from each tumor tissue sample using the Qiagen RNeasy Mini Kit according to the manufacturer's protocol (Qiagen, Valencia, CA). The integrity of the RNA was determined using the RNA 6000 Nano LabChip Kit in an Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Only samples with RNA integrity number (RIN) scores ≥ 8 were used, and were then quantified using a NanoDrop 1000 Spectrophotometer (Thermo Scientific, USA).

Labeling of RNA was performed using Low Imput Quick Amp Labeling Kit, One Color (Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's instructions. 100 ng of total RNA was amplified and labeled with Cy3 dye. This amplification product was measured for quantity and dye incorporation using the Nanodrop 1000 and then hybridized to Agilent SurePrint G3 Human Genome Expression microarray V2, 8x60K (Agilent Technologies, Santa Clara, CA, USA). All

microarray hybridizations were performed according to the manufacturer's instructions in the One-Color Microarray-Based Gene Expression Analysis manual. The fluorescently labeled RNA was hybridized to the microarray at 65 °C in a rotating oven (10 rpm). After 17 hours, the arrays were washed consecutively in gene expression wash buffers. Fluorescence intensities were measured with an Agilent microarray scanner and extracted using Agilent Feature Extraction 10.7.3.1.

Chapter 3.3.1 - Table 1. Clinic-pathologic characteristics of study population.

Patients (n = 48)			
	n (%)		n (%)
Gender		Age at diagnosis (Years)	
Male	48 (100)	**<60**	25 (52.1)
Female	0 (0)	**≥60**	23 (47.9)
Anatomic site		Invasion peri(neural)	
Tongue	34 (70.8)	**Yes**	21(43.8)
Floor of the mouth	14 (29.2)	**No**	21 (43.8)
Tobacco		NA	6 (12.5)
Yes	40 (83.3)	Differentiation	
No	8 (16.7)	**Well**	35 (72.9)
Alcohol		**Moderate**	11 (22.9)
Yes	39 (81.3)	**Poor**	2 (4.2)
No	8 (16.7)	**NA**	0 (0)
NA	1 (2.1)	Margins status	
TNM stage		**R0**	31 (64.6)
I	8 (16.7)	**R1**	10 (20.8)
II	16 (33.3)	**NA**	7 (14.6)
III	13 (27.1)	HPV	
IV	11 (22.9)	**Positive**	0 (0)
Treatment		**Negative**	42 (87.5)
surgery alone	15 (31.3)	**NA**	6 (12.5)
Surgery + RT	21 (43.8)	Vital status	
Surgery + RT + QT	6 (12.5)	**Relapses/Metastasis in follow-up**	16 (33.3)
Surgery + QT	1 (2.1)	**Dead - OSCC**	11 (22.9)
RT + QT	4 (8.3)	**Dead-non-OSCC**	2 (4.2)
RT alone	1 (2.1)	**Alive**	35 (72.9)

NA- Not Available; **RT -** Radiotherapy; **QT –** Chemotherapy

Statistical Analysis of microarray data

The raw gene expression dataset were input into BRB-array tools (Version 4.5.1 Stable Release), transformed to log2 format and normalized by providing a set of 11 housekeeping genes selected according to [Eisenberg and Levanon 2013].

The normalized dataset was then screened with a minimum fold change filter: if under 20% of the expression values did not have, at least, a minimum fold change of 3.0 in either direction from the median value of the expression for that gene, the gene would be excluded. Additionally, it was established that if a gene had over 50% of missing

values it would also be excluded from analysis. When a set of probes corresponded to the same gene, only the most variable probe determined by Inter- Quartile Range (IQR), across all the arrays, would remain. After this filtering step, a dataset of 3028 genes was generated.

Class Comparison

BRB ArrayTools offers an option to compare two predefined classes using univariate permutation tests. The compared classes were the location of the tumor: floor of mouth vs. tongue. BRB ArrayTools provided a list of the genes that best differentiated between the classes sorted by p-value of the permutation test ($p<0.001$).

Survival Analysis

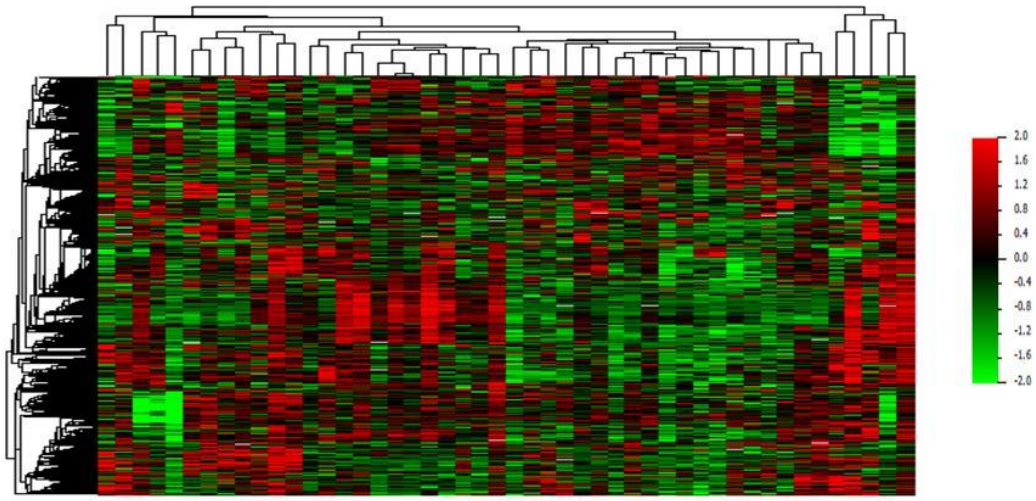
The survival analysis was performed considering the survival time (in months) for every patient and their vital status. When the subject was still alive or the outcome (alive or dead) was unknown, the latest follow-up date was used as reference for the survival time and the patients were censured. Resorting to IBM SPSS, a Kaplan-Meier method was used to obtain both the survival curves considering the location of the tumors (floor of the mouth or tongue) and the median, in months, of the survival time for both groups of patients.

Results

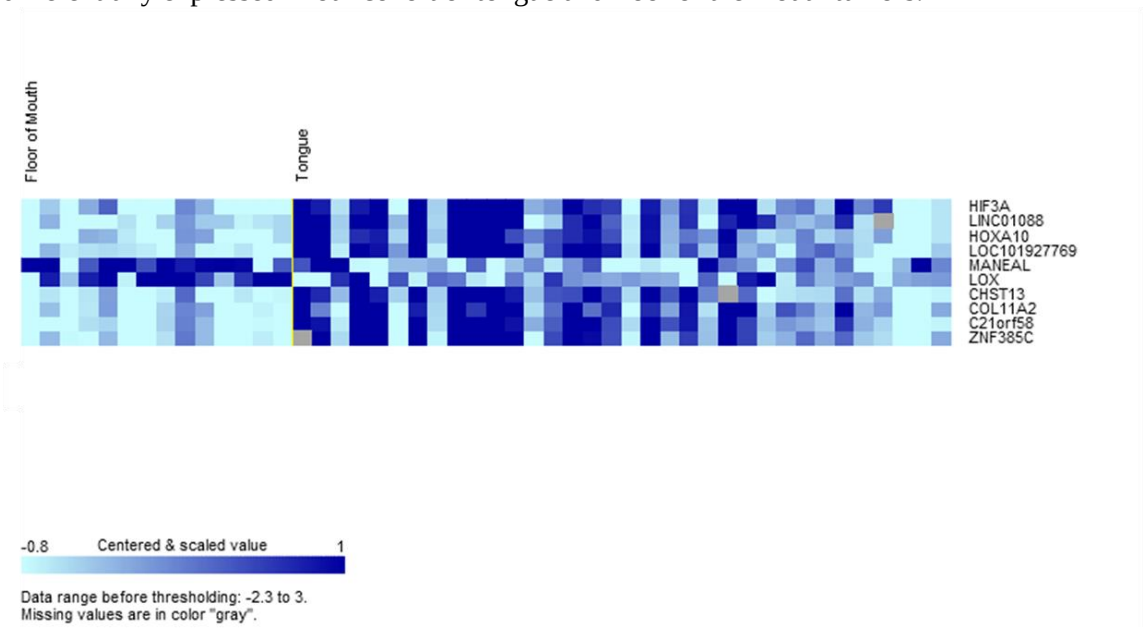
The high-throughput gene-expression microarray results from 48 male OSCC patients (34 tongue and 14 floor of the mouth tumors) exhibited, after a normalization step, a dataset of 3028 genes with deregulated expression (Figure 1).

Using class comparison of the BRB ArrayTools we identified a gene expression signature that molecularly discriminated tongue from floor of the mouth tumors (Figure 2).

This signature that best differentiated between tongue and floor of the mouth tumors contains 10 genes whose aberrant expression was meaningful when we compared the two tumor origins ($p<0.001$) (Table 2).



Chapter 3.3.1 - Figure 1. Cluster heatmap of gene expression profile with the 3028 genes differentially expressed in our cohort of tongue and floor of the mouth tumors.

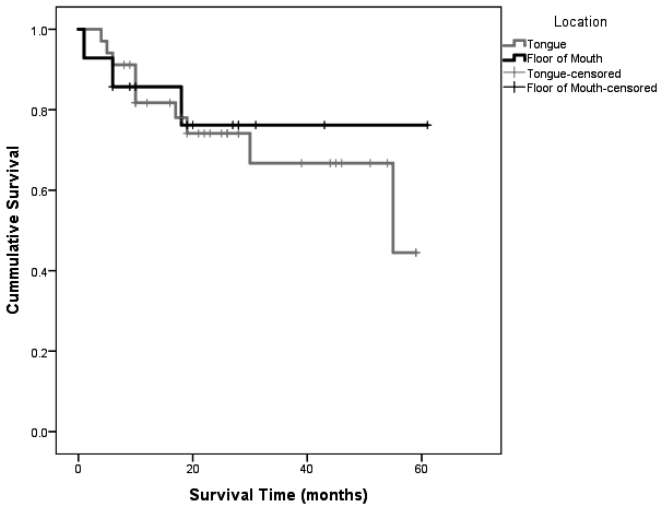


Chapter 3.3.1 - Figure 2. Heatmap of gene expression signature that molecularly discriminate tongue from floor of the mouth tumors, $p<0.001$.

Considering the survival rate, Kaplan-Meier curves for the patients with tongue and floor of the mouth tumors showed that the average survival in tongue tumors was 43.4 months (CI95% [35.5; 51.5]) and in floor of the mouth tumors was 48.7 months (CI95% [36.4; 61.0]) (Figure 3). We did not reach statistical significance for the survival differences (Log-Rank: $p = 0.728$).

Chapter 3.3.1 - Table 2. List of the genes differentially expressed between the tongue and floor of the mouth tumors, ordered by the parametric p-value

Parametric p-value	Gene Symbol	Gene Name
8.83E-05	*C21orf58*	chromosome 21 open reading frame 58
0.0002006	*LOC101927769*	uncharacterized LOC101927769
0.0003781	*LINC01088*	long intergenic non-protein coding RNA 1088
0.0004332	*HIF3A*	hypoxia inducible factor 3, alpha subunit
0.0005487	*CHST13*	carbohydrate (chondroitin 4) sulfotransferase 13
0.0008539	*COL11A2*	collagen, type XI, alpha 2
0.0009333	*HOXA10*	homeobox A10
0.0009637	*ZNF385C*	zinc finger protein 385C
0.0009655	*MANEAL*	mannosidase, endo-alpha-like
0.000992	*LOX*	lysyl oxidase



Chapter 3.3.1 - Figure 3. Kaplan-Meier curves for survival of our cohort of tongue and floor of the mouth tumors.

Discussion

We identified an OSCC anatomic site-specific gene expression signature that seems to capture the transcriptional transition of tumors from the tongue to those of the floor of the mouth, which could eventually justify the differences in the tumor behavior, prognosis and consequently in the therapy response. This gene expression signature with ten genes might help not only in the understanding of the tongue and floor of the mouth carcinogenesis process, but also in the prediction of the aggressiveness and survival rate of these patients. Among the 10 genes of the gene expression signature that discriminate the anatomic origin of both oral tumors (Figure 2), *LOX*, lysyl oxidase, mapped at 5q23.1, was up-regulated in our tumors from floor of the mouth comparatively to those from tongue. The covalent cross-link of collagens and/or elastin in the extracellular matrix (ECM) is the main function of this gene [Kagan and Li 2003]. *LOX*

expression is regulated by hypoxia-inducible factor (HIF) and therefore is associated with hypoxia in human breast and head and neck tumors [Erler et al. 2006]. It has been demonstrated that *LOX* could be a player in the metastatic tumor cell growth, since *LOX* is not only required for focal adhesion kinase (FAK), a mediator of cell proliferation and survival, but also could regulate fibronectin activity through FAK activation, providing an advantageous environment to metastatic disease development [Kaplan et al. 2005]. Thus, *LOX* is pointed out as vital for hypoxia-induced metastasis and a good therapeutic target for preventing and treating metastases [Erler et al. 2006]. In our cohort, the up-regulation of this gene could be related to a more aggressive behavior of floor of the mouth tumors in comparison with tongue tumors. This gene could be a biomarker of prognosis and survival in these patients as well as guide the development of specific treatment strategies. Another gene related to hypoxia and differently expressed in our two groups of tumors was *HIF3A*, hypoxia inducible factor 3 alpha subunit, mapped at 19q13.32. Down-regulation of this gene in tumors of floor of the mouth comparatively to those of tongue seems also be related with a more aggressive phenotype in the floor of the mouth tumors. HIF is a heterodimeric complex composed by three HIF-α genes, whereas HIF-1α and -2α are related to cellular and systemic adaptation to hypoxia, little is known about the regulation and function of HIF-3α [Tanaka et al. 2009]. In renal carcinomas HIF-3α down-regulation seems to enhance hypoxic gene induction and consequently promoting tumorigenesis [Maynard et al. 2005].

Down-regulation of *HOXA10*, homeobox A10, mapped at 7p15.2, was also observed in floor of the mouth comparatively to tongue tumors. HOX genes are members of the superfamily of homeobox genes that encode transcription factors associated with control of cell growth, cell-cell and cell-extracellular matrix interactions [Bei et al. 2007; Yoshida et al. 2006]. Deregulated expression of HOX genes has been described in several human malignancies including OSCCs [Bitu et al. 2012; Liborio-Kimura et al. 2015; Yamatoji et al. 2010]. High levels of HOXA10 protein expression in OSCCs were significantly correlated with advanced TNM stage, and HOXA10 up-regulation was linked to worse overall survival rate, indicating a higher aggressiveness and progression propensity of these tumors [Yamatoji et al. 2010]. Overall, it has been demonstrated that HOXA10 modulates important cellular events for the development and progression of OSCCs, such as those related to metastases, namely adhesion, epithelial–mesenchymal transition, migration and invasion and that its expression may be associated with a less aggressive tumor phenotype [Carrera et al. 2015]. Thus, the expression of this gene could be a prognostic and

diagnostic marker and a potential therapeutic target. However, further studies are required since the role of this gene is still controversial in different tumor types, e.g., in glioblastoma where the up-regulation of HOX genes was linked to tumor stem-like cell phenotype and high HOXA10 protein expression was associated with resistance to chemotherapy [Murat et al. 2008].

In the present study, the identified gene expression signature where *HIF3A* and *HOXA10* are down-regulated and *LOX* is up-regulated in floor of mouth tumors, appears to play an important role in the aggressive behavior and worse survival of these tumors in comparison to tongue tumors, which open new opportunities to treat differently these tumors according to their anatomic sites and gene expression profile.

Although accordingly to the identified gene expression signature it was expected that the patients with floor of the mouth tumors exhibited a worse survival, our results of survival did not confirmed this hypothesis (Figure 3), possibly, due to the reduced cohort size and follow-up time. Likewise, we could not find a discriminative and predictive gene expression signature for the different tumor stages analyzed and for the patients that already developed metastases and relapses during this period of follow-up. Increasing this cohort and the inclusion of new anatomic sites of the oral cavity, as well as a larger clinical follow-up period of these patients will be vital to confirm our findings. Further studies to calculate the accuracy of the classification using this gene expression signature in another independent dataset is required in order to validate this set of genes as biomarkers of clinical outcome and prognosis prediction and consequently with clinical utility to patient management.

Conclusions

With this preliminary study we identified a gene expression signature with different expression of *HIF3A, HOXA10* and *LOX* in tongue and floor of the mouth tumors, which seems to be related to prognosis and clinical outcome of these tumors. Floor of the mouth tumors presented a gene expression signature linked with a high aggressive behavior. In this gene expression signature the microenvironmental hypoxia seems to have a site-specific influence in the behavior of these tumors and consequently in the survival and prognosis. A molecular heterogeneity in the oral cavity tumors was demonstrated, which could justify the clinical variability in the patients' outcome, highlighting the importance of gene expression profiling to identify biomarkers of

prognosis and guide target therapies development. The expression signature identified here should be validated in larger and independent cohorts.

Acknowledgments

Ribeiro I.P. is a recipient of a PhD fellowship (SFRH/BD/52290/2013) from the Portuguese Foundation for Science and Technology. This work was in part supported by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra) and by the Portuguese Foundation for Science and Technology (grant: UID/NEU/04539/2013).

Conflict of interest

The authors have no conflict of interest to disclose.

3.4 Clinical Case Studies

3.4.1 Genomic profile of oral squamous cell carcinomas with an adjacent leukoplakia or with an erythroleukoplakia that evolved after the treatment of primary tumor: A report of two cases

Ilda P. Ribeiro, Francisco Marques, Leonor Barroso, Joana Rodrigues, Francisco Caramelo, Joana B. Melo, Isabel M. Carreira

Molecular Medicine Reports (Accepted)

Abstract

Oral leukoplakia and erythroleukoplakia are common oral potentially malignant disorders diagnosed in the oral cavity. The specific outcome of these lesions remains to be elucidated, as their malignant transformation rate exhibits great variation. The ability to predict which of those potentially malignant lesions are likely to progress to cancer would be vital to guide their future clinical management. The present study reported two patients with tongue squamous cell carcinoma: Case study 1 was diagnosed with a simultaneous leukoplakia and case study 2 developed an erythroleukoplakia following the primary tumor treatment. Whole genome copy number alterations were analyzed using array comparative genome hybridization. The present study determined more genomic imbalances in the tissues from leukoplakia and erythroleukoplakia compared with their respective tumors. The present study also identified in tumor and potentially malignant lesions common alterations of chromosomal regions and genes, including *FBXL5*, *UGT2B15*, *UGT2B28*, *KANSL1*, *GSTT1* and *DUSP22*, being some of these typical aberrations described in oral cancer and others are linked to chemoradioresistance. Several putative genes associated with hallmarks of malignancy that may have an important role in predicting the progression of leukoplakia and erythroleukoplakia to squamous cell carcinoma, namely gains in *BNIPL*, *MCL1*, *STAG2*, *CSPP1* and *ZNRF3* genes were also identified.

Keywords: Oral leukoplakia; Oral erythroleukoplakia; Genomic imbalances; Malignant transformation; Oral carcinoma; Biomarkers

Introduction

Oral potentially malignant lesions (OPMLs) are often clinically categorized as leukoplakia or erythroplakia, with being leukoplakia the most common, accounting for 85% of all these lesions [Vigneswaran and Williams 2014; Warnakulasuriya 2000]. While oral leukoplakia is defined as a white plaque without immediate apparent cause, erythroplakia is a bright red patch, which is rarely characterized as another definitive disease [Dionne et al. 2015]. Erythroleukoplakia has a mixed red and white appearance.

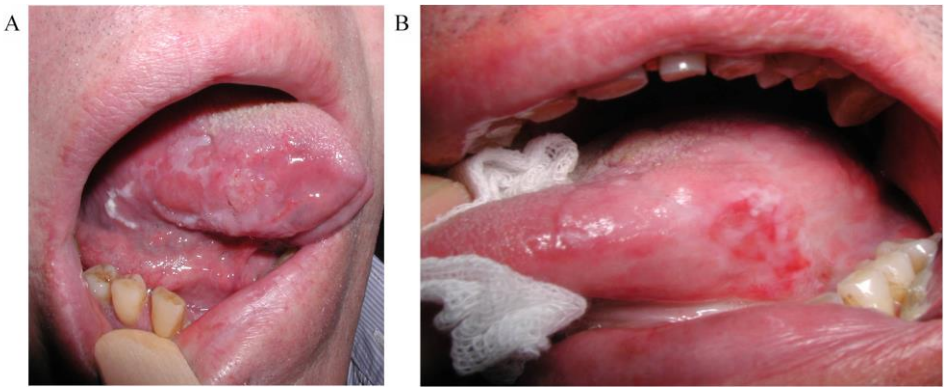
The diagnosis of these lesions is frequently made excluding known diseases or disorders lacking increased risk for cancer [Dionne et al. 2015]. These lesions precede malignant development in 0.13 to 34% oral squamous cell carcinoma (OSCC) cases [Warnakulasuriya and Ariyawardana 2016]. Histologically, leukoplakia with dysplasia is often associated with a high risk of malignant transformation [Kobayashi et al. 2012], dysplasia is currently the principal predictor of tumor development. Oral leukoplakia is more frequent in males; however, the malignant transformation is significantly higher in females [Warnakulasuriya and Ariyawardana 2016]. In addition to the presence of OPMLs being a risk factor for OSCC, its malignant transformation may be dependent on clinical, demographic, etiologic, histological and/or molecular features [William 2012]. Co-incidence of leukoplakia at the time of diagnosis of OSCC was demonstrated in up to 60% of cases [Bouquot et al. 1988; Gupta et al. 1980; Reibel 2003]. Patients with leukoplakia suffer frequently with recurrence and development of new leukoplakias after the primary treatment. The OPMLs may appear at any time, remaining stable for a considerable length of time or may progress into malignant tumors [Neville et al. 2009]. The molecular mechanism underlying malignant transformation of OPMLs remains to be elucidated and biomarkers which may predict this risk have not been identified. OPMLs have revealed several genetic alterations associated with OSCC [Califano et al. 2000; Ha and Califano 2003].

The present case report described one case of simultaneous OSCC and adjacent oral leukoplakia and another with erythroleukoplakia that evolved following treatment of primary OSCC. The patients were clinically followed for ~48 months. During this time, the patients developed local relapses of leukoplakia and erythroleukoplakia. The genomic analysis of the tumors and OPMLs allowed for the identification of some putative biomarkers of malignant transformation.

Case report

Case 1

A Caucasian 59-year-old man, drinker, heavy smoker (≥20 cigarettes/day) and negative for human papillomavirus infection, was diagnosed at the Maxillofacial Surgery Unit, Coimbra Hospital and University Centre, EPE (Coimbra, Portugal) with a simultaneous primary squamous cell carcinoma in the right side of the tongue and a leukoplakia with severe dysplasia (Figure 1A). The diagnosis was confirmed by a biopsy and the well differentiated tumor was classified as early stage (I), pT1, pN0, pMx, without compromised margins. The primary tumor and leukoplakia were simultaneously removed by surgery and the leukoplakia reached the surgical margins. Nine months after the initial diagnosis and surgery, the patient presented a leukoplakia, histopathologically classified as severe dysplasia. A local relapse of squamous carcinoma was diagnosed 28 months after the primary tumor diagnosis and total surgical excision was performed. The patient is alive and without signs of disease 48 months after the primary diagnosis.



Chapter 3.4.1 - Figure 1. A) Patient 1presents tumor and simultaneous leukoplakia in the right side of the tongue. **B)** Patient 2 was diagnosed with erythroleukoplakia after the treatment of primary tumor**.**

Case 2

A Caucasian 66-year-old man, drinker, smoker (<20 cigarettes/day) and negative for human papillomavirus infection, was diagnosed at the Maxillofacial Surgery Unit, Coimbra Hospital and University Centre, EPE, with a primary squamous cell carcinoma in the left side of the tongue (photography unavailable). The diagnosis was confirmed by a biopsy and the well-differentiated tumor was classified as early stage (II), pT2, pN0, pMx, with close but non-compromised margins. The patient underwent

glossectomy and ipsilateral neck dissection, followed by braquitherapy, 60 Gy. The patient presented an erythroleukoplakia without dysplasia 15 months after the primary tumor diagnosis (Figure 1B). A local relapse was suspected 46 months after the primary diagnosis; however, the patient refused biopsy and was lost to follow-up.

Genomic study

The present study was approved by the Ethics in Research Committee of the Faculty of Medicine of the University of Coimbra (ref: 030-CE-2015) and written informed consent from the patients was obtained. All the experiments were performed according to the regulations of Helsinki Declaration. From patient 1, samples of tumor tissue, macroscopically tumor-free tissue from surgical margins and from leukoplakia were simultaneously collected during surgery. From patient 2 samples of tumor tissue and macroscopically tumor-free tissue from surgical margins were simultaneously collected during surgery. Sample from the erythroleukoplakia that evolved after the treatment of primary tumor was obtained during the biopsy. All samples were immediately stored in liquid nitrogen until use. DNA extraction was performed using a High Pure PCR Template Preparation Kit (Roche GmbH, Mannheim, Germany) according to the manufacturer's protocol.

The genomic profile of all samples was analyzed through array Comparative Genomic Hybridization (aCGH) using Agilent SurePrint G3 Human Genome microarray 180K, (Agilent technologies, Santa Clara, CA, USA) as previously described [Pinto-Leite et al. 2014]. Reference male DNA provided by Agilent was used as control.

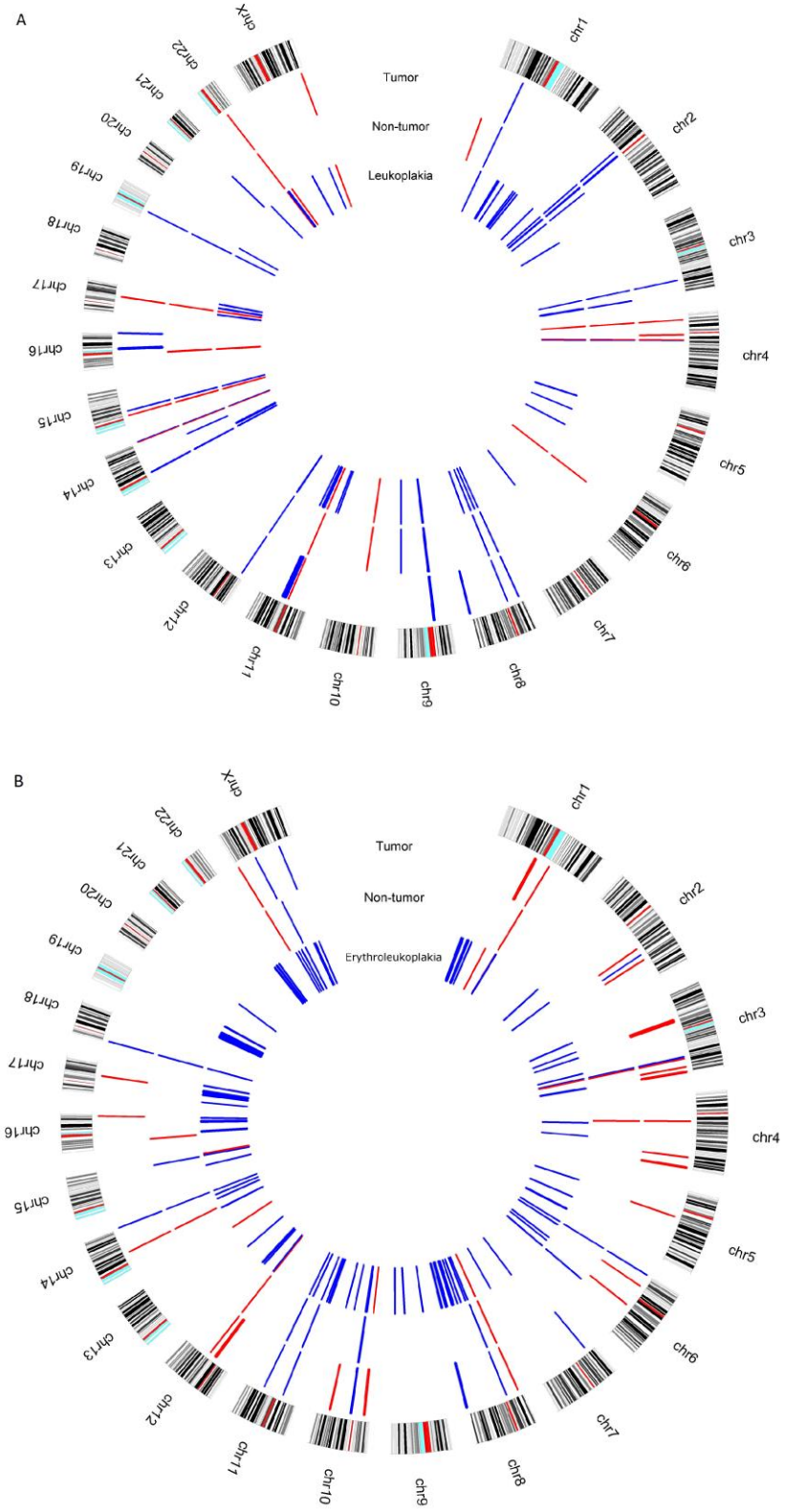
Discussion

In case 1 it was observed that samples from tumor, non-tumor and simultaneous leukoplakia shared copy number alterations (CNAs) in some chromosomal regions, gains were identified at 1p31.1; 2p12, 2p11.2-p11.1; 3q26.1; 4q13.2; 8p23.1; 8p11.22; 9p13.1p11.2; 12p13.31; 14q11.2; 14q32.33; 15q14; 19p12, and losses at 4p15.32; 4q13.2; 11q11; 14q32.33; 15q11.2; 17q21.31 and 22q11.23 (Figure 2A). In several of these regions the present study identified some relevant genes regarding their known function described in the University of California, Santa Cruz (UCSC) genome browser

(https://genome.ucsc.edu/) and their relationship with oral carcinogenesis process, including *FBXL5* (4p15.32), *UGT2B15* and *UGT2B28* (4q13.2), *KANSL1* (17q21.31) and *GSTT1* (22q11.23). Additionally, gains at 11q13.3-q13.4 were specifically observed in tumor and leukoplakia samples, which harbor the *CCND1*, *ORAOV1*, *FGF19*, *FGF4*, *FGF3*, *FADD*, *CTTN*, *FOLR3*, *INPPL1*, *ARAP1* and *P2RY2* genes. These chromosomal regions and genes have been associated with cancer, apoptosis, proliferation and OSCC [C. Jin et al. 2006; Ribeiro et al. 2014b; Ribeiro et al. 2014a]. Regarding the clinical outcome, gains at *CTTN* and *FADD* were associated with lymph node metastasis [Pattje et al. 2013; Rothschild et al. 2006] and reduced overall survival [Freier et al. 2010]. Additionally, gain in the *CCND1* was associated with poor prognosis, recurrence and metastasis [Ruiz et al. 2012]. *CCND1* amplification in OPMLs was associated with 8-fold increase in malignant transformation risk [Poh et al. 2012]. In the tumor sample a gain was observed at 16q24.3 harboring the *FANCA* gene, which is associated with locoregional progression- free survival [Bauer et al. 2008]. The presence of these CNAs in tumor or in leukoplakia may suggest poor prognosis and a recurrence/metastasis risk. Additionally, gain at the 3q29 that mapped the *PAK2* gene and loss at 6q14.1, 10q11.22 and 16p11.2 were specifically identified in the non-tumor and leukoplakia samples. The present study specifically identified some CNAs in the non-tumor sample, namely loss at 1p36.33, mapped the *CDK11B* and *CDK11A* genes and gains at 2q13, mapped *ZC3H8* gene, 14q24.3 and 21q11.2. The genomic imbalances observed in non-tumor tissue, similar to the ones observed in the tumors, may explain the local relapse observed, since these morphological normal cells presented already genomic aberrations characteristic of malignancy process; therefore, being able to suffer clonal expansion during the multistep process of malignization. This is particularly relevant for the oral carcinogenesis, since this patient has the traditional risk factors, tobacco and alcohol abuse, which may lead to local relapse disease through the field cancerization concept [Rettori et al. 2013; Slaughter et al. 1953]. It is on note that the present study observed more genomic imbalances in the leukoplakia compared with the tumor sample (Figure 2A). Chromosomes 5, 7 and 10 did not presented genomic imbalances in this tumor sample and chromosomes 13, 18 and 20 did not presented imbalances in both samples, tumor and leukoplakia.

In this case, distinct chromosomal changes and genes have been described that may have a strong prognostic potential to predict patients' outcome and tumor

transformation. However, further studies in larger cohorts are required in order to validate these findings.



Chapter 3.4.1 - Figure 2. Circus plot with aberration pattern identified by array-CGH technique for **A)** case report 1 and **B)** case report 2. Blue color represents copy number gains and red copy number losses.

In case 2 it was observed that tumor, non-tumor and the erythroleukoplakia that developed following primary tumor treatment shared CNAs in some chromosomal regions, particularly gains at 3q25.32, 8p11.22, 10q11.22, 11p12, 11q14.3, 14q32.33 and 18p11.32, and losses at 3q26.1, 6p25.3, 8p23.1, 12q14.2 and Yq11.223 (Figure 2B). Taking the known function of the genes mapped in these regions into account *DUP22* (6p25.3) may be the stronger putative candidate gene for the oral carcinogenesis process. Down-regulation of *DUP22* suggests its putative role as tumor suppressor gene; however, its function in cancer remains to be elucidated. This gene has been previously identified as a regulator of JNK signaling and with ability to dephosphorylate MAPKs [A. J. Chen et al. 2002]. *DUSP22* may also regulate metastasis, as its overexpression inhibited cell migration and reduced FAK phosphorylation [J. P. Li et al. 2010]. In the present study, losses at 1q21.2, 4q13.2, 14q21.1, Xp22.33 and Yp11.32 in tumor and non-tumor samples were observed and in non-tumor and erythroleukoplakia samples gain was observed at 6q22.1 and Xq11.1. Tumor and erythroleukoplakia samples shared copy number gains at 8q24.3, which harbored *GSDMD*, *SCRIB*, *PUF60*, *GRINA*, *SHARPIN* and *SCRT1*. The 8q24.3 chromosomal band is frequently amplified in OSCC [Ribeiro et al. 2014b; Ribeiro et al. 2014a], which may suggest its importance for malignant transformation. These genes may be important predictors of tumor transformation risk, warranting further investigation. Copy number losses in Yq11.23 were also observed in both of these samples. *MYC* (8q24.21) and *PTK2* (8q24.3) were detected only in the erythroleukoplakia sample and have already been associated with OSCC. A previous study revealed that overexpression of *MYC* was linked with malignant transformation and poor survival [Susanne M. Gollin 2001]. *PTK2* has been associated with resistance to radiotherapy [Skinner et al. 2016]. The presence of these genomic imbalances in the erythroleukoplakia, diagnosed 15 months following the primary tumor, may be the trigger for the development of relapse. There is no histological confirmation of relapse in this patient, as he refused to performed biopsy to verify the clinical suspicion. Additionally, these genes seem to be important candidates for the OSCC prognosis and especially for the prediction of the risk of relapse. The present study observed in the tumor, non-tumor and erythroleukoplakia collected from this patient, that was treated by surgery and braquitherapy several genomic alterations which were previously identified by Van den Broek and colleagues [van den Broek et al. 2007] with chemoradioresistance and some with chemoradiosensitivity (Table 1). Additionally, the present study highlighted some putative genes for these regions based in its known function described at the

UCSC, including *DUSP22* and *JARID2* (Table 1). Genomic imbalances in non-tumor tissues may indicate the presence of altered clones of cells even in the surrounding clinically and histologically normal oral mucosa, originating a progression to malignancy. Additionally, the present study detected more imbalances associated with chemoradiotherapy in erythroleukoplakia compared with primary tumor sample, which may suggest that these alterations occurred following the treatment for the primary tumor or there was a selection of radioresistant cell populations due to the treatment.

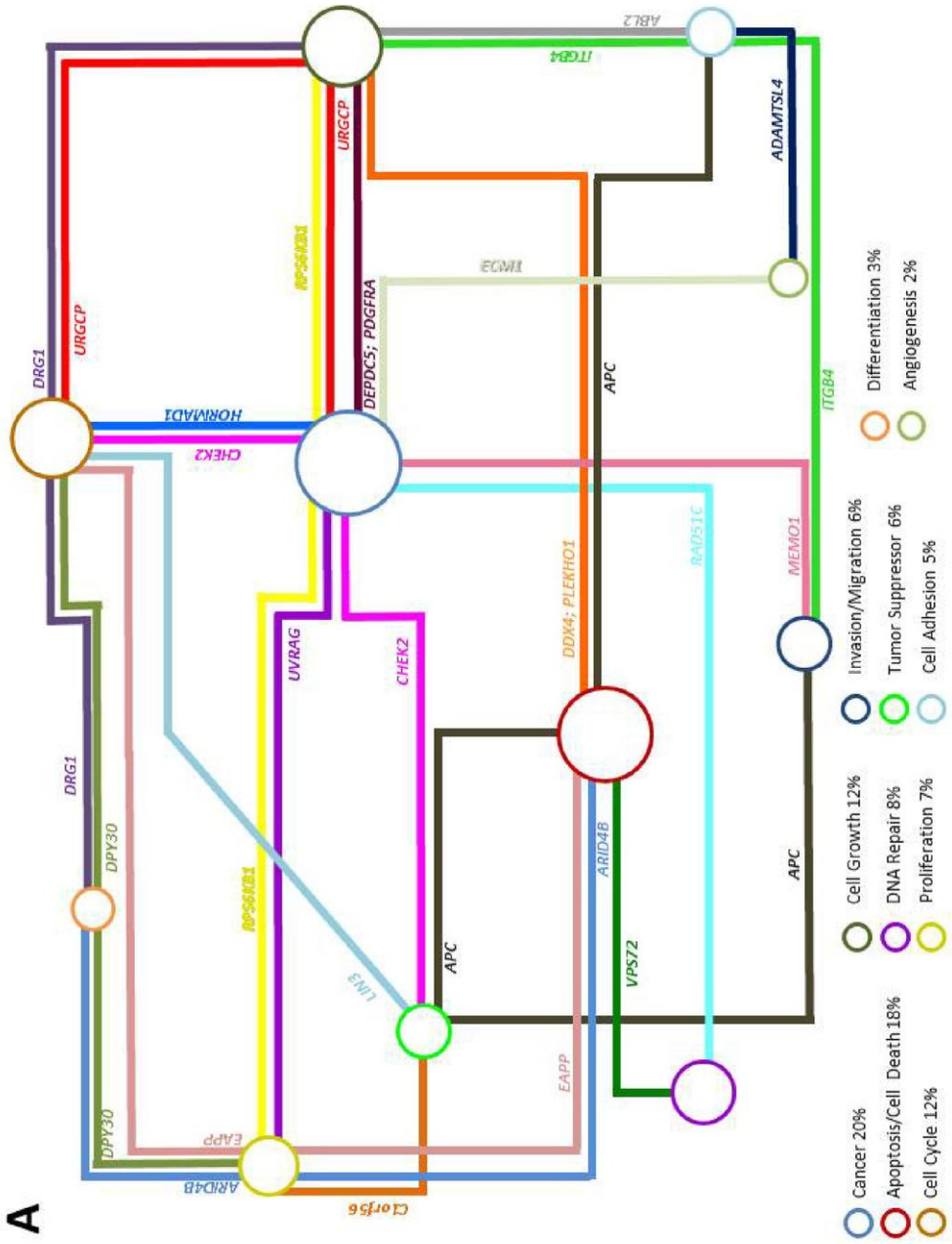
Overall, leukoplakia and erythroleukoplakia samples of the two patients presented more CNAs than the respective primary tumor. The erythroleukoplakia sample presented more CNAs than leukoplakia one (Figure 1A, B), which may be due to the fact that erythroleukoplakia is associated with significantly higher rates of dysplasia, carcinoma *in situ* and invasive carcinoma compared with leukoplakia [G. Thomas et al. 2003]. As this patient presented an erythroleukoplakia without dysplasia, it is possible that erythroleukoplakia occurred following braquitherapy, which may induce several of the genomic alterations detected. CNAs were identified in several genes in these two OPMLs, which were associated with cancer and hallmarks of the carcinogenesis process, including cell cycle, cell growth, proliferation, differentiation, angiogenesis, apoptosis/cell death, DNA repair and invasion/migration (Figure 3 A, B). The present study observed common genes altered in both samples, namely gains in *BNIPL, MCL1*, *CSPP1* and *ZNRF3* genes. These specific genes may represent important players in the malignant transformation of OPMLs into tumors, as these two patients relapsed. Further investigation is required to validate these results. Additionally, in the non-tumor sample of patient 1 and in the erythroleukoplakia sample of patient 2 a loss at 1p36.33 was identified, where the *CDK11B* and *CDK11A* genes, which are associated with cell cycle and apoptosis, are mapped.

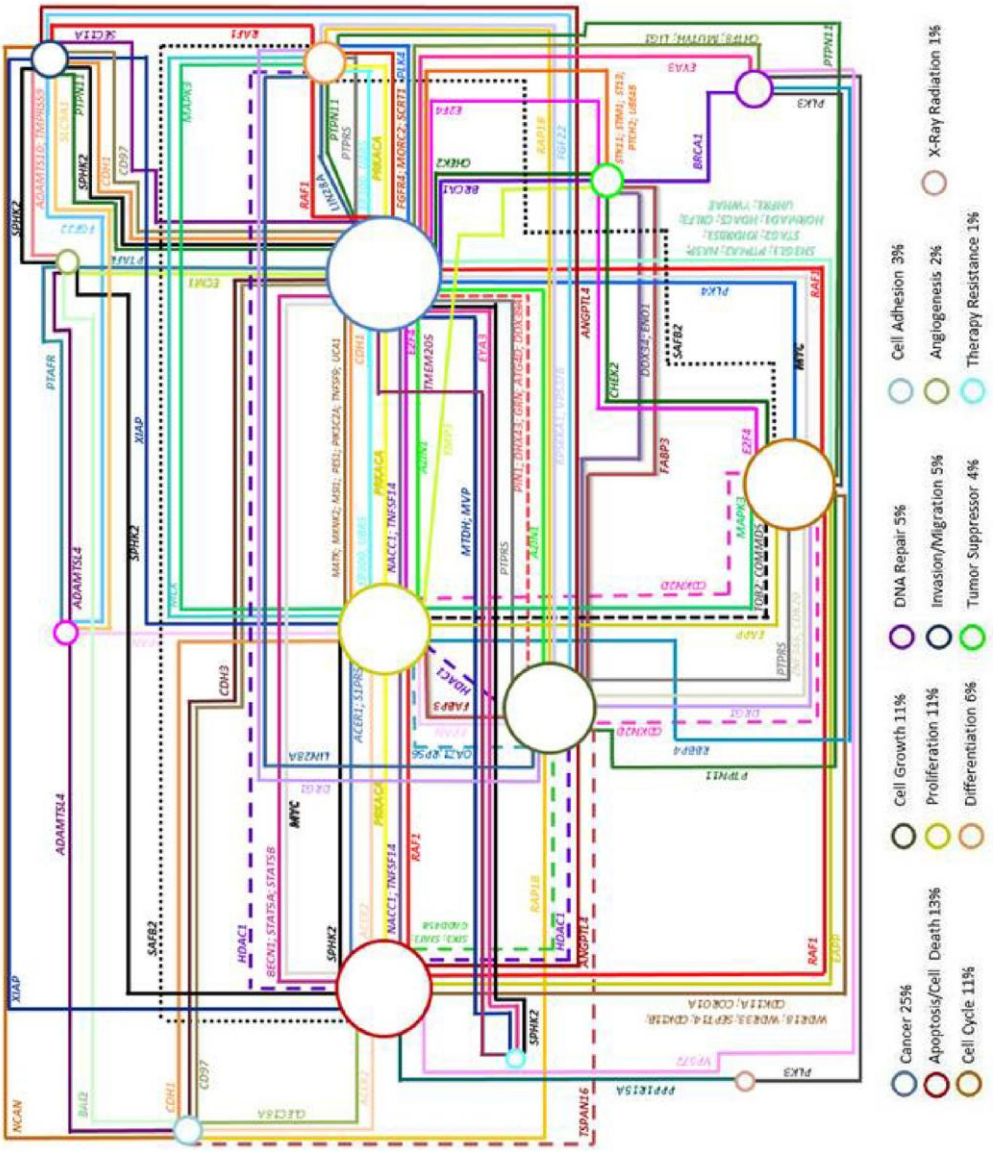
Identifying accurately and prospectively the OPMLs likely to progress to tumor is of paramount clinical significance. The present study identified the chromosomal regions and genes with CNAs in OSCCs and in OPMLs, such as *FBXL5*, *UGT2B15*, *UGT2B28, KANSL1, GSTT1* and *DUSP22* in the two patients presented. Leukoplakia and erythroleukoplakia had a high genomic heterogeneity with several genes altered, being some of those common in these two samples, specifically gains in *BNIPL, MCL1*, *STAG2*, *CSPP1* and *ZNRF3*. Therefore, the current study identified several genes that may be associated with malignant transformation. The presented study also highlighted

several putative genes that may be associated with chemoradioresistance, particularly *DUSP22* and *JARID2*.

Chapter 3.4.1 - Table 1. Chromosomal regions described in the study of Van den Broek et al as associated with chemoradioresistance and chemoradiosensitivity and some putative candidate genes for these regions identified in the patient 2.

Chromosomal Region	Type of Alteration	Clinical Association	Patient 2 - Tumor	Patient 2 - Non-Tumor	Patient 2 - Leukoplakia	Putative Candidate Genes
3q21-q26.1	**Gain**	**Chemoradioresistance**			3q25.32	
6p11-pter			6p25.3		6p25.3; 6p23 - p22.3	*DUSP22; JARID2*
6q22-q27				6q22.1	6q22.1	
Xq11-qter				Xq11.1	Xq11.1	
			Xq25		Xq25	
7p11.2-12	**Amplification**				7p11.2	*SEPT14; CHCHD2*
8p11.1-12			8p11.22	8p11.22	8p11.22	
12q15					12q15	*RAP1B; MDM2*
15q21					15q21.2	
18p11.3			18p11.32	18p11.32	18p11.32	
3p11-pter	**Loss**		3p14.2 - p14.1			*FHIT; PTPR6; ADAMTS9*
5q11-q12	**Gain**	**Chemoradiosensitivity**			5q11.2	
10q11-q22			10q11.22	10q11.22	10q11.22; 10q11.21	
14q distal			14q32.33	14q32.33	14q32.33	
14q13	**Amplification**				14q13.1 - q13.2	*EAPP*
2q22-q25	**Loss**		2q23.3 - q24.1			


A
ABL2
ITGB4
ADAMTSL4
URGCP
RPS6KB1
DEPDC5; PDGFRA
ECM1
DRG1
URGCP
HORMAD1
CHEK2
APC
ITGB4
Differentiation 3%
Angiogenesis 2%
DDX4; PLEKHO1
RAD51C
MEMO1
UVRAG
CHEK2
Invasion/Migration 6%
Tumor Suppressor 6%
Cell Adhesion 5%
DRG1
DPY30
RPS6KB1
ENT1
APC
ARID4B
VPS72
APC
Cell Growth 12%
DNA Repair 8%
Proliferation 7%
EAPP
DPY30
EAPP
ARID4B
C1orf56
Cancer 20%
Apoptosis/Cell Death 18%
Cell Cycle 12%


X-Ray Radiation 1%
Cell Adhesion 3%
Angiogenesis 2%
Therapy Resistance 1%
DNA Repair 5%
Invasion/Migration 5%
Tumor Suppressor 4%
Cell Growth 11%
Proliferation 11%
Differentiation 6%
Cancer 25%
Apoptosis/Cell Death 13%
Cell Cycle 11%

B

Chapter 3.4.1 - Figure 3. Genes associated to diverse biological processes with impact on cancer development according to UCSC genome browser. The percentage represents the number of genes associated with each biological process. **A)** Genes identified specifically in leukoplakia of patient 1. The following genes are associated only with one specific biological process: **DNA Repair:** *C11orf30; CHAF1B; GTF2H1; POLD3; RNF169.* **Apoptosis/Cell Death:** *ANO1; BIRC6; BNIPL; DAP3; MCL1; SAP30BP; SUMO2; PAK2; PTRH2.* **Cancer:** *SERPINH1; STAG2; TRIM37; WNT11; GSTT1; PPME1; PRUNE.* **Cell Cycle:** *CCNO; COPS5; CSPP1.* **Invasion/Migration:** *LLGL2; RAB25.* **Tumor Suppressor:** *ZNRF3.* **X-Ray Radiation:** *XRRA1.* **B)** Genes identified specifically in erythroleukoplakia of patient 2. The following genes are associated only with one specific biological process: **DNA Repair:** *KIAA0146; MCM4; NSMCE2; PRKDC; RNF168; SF3B3; SIRT6; UIMC1; XRCC2; XRCC6.* **Apoptosis/Cell Death:** *ATAD5; BAX; BBC3; BCL2L13; BNIPL; CARD8; DAPK3; GRINA; IFI6; MCL1; PYCARD; RERE; TAOK2; TMEM161A; TRADD; TRIAP1.* **Angiogenesis:** *CCR10.* **Cancer:** *ALDH2; CREB3L3; CTCF; DPP9; DUSP22; ELAVL1; EWSR1; FAM83A; GAL3ST1; HIGD1B; MDM2; MTSS1; NQO1; PGPEP1; PIK3CB; PRDX1; PRDX2; PVT1; RFX1; RNF139; RPS15; RPS8; S100PBP; SAFB; SELENBP1; SPIN1; WWP2;.* **Proliferation:** *ANGPTL6; CIB1; DLG1; FUT3; PLA2G1B; RASL10A; S1PR2.* **Cell Cycle:** *ARID3A; BRD4; CDK2AP1; CSPP1; FZR1; SGIP1; NAE1; NPEPPS; PPP6C;* **Invasion/Migration:** *ELMO3; MRI1; PARD6A; SCAI; ZRANB1.* **Cell Growth:** *ARHGEF18; CD37; CDIPT; DDX19B; DDX20; DDX28; PPAN-P2RY11; PTK2; SESN2.* **Tumor Suppressor:** *ADAM11; APC2; MAPKAPK5; ZNRF3.* **Tumor Growth:** *ANXA13; KLF10; PDF.* **Differentiation:** *PUM1.*

Acknowledgements

This work was in part supported by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra). Ribeiro I.P. has a PhD fellowship (SFRH/BD/52290/2013) supported by the Portuguese Foundation for Science and Technology.

Conflict of interest statement

The authors have no conflict of interest to disclose.

3.4.2 Genomic and epigenetic characterization for the comparison of synchronous bilateral tongue squamous cell carcinomas - a case report

Ilda P. Ribeiro, Leonor Barroso, Francisco Marques, André Santos, Francisco Caramelo, Maria J. Julião, Joana B. Melo, Isabel M. Carreira

Submitted

Abstract

Introduction The tongue is the most common and aggressive site for tumors in the oral cavity. These tumors are usually located in the lateral border of the tongue and are often related to the use of tobacco and alcohol. Clinical management of these tumors is predominantly based on anatomic location and TNM classification. The identification of molecular signatures with ability to explain the different outcomes observed in these patients is of paramount importance to guide and help their management. ***Case Presentation*** We herein describe an 88-year-old woman diagnosed with synchronous bilateral tongue carcinoma. This woman did not present the traditional risk factors related to oral cancer - alcohol, tobacco or presence of human papiloma virus (HPV). Both tumors were classified by a pathologist as pT2. This patient was submitted to surgery, six months later was diagnosed with cervical metastasis and in the following two months died. Copy number alterations and methylation status of these two simultaneous tumors were analyzed using array Comparative Genomic Hybridization, Multiplex Ligation-dependent Probe Amplification (MLPA) and Methylation Specific MLPA. ***Discussion and Conclusion*** In both tumors we identified several molecular traits usually found among oral cavity tumors and some of those have been associated with clinical outcome, reinforcing their importance to accurately establish biomarkers with clinical applicability. Specific genomic and epigenetic signatures for each of these two tumors were also observed allowing their molecular discrimination. The tumor of the right side of the tongue exhibited more copy number gains than the tumor of the left side. In the left side tumor less and smaller copy number alterations and more methylated genes were observed, which could be indicative of an early phase of tumor development. This case shows the molecular heterogeneity of oral cavity tumors even in the same patient and anatomic site, which could be the key to explain the different outcomes of oral tumor patients.

Keywords: Tongue squamous carcinoma; Genomic and epigenetic signatures; Methylation patterns; Bilateral tumors.

Introduction case report

Oral cavity cancer is ranked in the top 10 leading cause of deaths in males worldwide [H. C. Chen et al. 2015a]. Tongue cancer is considered the most aggressive tumor and with the worst prognosis of all oral cavity neoplasms, presenting a biological distinct entity when compared with other oral tumors [Bello et al. 2010]. Lymphatic metastases are frequently detected in tongue cancer, occurring in about 15-75% of cases, decreasing significantly the survival rate [Aksu et al. 2006]. Tongue cancer is typically diagnosed in older people, mostly males with heavy tobacco and alcohol habits [Bello et al. 2010]. Clinical management of oral cancer is predominantly based on anatomic location and TNM classification, surgical excision is the mainstay of treatment for tongue tumors. Identification of molecular signatures in oral cavity cancers that can explain the different outcomes observed in these patients is of paramount importance to guide and help clinical management. The understanding of the biological and clinical heterogeneity of these tumors, highlighting biomarkers of diagnosis and prognosis, opens new doors for the choice of treatment modalities and for the clinical routine follow up of the patients. In the present study, we report an 88-year-old woman diagnosed with synchronous bilateral tongue carcinomas. Specific genomic and epigenetic signatures for each of these two tumors were observed allowing its molecular discrimination, highlighting specific molecular signatures for the right and left tongue tumors, including putative biomarkers of tongue site-specific location and prognosis, allowing for the genetic discrimination between the two lesions.

Case report

In April 2013, a Caucasian 88-year-old woman, without the traditional risk factors related to oral cancer - alcohol, tobacco and presence of human papiloma virus (HPV) - was diagnosed at the Maxillofacial Surgery and Stomatology Units, of the Coimbra Hospital and University Centre, EPE, Portugal, with two simultaneous primary tumors, in the left and right side of the tongue (Figure 1A, B). Due to advanced age and comorbidities (heart failure, serious chronic obstructive pulmonary disease) the patient was deemed unfit for major surgery, and was not submitted to cervical lymphadenectomy. Both tumors were classified by a pathologist as pT2.

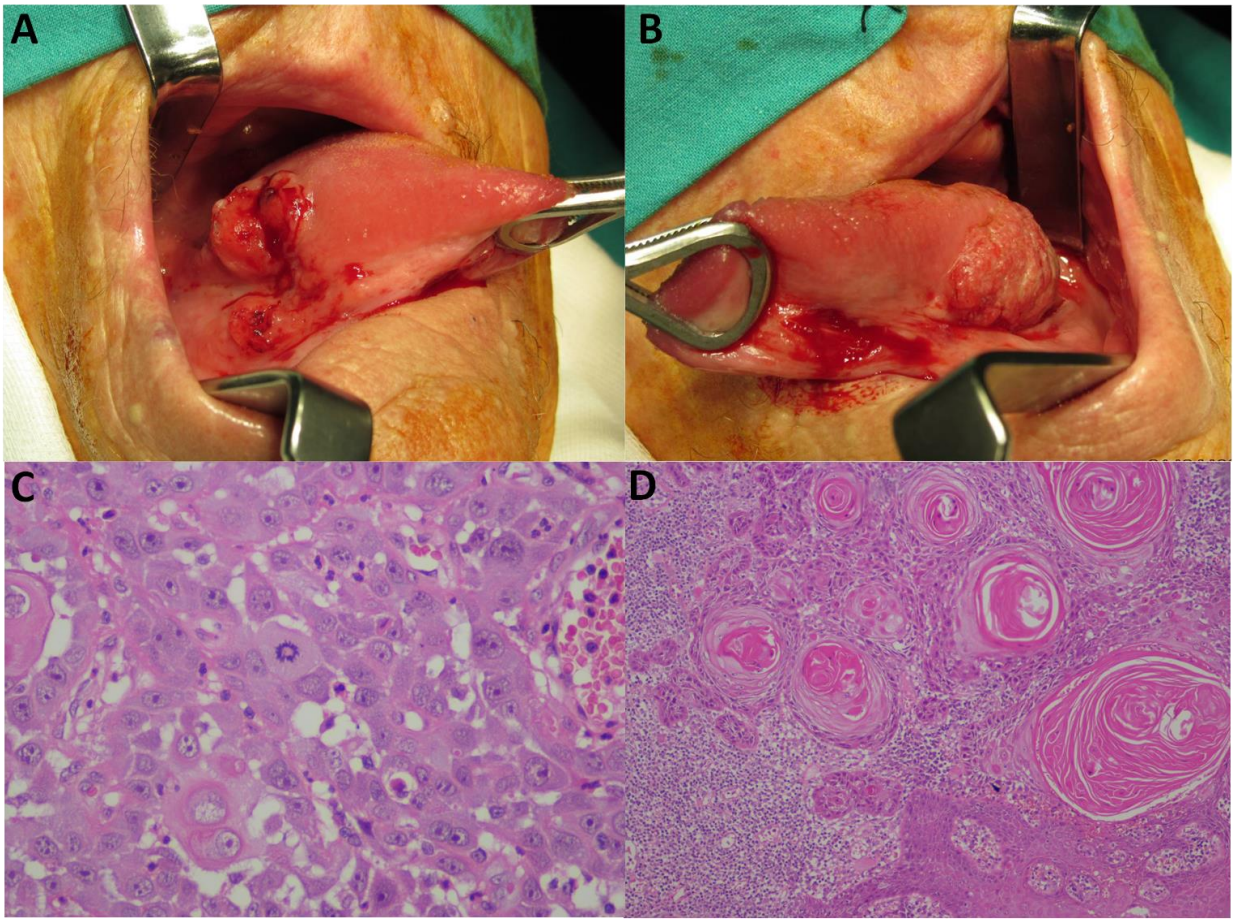
In the right tongue tumor the surgical resection margins were compromised with carcinoma *in situ*. This was a moderately/well differentiated tumor with neural invasion.

Microscopically, the hematoxylin-eosin staining demonstrated areas of keratinization and areas with pleomorphism and numerous mitoses (Figure 1C). There was inflammatory infiltrate.

The left tongue tumor was well differentiated with keratinization and exophytic growth. The histological evaluation of this tumor also showed the presence of neural invasion but the surgical resection margins were free of tumor (Figure 1D).

In spite of these adverse pathological findings (compromised margins, neural invasion), the Radiation Oncology team decided that, due to advanced age, comorbidities and general frailty, the deleterious effects of radiation therapy would be superior to its potential benefits, and the patient was not submitted to radiation therapy.

The tongue was clinically free of tumor in the subsequent consultations, including on the consultation of October of 2013, when she was diagnosed with right cervical metastasis. The patient died two months later.



Chapter 3.4.2 - Figure 1. Synchronous bilateral tongue tumors. **A)** Photography of tumor in the right side of the tongue. **B)** Photography of tumor in the left side of the tongue. **C)** Hematoxylin-eosin stains showing the morphology of tumor in the right side of tongue (x200). **D)** Hematoxylin-eosin stains showing the morphology of tumor in the left side of tongue(x200).

Genomic and epigenetic study

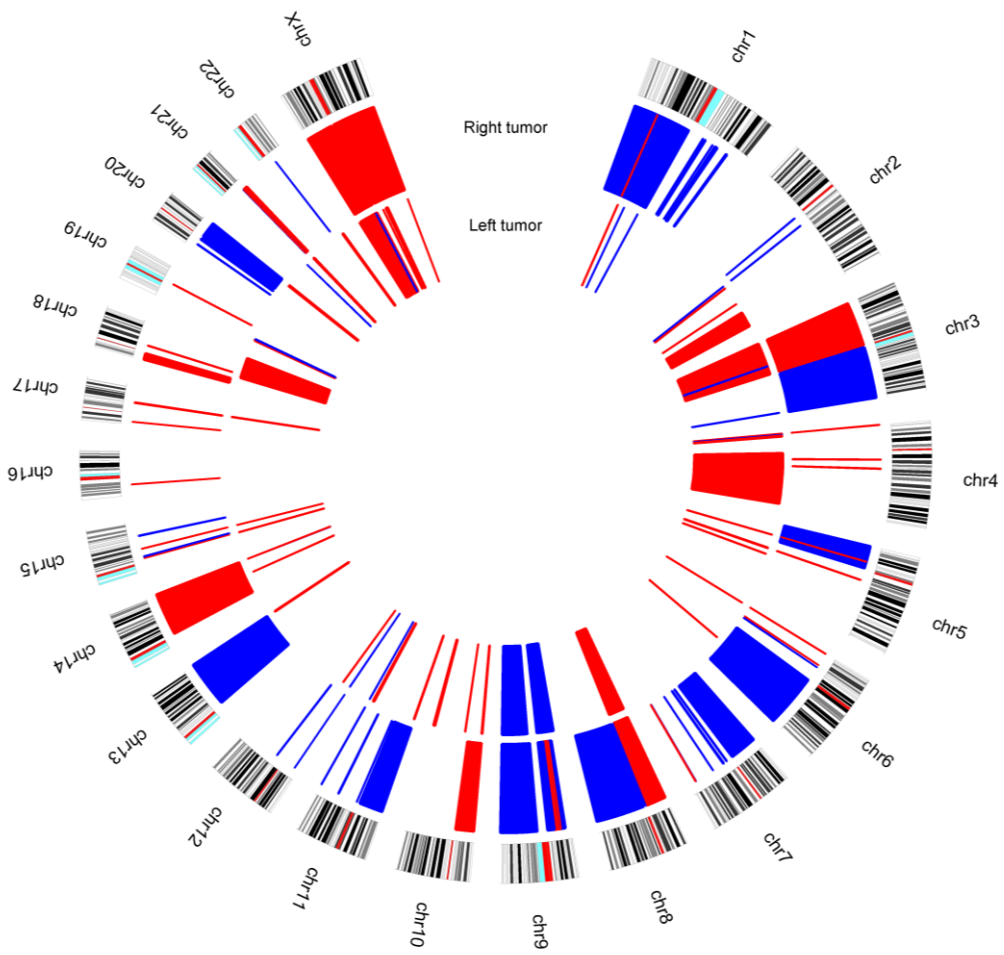
This study was approved by the Committee on Ethics in Research of the Faculty of Medicine of the University of Coimbra and we obtained written informed consent from the patient, performing all the experiments according to the regulations in the Declaration of Helsinki. Tumor tissue and macroscopically tumor-free tissue from surgical resections were obtained from both tumors. The tissue samples were snap-frozen in liquid nitrogen and stored at -80ºC until use. DNA extraction was performed using a High Pure PCR Template Preparation Kit (Roche GmbH, Mannheim, Germany), according to the manufacturer's instructions. We analyzed copy number alterations (CNAs) of both tumor samples through array Comparative Genomic Hybridization (aCGH) using Agilent SurePrint G3 Human Genome microarray 180K, (Agilent technologies, Santa Clara, CA) as we previously described [Pinto-Leite et al. 2014].

Multiplex Ligation-dependent Probe Amplification (MLPA) and Methylation Specific MLPA (MS-MLPA) using the P248 and ME002 SALSA probemixes (MRC-Holland, Amsterdam, The Netherlands) were performed in tumor and non-tumor tissue samples of both tumors in order to simultaneously evaluate the CNAs and methylation patterns in a specific set of genes as we previously described [Ribeiro et al. 2014; Ribeiro et al. 2016]. DNA from gender-matched gingival tissue was used as controls.

Results and Discussion

Whole genome approach, aCGH, revealed more imbalances and larger in size in tumor of the right side of the tongue than in the left side (Figure 2). Moreover, the tumor of the right side presented more copy number gains than the tumor of the left side. Several questions can be raised: 1) could genomic and epigenetic signatures help to discriminate these two tumors? 2) Could the oral tumor laterality be explained by a different molecular profile in each tumor? 3) Could genomic imbalances explain the synchronism of oral tumors? 4) Should molecular findings be used in clinical routine management of the oral tumor patients in order to improve survival?

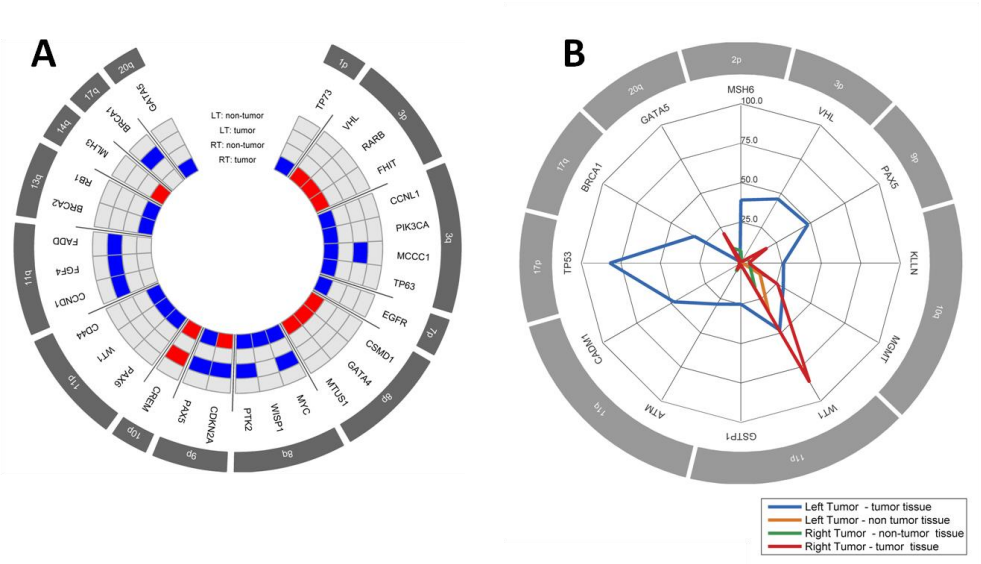
These two tongue tumors presented a specific genomic signature that allows its molecular discrimination and the identification of genes that could explain the laterality of the tongue tumors. Thus, the specific genomic profile of the right side tongue tumor was gains at chromosomes 1p, 3q, 5p, 6q, 7p, 8q, 13q, 20q11.21-q13.33 and losses at chromosomes 9p21.3-p21.1, 11p, 14q, 18p, 21q22.13-q22.2 and Xq. Similarly, the specific genomic profile of the left side tongue tumor was gain at 9p21.3-p21.1 and losses at 2q terminal, 4q, 10q terminal, 11q23.3-q24.1, 13q terminal, 18q, 19p13.1 and 22q terminal.



Chapter 3.4.2 - Figure 2. Circus plot with aberration pattern of tumors in the right and left side of the tongue. Blue represents copy number gains and red copy number losses.

These results were corroborated by MLPA and MS-MLPA techniques as depicted in figure 3A, where specific CNAs were identified in the right and left tumors. Gains at *TP73*, *CCNL1*, *PIK3CA*, *TP63*, *EGFR*, *WISP1*, *PAX6*, *WT1*, *CD44*, *BRCA2*, *RB1* and *GATA5* and losses at *VHL*, *RARB*, *FHIT*, *CSMD1*, *GATA4*, *MTUS1*, *CDKN2A* and *MLH3* were specifically detected in right tumor. Gains at *CDKN2A*, *CCND1*, *FGF4*, *FADD* and *BRCA1* were specifically detected in left tumor. Besides these genomic differences observed, in common both tumors shared losses at 3p, 8p, 21q22.3, Xp and Xq21.2-q21.32 and gain at chromosome 9, namely in 9p23-p21.3, 9p21.1-p11.2 and 9q. In the development of both tumors the chromosomes 12, 15, 16 and 17 seem to do not have a significant role since they are normal regarding CNAs. These common imbalances identified could be site-specific and in this way determine the anatomic location of the tumors in the tongue. Likewise, some of these common imbalances detected could also be the key for the simultaneous development of these tumors,

without traditional risk factors and respective cancer cancerization. Smoking is an epidemiological risk factor for squamous cell carcinoma of the tongue, but its functional impact on carcinogenesis mechanism is still poorly understood and even the carcinogenesis may differ among anatomic sites [Pickering et al. 2014]. Using MLPA and MS-MLPA, we found in both tumors gains at *MCCC1*, *MYC*, *PTK2*, *PAX5* and loss at *CREM* genes (Figure 3A), which make these genes promising targets for further studies related to tongue site-specific tumors development and also to development of tumors simultaneously. These hypotheses could be addressed in large cohorts.



Chapter 3.4.2 - Figure 3. Results from MLPA and MS-MLPA. **A)** Radar chart with copy number alterations detected by P248 and ME002 probemixes. Blue represents copy number gains and red copy number losses. **B)** Radar chart with methylation status detected by ME002 probemix.

Furthermore, the overall imbalances detected in these two tumors are in common with the ones more usually identified in oral carcinomas, i.e., the gains at 3q, 5p, 7p, 8q and losses at 3p, 8p, 9p, 11q, 18q and 21q [Susanne M. Gollin 2001; S. M. Gollin 2014], which reinforce the presence of genomic patterns in these carcinomas despite the different anatomic locations, tumor stage and clinical outcome. Regarding the presence of these patterns, the identification of specific signatures in each patient or group of similar patients could explain the clinical and biological heterogeneity of the patients and guide the treatment choices and clinical follow-up management. Thus, in both tumors we observed genomic aberrations related to presence of metastasis, poor survival and unfavorable outcome, such as loss at 3p (*SEMA3F* gene), at 8p21-22 (*NEFL* gene) and gain at 9q (Table 1).

Chapter 3.4.2 - Table 1. Several genomic imbalances identified in the both tumor samples and their clinical association described in the literature.

Chr.	Genes	Type of alteration	Clinical association	References	Right tumor	Left Tumor
3p21.31	***SEMA3F***	Loss	High metastasis and poor survival	[Doci et al. 2015]	Yes	Yes
3p22.2	***MLH1***	Loss	Early stages of disease	[Ghosh et al. 2010]	Yes	Yes
3p14.2	***FHIT***	Loss	Early Event	[C. Garnis et al. 2003]	Yes	Yes
3q26.32	***PIK3CA***	Gain	Poor prognosis	[Brauswetter et al. 2016; Suda et al. 2012]	Yes	No
3q26.33	***SOX2***	Gain	Metastasis , worse outcome, resistance to cisplatin	[Schrock et al. 2014]	Yes	No
4q32.3	***PALLD, DDX60L, MAML3***	LOH	Poor survival	[X. Chen et al. 2015b]	No	Yes
4q35.2	***ING2***	LOH	Advanced stage	[Borkosky et al. 2009]	No	Yes
4q35.2	***FAT1***	LOH	Advanced tumor stage	[Pickering et al. 2013]	No	Yes
7p11.2	***EGFR***	Gain	Poor prognostic	[Chung et al. 2004]	Yes	No
7p		Gain	Nodal metastases, poor prognosis	[Pathare et al. 2011]	Yes	
8p21.2	***NEFL***	LOH	Increased mortality risk	[Coon et al. 2004]	Yes	Yes
8p21- 22		Loss	Advanced tumor stages and poor survival	[Bockmuhl et al. 2000]	Yes	Yes
8p		Loss	Nodal metastases	[Pathare et al. 2011]	Yes	Yes
8q24.21	***MYC***	Amplification	Advanced primary tumors/ late event in the tumorigenesis	[Rodrigo et al. 1996]	Yes	No
9p21.3	***CDKN2A***	Loss	High frequency of recurrences; early event in HNSCC progression	[Danahey et al. 1999]	Yes	Gain
9p21.3	***CDKN2B***	Loss	Early event in HNSCC progression	[Worsham et al. 2006]	Yes	Gain
9q		Gain	Unfavorable outcome	[Pathare et al. 2011]	Yes	Yes
18q21.2	***SMAD4***	Loss	Advanced stage and poor prognosis	[Bornstein et al. 2009]	No	Yes
18q		LOH	Advanced stage, aggressive tumor behavior	[Takebayashi et al. 2004]		Yes
22q13.2	***CYB5R3***	Loss	Worse prognosis, decreased survival	[Reis et al. 2002]	No	Yes

These genes could be important biomarkers of prognosis; however, more studies are needed. This predictive knowledge could change the clinical management of the patients increasing their survival and quality of life.

The left side tumor was not only the tumor that presented less and smaller CNAs but also the one that presented more gene promoter methylation (Figures 2, 3B). Thus, we found more methylated genes in the left than in the right tongue tumor (Figure 3B). Despite, both tumors have been diagnosed at the same time, we could hypothesize that the left side tumor of the tongue was in an early phase of development comparatively to the tumor in the right side. This assumption was based in the fact that gene promoter methylation is considered as an early event for solid tumors development [Maruya et al. 2004] and also that malignant cells will be progressively acquired CNAs during the carcinogenesis process. Thus, the left tumor presented more methylated genes and the right tumor exhibited larger CNAs. In both tumor samples methylation of *MGMT* and *WT1* were found (Figure 3B). *WT1* gene promoter methylation was also found in the non-tumor sample belonging to left tongue tumor. In our previous work [Ribeiro et al. 2016]

we showed that *WT1* gene promoter methylation was associated with early tumor stage, which is in agreement with this reported patient that had two tumors with pT2. *WT1* gene promoter methylation in non-tumor sample could be indicative of the presence of morphologically normal cells already with some molecular features of malignancy.

MGMT gene promoter methylation was significantly correlated with increased tumor recurrences and decreased patient survival, independent of other risk factors and chemoradiation therapy [Zuo et al. 2004]. Our patient exhibited *MGMT* methylated, developed metastases and died in a short period of time, thus, this gene could be explored as a poor prognosis biomarker with clinical application.

Moreover, we found that none of the genes evaluated exhibited both copy number loss and methylation. In contrast, we observed copy number gain and methylation in *WT1* and *GATA5* in the right side tumor and in *PAX5* and *BRCA1* in the left side tumor. We reported [Ribeiro et al. 2016] that when *PAX5* gene is methylated the tongue is the most likely site of tumor origin. However, considering that in the reported patient, *PAX5* gene promoter methylation was only found in the left tongue tumor, could be interesting to verify if there are differences in the methylation pattern of this gene regarding the left and right side of tongue. This could be an important gene to explain the differences in tongue tumor laterality.

Contrarily to what could be expected, the non-tumor sample of the right tumor, which belonged to a compromised resection margin, did not present any genomic or epigenetic alterations. Explanations for that could be the sampling of a portion of the margin truly non-compromised, the presence of alterations in other genes rather than those of the probemixes used or a sample with higher normal cells admixture, which diluted the genomic and epigenetic alterations and made its detection impossible with the molecular techniques used.

Conclusions

The analysis of the patient here reported highlighted genomic and epigenetic differences in synchronous tumors of the right and left side of the tongue. We identified a specific genomic and epigenetic signature for each of these tumors, which could be associated with the laterality of these tumors, the synchronism and eventually the tongue site-specific anatomic site. CNAs and methylation patterns identified in the tumors of this patient have been associated with clinical outcome, reinforcing their

importance to accurately establish diagnostic, prognostic and, ultimately patient selection biomarkers potentially useful to identify the optimal treatment and clinical management. Examples of such potential biomarkers are: *CREM*, *GATA5*, *MCCC1*, *MGMT*, *MYC*, *NEFL*, *PAX5*, *PTK2*, *SEMA3F* and *WT1*. More studies addressing the role of these genes in oral cancer are needed. This case report also shows the molecular heterogeneity of these tumors even in the same patient and anatomic site.

Acknowledgments

Ribeiro I.P. is a recipient of a PhD fellowship (SFRH/BD/52290/2013) from the Portuguese Foundation for Science and Technology. This work was in part supported by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra) and by the Portuguese Foundation for Science and Technology (grant: UID/NEU/04539/2013).

Conflict of interest

The authors have no conflict of interest to disclose.

3.4.3 Genetic and epigenetic characterization of the tumors in a patient with a tongue primary tumor, a recurrence and a pharyngoesophageal second primary tumor

Ilda P. Ribeiro, Francisco Marques, Leonor Barroso, Jorge Miguéis, Francisco Caramelo, André Santos, Maria J. Julião, Joana B. Melo, Isabel M. Carreira

Molecular Cytogenetics
Volume 10, Issue 13, Pages 2-9 (2017)
doi: 10.1186/s13039-017-0310-z

Abstract

Background The choice of therapeutic modality for oral carcinoma in recurrent or second primary tumors remains controversial, as the treatment modalities available might be reduced by the treatment of the first tumor, and the overall survival is lower when compared with patients with a single or first tumor. Identifying biomarkers that predict the risk of relapse and the response to treatment is an emerging clinical issue. ***Case presentation*** A Caucasian 49-year-old man was treated with chemotherapy followed by chemoradiotherapy for a primary left side tongue tumor, achieving a complete response. After 49-months of follow-up, a local recurrence was diagnosed. After three months, a second primary tumor at the pharyngoesophageal region was detected. Genomic and epigenetic characterization of these three tumors was performed using array Comparative Genomic Hybridization, Multiplex Ligation-dependent Probe Amplification (MLPA) and Methylation Specific MLPA. ***Results*** The three tumors of this patient shared several imbalances in all chromosomes excluding chromosomes 9, 20 and 22, where genes related to important functional mechanisms of tumorigenesis are mapped. The shared genomic imbalances, such as losses at 1p, 2p, 3p, 4q, 5q, 6q, 7q, 8p, 10p, 11q, 12p, 12q, 13q, 15q, 16p, 16q, 17p, 17q, 18q, 19p, 19q, 21q and Xp and gains at 3q, 7q, 14q and 15q showed a common clonal origin for the diagnosed relapses. We identified some chromosomal imbalances and genes mapped in the chromosomes 2, 3, 4, 6, 7, 11, 14, 17, 18 and 22 as putative linked to chemoradioresistance and chemoradiosensitivity. We also observed that gains in short arm of chromosomes 6, 7, 8 and 18 were acquired after treatment of the primary tumor. We identified losses of *VHL* gene and promoter methylation of *WT1* and *GATA5* genes, as predictors of relapses. ***Conclusions*** A common clonal origin for the diagnosed relapses was observed and we identified some putative candidate biomarkers of prognosis, relapse risk and treatment response that could guide the development of management strategies for these patients.

Keywords: Recurrence, Second primary tumor, Genetic and epigenetic profile, Oral cancer, Chemoradioresistance

Background

Oral squamous cell carcinoma (OSCC) is the most common malignant tumor of the head and neck [Bagan and Scully 2008]. These tumors are associated with high morbidity and mortality and their incidence is increasing in the younger population [Cabanillas et al. 2013; Markopoulos 2012]. Considering all tumors that arise in the head and neck region, tongue tumors are among the worst in terms of prognosis [Sano and Myers 2007]. Treatment of OSCC is predominantly based on tumor location and TNM classification and includes surgery, radiotherapy and chemotherapy, either individually or in combination [van Kempen et al. 2015]. These treatment modalities do not benefit patients equally and are often associated with side effects that reduce compliance and prevent timely completion of therapy [Bossi et al. 2013]. OSCC survivors have a high risk of developing relapses (tumor recurrences or second primary tumors (SPT)) and also distant metastasis, which leads to treatment failure and hampers the overall survival [Gonzalez-Garcia et al. 2009]. The 5-year survival rate and disease-free survival of OSCC patients are negatively affected by the presence of recurrences, which lead to a poor prognosis and a poor quality of life. Since local recurrence and treatment resistance are the major obstacles in achieving a cure in this neoplasm, the identification of molecular markers to predict the risk of relapse development and the response to the treatment is important in the management of these patients. We report a Caucasian 49-year-old man diagnosed with a primary squamous cell carcinoma in the left side of the tongue. 49-months after the completion of treatment a local recurrence was diagnosed followed in the next three months by a second primary tumor at the pharyngoesophageal region. Genomic and epigenetic studies were conducted allowing the identification of shared imbalances by these three tumor samples in several chromosomal regions and genes, which could indicate a common clonal origin.

Case Presentation

Sample 1 - Primary Tumor

In January 2011, a Caucasian 49-year-old man, drinker and heavy smoker (≥20 cigarettes/day), was diagnosed at the Maxillofacial Surgery and Stomatology Unit, of the Coimbra Hospital and University Centre, CHUC, EPE, Portugal, with a primary squamous cell carcinoma in the left side of the tongue. The diagnosis was confirmed by

a biopsy and the well differentiated tumor was classified in advanced stage (IVa), cT4, cN2, cM0. Microscopically, the hematoxylin-eosin staining demonstrated the tumor formed by polygonal-shaped cells with eosinophilic cytoplasm and mild nuclei showing pleomorphism and hyperchromatic chromatin (photomicrograph unavailable). The treatment was three-cycles of chemotherapy (cisplatin 75mg/m2 x 1day, docetaxel 75mg/m2 x 1 day and 5-fluorouracil 1000mg/m2 x 5 days, on days 1, 22 and 43)) followed by three-cycles of chemoradiotherapy (cisplatin 75mg/m2 before the radiation – 60Gy/30 fractions on days 1, 22 and 43). One month after the completion of treatment the patient did not have any sign of neoplasm, achieving an apparent clinical and radiological complete response as evaluated by computed tomography scan. In July 2015, 49-month after the conclusion of treatment, the clinician observed a small suspicious lesion (5mm) in the left side of the tongue, in the same localization of the primary tumor.

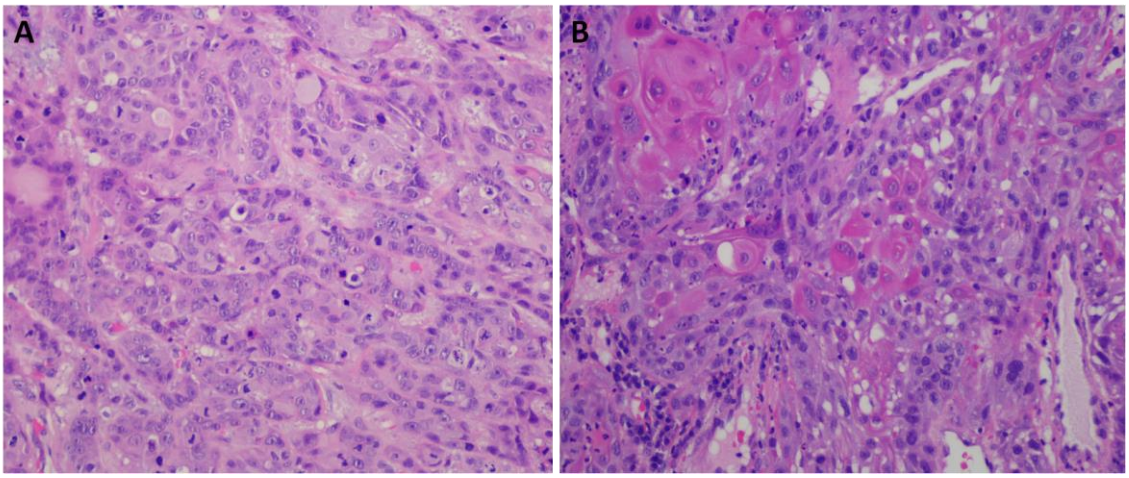
Sample 2 - Recurrence of the primary tumor

In September 2015, a squamous cell carcinoma recurrence of the primary tumor in the left side of the tongue was diagnosed in this patient, according to the Warren and Gates criteria [Warren S and Gates O 1932]. In October 2015, the patient underwent surgery and the tumor was completely removed. Microscopically, the hematoxylin-eosin staining showed tumor cells with eosinophilic cytoplasm, pleomorphic nuclei and numerous mitoses (Figure 1A). The tumor showed a high mitotic index and areas of necrosis with vascular and neural invasion. All resection margins were negative for neoplastic involvement. The tumor was classified as early stage (II), rpT2, cN0, cM0. The patient stopped his smoking habits after the primary tumor diagnosis but kept the alcohol consumption.

Sample 3 - Second primary tumor

In December 2015, a second primary tumor at the posterior wall of the pharyngoesophageal junction was diagnosed at the Department of Otorhinolaryngology - Head and Neck Surgery, Coimbra Hospital and University Centre, CHUC, EPE, Coimbra, Portugal. The diagnosis was confirmed by a biopsy and the well differentiated squamous cell carcinoma was classified in advanced stage (IVa), cT4b. Microscopically, the hematoxylin-eosin stain demonstrated a presence of dyskeratotic

cells, polygonal-shaped cells with eosinophilic cytoplasm and nuclei showing mild to moderate atypia (Figure 1B). The therapeutic decision was palliative care. In April 2016 the patient died.



Chapter 3.4.3 - Figure 1. Hematoxylin-eosin stains showing the morphology of tumor cells (H&E 200x). **A)** recurrence, **B)** Second primary tumor (SPT).

Genomic and epigenetic study

This study was approved by the Committee on Ethics in Research of the Faculty of Medicine of the University of Coimbra and written informed consent from the patient was obtained, performing all the experiments according to the regulations in the Declaration of Helsinki. Tumor tissue samples were obtained of the primary tumor and SPT from biopsies and of recurrence from the surgical resection. Additionally, macroscopically tumor-free tissue was also obtained from recurrence and SPT. This tissue, in the case of recurrence was collected from surgery resection margin and from an identical distance of the tumor in the case of SPT. The tissue samples were immediately snap-frozen in liquid nitrogen after resection and stored at −80 °C until use. DNA from fresh frozen tissues was extracted using a High Pure PCR Template Preparation Kit (Roche GmbH, Mannheim, Germany), according to the manufacturer's instructions. We analyzed copy number alterations (CNAs) of the three tumor samples through array Comparative Genomic Hybridization (aCGH) using Agilent SurePrint G3 Human Genome microarray 180K, (Agilent technologies, Santa Clara, CA) as previously described [Pinto-Leite et al. 2014]. Multiplex Ligation-dependent Probe Amplification (MLPA) and Methylation Specific MLPA (MS-MLPA) using the P248 and ME002 SALSA probemixes (MRC-Holland, Amsterdam, The Netherlands) were

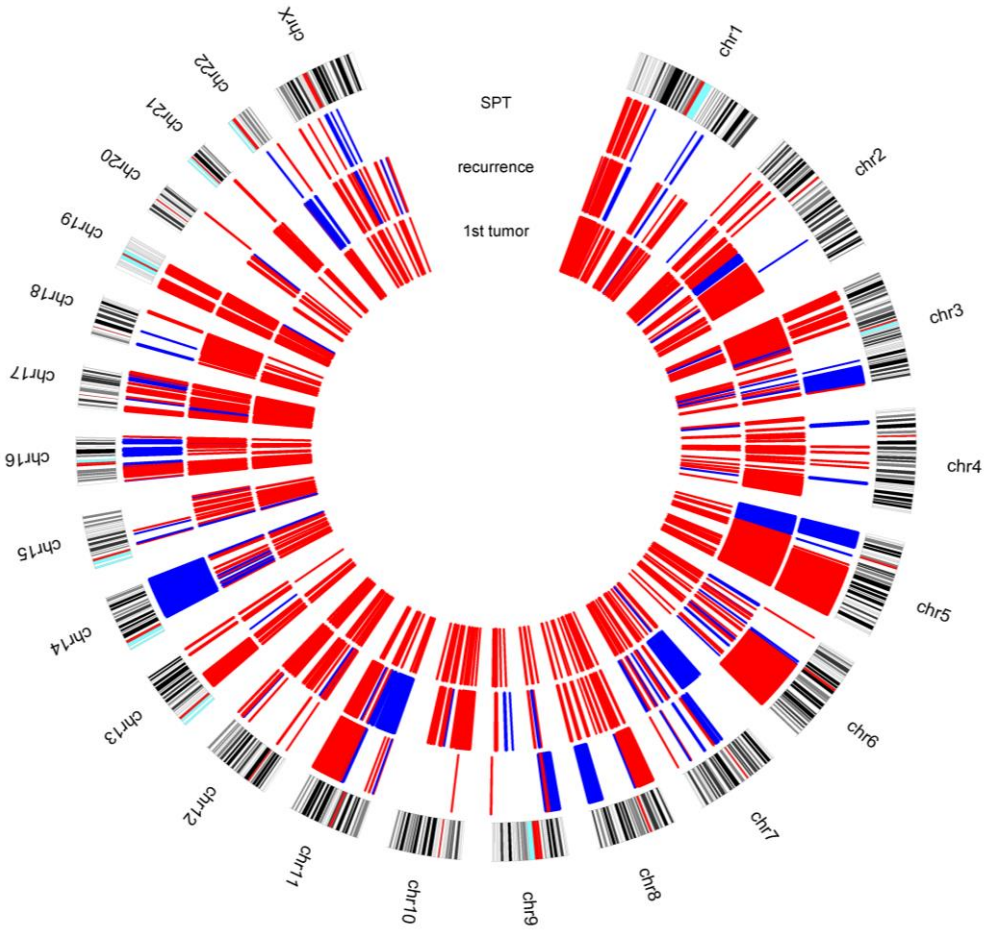
performed in tumor and non-tumor tissue samples in order to simultaneously evaluate the CNAs and methylation patterns in a specific set of genes as we previously described [Ribeiro et al. 2014; Ribeiro et al. 2016]. DNA from gender-matched gingival tissue of healthy subjects submitted to wisdom teeth removal was used as controls. The same three controls were used for MS-MLPA and MLPA techniques and one of those controls was used for aCGH technique.

Results and Discussion

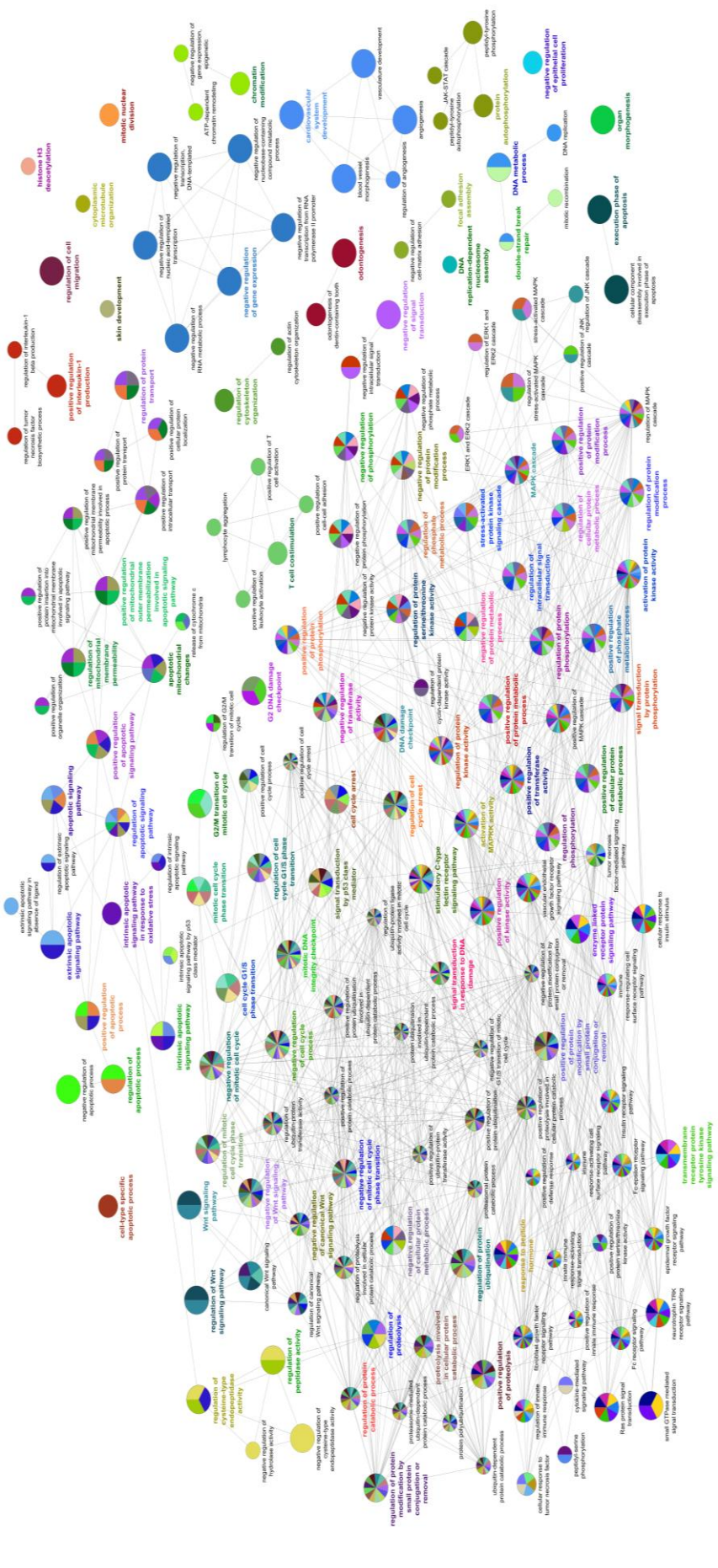
When we compared the primary tongue tumor with recurrence and SPT, we found that these three tumors shared several genomic imbalances in almost all chromosomes, excluding chromosomes 9, 20 and 22 (Figure 2).

The simultaneous altered genes identified in these three tumor samples are linked with several cellular processes, namely regulation of apoptosis, cell cycle, cell proliferation, cell migration, angiogenesis, chromatin remodeling, DNA repair and ubiquitination (Figure 3). These shared genomic imbalances seem to indicate that these three tumors have arisen from a common cell clone. Regarding the imbalances associated with chemoradioresistance and chemoradiosensitivity by Van den Broek and colleagues [van den Broek et al. 2007], we observed that in terms of size our samples presented some smaller chromosomal imbalances than those described (Table 1), which allow us to suggest some putative candidate genes in these specific chromosomal regions with an apparent link to radiotherapy response (Table 1). However, studies addressing the role of these genes in the chemoradioresistance and chemoradiosensitivity are needed. Our samples exhibited more imbalances associated with chemoradioresistance than chemoradiosensitivity. We could hypothesize that the recurrence evolved after the chemoradiotherapy treatment of the primary tumor, due to the presence of chemoradioresistance clone cells. Additionally, some imbalances of the recurrent and SPT cells seem to be further acquired after treatment, such as in short arm of chromosomes 6, 7, 8 and 18. Only the tumor from recurrence, diagnosed at stage II, presented simultaneous losses at 3p, 9p and 17p, the first imbalances described as associated with early tumor stage, by Califano and colleagues [Califano et al. 1996] (Table 2). All three tumor samples presented imbalances linked to poor patient outcome, being in the recurrence and SPT samples more evident the putative biomarkers of poor survival and of the presence of metastasis, namely gains in *EGFR* gene and at 11q13.3 (Table 2).

The MLPA and MS-MLPA results demonstrated that for all the analyzed genes the three tumors shared gain at 11q in the *GSTP1* gene (Figure 4A). Considering recurrence and SPT we observed several shared imbalances, namely at 3p, 5q, 7p, 8p, 11q and 13q. We observed some different results between aCGH and MLPA/MS-MLPA due to the different sensibility of these techniques to detect low-level imbalances. The non-clonal chromosome aberrations evidenced the genomic heterogeneity and complexity that is the reflex of chromosomal instability in the cellular population; however its frequency is relatively low, being for that often reported in the literature only the clonal chromosome aberrations [Heng et al. 2016].



Chapter 3.4.3 - Figure 2. Circus plot revealing aberration pattern differences among the three tumor samples: primary tumor, recurrence and second primary tumor (SPT). Blue represents copy number gains and red copy number losses.



Chapter 3.4.3 - Figure 3. Genes simultaneously altered in the three tumor samples grouped in a functionally network that were linked to their biological function, highlighted regulation of apoptosis, cell cycle, cell proliferation, cell migration, angiogenesis, chromatin remodeling, DNA repair and ubiquitination (ClueGO analysis using Cytoscape). The ClueGO network is created with kappa statistics and reflects the relationships between the terms based on the similarity of their associated genes.

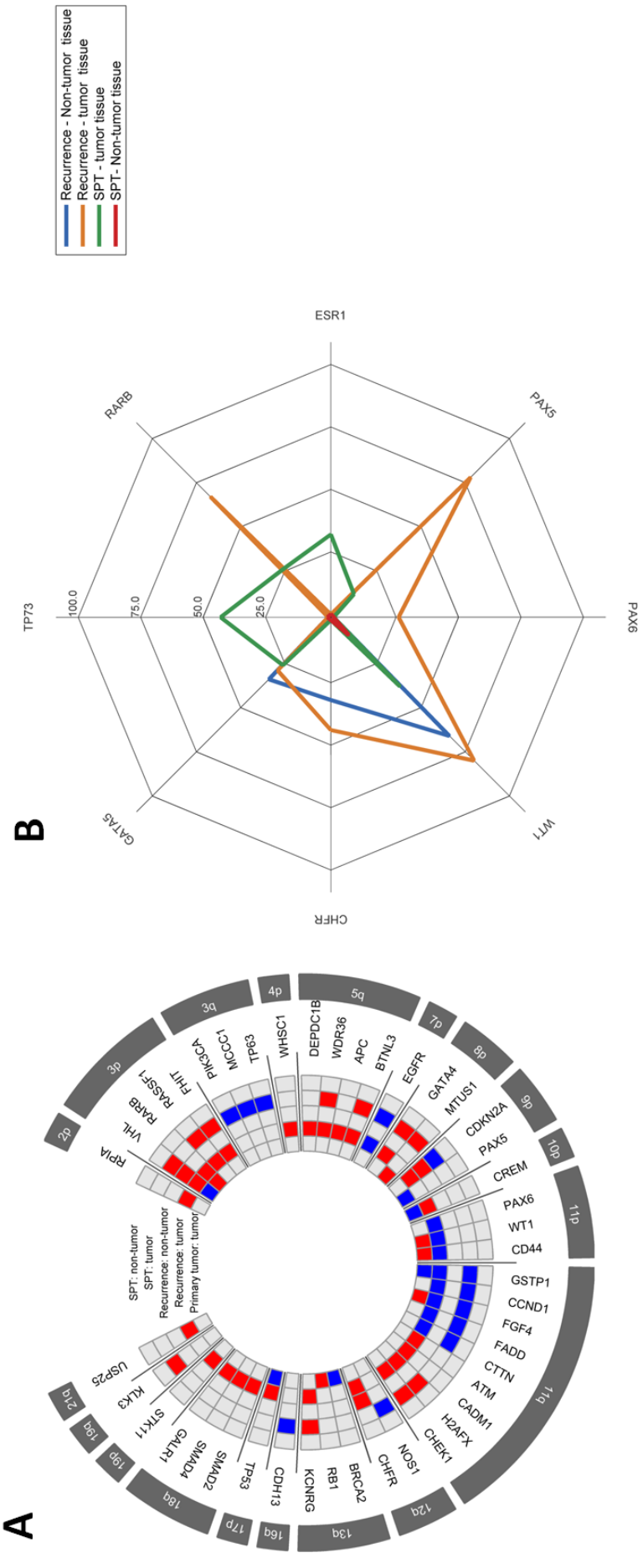
Chapter 3.4.3 - Table 1. Chromosomal regions described in the study of Van den Broek et al as associated with chemoradioresistance and chemoradiosensitivity and the specific alterations and putative candidate genes identified in the present study related to those already described.

			Present study			
Chromosomal region	Type of alteration	Clinical association	Primary tumor	Recurrence	SPT	Possible candidate genes
3q21-q26.1	Gain	Chemoradioresistance	3q26.1	3q26.1	3q26.1	*ZIC1, ZIC4*
6p11-pter				6p25.3-p25.2		*IRF4, HUS1B, FOXQ1, FOXC1, NQO2*
				6p21.33		*CYP21A2, TNXB, STK19*
				6p12.1		*GFRAL, HMGCLL1, BMP5*
					6p11.2	*PRIM2*
3q24	Amplification		3q24	3q24	3q24	-
7p11.2-12				7p12.1-p11.2	7p12.1-p11.2	*EGFR*
8p11.1					8p11.1	-
18p11.3					18p11.31-p11.23	*LAMA1, PTPRM*
3p11-pter	Loss		3p26.1	3p26.1	3p26.1	*ARL8B*
			3p25.3	3p25.3	3p25.3	*BRPF1, CIDEC, FANCD2, IRAK2, SEC13, TADA3, VHL*
			3p24.3	3p24.3		*EFHB, RAB5A, C3orf48*
			3p22.3	3p22.3		*TRIM71, CCR4*
			3p22.2	3p22.2	3p22.2	*MLH1*
			3p21.31	3p21.31	3p21.31	*DHX30, CDC25A, ATRIP, TREX1, SHISA5, PFKFB4, NCKIPSD, IP6K2, ARIH2, KLHDC8B, APEH, UBA7, RBM5*
			3p14.3	3p14.3	3p14.3	*APPL1, ARF4, ARHGEF3, DNASE1L3, FLNB, HESX1, IL17RD, LRTM1, WNT5A*
				3p13	3p13	*FOXP1, PPP4R2*
4p11-pter			4p16.3	4p16.3		*MXD4, TNIP2, NOP14*
			4p14	4p14		*UGDH, UBE2K*
11q distal			11q22.3	11q22.3	11q22.3	*ATM*
			11q23.3	11q23.3	11q23.3	*TAGLN, KMT2A, CBL, H2AFX*
17p13.1 (*TP53)*			17p13.1	17p13.1	17p13.1	*TP53*
14q distal	Gain	Chemoradiosensitivity	14q32.33	14q32.33	14q32.33	-
17q					17q24.1-q24.2	*PRKCA*
					17q25.1	*TTYH2*
22			22q11.23	22q11.23	22q11.23	*GSTT1*
2q31	Amplification		2q31.1		2q31.1	-
7q21				7q21.3	7q21.3	*PEG10*
14q13				14q13.3	14q13.3	*PAX9*
2q22-q25	Loss		2q24.2	2q24.2		*MARCH7, CD302*
7q11-q22			7q11.22 - q11.23	7q11.22 - q11.23		*WBSCR22, CLDN3*

Chapter 3.4.3 - Table 2. Several genomic imbalances identified in the three tumor samples and their clinical association described in the literature.

					Present Study		
Chr.	***Genes***	**Type of alteration**	**Clinical association**	**References**	**Primary Tumor**	**Recurrence**	**SPT**
3p21.31	*SEMA3F*	Loss	High metastasis and poor survival	[Doci et al. 2015]	No	Yes	Yes
3p22.2	*MLH1*	Loss	Early stages of disease	[Ghosh et al. 2010]	Yes	Yes	Yes
3p14.2	*FHIT*	Loss	Early Event	[C. Garnis et al. 2003]	No	Yes	No
3q26.32	*PIK3CA*	Gain	Poor prognosis	[Brauswetter et al. 2016; Suda et al. 2012]	No	No	Yes
3q26.33	*SOX2*	Gain	Metastasis ,worse outcome, resistance to cisplatin	[Schrock et al. 2014]	No	Yes	No
4q32.3	*PALLD*	LOH	Poor survival	[X. Chen et al. 2015b]	No	Yes	No
4q32.3	*DDX60L*	LOH	Poor survival	[X. Chen et al. 2015b]	No	Yes	No
4q35.2	*ING2*	LOH	Advanced stage	[Borkosky et al. 2009]	Yes	Yes	No
4q35.2	*FAT1*	LOH	Advanced tumour stage	[Pickering et al. 2013]	No	Yes	No
7p11.2	*EGFR*	Gain	Poor prognostic	[Chung et al. 2004]	No	Yes	Yes
9p21.3	*CDKN2A*	Loss	High frequency of recurrences; early event in HNSCC progression	[Danahey et al. 1999]	No	Yes	Gain
9p21.3	*CDKN2B*	Loss	Early event in HNSCC progression	[Worsham et al. 2006]	No	Yes	Gain
11q13.3	*FADD*	Gain	Worse disease-specific survival	[Gibcus et al. 2007]	No	Yes	Yes
11q13.3	*ANO1*	Gain	Poor overall survival, metastases	[Ayoub et al. 2010; Ruiz et al. 2012]	No	Yes	Yes
11q13.3	*CTTN*	Gain	Lymph node metastasis	[Rothschild et al. 2006]	No	Yes	Yes
11q13.3	*CCND1*	Gain	Presence of occult lymph node metastases, advanced clinical stage and shorter survival	[Capaccio et al. 2000; Noorlag et al. 2015]	No	Yes	Yes
11q21	*MRE11A*	Loss	Reduced sensitivity to ionizing radiation in HNSCC	[Parikh et al. 2007]	No	No	Yes
11q22.3	*ATM*	Loss	Reduced sensitivity to ionizing radiation in HNSCC	[Parikh et al. 2007]	Yes	Yes	Yes
11q23.3	*H2AFX*	Loss	Reduced sensitivity to ionizing radiation in HNSCC	[Parikh et al. 2007]	Yes	Yes	Yes
13q13.1	*BRCA2*	LOH	Poor patient outcome	[Sabbir et al. 2006]	Yes	Yes	Yes
13q14.2	*RB1*	LOH	Poor patient outcome	[Sabbir et al. 2006]	Yes	Yes	Yes
17p13.1	*TP53*	Loss	Nonresponse to neoadjuvant chemotherapy	[Cabelguenne et al. 2000]	Yes	Yes	Yes
18q21.2	*SMAD4*	Loss	Advanced stage and poor prognosis	[Bornstein et al. 2009]	No	Yes	No
18q23	*GALR1*	Methylation	Advanced stage and poor prognosis	[Misawa et al. 2008; Misawa et al. 2016]	No	Loss	No
22q13.2	*CYB5R3*	Loss	Worse prognosis, decreased survival	[Reis et al. 2002]	Yes	No	No

Ribeiro *et al. Molecular Cytogenetics (2017) 10:13* doi: 10.1186/s13039-017-0310-z



Chapter 3.4.3 - Figure 4. Results from MLPA and MS-MLPA. **A)** Radar chart with copy number alterations detected by P248 and ME002 probemixes. Blue represents copy number gains and red copy number losses. **B)** Radar chart with methylation status detected by ME002 probemix, highlighted the eight genes methylated in the samples of this patient. The scale represents the percentage of methylation detected.

Non-tumor sample of the recurrence presented genomic imbalances, namely losses at *VHL*, *CDKN2A* and *CHFR* genes (Figure 4B). Primary tumor did not present methylation in any of the evaluated genes. Both recurrence and SPT presented *RARB*, *PAX5*, *WT1* and *GATA5* methylated (Figure 4B). The highest number of gene promoter methylation was observed in the recurrent tumor sample. Tumor recurrence and the corresponding non-tumor sample exhibited *WT1*, *CHFR* and *GATA5* methylated (Figure 4B). The presence of several genetic and epigenetic imbalances in both tumor and macroscopically non-tumor samples is indicative of the dissemination of cells with malignant features even without visible morphologic changes, remaining these cells after the resection of the tumor, and consequently increasing the risk of relapse. Thus, loss at *VHL* gene and *WT1* and *GATA5* gene promoter methylation seem to be important in the observed relapses, since these genetic and epigenetic imbalances were observed in both tumor and non-tumor tissue of recurrence and also in SPT tissue. These specific alterations have a role in the prognosis, relapse prediction and in the therapeutic response; however, a validation in a cohort of patients is needed. Clinical examination and histological assessment of surgical margin status alone have been considered not enough to predict the risk of recurrence [Tabor et al. 2002], which was corroborated by our data of this patient. Moreover, a significant correlation between epigenetic profiling of clinically and histologically negative surgical margins and the development of SPT was also reported [de Carvalho et al. 2012].

Our results are in agreement with the field cancerization theory, described in 1953 by Slaughter et al [Slaughter et al. 1953] and with the cancer stem cell network model [Cabanillas and Llorente 2009; Gutierrez et al. 2012], since the synergetic effect of alcohol and tobacco abuse lead to cumulative DNA alterations with higher progression to malignancy in the left side of the tongue (primary tumor), consequently, the genomic imbalances related to therapeutic resistance guarantees the persistence of cells with malignant features even after treatment which culminated in a local recurrence and also in the development of a SPT in a distant anatomic site (pharyngoesophageal region). The patient here described is a good example that genetic and epigenetic signatures should be taking into account in order to help in clinical management of OSCC patients.

Conclusions

The clinical management of OSCC patients is complex and challenging. In the reported patient, we verified shared genomic imbalances, namely losses at 1p, 2p, 3p, 4q, 5q, 6q, 7q, 8p, 10p, 11q, 12p, 12q, 13q, 15q, 16p, 16q, 17p, 17q, 18q, 19p, 19q, 21q and Xp and gains at 3q, 7q, 14q and 15q, which are indicative of a common clonal origin for the relapses diagnosed. In the recurrent and SPT cells some imbalances seem to be acquired after treatment, such as in short arm of chromosomes 6, 7, 8 and 18. Losses at *VHL* gene and promoter methylation of *WT1* and *GATA5* genes seem to be important predictors of relapses. Further studies are needed in order to validate the putative biomarkers of diagnostic and prognostic highlighted with this patient.

Funding

Ribeiro I.P. is a recipient of a PhD fellowship (SFRH/BD/52290/2013) from the Portuguese Foundation for Science and Technology. This work was in part supported by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra) and by the Portuguese Foundation for Science and Technology (grant: UID/NEU/04539/2013).

Authors' contributions

IPR was the major contributor in writing the manuscript, analyzed and interpreted the data and performed the literature review; FM, LB and JM collected the samples and performed the critical interpretation of the data from the clinical point of view. FC and AS helped in data analysis, MJJ performed the anatomopathological analyses and data interpretation**;** JBM and IMC performed the critical review of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank MRC-Holland for the P248 SALSA probemix.

3.5 Cell Line and Primary Cell Culture Studies

3.5.1 Cytogenetic, genomic and epigenetic characterization of HSC-3 tongue cell line, with lymphnode metastasis

Ilda P. Ribeiro, Joana Rodrigues, Alexandra Mascarenhas, Nadezda Kosyakova, Francisco Caramelo, Thomas Liehr, Joana B. Melo, Isabel M. Carreira

Journal of Oral Sciences (Accepted)

Abstract

Oral cavity carcinoma develops from squamous epithelial cells by acquisition of multiple (epi)genetic alterations that target different genes and molecular pathways. We performed a comprehensive genomic and epigenetic characterization of HSC-3 cell line, through karyotyping, multicolor Fluorescence In Situ Hybridization, array Comparative Genomic Hybridization and Methylation-Specific Multiplex Ligation-dependent Probe Amplification. HSC-3 turned out to be a near-triploid cell line with a modal number of 61 chromosomes. Banding and molecular cytogenetic analysis revealed that non-random gains of chromosomal segments occurred more frequently than losses. Overall, gains of chromosome 1, 3q, 5p, 7p, 8q, 9q, 10, 11p, 11q13, 12, 13, 14, 17, 18p, 20, Yp and Xq were observed. The largest region affected by copy number loss was observed at chromosome 18q. Several of the observed genomic imbalances and their mapped genes were already associated with oral carcinoma and adverse prognosis, invasion and metastasis. The most common observed rearrangements are translocations in centromeric/near-centromeric regions. *RARB*, *ESR1* and *CADM1* genes were both methylated and showed copy number loss, while *TP73* and *GATA5* presented both methylation and copy number gain. The here reported comprehensive characterization of HSC-3 enriches the resources available for oral cancer research using this cell line, especially for testing therapeutic agents.

Keywords: Oral cavity cell lines; HSC-3, Chromosomal rearrangements, Translocation, Copy number alterations; Methylation

Background

Oral squamous cell carcinoma (OSCC) is ranked among the top 10 malignancy incidences worldwide and includes lips, tongue, gingival tissue, floor of the mouth, parotid and salivary glands tumors [Rivera 2015]. Carcinoma of the tongue is the most common type of malignant tumors in the oral cavity [Fang et al. 2014]. Despite developments in tumor therapy, the survival rates and overall prognosis of OSCC patients have not improved significantly in the last years, with a 5-year survival rate of around 50-60% [Pulte and Brenner 2010]. These tumors often invade adjacent bones and metastasize to cervical lymphnodes or distant organs [Erdem et al. 2007]. Thus, tumor recurrence and distant metastasis are the main contributors for the poor overall survival due to treatment failure and ultimately death [Gonzalez-Garcia et al. 2009]. Importantly, the rate of OSCC metastasis after recurrence is high, and relapse/recurrence is associated with increased treatment resistance [Z. G. Chen 2009]. Nowadays, there is no prognostic model available to predict disease outcome, treatment response or risk of metastasis development after primary treatment [G. R. Thomas et al. 2005].

OSCC results from accumulation of genetic alterations, where numeric and structural chromosome changes play an important role [Saunders et al. 2000]. This carcinoma is heterogeneous at clinical and molecular levels, having different aberrant pathways that contribute to the maintenance of a cancer phenotype. Understanding OSCC metastasis on a genetic and epigenetic level would be of significance to provide new possibilities of treatment and improve clinical management. Oral cavity cell lines represent useful in vitro models for studies related to oral carcinogenesis, allowing functional analysis of interesting candidate genes, proteins and drugs. The value of cell lines as models of disease is controversial, but it has been shown very clearly that OSCC cell lines are accurate reflections of the original tumor [Worsham et al. 1999]. Well characterized cell lines are pivotal for the establishment of reliable cell culture models used in pharmacogenomic studies, being available a wide variety of primary and metastatic OSCC cell lines [C. J. Lin et al. 2007b] which need an accurate molecular characterization. In general, OSCC cell lines present ploidy levels that ranged from diploid to pentaploid, including high frequencies of aneuploidy with multiple chromosomal rearrangements, such as unbalanced and balanced translocations, deletions, isochromosomes, and homogeneously staining regions (hsrs) [Martin et al. 2008]. Here, we present the comprehensive characterization of a metastatic OSCC cell line

from the tongue (HSC-3) and discuss the biological significance of the imbalances identified in terms of disease outcome. A combination of G-banding techniques and molecular technologies such as array Comparative Genomic Hybridization (aCGH), multicolor Fluorescence In Situ Hybridization (mFISH) and Methylation-Specific Multiplex Ligation-dependent Probe Amplification (MS-MLPA) allowed detailed identification and characterization of numerical and structural aberrations as well as methylation patterns in this cell line.

Methods

Cell lines and culture conditions

HSC-3 cell line (JCRB Cell Bank: JCRB0623), isolated from a 64-year-old Japanese man, is a moderately differentiated cell line derived from a primary tumor of the tongue with lymphnode metastasis. HSC-3 was cultured in DMEM medium containing 10% of fetal bovine serum (FBS) and 1% of penicillin and streptomycin (all reagents from Gibco, Life Technologies, California, USA). The cell line was validated by STR DNA fingerprinting using the AmpFlSTR Identifiler PCR Amplification Kit according to manufacturer instructions (Applied Biosystems, Foster City, California, USA). Conventional PCR was used to detect the presence of mycoplasma in samples using primers specific: MycoF-GTAATACATAGGTCGCAAGCGTTATC and MycoR-TGCACCATCTGTCACTCTGTTAACCTC [Shipitsyna et al. 2010]. HSC-3 cell line infected with mycoplasma was treated with BM-Cyclin (Sigma-Aldrich, USA) according to the supplier's recommendations.

Controls were obtained from healthy patients undergoing surgical removal of wisdom teeth. They were cultured in DMEM medium containing 1% of FBS, 1% of hydrocortisone (Sigma-Aldrich, Missouri, USA) and 1% of penicillin and streptomycin. Both cultures and controls were maintained in an incubator at 37ºC in a 5% CO_2 atmosphere. After reaching confluence, the samples were subcultured by incubation with a solution of trypsin (Gibco, Life Technologies, California, USA) and seeded into fresh tissue culture flasks.

Karyotyping

HSC-3 cell line from passages 45 and 52 were studied. Metaphase chromosomes were prepared and analysed by GTG-banding using standard protocols. Chromosomes

of overall 32 metaphases were analyzed. Metaphases were digitally imaged and karyotyped resorting to a microscope (Eclipse-400, Nikon, Tokyo, Japan) and Cytovision software (Applied Imaging System, USA). Karyotype description followed the International System for Human Cytogenetic Nomenclature 2013 recommendations.

mFISH

mFISH using all 24 human whole chromosome paints as probes was performed as described previously [Liehr 2008]. 30 metaphases from the passage 45 were analyzed.

DNA extraction

DNA from passages 45, 50 and 52 of HSC-3 cell line and from three controls cultured cells was extracted using High Pure PCR Template Preparation Kit (Roche GmbH, Manmheim, Germany), according to the manufacture's recommendations. DNA concentration and purity were measured using a NanoDrop1000 Spectrophotometer (Thermo Scientific, Waltham, USA).

Array CGH

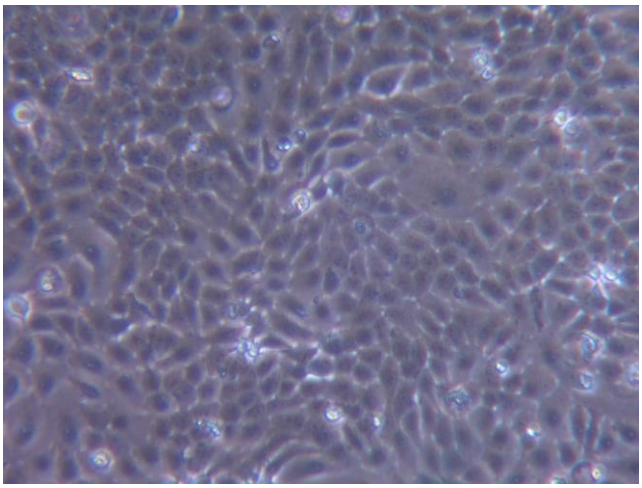
High-resolution whole genome analyses were performed using Agilent SurePrint G3 Human Genome microarray 180 K (Agilent Technologies, Santa Clara, CA, USA), according to [Pinto-Leite et al. 2014]. DNA of HSC-3 cell line from passage 45 was labelled with Cy5 by random primer labelling. DNA from control was labelled with Cy3. Results were analysed using Agilent Genomic Workbench v6.5 software with the following settings: ADM2 as aberration algorithm, threshold of 6.0, moving average 2 Mb. The results are according to Human Genome build 19 and include imbalances with at least three consecutive probes with abnormal $\log_2$ ratios. Results are given accordingly to GRCh37/hg19.

MS-MLPA

MS-MLPA analyses were performed using ME002 probemix (MRC-Holland, Amsterdam, The Netherlands), which can simultaneously detect copy number alterations (CNAs) in 38 tumor suppressor genes and aberrant methylation patterns in a subset of 25 of these genes (Supplementary Figure 1). All MS-MLPA reactions were performed according to [Ribeiro et al. 2016] using DNA from passages 45, 50 and 52.

Results

Morphologically, HSC-3 cells presented a shape between round to polygonal with small granules in the cytoplasm (Figure 1). These cells formed a monolayer and piled up after reaching confluence.



Chapter 3.5.1 - Figure 1. Morphology of the HSC-3 cell line, photographed by phase- contrast inverted microscopy (Axio Cam ERc55 camera, Zeiss, Germany) (x100).

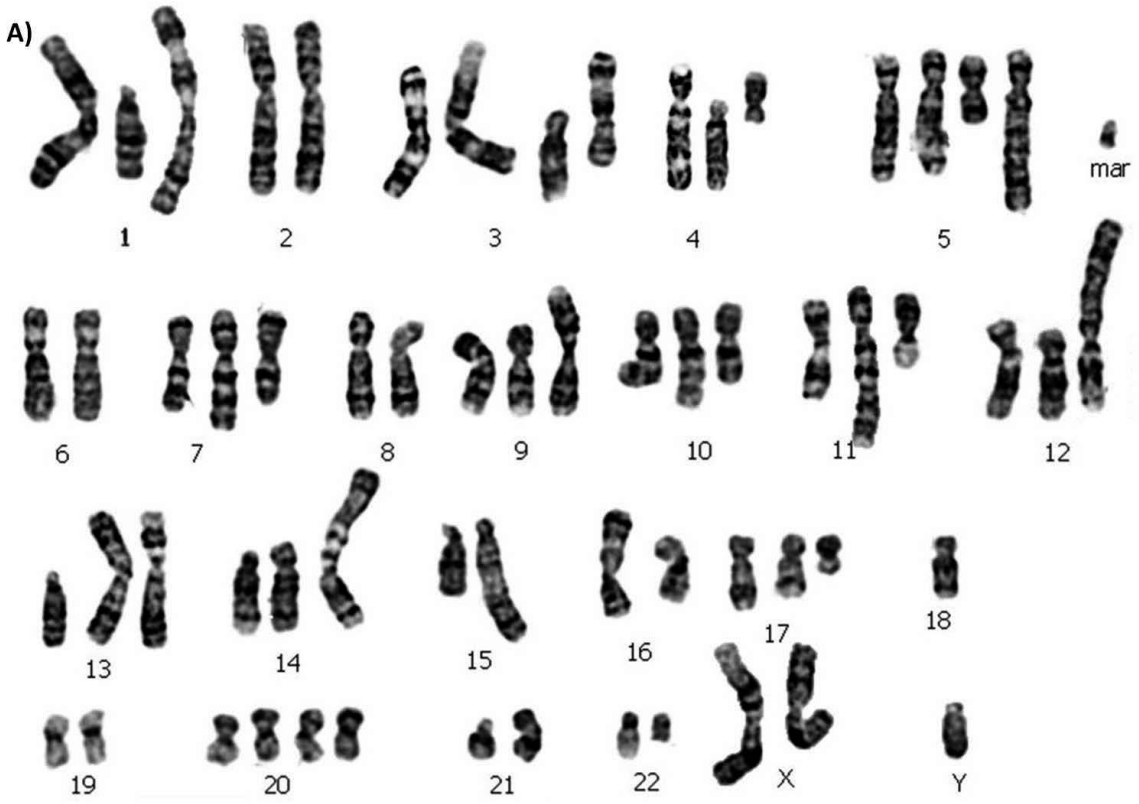
The identity of this cell line as confirmed by STR DNA fingerprinting (Supplementary Table 1), was in agreement with that described in the Japanese Collection of Research Bioresources Cell Bank (JCRB cell bank).

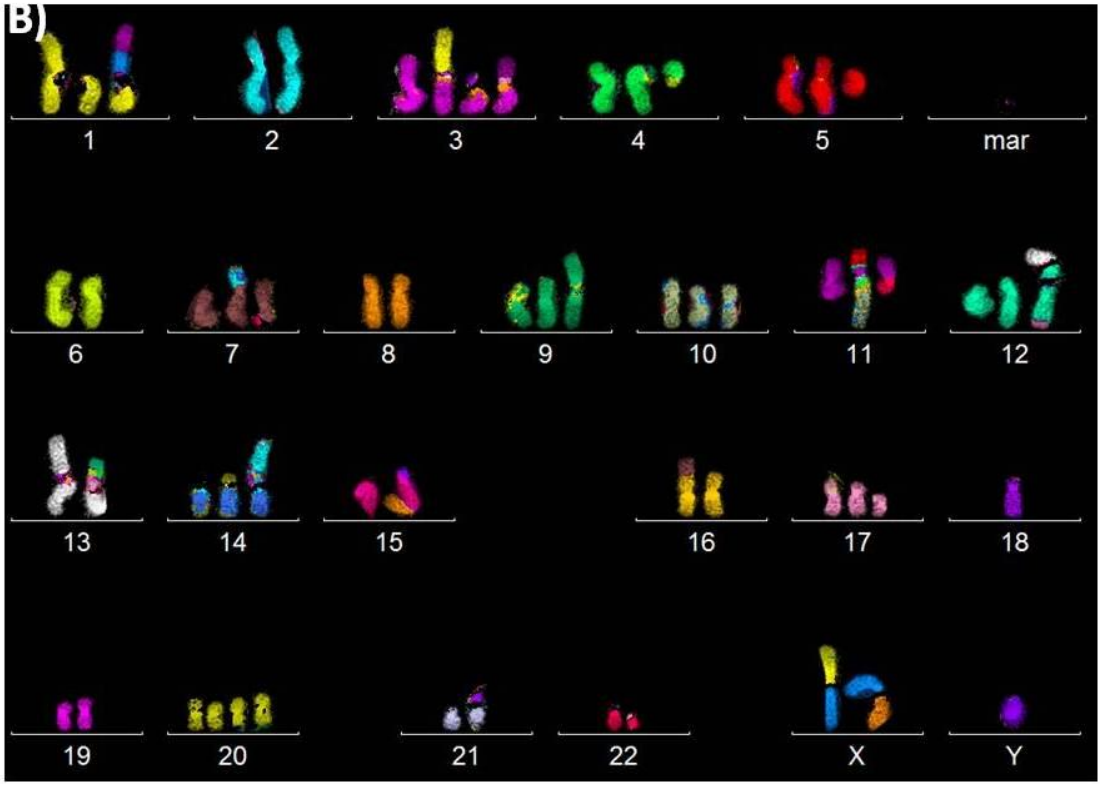
Banding cytogenetic analysis

Numerical and structural chromosomal abnormalities were found in HSC-3 cell line. HSC-3 is a near-triploid cell line with an average number of 61 chromosomes (range from 58 to 63). The most common rearrangements observed are derivatives chromosomes from translocations, in which the breakpoints often happen in centromeric/near-centromeric regions. The composite karyotype of HSC-3 cell line can be described as follows (Figure 2A):

58~63,der(X)t(X;1)(q10;p10)[31],der(X)t(X;8)(q10;q10)[32],Y,-Y[4],-1[3],
+del(1)(q10)[30],der(1)t(1;?;11)(p10;?;q10)[29], -2[4], -2,-2[1],-
3[1],der(3)t(1;3)(p10;q10)[32],+der(3)t(3;11)(p10;p10)[29],
+der(3)t(3;11)(p10;p10)x2[1],+der(3)t(3;18)(p10;q10)[32],-
4[1],+4[3],del(4)(p12)[30],+del(4)(q12)[32],-
5[3],+5[2],+add(5)(q?)[7],+i(5)(p10)[30],+ i(5)(p10)x2[1], -6[2],-
7[1],+7[2],del(7)(q31?)[31], del(7)(q31?)x2[1],der(7)(p11;q22?)[29],-

8[27],+8[1],+i(9)(q10)[31],-10[1] ,+del(10)q(24)[31],der(10)t(Y;10)[1],-11[1], +11[1],der(11)t(11;22)(p10;q10)[32],+der(11)(5;11;10)(p13;p11.2;q11.2)[29],-12[2],+12[1],+der(12)add(p11)(2p?)(13?),+13[4],der(13)t(11;13)(q23?;q11)[32], idic(13)(p11.1)[31], idic(13)(p11.1)x2[1], -14[1],+14[1], add(14)(p?)[31],+der(14)t(2;14)(q21;p10)[29],-15[1],der(15)t(8;15)(q13q26)[29], +16[2],der(16)t(7;16)(p13;p13.3)[30],+17[2],del(17)(p13)[28],+del(17)(q12)[27],-18[32],+20,+20[26],+20[5],+20,+20,+20[1],+21[1],der(21)t(18;21)(q10;p10)[31],-22[2],del(22)(q11.2)[25],+mar[30],+2mar[2][cp32].





Chapter 3.5.1 - Figure 2. . HSC-3 cell line. **A)** karyogram of a representative G-banded metaphase cell. **B)** Representative Multicolor fluorescence in situ hybridization (mFISH) karyogram

We observed several simple rearrangements involving just two chromosomes and also complex rearrangements involving multiple chromosomes (Table 1). Isochromosomes 5p and 9q were identified in almost all metaphases. Regarding aneuploidies, only chromosome 18 was lost. However, partial deletions in chromosomes 1, 7, 10, 17p, 17q and 22 were observed. Chromosome 20 presented the highest number of extra copies. In 14p we observed an insertion of a portion from an unidentified chromosome.

Molecular cytogenetic analysis based on mFISH

Based on mFISH results, HSC-3 cell line had a complex karyotype with multiple chromosomal aberrations. The most common type of the observed rearrangements involved peri-centromeric regions and resulted in whole-arm chromosome translocations. The stemline karyotype of the cell line is near-triploid and can be described as follows (Figure 2B):

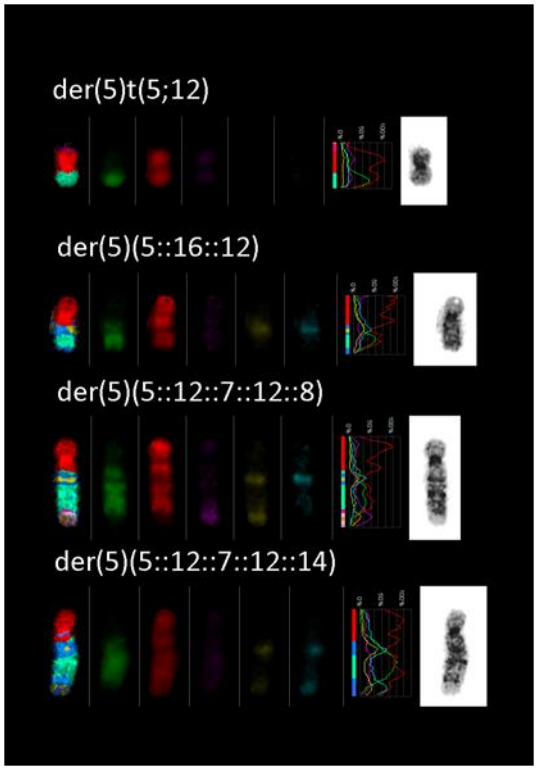
61,der(X)t(X;1),der(X)t(X;8),Y,del(1),der(1)(1::11::X::11),
-2,der(3)t(1;3),der(3)t(3;11),+der(3)t(3;18),der(4)t(4;6),der(4)t(4;9),der(5),-

6,der(7)t(7;14),dic(7;22),-8,idic(9)(q10),del(10),
der(11)(5::11::4::10),der(11)t(11;22),dic(12;13)(13::12::7::12::17),idic(13;13),
der(13)(5or12::11::16::11::8or17::13),14p+?,der(14)t(14;?),der(14)(2::4::2::4::14),-
15,der(15)t(8;15),-16,der(16)t(7;16),der(17),-18,-18,-19,+20,
-21,der(21)t(18;21),del(22),+mar.

However, the stemline karyotype was observed only in 4 out of 30 analyzed metaphases (~13%) with other metaphases showing additional non-recurrent chromosome aberrations. The composite karyotype of HSC-3 cell line based on the analysis of 30 mFISH metaphases is as followed:
58~63,der(X)t(X;1)[28],der(X)t(X;8)[29],der(X)t(X;11)[1],
der(X)(12::12?::1::X::13)[1],Y,-Y[1],del(1),del(1)[1],del(1)[1],
der(1)(1::11::X::11)[27],der(1)(1::11::X::11::2)[1],der(1)(1::10::3)[1],
dic(1;5)(5::?::12::1)[1],-2,-2[1],der(2)t(2;11)[1],t(2;3)[1],
+3[1],der(3)t(1;3)[29],der(3)(1::3::1)[1],der(3)t(3;11)[29],+der(3)t(3;18),
dic(3;5)(5::12::7?::4::11::3)[1],+der(3)t(X;3)[1],der(4)t(4;6)[29],
+der(4)t(4;6)[1],der(4)t(4;9),der(4)t(4;?)[1],del(4)[1],-5[2],der(5),del(5)[1],
del(5)[1],+der(5)t(5;12)[2],+der(5)(5::12::9)[1],+der(5)(5::16::12)[1],
+der(5)(5::16::12::16?)[1],+der(5)(5::16::12::5)[1],+der(5?)(5?::12::7?::12::8)[1],
+der(5?)(5?::12::?::5)[1],+der(5)(5::12::7?::12)[1],+der(5)(5::12::7::12::5)[1],
+der(5)(5::?::12::2)[1],+der(5)(5::12::16?::12::2)[1],+der(5)(5::12::7::12::14?)[1],
-6,der(7)t(7;14)[29],dic(7;22)[28],dic(7;22)(20::7::22)[1],
dic(7;12)(7::14::13::12::7?::12::17)[1],der(7)t(7;13)[1],der(7?)t(5;7)[1],
-8,-8[2],t(8;10)[1],-9[1],der(9)t(9;14)[1],idic(9)(q10)[29],del(9)[1],del(9)[1],
del(10),der(10)t(Y;10)[1],-11[1],der(11)(5::11::4::10)[29],der(11)(5::11::4::10::5)[1],
der(11)t(11;22),dic(12;13)(13::12::7::12::17)[28],der(12)t(8;12)[1],idic(13;13)[29],
dic(13;13)(13::13::7)[1],der(13)(5or12::11::16::11::8or17::13)[27],
der(13)(5or12::11or17::16::8or17::13::X)[1],der(13)(11::12::11:16::11?::13)[1],
14p+?,der(14)t(14;?),der(14)t(5;14)[1],der(14)(2::4::2::4::14)[27],
-15,der(15)t(8;15),-16,der(16)t(7;16),der(17),der(17)[1],-18,-18,-19,+20[28],
-21,der(21)t(18;21),der(21)t(5;21)[1],-22[1],del(22),der(?)(?::X::11)[1],
der(?)(?::11::X::11)[1],+mar[29],+mar[1][cp30].

Except for chromosomes 19, 20 and Y, all other autosomes and X chromosome were involved in structural chromosomal aberrations. Interestingly, almost 50% of cells

had an additional derivative chromosome (different in all the analyzed metaphases) which seemed to be "based" on der(5)t(5;12) – perhaps representing a jumping translocation hotspot (Figure 3).



Chapter 3.5.1 - Figure 3. Different derivative chromosomes based on der(5)t(5;12) identified in almost 50% of cells observed by mFISH.

Some chromosomal rearrangements previously observed by GTG-banding could be confirmed, and others could be redefined by mFISH (Table 1).

However, mFISH approach cannot detect any intrachromosomal rearrangements or identify the exact breakpoints; the complexity of combinatorially labeled probes means that small translocations may be even more difficult to detect than with singles color paints [Kearney 2006]. Also, mFISH technique cannot reliably identify small translocations/rearrangements involving peri-centromeric and near-telomeric regions. In our study, mFISH could not reveal the origin of the small marker chromosome in HSC-3 cell line and we cannot reliably assign the centromeric region in derivative chromosomes resulting from the whole-arm translocations.

Chapter 3.5.1 - Table 1. Comparison among the results obtained by the four techniques used in this study and the previous published SKY results for HSC-3 cell line.

Chr	Uchida et al (2006)		Own data							
	SKY		mFISH	aCGH			MS-MLPA			
	Karyotype			Gains	Losses	Comments	Gains		Losses	
	Alteration		Alteration							
1	del(1)(q10)		del(1)	p36.33-p12		see chr X	*TP73* (1p36.32)			
	der(1)t(1;?;11)(p10;?;q10)		der(1)(1::11::X::11)	q21.1-q44		see chr 3				
		i(1)(q10)								
2	+fis(2)(p10)			p21						
				p12-p11.1		see chr 14				
				q11.1-q31.1						
	del(2)(q21)		-2							
	-2									
3	der(1;3)(p33;q10)		der(3)t(1;3)(p10;q10)	p12.3			RARβ (3p24.2)			
	der(3)t(3;11)(p11;q13)		der(3)t(3;11)(p10;p10)	q11.1-q29			*CASR* (3q21.1)			
	der(3;18)(q10;p10)		der(3)t(3;18)(p10;q10)							
	der(3;5)(p10;p10)									
	der(3)t(3;22;11;13)(p12;q?;?;q?)									
	dup(3)(q22q26)									
4	der(4)t(4;6)(q13;q?)		der(4)t(4;6)	p16.3 - p15.31		see chr 11 and 14				
	del(4)(q12)									
	der(4;5)(q22;q22)			p12-p11						
	der(4)t(4;9)(p13;p21)		der(4)t(4;9)	q13.1						
	del(4)(p12)									
	der(4;5)(q22;q22)			q13.3						
									IL2 (4q27)	
5	i(5)(p10)		der(5)	p15.33 - p11		see chr 11				
	add(5)(q?)			q31.2 - q35.3						

Chr	Uchida et al (2006)	Own data								
	SKY	Karyotype	mFISH	aCGH			MS-MLPA		
		Alteration	Alteration	Gains	Losses	Comments	Gains	Losses	
6	del(6)(q21)	-6	-6	p25.3 - p21.32		see chr 4			*APC* (5q22.2)
				p21.1 - p12.3					
				q11.1 - q12					
				q22.31					
									ESR1 (6q25.1)
7	-7	del(7)(q31?)	dic(7;22)	p22.3 - p11.2		see chr 16			
				q11.21 - q22.3					*CFTR* (7q31.2)
		der(7)(p11;q22?)	der(7)t(7;14)						
8		-8	-8	q11.1 - q24.3		see chr 15 and X			
9	der(9;16)(p10;q10)	i(9)(q10)	idic(9)(q10)	q21.11 - q34.3					
10	-10	del(10)(q24)	del(10)	p15.3 - p11.1		see chr 11			*KLLN* (10q23.31)
				q11.21 - q26.3					*PTEN* (10q23.31)
11	der(11)t(10;11)(q11;q13)	der(11;22)(p10;q10) der(11)(5;11;10)(p13;p11.2;q11.2)	der(11)t(11;22) der(11)(5::11::4::10)	p15.5 - p11.12		see chr 3	*CD44* (11p13)		
				q12.3 - q13.4		see chr 1 and 13	*GSTP1* (11q13.2)		
				q14.1					
				q14.3					
				q22.1 - q22.3					
	der(11)t(11;15)(p13;p11)				q12.1 - q12.2				
	ins(11:X)(q13;p11p22)								*ATM* (11q22.3)

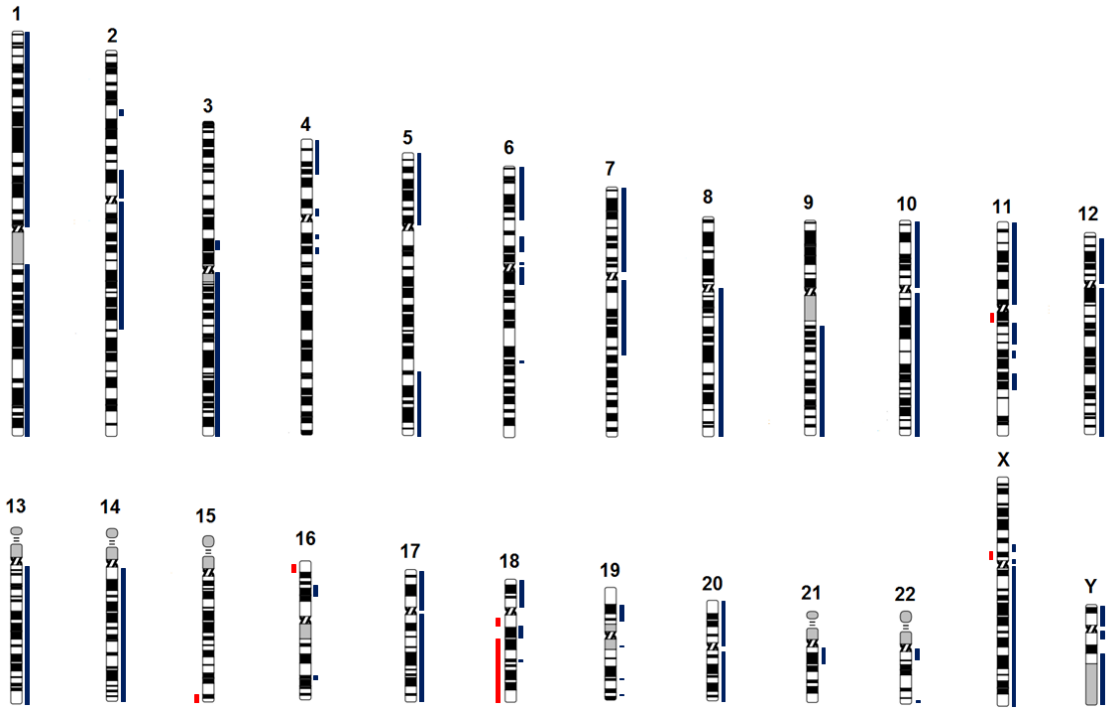
Chr	Uchida et al (2006)	Own data						
	SKY	Karyotype	mFISH	aCGH			MS-MLPA	
		Alteration	Alteration	Gains	Losses	Comments	Gains	Losses
12		der(12)add(p11)(2p?)(13?)	dic(12;13)(13::12::7?::12::17)	p13.31 - p11.1 q12 - q24.33				*CADM1*(11q23.3)
13	der(13;13)(q10;q10)	idic(13)(p11.1)	idic(13;13)	q12.11 - q34			*RB1* (13q14.2)	
		der(13)t(11;13)(q23?;q11)	der(13)(5or12::11::16::11::8or17::13)					
14	+14	add(14)(p?)	14p+?					
		der(14)t(2;14)(q21;p10)	der(14)(2::4::2::4::14)	q11.2 - q32.33			*MLH3* (14q24.3)	
			der(14)t(14;?)					
15	-15	-15	-15		q26.2 - q26.3			
	der(15)t(8;15)(q13;q26)	der(15)t(8;15)						
16	-16	-16	-16		p13.3			
	der(16)t(7;16)(p13;p13.3)	der(16)t(7;16)		p13.11 - p12.3				*CDH13* (16q23.3)
				q23.1				
17	der(17)t(17;14;8;12)(p11.?;?;?)	del(17)(p13)	del(17)	p13.3 - p11.1		see chr 12	*TP53* (17p12)	
		del(17)(q12)		q11.1 - q25.3			*BRCA1* (17q21.31)	
18	-18	-18	-18	p11.32 - p11.21		see chr 21		
				q21.2				
					q11.2			
					q12.1 - q23			
					q21.2			
19	-19	-19	-19					
	der(19)t(19;11;20)(p13;?;?)			p13.2 - p13.12				

Chr	Uchida et al (2006)	Own data								
	SKY	Karyotype	mFISH	aCGH				MS-MLPA		
		Alteration	Alteration	Gains	Losses	Comments	Gains	Losses		
20	+20	+20	+20	q12			*GATA5* (20q13.33)			
				q13.11						
				q13.42						
				p13 - p11.1						
				q11.21 - q13.33						
21	-21	-21	-21							
	der(21)t(18;21)(p10;p10)	der(21)t(18;21)(p10;p10)	der(21)t(18;21)	q11.2 - q21.1						
22	del(22)(q11.2)	del(22)(q11.2)	del(22)	q11.1 - q11.21		see chr 7 and 11				
	der(22)t(22;6;16)(p11;?p11)t(4;22)(q22;q13)			q13.33						
	der(22)t(6;22)(p21;q13)									
X	der(X;1)(q10;p10)	der(X)t(X;1)(q10;p10)	der(X)t(X;1)	p11.23	p11.23 - p11.22					
				p11.22 - p11.1	p11.22 - p11.1					
	der(X;8)(q10;q10)	der(X)t(X;8)(q10;q10)	der(X)t(X;8)	q11.1-q28						
Y	-Y	-Y								
				p11.32-p11.2						
				q11.1-q11.21						
				q11.23-q12						

Copy number alterations detected by aCGH

Results from the whole genomic approach, i.e. aCGH helped establishing breakpoints as well as the copy number gains and losses in HSC-3 cell line (Figure 4 and Table 1). Gains of the entire chromosomes 1, 10, 12, 13, 14, 17 and 20 were observed. Gains of the whole short arms were identified in chromosomes 5, 7, 11, 18 and Y. On the other hand, whole long arm gain was detected in chromosomes 3, 8, 9 and X. Copy number losses were observed at 11q12.1-q12.2 (2.6 Mb), 15q26.2-q26.3 (5.9 Mb), 16p13.3 (267 Kb), 18q11.2 (3.2 Mb), 18q12.2-q21.1 (50.6 Mb) and Xp11.23-p11.22 (2.7 Mb), being suggestive of the breakpoints of rearrangements involving these chromosomes.

It is important to note that aCGH results were considered based on a diploid cell line, since HSC-3 is a near-triploid cell line and this technique is not able to detect ploidy changes.



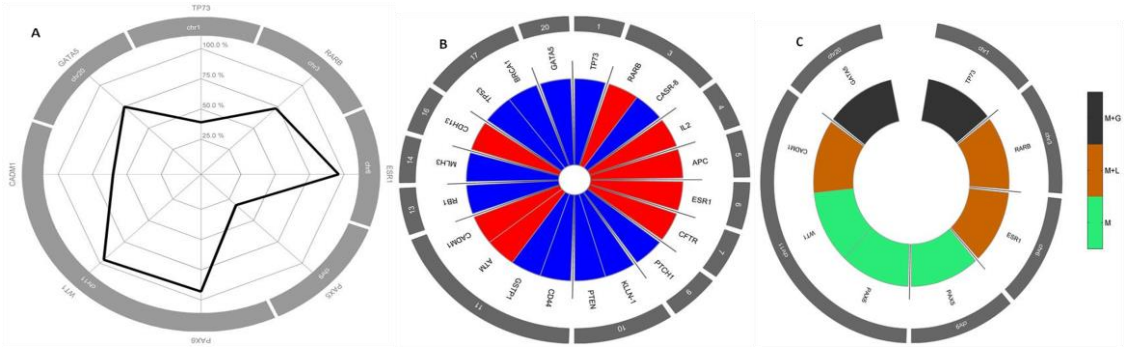
Chapter 3.5.1 - Figure 4. Copy number alterations detected by aCGH in HSC-3 cell line. Losses are represented in red and gains in blue.

Copy number alterations and methylation signature based in MS-MLPA

From the 25 genes analyzed for methylation signature, we found that *TP73*, *RARB*, *ESR1*, *PAX5*, *PAX6*, *WT1*, *CADM1* and *GATA5* genes were methylated. The levels of methylation in these eight genes vary from 32 to 96% (Figure 5A). In opposite,

MSH6, *VHL*, *CDKN2A*, *KLLN*, *MGMT*, *CD44*, *GSTP1*, *ATM*, *CHFR*, *BRCA2*, *RB1*, *THBS1*, *PYCARD*, *CDH13*, *TP53*, *BRCA1* and *STK11* genes did not present methylation (data not shown).

Copy number alterations in 20 of the 38 genes analyzed were detected (Figure 5B). Copy number gains were observed in chromosomes 1 (*TP73*), 3 (*CASR*), 9 (*PTCH1*), 10 (*KLLN* and *PTEN*), 11 (*CD44* and *GSTP1*), 13 (*RB1*), 14 (*MLH3*), 17 (*TP53* and *BRCA1*) and 20 (*GATA5*). Copy number losses were identified in chromosomes 3 (*RARB*), 4 (*IL2*), 5 (*APC*), 6 (*ESR1*), 7 (*CFTR*), 11 (*ATM* and *CADM1*) and 16 (*CDH13*). Considering simultaneous copy number alterations and methylation status, we verified that *TP73* and *GATA5* presented both methylation and copy number gains. Likewise, *RARB, ESR1* and *CADM1* presented both methylation and copy number losses (Figure 5C).



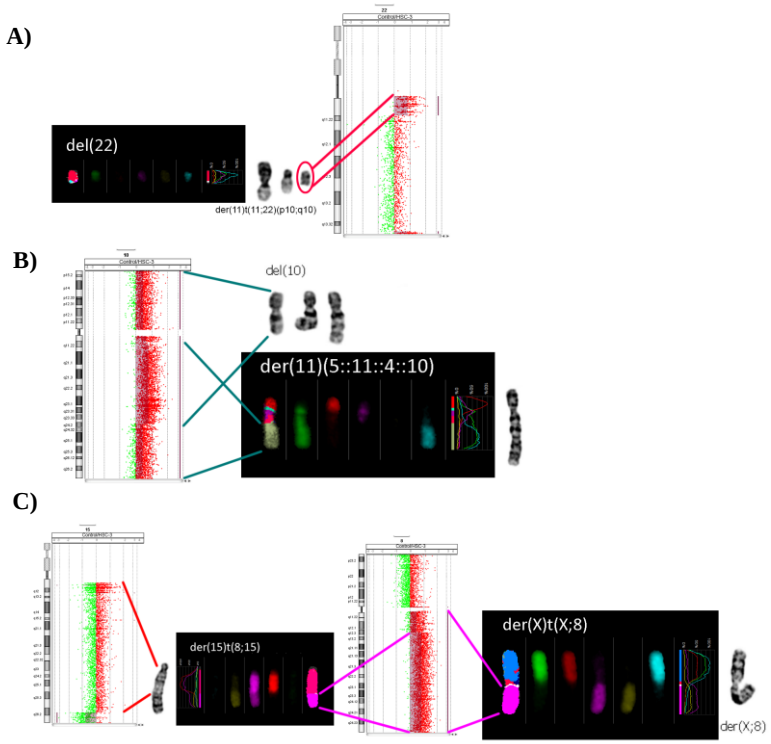
Chapter 3.5.1 - Figure 5. MS-MLPA results for HSC-3 cell line using ME002 probe panel. **A)** Methylation signature of HSC-3, showing the percentage of methylation for the eight methylated genes identified. **B)** Copy number variations detected in this cell line. Losses are represented by red and gains by blue. **C)** Simultaneous visualization of copy number alterations and methylation signature of HSC-3. Legend: M+G - methylation + copy number gains; M+L - methylation + copy number losses; M- methylation; G - gains; L- losses.

Comparison of banding and molecular (cyto)genetics with MS-MLPA results

The simultaneous analysis of the results from these four techniques allowed a comprehensive characterization of HSC-3 cell line (Table 1).

Considering the der(11)t(11;22) observed in all metaphases analyzed in both banding cytogenetic and mFISH techniques, the entire long arm of chromosome 22 was translocated with chromosome 11. Additionally, these two techniques identified a chromosome 22 with a partial deletion (del(22)(q11.2)). aCGH technique demonstrated an amplification (four copies) of ~4.7 Mb at the 22q11.1-q11.21 (17.430.492-

22.186.389), which seems to represent the portion of chromosome identified as (del(22)(q11.2)) (Figure 6A).



Chapter 3.5.1 - Figure 6. Example of imbalances obtained in HSC-3 cell line by the different techniques: conventional cytogenetics, mFISH and aCGH. **A)** del (22); **B)** der (11) (5::11::4::10) and del(10); **C)** der(15)t(8;15) and der(X)t(X;8).

In the der(11)(5::11::4::10) it was not possible to determine the specific breakpoints of these rearrangements, since chromosomes 4, 5 and 11 are also involved in other derivative chromosomes. We verified by aCGH, four regions of gain in chromosome 4: 4p16.3-p15.31 (19.5 Mb) and 4p12-p11 (1.7 Mb) with three copies; 4q13.1 (1.4 Mb) and 4q13.3 (5.2 Mb) with four copies; two regions in chromosome 5: 5p15.33-p11 (46 Mb), with four copies and 5q31.2-q35.3 (42 Mb) with three copies. Whole chromosome 10 presented gain by aCGH, 10p15.3-p11.1 (with three copies), 10q11.22-q24.2 (with four copies) and 10q24.2-q26.3 (with three copies). Banding and molecular cytogenetics allowed the clarification of this imbalance as not only a numeric but also a structural rearrangement. Since, cytogenetically we have two normal chromosomes 10, one more with a deletion (defined by aCGH as 10q24.2-q26.3, due to the presence of only 3 copies) and an entire 10q arm that was inserted in chromosome 11, involved in the aforementioned der(11)(5::11::4::10). We verified by aCGH, six regions of gain in chromosome 11: 11p15.5-p13 (33.3 Mb) with three copies; 11p13-

p11.2 (13.3 Mb) 11q12.3-q13.4 (9.9 Mb) and 11q14.1 (600 Kb) with four copies; 11q14.3 (43 Kb) and 11q22.1-q22.3 (3.9 Mb) with five copies (Figure 6B). In agreement with aCGH, MS-MLPA technique highlighted gain in 11p13 (*CD44*) and 11q13.2 (*GSTP1*) and loss at 11q22.3 (*ATM* and *CADM1*). This loss could map the specific breakpoint in one rearrangement involving this chromosome.

Another example was chromosome 15, in which, by both banding cytogenetic and mFISH, it was possible to observe a der(15)t(8;15). Additionally, aCGH demonstrated on one hand a whole 8q (8q11.1-q24.3) gain and on the other hand a 5.9 Mb distal 15q (15q26.2 -q26.3) loss, suggesting the breakpoints of this rearrangement (Figure 6C). It is important to refer that through the analysis of aCGH log ratios we identified two regions of amplification: 8q11.1-q12.3 (16.8 Mb) with three copies and 8q12.3-q24.3 (77.6 Mb) with four copies. Considering the visual analysis of this derivative chromosome we hypothesized that the translocated segment for chromosome 15 were the 8q12.3-q24.3, being the following probable breakpoints: 64.408.653-142.041.711 at chromosome 8 and 96.488.170-102.383.473 at chromosome 15. The entire 8q arm seems to be fused at the X chromosome, der(X)t(X;8) (Figure 6C). In this sense, one X chromosome presented unbalanced translocation with chromosome 8 (der(X)t(X;8)) and the other with chromosome 1 (der(X)t(X;1)). Array-CGH demonstrated a gain in the Xq, which seems to be due to insertion in chromosome 1 (der(1)t(1::11::X::11)).

Another example of the agreement between recurrent chromosomal structural rearrangements (conventional cytogenetics and mFISH) and DNA copy number aberrations (aCGH and MS-MLPA) is the identification by conventional cytogenetic of i(5)(p10) and i(9)(q10), mFISH showed der(5) and idic(9)(q11), and by aCGH it was observed four copies of 5p and 9q, respectively. MS-MLPA has only one probe for 9q, *PTCH1* gene, which also presented gain.

Furthermore, deletions in chromosomes 1, 10, 17 and 22 were observed in all evaluated metaphases and these chromosomes presented only gains in aCGH. This result apparently contradictory is explained by the insertion of portions of these chromosomes in other chromosomes, originating derivate chromosomes (Table 1).

Using these combined techniques, we highlighted in this HSC-3 cell line several chromosomal aberrations and genes associated with therapy response, disease progression and outcome (Table 2). This association with clinic features seems to be

vital in order to choose the most suitable in vitro models for each study allowing functional analysis of interesting candidate genes, proteins and drugs.

Chapter 3.5.1 - Table 2. Chromosomal aberrations and genes identified in HSC-3 cell line that are associated with disease outcome.

Cancer	Sample Type	n	Technique	Chromosomal region	Genes	Alteration	Association	References
HNSCC	FFPE	52	aCGH	1p33-34, 7q21, 10q11-q22, 12q11, 14q13, 17q		Gain	Chemoradiosensitivity	[van den Broek et al. 2007]
HNSCC	Fresh-Frozen	54	CGH	1q21-q22		Gain	Metastasis	[Bockmuhl et al. 2002]
				3q		Gain	Early Marker for Tumor Invasion	
				7q11.2		Gain	Lymph Node Metastasis	
				11q13		Gain	Metastasis	
HNSCC			Karyotype, SKY, FISH, CGH	1q21-q24, 1q32-q41, 3q26, 11q13, 12q13		Gain	Metastasis	[Bockmuhl et al. 2002]
HNSCC	Fresh-Frozen	280	FISH	1q21-q22, 7q11.2, 11q13		Gain	Metastasis	[Sticht et al. 2005]
HNSCC	Fresh-Frozen	113	CGH	2q12, 3q21-29, 6p21.1, 11q13, 14q23, 14q24, 14q31, 14q32		Gain	Shorter Disease-free Interval; Disease-specific Survival	[Bockmuhl et al. 2000]
				18q11.2		Loss		
				3q		Gain	Worse prognosis	
				11q13		Gain	Worse prognosis	
				18q		Loss	Advanced Clinical Stage and Poor Survival	
HNSCC	FFPE		YAC clones; FISH	3q		Gain	Transition to invasive squamous cell carcinoma; Local Recurrences; Cancer-related Death	[Singh et al. 2001b]
HNSCC	Fresh-Frozen	45	CGH	3q25-q27		Gain	Reduced disease-specific Survival	[Ashman et al. 2003]
				11q13; 17q; 20q		Gain	Reduced disease-free Survival	
OSCC	Cell line	14	Real-time RT-PCR	3q26.32	*PIK3CA*	Gain	Advanced Clinical Stage	[Kozaki et al. 2006]
HNSCC	Cell line	20	cCGH, FISH, Immunohistochemistry	3q26.32	*PIK3CA*	Gain	Adverse Outcome	[Freier et al. 2010]
				11q13.3	*CCND1,*	Gain	Reduced Overall Survival	
					ORAOV1			
					PPFIA1			
					CTTN			
HNSCC	Fresh-Frozen	115	qPCR; aCGH	3q26.32	*PIK3CA*	Gain	Poor Prognosis	[Suda et al. 2012]
LSCC	Fresh-Frozen	79	MS-MLPA	6q25.1-25.2	*ESR1*	Methylation	Advanced Clinical Stage	[Stephen et al. 2010]
HNSCC	FFPE	84	Immunohistochemistry	8q24.21	*MYC*	Gain	Poor Survival Rate	[Vora et al. 2003a]

Cancer	Sample Type	n	Technique	Chromosomal region	Genes	Alteration	Association	References
OSCC	Fresh-Frozen	97	CGH	7p		Gain	Lymph Node Metastasis, Poor Prognosis	[Pathare et al. 2011]
				9q		Gain	Unfavorable Outcome	
				11q13		Gain	Poor Prognosis	
				18q		Loss		
OSCC	FFPE	43	PCR	7p11.2	*EGFR*	Gain	Invasion	[Nagatsuka et al. 2001]
HNSCC	FFPE	86	FISH	7p11.2	*EGFR*	Gain	Poor Prognosis	[Chung et al. 2004]
HNSCC	Cell lines	4	CGH	10p11-12, 11p		Gain	Metastasis	[Wreesmann et al. 2004]
HNSCC	FFPE	47	Immunohistochemistry	11q13.3	*CCND1*	Overexpression	Poor Prognosis; Recurrence	[Michalides et al. 1995]
LSCC	Fresh-Frozen	58	Southern Blot	11q13.3	*CCND1*	Gain	Lymph Node Metastasis; Advanced Clinical Stage	[Fracchiolla et al. 1997]
HNSCC	FFPE, Cell line	39+5	FISH, Immunohistochemistry	11q13	*CTTN*	Gain	Lymph Node Metastasis	[Rothschild et al. 2006]
HNSCC	?	365	FISH, Immunohistochemistry	11q13.3	*ANO1*	Gain	Poor Overall Survival; Metastasis	[Ruiz et al. 2012]
HNSCC	Cell lines, FFPE	11 + ?	FISH, Immunobloting, RT-PCR	distal 11q	*ATM*	Loss	Reduced sensitivity to Ionizing Radiation	[Parikh et al. 2007]

cCGH - Chromosomal comparative genomic hybridization; **FFPE** - Formalin-fixed, paraffin-embedded; **HNSCC** - Head and neck squamous cell carcinoma; **LSCC** - Laryngeal squamous cell carcinoma; **RT-PCR** - Reverse transcription PCR; **YAC** - Yeast artificial chromosomes;

Discussion

HSC-3 has been used in several studies [Cha et al. 2011; C. C. Lin et al. 2007a; Neves et al. 2006] and to the best of our knowledge, only one study [Uchida et al. 2006] evaluated this cell line at genetic level. Additionally, this cell line was included at COSMIC-Cell Lines Project [COSMIC 2016]. The present comprehensive and complementary characterization will be extremely useful to accurate design basic and translation studies in future and maybe to interpret better previous undertaken ones.

In our study, banding and molecular cytogenetic analyses characterized the overall genome-wide alterations occurring in OSCC, showing that the molecular genetic composition of the HSC-3 cell line mirrors that of primary tumors. Example of these alterations is 3q, 8q, 11q13 gains and 18q losses, which have been previously reported in various studies [Freier et al. 2010; Susanne M. Gollin 2001; C. Jin et al. 2006; S. C. Lin et al. 2002; Parikh et al. 2007; Ribeiro et al. 2014b; Ribeiro et al. 2014a].

The karyotype of HSC-3 is highly complex and contains numerous structural and numeric aberrations. It was possible to identify different translocations, being described as a underlying mechanisms of oncogenes activation by gene fusion [Haluska et al. 1987]. The high frequency of CNAs and structural rearrangements, such as derivative, dicentric or isochromosomes, observed in this cell line could suggest that combinations of these chromosomes may occur non-randomly and bear a relationship with OSCC carcinogenesis. Non-random chromosomal loss of chromosome 18q has frequently been reported in OSCC, whereas three distinct breakpoints (18q11, 18q12, and 18q22) were identified by Martin and colleagues [Martin et al. 2008] using CGH technique. In our study we further confirmed these findings by identification of losses at 18q11.2 and 18q12.1-q23. In HSC-3 the most common chromosomal breakpoints were detected around centromeric/near-centromeric regions, which are typical cytogenetic features of OSCC [Singh et al. 2001b]. The chromosomal breakpoints can affect genes that may be disrupted and thereby either cause loss of function or translocation to another part of the genome leading to altered gene function [Zitzelsberger et al. 2010]. By banding cytogenetics and MS-MLPA the 45 and 52 cell line passages were studied and similar results were obtained by both techniques. Comparing our results with those previously published, in 2006 by Uchida and colleagues [Uchida et al. 2006], we verified several common rearrangements (Table 1). Intriguingly, in opposite to our data, these authors did not identified i(5p), which is very common in OSCC [Y. Jin et al. 1995]. It is also important to state that Uchida

and collaborators [Uchida et al. 2006] identified 61~63 chromosomes and the cell line passage was not mentioned. Some additional and unique aberrations can emerge during in vitro cultivation [Kudo et al. 2001; Tabor et al. 2001]. Since, in spectral karyotyping (SKY) and CGH analyses small copy number alterations are not detectable, our integrated study allowed to deeper insights into specific complex chromosomal rearrangements, defining the size and breakpoints of the induced imbalances. The progression of cancer is often related with genomic and epigenetic instability that lead to an advantageous adaptive evolution of malignant clones [Prasad et al. 2008], which justify the highly OSCC heterogeneity and consequently differential responses to therapy. Genetic and epigenetic mechanisms are correlated during tumorigenesis, since on one hand alterations in epigenetic mechanisms can originate genetic aberrations and on the other hand genetic aberrations in epigenetic regulators can cause an altered epigenome [You and Jones 2012]. Thus, CNAs and gene promoter methylation can point out interesting candidate genes that may play a role in tumor development and/or progression. The development of nodal metastasis is the most significant clinical prognostic factor in OSCC. Wreesmann et al [Wreesmann et al. 2004] observed on one hand a high degree of concordance between genetic abnormalities in cell lines derived from primary tumors and their corresponding nodal metastases, which reflects a common clonal origin, and on the other hand to molecularly differentiate the metastases samples through the identification of several genetic aberrations that seem to be involved in the onset and/or progression of a metastatic phenotype. Amplification of 11q13 correlates with increased invasive disease, regional metastasis and more frequent recurrence, being amplification and overexpression of cortactin gene (*CTTN*) implicated in tumor cell motility and invasion, which seems to be due to increased binding between cortactin and Arp2/3 complex [Rothschild et al. 2006]. Luo et al [Luo et al. 2006] showed that both *CTTN* amplification and overexpression are associated with lymph node metastasis in esophageal squamous cell carcinoma since these aberrations contribute to tumor aggressiveness through increasing the capacity of cell migration and anoikis resistance. Ruiz et al [Ruiz et al. 2012] suggested that *ANO1* gene facilitates regulation of the cell volume and leads to cell migration, which contribute to metastatic progression in head and neck cancer. We highlighted in table 2 several chromosomal regions and genes associated with OSCC progression and outcome.

Since we were studying a cell line with lymphnode metastasis, it was not surprising to verify that we found several genetic aberrations previously related with advanced clinical stage, adverse prognosis and presence of metastatic disease. Several chromosomal aberrations and genes that we found have already been identified presenting promising diagnostic and prognostic biomarkers, without being, so far, applicability in the clinical practice [Chang and Califano 2008]. Thus, HSC-3 cell line is, indeed, a good in vitro model, not only to understand the advanced clinical stage of OSCC, which represents the most frequent diagnoses in these patients, but also to study within the well-characterized molecular profile of HSC-3 the effect of potential therapeutic modalities. Regarding the distribution of the genes highlighted with MS-MLPA technique by their biological processes, we verified that in general, the cellular functions of all reported genes were related to apoptosis, transcription regulation, cell growth and differentiation, cell cycle control and cell adhesion, all hallmarks of malignant transformation and cancer progression [Ribeiro et al. 2016].

All techniques used in this study introduced advantages and an improvement in the characterization of the HSC-3 cell line. This cell line showed aneuploidy and several structural rearrangements after G-banding analyses. Adding the mFISH technique allowed a precisely identification of the majority of the chromosomes involved in each rearrangement. Both techniques enabled the examination of individual tumor cells, which is very informative due to the high percentage of intratumoral heterogeneity. aCGH approach highlighted the specific chromosomal bands with copy number gains and losses. It identified several rearrangements, specific chromosomal regions and helped in the delimitation of the size of the imbalances, allowing the establishment of specific breakpoints. MS-MLPA provided a methylation signature of this cell line, highlighting several genes with loss and gain and also genes where co-occurred methylation and CNAs.

Conclusions

HSC-3 cell line presents a complex karyotype with multiple rearrangements involving the majority of all human chromosomes, which reflects the general OSCC signature. Simultaneously, at genetic and epigenetic level several altered genes associated with malignant transformation and cancer progression hallmarks were identified. This metastatic cell line presents several genetic imbalances associated with

adverse disease outcome. With this genetic and epigenetic characterization, HSC-3 cell line could be a more attractive and helpful resource to understand the evolution of OSCC and to study drug-therapeutic approaches for the treatment of OSCC patients, since they are frequently diagnosed in advance stages.

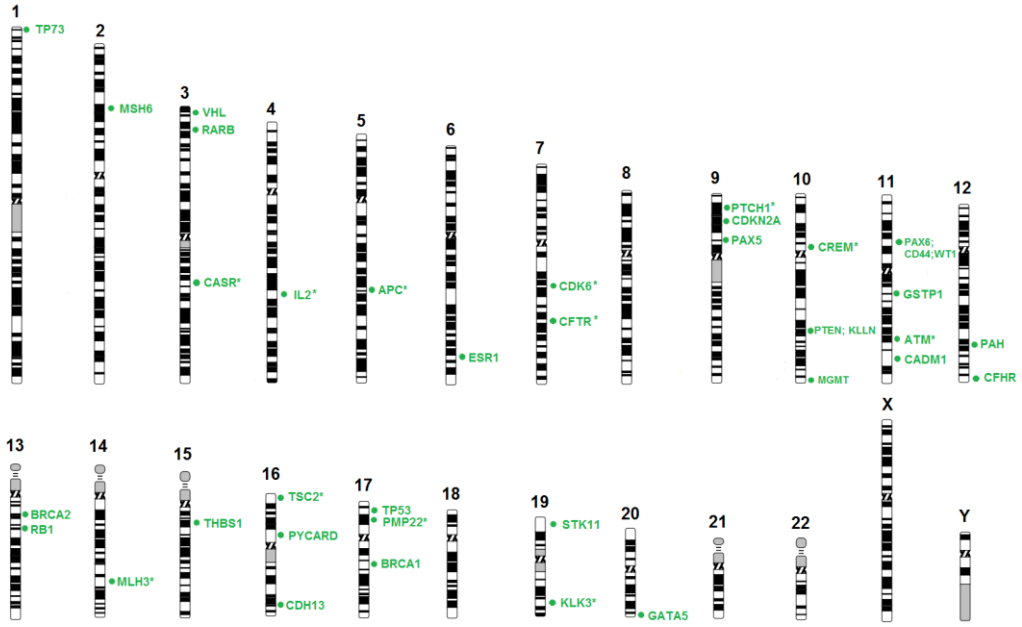
Funding

This work was in part supported by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra). Ribeiro I.P. has a PhD fellowship (SFRH/BD/52290/2013) supported by the Portuguese Foundation for Science and Technology.

Acknowledgments

We thank Professor Dr. N. Düzgünes (University of the Pacific, San Francisco, CA, USA) and Professor Dr. Marilia Dourado (Faculty of Medicine of University of Coimbra) for kindly provide the HSC-3 cell line.

We thank Joana Jorge and Professor Dr. Ana Bela Sarmento-Ribeiro (Faculty of Medicine of University of Coimbra) for the mycoplasma testing.



Chapter 3.5.1 - Supplementary Figure 1. All genes analyzed for copy number alterations and methylation status in this study using the ME002 probemix. The asterisk means that these genes were only analyzed for copy number alterations.

Chapter 3.5.1 - Supplementary Table 1. STR DNA fingerprinting of HSC-3 cell line. The asterisk means that these STRs are not described in the Japanese Collection of Research Bioresources Cell Bank (JCRB cell bank).

STR Marker	DNA profile	Fragment size (pb)
Amelogenin X:	X	107
Amelogenin Y:	Y	112
TPOX	(8;8)	(230;230)
D2S1338*	(19;23)	(324;341)
Vwa	(14;17)	(167;179)
D3S1358*	(17;17)	(132;132)
FGA*	(23;27)	(239;255)
CSF1PO	(11;11)	(326;326)
D5S818	(11;13)	(151;160)
D7S820	(13;13)	(283;283)
D8S1179*	(13;13)	(144;144)
TH01	(6;9.3)	(171;187)
D13S317	(12;12)	(233;233)
D16S539	(9;9)	(268;268)
D18S51*	(16;16)	(299;299)
D19S433*	(14.2;14.2)	(123;123)
D21S11*	(30;32)	(208;216)

3.5.2 (Cyto)genomic and epigenetic characterization of BICR 10 cell line and three new established primary human head and neck squamous cell carcinoma cultures.

Ilda P. Ribeiro, Joana M. Rodrigues, Alexandra Mascarenhas, Vanessa Marques, Nadezda Kosyakova, Francisco Caramelo, Maria J. Julião, Thomas Liehr, Joana B. Melo, Isabel M. Carreira

Submitted

Abstract

Head and neck squamous cell carcinoma cell lines are useful preclinical models to understand the molecular processes behind these tumors development and to search targeted therapies. We performed a comprehensive (cyto)genomic and epigenetic characterization of three new established primary human head and neck squamous cell carcinoma cultures and a BICR 10 cell line, through karyotyping, multicolor Fluorescence In Situ Hybridization, array Comparative Genomic Hybridization and Methylation-Specific Multiplex Ligation-dependent Probe Amplification. The three primary cultures turned out to be a near-triploid and BICR 10 is near-diploid. Banding and molecular cytogenetic analysis revealed non-random numerical and structural aberrations. The most common rearrangements identified in BICR 10 cell line were non-complex derivatives of reciprocal translocations, in which the breakpoints often happen in centromeric/near-centromeric regions. In primary cell cultures the most common rearrangements observed were isochromosomes and derivatives chromosomes from translocations. Overall, gains of 7p, 8q and losses at 3p, 8p, 9p, 18q and Xp were observed in all primary cultures and in the cell line. Some of the observed genomic imbalances and their mapped genes observed in all cell cultures were already associated with head and heck carcinoma and clinical outcome. From the analyzed gene, BICR 10 cell line exhibited higher number of gene promoter methylation, being *PAX5*, *WT1* and *GATA5* methylated in all primary cells and BICR 10 cell line. The here reported comprehensive characterization of BICR 10 cell line and the new established cultures enriches the resources available for head and cancer research, mainly for testing therapeutic agents.

Keywords: Primary cell cultures; BICR 10 cell line, Head and Neck cancer; Chromosomal rearrangements; Copy number alterations; Karyotyping; Methylation

Introduction

Cell lines derived from tumors are extensively used as cancer models in research allowing to increase the understanding of cancer biology [Domcke et al. 2013]. Ironically, human carcinomas exhibited a wildly growth in the body but in culture they are frequently difficult to establish and grow [Ince et al. 2015]. The difficulty in the establishment of primary human tumor cell cultures is cross-cutting to a wide range of tumor types. Only a small number of primary cells have been established in culture, being reported in the literature a low success rate [Gazdar et al. 1998; Heo et al. 1989; White et al. 2007], which suggest a lack of representativeness of the patients heterogeneity in the available tumor cell line collections [Ince et al. 2015]. An emerging question in cell culture studies is to understand if the established cell lines keep the characteristics of the original tumor, since these cells have been cultured in the absence of stroma, lacking the original microenvironment and tissue architecture as well as the tumor heterogeneity present in the primary cancer [Webster et al. 2010]. However, it has been demonstrated that at genetic level cell lines accurately represent the tumors of origin [Douglas et al. 2004; Greshock et al. 2007; Ottaviano et al. 2010; Ross et al. 2000; Stratton et al. 2009]. The genetic characterization of cell lines helps in the optimized choice of tumor model cells for each specific research questions, increasing the value of preclinical studies. Cell line model systems with the capability to predict the patient response to drugs enable the development and implementation of new personalized treatments [Ince et al. 2015]. Head and neck cancer (HNC) is the sixth most common type of cancer worldwide and, despite advances in diagnosis and treatment, the mortality and morbidity rates are still high, which could be the reflection of a high variation of genetic instability and molecular heterogeneity [Singchat et al. 2016], reinforcing the need of investigating novel and targeted therapies. Immortalized HNC cell lines have been an useful resource for molecular, biochemical, genetic, and immunological studies of these tumors [C. J. Lin et al. 2007b]. Available well characterized HNC cell lines are crucial in order to allow the selection of the most appropriate preclinical models for HNC tumor biology or therapeutic targets studies. In the present study, we reported the establishment and (epi)genomic characterization of three new primary human head and neck squamous cell carcinoma (HNSCC) cultures from surgical tumor samples of patients with different phenotypes. Additionally, a commercial recurrent buccal mucosa cell line, BICR 10, was also characterized, which could improve the translational research and preclinical optimization of personalized drug therapy using this line. This comprehensive

characterization showed the high heterogeneity of HNSCC, emphasizing the needed of understanding the molecular carcinogenesis mechanisms behind this neoplasm to truly treat the patients.

Material and methods

Cell lines and culture conditions

BICR 10 cell line (ECACC: 04072103), isolated from a Caucasian female, was kindly provided by Professor Marilia Dourado (Faculty of Medicine of University of Coimbra). This adherent cell line is derived from a recurrent squamous cell carcinoma of the buccal mucosa.

Three HNSCC primary cell cultures (PC) were established at the Cytogenetics and Genomics Laboratory of the Faculty of Medicine, University of Coimbra from surgical tongue and retromolar trigone tumor resected samples. PC1 was established from a stage II tongue tumor (pT2) of a Caucasian 88-year-old woman, without the traditional risk factors related to oral cancer, such as alcohol, tobacco and presence of human papiloma virus (HPV). Due to advanced age and comorbidities the patient was deemed unfit to major surgery, and was not submitted to cervical linfadenectomy. Six months after the diagnosis, this patient developed right cervical metastasis and died two months later. PC2 was established from a stage II tongue tumor (pT2) of a Caucasian 84-year-old man, only with alcohol abuse as risk factor. PC3 was established from a stage IV retromolar trigone tumor (pT3, pN2) of a Caucasian 46-year-old man, with alcohol and tobacco consumption as risk factors. Thirty months after the diagnosis, these two man patients (PC2 and PC3) that have gone through surgery and radiotherapy, are still alive and without signs of disease.

This study was approved by the Committee on Ethics in Research of the Faculty of Medicine of the University of Coimbra and we obtained written informed consent from the patient, performing all the experiments according to the regulations in the Declaration of Helsinki.

BICR 10 cell line was cultured in DMEM medium (Gibco, Life Technologies, California, USA) containing 10% of Fetal Bovine Serum (FBS) (Gibco, Life Technologies, California, USA), 1% of hydrocortisone (Sigma-Aldrich, Missouri, USA) and 1% of penicillin and streptomycin (Gibco, Life Technologies, California, USA). This cell line was validated by STR DNA fingerprinting using the AmpFlSTR

Identifiler PCR Amplification Kit according to manufacturer instructions (Applied Biosystems, Foster City, California, USA).

Primary cell cultures from tumor and control samples (obtained from healthy patients undergoing surgical removal of wisdom teeth) were cultured in DMEM medium containing 1% of FBS, 1% of hydrocortisone and 1% of penicillin and streptomycin.

All cell cultures were maintained in an incubator at 37ºC in a 5% CO_2 atmosphere. After reaching confluence, the samples were subcultured by incubation with a solution of trypsin (Gibco, Life Technologies, California, USA) and seeded into fresh tissue culture flasks.

Conventional PCR was used to detect the presence of mycoplasma using primers specific: MycoF-GTAATACATAGGTCGCAAGCGTTATC and MycoR-TGCACCATCTGTCACTCTGTTAACCTC [Shipitsyna et al. 2010]. Cell line and primary cultures were treated for mycoplasma infection with BM-Cyclin (Sigma-Aldrich, Missouri, USA) according to the supplier's recommendations.

Cell cultures were keep in liquid nitrogen for future studies.

Karyotyping

BICR 10 cell line from passages 30, 34, 35, and 36 and primary cell cultures were studied. Metaphase chromosomes were prepared and analyzed by GTG-banding using standard protocols. Chromosomes of 30 metaphases were analyzed for PC1 and PC3 and 10 metaphases for PC2. Metaphases were digitally imaged and karyotyped resorting to a microscope (Eclipse-400, Nikon) and Cytovision software (Applied Imaging System). Karyotype description followed the International System for Human Cytogenetic Nomenclature 2013 recommendations.

mFISH

mFISH was performed in BICR 10 cell line as described previously in Liehr, 2008 [Liehr 2008]. 30 metaphases from the passage 30 were analyzed.

DNA extraction

DNA from passages 30, 34 and 35 of BICR 10 cell line, from three primary cell cultures and from three cultured control cells was extracted using High Pure PCR

Template Preparation Kit (Roche GmbH, Manmheim, Germany), according to the manufacture's recommendations. DNA concentration and purity were measured using a NanoDrop1000 Spectrophotometer (Thermo Scientific, Waltham, USA).

Array CGH

High-resolution whole genome analyses were performed using Agilent SurePrint G3 Human Genome microarray 180 K (Agilent Technologies, Santa Clara, CA, USA), according Pinto-Leite et al 2014 [Pinto-Leite et al. 2014]. DNA of BICR 10 cell line from passage 30 and from the three primary cell cultures was labelled with Cy5 by random primer labelling. DNA from control was labelled with Cy3. Results were analysed using Agilent Genomic Workbench v6.5 software with the following settings: ADM2 as aberration algorithm, threshold of 6.0, moving average 2 Mb. The results are according to Human Genome build 19 and include imbalances with at least three consecutive probes with abnormal $\log_2$ ratios.

MS-MLPA

MS-MLPA analyses were performed using ME002 probemix (MRC-Holland, Amsterdam, The Netherlands) to detect aberrant methylation patterns in a subset of 25 genes. All MS-MLPA reactions were performed according to [Ribeiro et al. 2016] using DNA from passages 30, 34 and 35 of BICR 10 and from the three primary cell cultures.

Results

The identity of BICR 10 cell line as confirmed by STR DNA fingerprinting (supplementary Table 1), was in agreement with that described in the European Collection of Authenticated Cell Cultures (ECACC).

Banding cytogenetic analysis

Numerical and structural chromosomal abnormalities were found in BICR 10 cell line and primary cultures. BICR 10 cell line is near-diploid with an average number of 41 chromosomes (range from 38 to 44). Primary cell cultures are near-triploid. The modal number of chromosomes were 70 (range from 62 to 72), 64 (range from 59 to 65), and 69 (range from 64 to 72) for PC1, PC2 and PC3, respectively.

The most common rearrangements observed in BICR 10 cell line were non-complex derivatives of reciprocal translocations, in which the breakpoints often happen in centromeric/near-centromeric regions.

The composite karyotype of BICR-10 cell line can be described as follows (Figure 1A):

38~44,-X[29],XX[1],-1[30],der(2)t(1;2)(q21;p23)[30],
der(2)t(1;2)(p13;q21)[28],der(2)t(1;2)(p13;q21)x2[2],der(3)t(3;7)(p10;p10)[25];
der(3)t(3;7)(p10;p10)x2[3];-4[24],-2x4[3],+4[3],-5[7],+5[1],der(5)t(3;5)(p13;p15)[27];
-6[2],-7[19],der(7)t(7;11)(q21.3;q11.1)[24];der(7)t(7;11)(q21.3;q11.1)x2[4];-8[11],
der(8)t(8;8)(q10;q?)[24],der(8)t(8;8)(q10;q?)x2[2],-9[6],+9[4],del(9)(p22)[24],-10[5],
-11[27],der(11)t(7;11)(p10;p10)[28],der(11)t(7;11)(p10;p10)x2[2],-12[2],-13[22],
+13[3],der(13)t(11;13)(q13;q10)[4],der(13;13)[1],-14[24],der(14)t(7;14)[3],-15[2],
der(15)t(X;15)(q10;q10)[27],-16[7],2x16[1],-17[9],der(17)t(7;17)(q10;q10)[21],
+18[1],der(18)t(14;18)(q10;q10)[24],der(18)t(17;18)[23],+19[3],-20[7],-2x20[1],
-21[4],-2x21[1],-22[11],+mar1[30],+mar2[10],+mar3[3][cp30].

In primary cell cultures the most common rearrangements observed were isochromosomes and derivatives chromosomes from translocations.

The composite karyotype of PC1 can be described as follows (Figure 1B):

72~62X,der(X)t(X;13)(q10;q10)[29];-X[30];-1[4];+1[1];+der(1)t(1;15)(q10;q10)[24];
+del(1)(q25?)[25];-2[3];+del(2)(p23)[4];-3[5];+i(3)(q10)[27];-4[9];+5[5];
+i(5)(p10)[28];+i(5)(p10)x2[2];-6[2];-7[5];+del(7)(q22)[28];+8[2];i(8)(q10)[23];
-9[7];del(9)(p22)[7];del(9)(p22)x2[22];der(9)t(2;9)(p10;q10)[23];
der(9)t(2;9)(p10;q10)x2[5];der(9)t(2;9)(p10;q10)x3[1];-10[24];-2x10[6];+11[23];
+2x11[1];-12[2];-13[6];der(13)t(13;18)[22];-14[25];-2x14[1];
der(14)t(8;14)(q10;q10)[30];-16[8];-17[5];-18[29];-19[1];+19[2];-20[5];+20[12];
i(20)(q10)[26];i(20)(q10)x2[2];-21[7];add(21)(p11.1?)[20];-22[5];+22[3];+2x22[1];
+mar1[13];+mar2[6][cp30].

The composite karyotype of PC2 can be described as follows (Figure 1C):

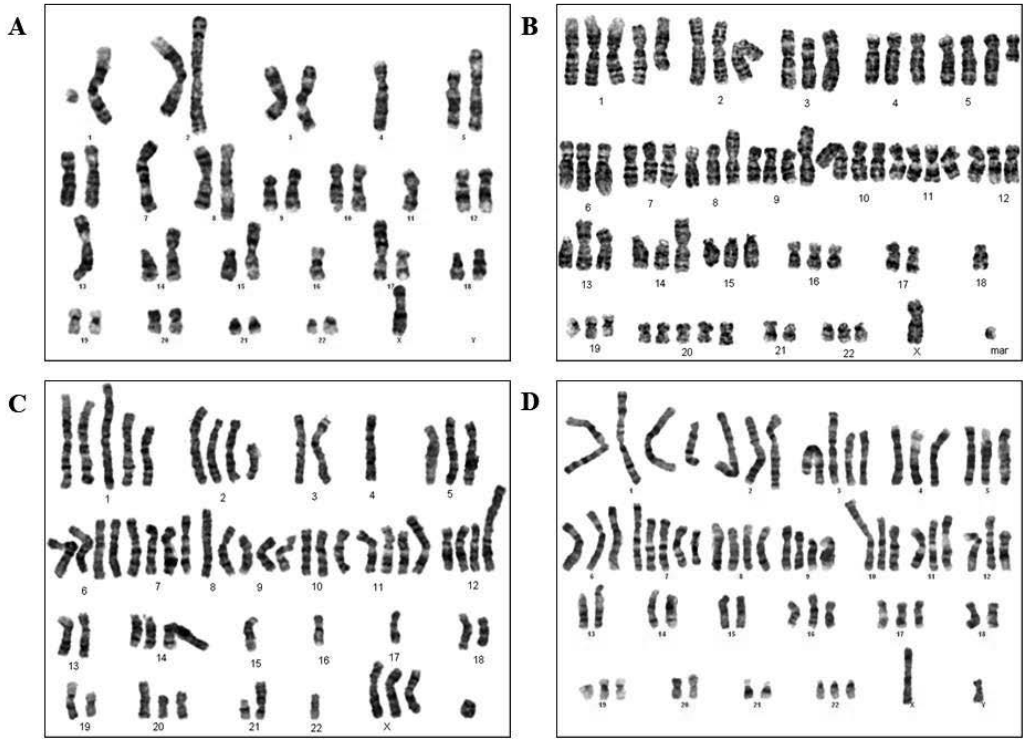
65~60, XY[1];+X[3];-Y[9];der(1)t(1;7)(q10;p10)[9];der(1)t(1;7)(q10;p10)x2[1];
+del(1)(p13?)[2];-2[4];i(2)(p10)[8];del(2)(p11.2?)[8];-3[6];der(3)t(2;3)(p10;p10)[10];
-4[10];-5[9];-7[3];+add(7)(q36?)[2];-8[4];+der(8)t(2;8)(p?;p23)[5];
-9[7];del(9)(p21?)[3];-10[4];+10[2];-11[6];+der(11)t(11;13)(p10;p10)[6];
+2xder(11)t(11;13)(p10;p10)[4];-12[5];+12[2];-3x13[7];add(13)(p11.2?)[3];-14[2];

der[14]t[3;14](p10;q10)[4];2xder(14)t(13;14)(p10;q10)[6];-16[6];
der(16)t(16;17)(q10;q10)[4];-17[2];-2x17[8];-18[2];-2x18[8];-19[2];add(19)(q13?)[6];
-20[4];+20[1];-2x21[8];-21[2];-3x22[4];-2x22[4];+mar[3];+der(?)t(?;?)[8] [cp10].
The composite karyotype of PC3 can be described as follows (Figure 1D):
64~72,XY,der(X;17)(p22.1;q12)[24];der(X;17)(p22.1;q12)x2[1];-X[5];-Y[24];
-1[2],+1[2]; i(1)(p10)[24];del(1)(p13)[24];del(1)(p13)x2[1];del(1)(p13)x3[2];
+i(1)(q10)[8]; del(1)(q21?)[4];del(1)(q21?)x2[1];-2[7];+del(2)(q13)[6];
+del(2)(q13)x2[1];-3[2]; del(3)(p14?)[11];del(3)(p14?)x2[11];del(3)(q13.3?)[9];
-4[6];del(4)(p14)[23];del(4)(p14)x2[1];-5[9];der(5)t(5;22)(p12;q11.2)[25];
der(5)t(5; 22)(p12;q11.2)x2[3];+der(5)t(5;9)(p10;q10)[14];+del(5)(p15.1?)[5];
+del(5)(q11.2)[3];-6[11];-7[5];+del(7)(q21)[15];+del(7)(q21)x2[16];-8[6];
del(8)(p23)[22];+9[5];der(9)t(9;19)(p12;q13?)[26];i(9)(p10)[26];del(9)(p21)[26];
del(9)(p21)x2[2];+i(9)(q10)[4];-10[5];-2x10[3];+11[6];der(11)t(9;11)(q22.3;q23.1)[25];
-12[4];-2x12[1];-13[13];-2x13[11];-3x13[2];-14[8];-2x14[2];-15[8];-2x15[4];
der(15)t(8;15)(q22;p11.2)[18];+16[11];-16[4];-17[6];-2x17[2];dic(7;17)(q31?;p13)[3];
-18[4];-2x18[26];-19[24];-2x(19)[1];-20[20];-2x20[7];-3x20[1];-21[7];-2x21[7];
add(21)(p11.2)[23];-22[21];-2x21[7];+mar1[5];+mar2[11];+mar3[7];+mar4[6];
+mar5[1];+mar6[1] [cp30].

All three primary cultures presented isochromosomes, being from the most frequently altered to the least frequently i(5)(p10), i(3)(q10), i(20)(q10), and i(8)(q10) observed in PC1. The i(2)(p10) was the most common isochromosome detected in PC2, whereas i(9)(p10) and i(1)(p10) were the ones most frequently observed in PC3.

In PC1 and PC2 it was frequently observed whole chromosomal arm translocations, whereas in PC3, youngest of the three patients, reciprocal translocations between chromosomal segments were more common.

From the most frequently altered to the least frequently, in PC1, der(14)t(8;14)(q10;q10), der(X)t(X;13)(q10;q10), der(1)t(1;15)(q10;q10), der(9)t(2;9)(p10;q10), der(13)t(13;18) were observed. From the most frequently altered to the least frequently, in PC2 der(3)t(2;3)(p10;p10), der(1)t(1;7)(q10;p10), der(11)t(11;13)(p10;p10), der(14)t(13;14)(p10;q10) and der(8)t(2;8)(p?;p23) were often found, whereas der(9)t(9;19)(p12;q13?), der(5)t(5;22)(p12;q11.2), der(11)t(9;11)(q22.3;q23.1) and der(X)t(X;17)(p22.1;q12) were common in PC3. However, in several cases it is difficult to distinguish which chromosomal material represents the centromere of the derivative chromosomes.



Chapter 3.5.2 - Figure 1. Karyogram of a representitive G-banded methaphase. **A)** BICR 10 cell line; **B)** PC1 cells; **C)** PC2 cells and **D)** PC3 cells.

Molecular cytogenetic analysis based on mFISH for BICR 10 Cell Line

Based on mFISH results, BICR 10 is not karyotypically stable, presenting some translocations in centromeric/near-centromeric regions (Figure 2). This cell line seems to exhibit two major clones: clone 1 has a der(15)t(X;15), clone 2 instead of it has a der(15)(3::X::15) (Figure 3).

Interestingly, 8 out of 19 metaphases (~42%) of clone 1 and 2 out of 10 metaphases (~20%) of clone 2 showed two copies of der(7)t(7;11).

Two metaphases out of clone 1 showed the presence of two identical aberrations: der(8)(8::5::11) and der(13)t(11;13). One metaphase had two copies of der (13)t(11;13).

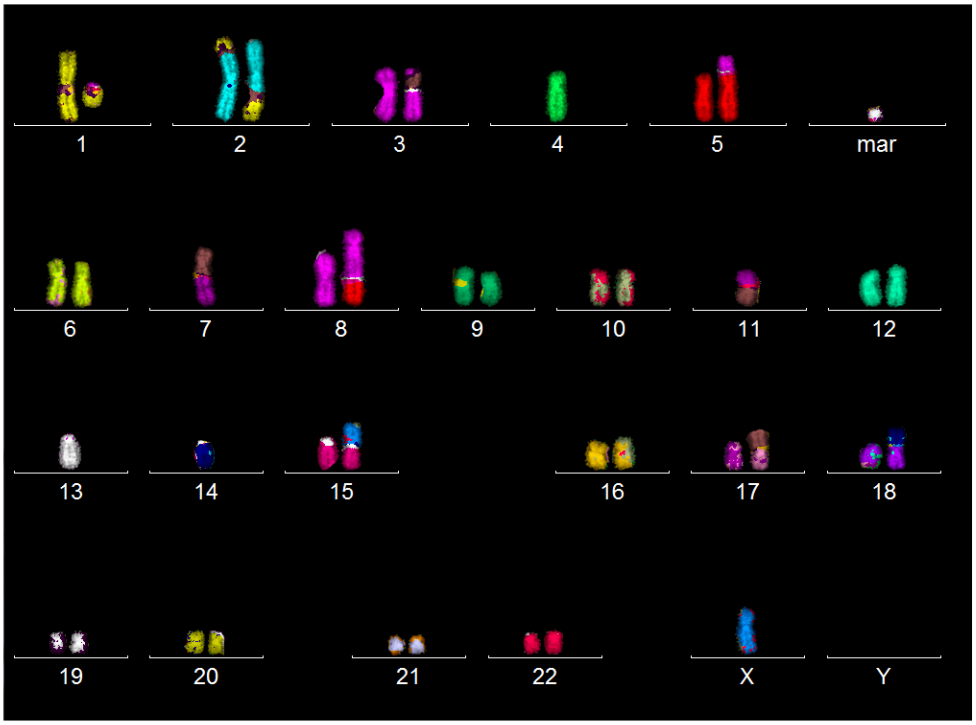
The stemline karyotype of clone 1 can be described as follows:

41,X,der(1)t(1;?),der(2)t(1;2),der(2)t(1;2),der(3)t(3;7),-4,der(5)t(3;5),-7,der(7)t(7;11), der(8),der(8)t(5;8),del(9),-11,der(11)t(7;11),-13,-14,der(15)t(X;15),der(17)t(7;17), der(18)t(14;18),der(18)t(17;18),+mar.

However, the intact stemline karyotype was observed only in one metaphase. The composite karyotype of the clone 1 can be described as follows:

38~42,X,-X[1],+del(X)[1],der(X)t(X;21)[1], der(1)t(1;?)[14],der(2)t(1;2)[16], der(2)(1::2::4)[1],der(2)t(1;2),-3[1],der(3)t(3;7),-4, del(4)[1], -5[1], der(5)t(3;5),

der(5)t(3;5)[1],-6[1],-7,der(7)t(7;11),+der(7)t(7;11)[8],der(8),der(8)t(5;8)[13], der(8)(8::5::11)[1],der(8)(17::18::8::5)[1],der(8)(8::5::15)[1],der(8)(8::5)[1], del(9)[16], -del(9)[1],der(9)t(X?;9)[1],+min(10),-11,der(11)t(7;11),-der(11)t(7;11)[1], +der(11)t(7;11)[1],der(11)t(11;22)[1],-12[1],-13,-13[1],-14,-14[1],der(15)t(X;15), -16[1],-17[1],der(17)t(7;17),der(18)t(14;18),der(18)t(17;18),-der(18)t(17;18)[2], der(18)t(8;18)[1],-20[1],-21[1],-22[4],der(22)t(9?::22)[1],der(22)t(7;22), der(22)t(9;22)[1],+mar[cp17].



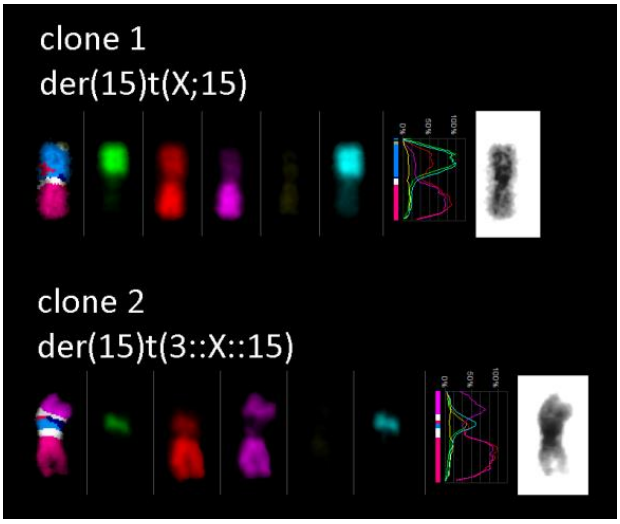
Chapter 3.5.2 - Figure 2. Representative Multicolor fluorescence in situ hybridization (mFISH) karyogram for BICR 10 cell line.

The stemline karyotype of clone 2 seems to be as follows; however, there were no metaphases truly identical:
41,X,der(1)t(1;?),der(2)t(1;2),der(2)t(1;2),der(3)t(3;7),-4,der(5)t(3;5),-7,der(7)t(7;11), der(8),der(8)t(5;8),del(9),-11,der(11)t(7;11),-13,-14,der(15)t(3::X::15),der(17)t(7;17), der(18)t(14;18),der(18)t(17;18),+mar

The composite karyotype of the clone 2 can be described as follows:
39~42,X,der(1)t(1;?),der(2)t(1;2),der(2)t(1;2),der(3)t(3;7),+der(3)t(3;7)[1],-4, der(4)t(4;9)[1],-5[3],der(5)t(3;5),-7,der(7)t(7;11),+der(7)t(7;11)[2],der(8), der(8)t(5;8)[9],der(8)(8::5)[1],del(9),-11,der(11)t(7;11),-13,-14,der(15)t(3::X::15),

der(17)t(7;17),+der(17)t(7;17)[1],der(18)t(14;18),-der(18)t(14;18)[1], der(18)t(17;18)[9], del(18)[1],idic(21;21)[1,]+mar[cp10].

With mFISH technique we could confirm some chromosomal rearrangements previously observed by GTG-banding and also redefine others. We are not able to identify the origin of the small marker chromosomes, assign the centromeric region in derivative chromosomes resulting from the whole-arm translocations and also detect any intrachromosomal rearrangements or identify the exact breakpoints [Kearney 2006].

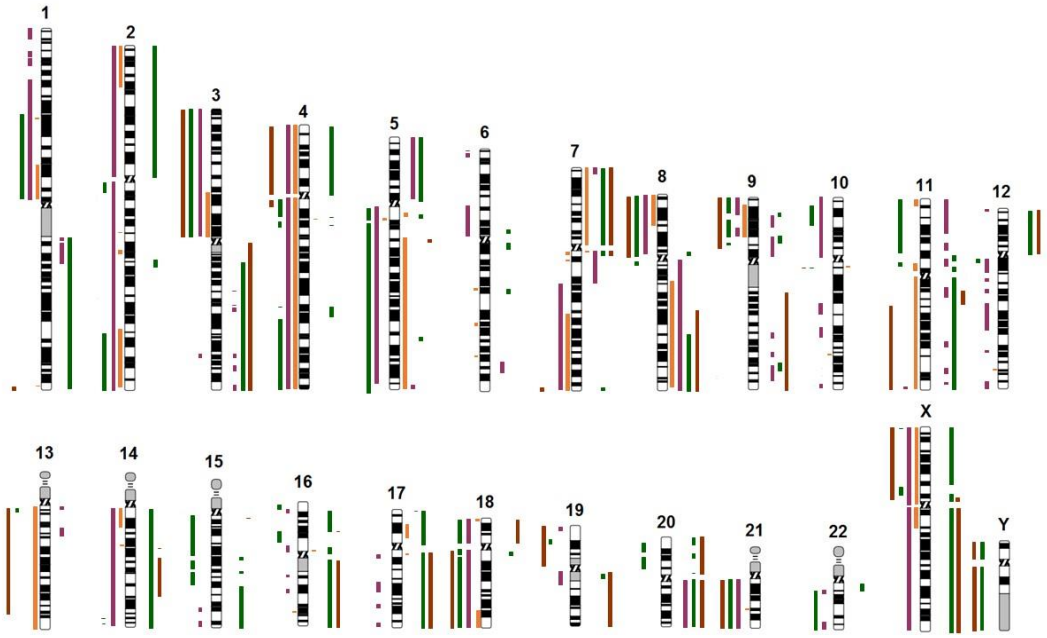


Chapter 3.5.2 - Figure 3. In BICR 10, der (15)t(X;15) and der(15)t(3::X::15) the typical rearrangement of clone 1 and 2, repectively.

Copy number alterations detected by aCGH

Results from aCGH helped establishing breakpoints as well as the copy number gains and losses. In BICR 10 cell line, whole short arm gain in chromosome 7, in the most part of long arm of chromosomes 5 (5q13.2-q35.3, 109 Mb) and 8 (8q13.1-q24.3, 77 Mb) was observed (Figure 4). Loss of whole chromosomes 4 and 13 was detected as well as loss of 11q and Xp. Terminal losses of chromosome 8 (8p23.3-p21.2, 27 Mb) and 11 (11p15.5-p15.4, 9.5 Mb) were identified. Interestingly, chromosomes 2 and 18 presented simultaneous loss of p-terminal and q-terminal, namely at 2p25.3-p22.3 (32 Mb), 18p11.32 (268 Kb), 2q33.1-q37.3 (37.6 Mb), and 18q22.3-q23 (8.3 Mb) respectively. Additionally, copy number losses were also observed at 1p21.1-p12 (16 Mb), 3p14.2-p11.1 (28.9 Mb), 7q22.3-q36.3 (54.4 Mb), 9p24.1-p21.1 (19.9 Mb), 14q11.2-q12 (12.4 Mb), being suggestive of the breakpoints of rearrangements involving these chromosomes. On the other hand, chromosomes 6, 10, 12, 15, 16, 19,

20, 21 and 22 did not exhibited major copy number alterations (Figure 4).



Chapter 3.5.2 - Figure 4. Copy number alterations detected by aCGH in BICR 10 cell line and in the three primary cultures. Each line represents one sample. BICR10 is represented by orange line, PC1 by violet line, PC2 by green line and PC3 by brown line. Losses are represented in the left and gains in the right side.

Regarding the three primary cultures, we observed that all presented gain of 20q and loss of 3p, 8p, 18q and 21q (Figure 5). Additionally, these three cultures also exhibited gain at 8q22.2-q24.3 (46.8 Mb), 11q13.1-q13.2 (2.4 Mb) and loss at 9p24.3-p23 (10.4 Mb) and 9p21.3-p21.1 (12 Mb).

Both cultures from stage II tongue tumors (PC1 and PC2) presented several common aberrations, namely, loss at 1p31.2-p12 (47, 3 Mb), 2q11.2-q12.1 (7.4 Mb), 2q33.3-q37.3 (37 Mb), 4q11-q13.1 (13.9 Mb), 4q28.3-q35.2 (57 Mb), 5q11.1-q35.2 (12.6 Mb), 10p15.1-p13 (9.7 Mb) and gain in 1q21.2-q23.1 (7 Mb) and whole 5p.

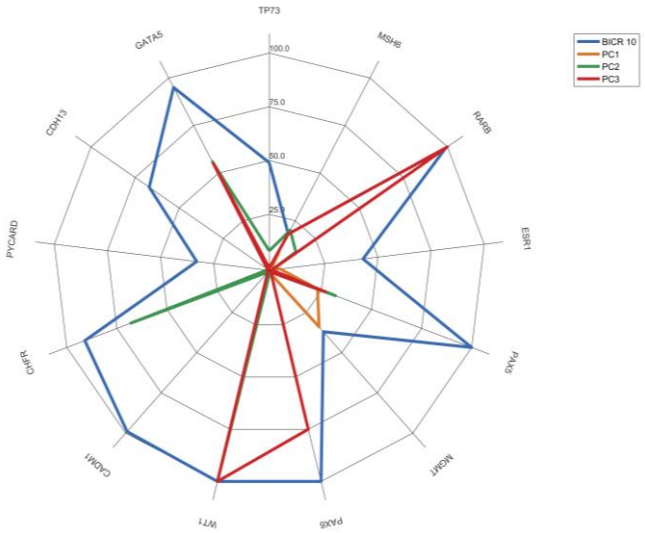
Both samples belonging to man patients (PC2 and PC3) presented gain in whole short arm of chromosomes 7 and 12, whole long arm of 16, 17 and X, as well as loss of entire Y chromosome. Moreover, these two samples also exhibited gain at 3q13.12-q29 (90 Mb) and 14q22.1-q24.3 (25 Mb).

Both samples PC1 and PC3 shared the losses of whole short arm of chromosomes 4 and X.

It is important to note that aCGH results were considered based on a diploid cell line, since this technique is not able to detect ploidy changes.

Methylation signature based in MS-MLPA

From the 25 genes analyzed for methylation signature, we found that *PAX5*, *WT1* and *GATA5* genes were methylated in all primary cells and BICR 10 cell line (Figure 5). Additionally, *TP73, ESR1, CADM1, PYCARD* and *CDH13* presented gene promoter methylation only in BICR 10 cell line (Figure 5).

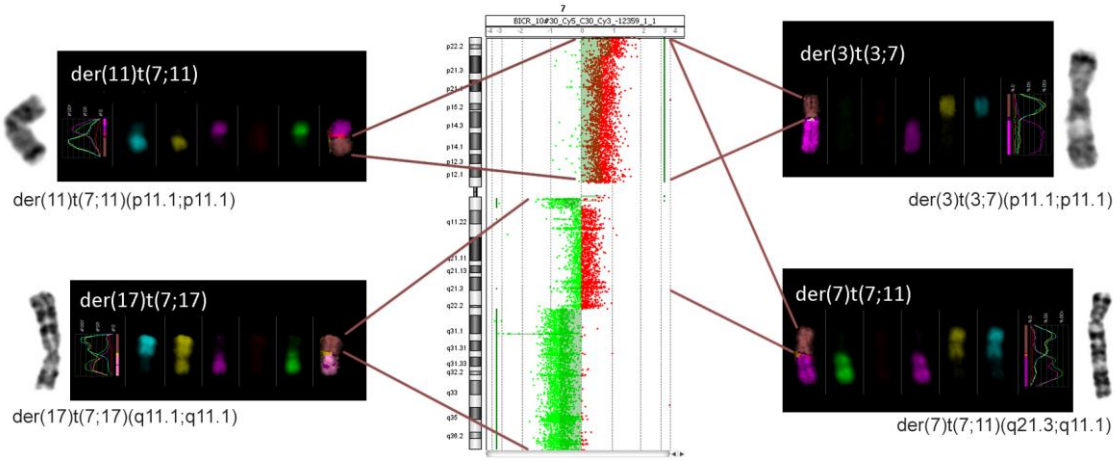


Chapter 3.5.2 - Figure 5. Methylation results using ME002 probe panel, for BICR 10 cell line and the three primary cultures.

Comparison of banding and molecular (cyto)genetics results

The simultaneous analysis of the results obtained with the different techniques allowed a comprehensive characterization of these cells. In BICR 10 cell line, considering for example the chromosome 7, we observed by both banding cytogenetic and mFISH its involvement in four derivative chromosomes, der(3)t(3;7), der(7)t(7;11), der(11)t(7;11) and der(17)t(7;17). Additionally, array-CGH demonstrated gain at 7p22.3-p11.2 (56.1 Mb) and loss at 7q22.3-q36.3 (54.4 Mb), suggesting the breakpoints of these rearrangements. Through the visual analysis of these derivative chromosomes (Figure 6), we hypothesized that the translocated segment for chromosomes 3 and 11 were 7p22.3-p11.2, being the following the plausible breakpoints: 45.130-56.174.888 at chromosome 7, 61.452.794-90.309.110 at chromosome 3 and 196.855-55.031.988 at chromosome 11. Moreover, the entire 7q arm seems to be fused at the chromosome 17 in the der(17)t(7;17)(q11.1; q11.1). Likewise, the 7p and 7q-proximal seems to be

involved in the der(7)t(7;11)(q21.3;q11.1), being the probable breakpoints 45.130-104.746.960 at chromosome 7 and 55.031.988-134.945.165 at chromosome 11.



Chapter 3.5.2 - Figure 6. Example of imbalances involving chromosome 7 in BIRCR 10 cell line, obtained by the different techniques: conventional cytogenetics, mFISH and aCGH.

Considering the BICR 10 cell line, we observed that chromosome 7 is the one more frequently involved in structural rearrangements (Table 1). Considering the primary cell cultures we observed that chromosomes 1, 2, 8, 9 and 13 are the ones more frequently involved in structural rearrangements (Table 2). The presence of some apparently discrepant results between the conventional cytogenetics and aCGH could be due to the fact that aCGH does not take into consideration the ploidy of the cells and also because tumor is constituted by different subclones with different structural and numerical aberrations (intra-tumor heterogeneity), having these cells great instability in their genetic content, which is frequently diluted in the molecular analysis.

Chapter 3.5.2 - Table 1. Numerical and structural rearrangements observed in BICR 10 cell line through conventional and molecular (cyto)genomics.

Chr	Karyotype	mFISH	aCGH	
	Alteration	**Alteration**	**Gains**	**Losses**
1	-1			p21.1-p12
		der(1)t(1;?)		
2	der(2)t(1;2)(q21;p23)	der(2)t(1;2)		p25.3-p22.3
	der(2)t(1;2)(p13;q21)	der(2)t(1;2)		q33.1-q37.3
3	der(3)t(3;7)(p10;p10)	der(3)t(3;7)		p14.2-p11.1
4	-4			p16.3-p11
				q11-q35.2
5	der(5)t(3;5)(p13;p15)	der(5)t(3;5)	q13.2-q35.3	
6				
7	-7			
	der(7)t(7;11)(q21.3;q11.1)	der(7)t(7;11)	p22.3-p11.2	
				q22.3-q36.3
8	-8			p23.3-p21.2
	der(8)t(8;8)(q10;q?)	der(8)	q13.2-q23.3	
		der(8)t(5;8)	q24.11-q24.3	
9	del(9)(p22)	del(9)		p24.1-p21.1
10				
11	der(11)t(7;11)(p10;p10)	der(11)t(7;11)		p15.5 - p15.4
				p11.2 - p11.12
	-11			q11 - q25
12				
13	-13			q12.11-q34
14	-14			q11.2-q12
15	der(15)t(X;15)(q10;q10)	der(15)t(X::15)		
		der(15)t(3::X::15)		
16				
17	der(17)t(7;17)(q10;q10)	der(17)t(7;17)		
18	der(18)t(17;18)	der(18)t(17;18)		q22.3-q23
	der(18)t(14;18)(q10;q10)	der(18)t(14;18)		
19				
20				
21				
22	-22			
X				p22.33-p11.1
				q11.1-q13.2
X	+mar1			
	+mar2			

Chapter 3.5.2 - Table 2. Numerical and structural rearrangements observed in the primary cultures through conventional cytogenetics and array-CGH.

Chr	PC1			PC2			PC3		
	Karyotype	aCGH		Karyotype	aCGH		Karyotype	aCGH	
	Alteration	Gains	Losses	Alteration	Gains	Losses	Alteration	Gains	Losses
1	der(1)t(1;5)(q10;q10) +del(1)(q25?)		p36.33 - p36.23 p33 - p12	der(1)t(1;7)(q10;p10)	q21.1 - q44	p31.1 - p12	i(1)(p10) del(1)(p13)		
2			p25.3 - p11.1 q11.1 - q37.3	i(2)(p10) del(2)(p11.2?)	p25.3 - p11.1	q11.2 - q12.1 q33.3 - q37.3			
3	i(3)(q10)		p26.3 - p11.1	-3 der(3)t(2;3)(p10;p10)	q13.12 - q29	p26.3 - p11.1	del(3)(p14?)	q11.1 - q29	p26.3 - p11.1
4			p16.3 - p11 q12 - q35.2	-4	p16.3 - p11	q11 - q13.1 q28.3 - q35.2	del(4)(p14)		p16.3 - p11 q11 - q12
5	i(5)(p10)	p15.33 - p11		-5	p15.33 - p11		der(5)t(5;22)(p12;q11.2) +der(5)t(5;9)(p10;q10)		
6			q11.1 - q35.2 p21.1 - p12.1			q11.2 - q35.3	-6		
7	del(7)(q22)	q25.2 - q25.3 q11.21 - q21.11			p22.3 - p11.2		+del(7)(q21)	p22.3 - p11.2	
8	i(8)(q10)	q11.1 - q24.3	q21.11 - q36.3 p23.3 - p11.21	+der(8)t(2;8)(p?;p23)	q22.2 - q24.3	p23.3 - p11.1	del(8)(p23)	q21.3 - q24.3	p23.3 - p11.23

Chapter 3.5.2 – Table 2. Numerical and structural rearrangements observed in the primary cultures through conventional cytogenetics and array-CGH (continued).

	PC1			PC2			PC3		
Chr	**Karyotype**	**aCGH**		**Karyotype**	**aCGH**		**Karyotype**	**aCGH**	
	Alteration	**Gains**	**Losses**	**Alteration**	**Gains**	**Losses**	**Alteration**	**Gains**	**Losses**
9	del(9)(p22)		p24.3 - p23	-9		p24.3 - p23	del(9)(p21)		p24.3 - p13.1
	der(9)t(2;9)(p10;q10)	p23 - p21.3				p23 - p21.1	i(9)(p10)	q21.11 - q34.3	
			p21.3 - p21.1		p21.1 - p13.2		der(9)t(9;19)(p12;q13?)		
		p21.1 - p11.2			q33.2 - q34.11				
		q33.3 - q34.13							
10	-10		p15.3 - p11.1			p15.1 - p13			
			q23.1 - q23.2						
			q26.2 - q26.3						
11	+11	p15.4 - p14.3		-11		p15.5 - p12	der(11)t(9;11)(q22.3;q23.1)		q13.3 - q25
		q12.2 - q13.2		der(11)t(11;13)(p10;p10)	q11 - q25				
12			q12 - q13.11	-12	p13.33 - p11.1			p13.33 - p11.1	
			q21.2 - q21.33						
			q24.32 - q24.33						
13	der(13)t(13;18)			-13			-13		q12.11 - q33.1
14	-14		q11.2 - q32.33	der(14)t(13;14)(p10;p10)	q11.2 - q32.33			q22.1 - q24.3	
	der(14)t(8;14)(q10;q10)								
15						q11.2 - q21.2	der(15)t(8;15)(q22;p11.2)		
						q21.3 - q22.2			
					q22.2 - q22.31				
16			q21	-16	p13.3 - p12.3		+16	q11.2 - q24.3	
						p12.3 - p11.2			
					q11.2 - q24.3				

Chapter 3.5.2 – Table 2. Numerical and structural rearrangements observed in the primary cultures through conventional cytogenetics and array-CGH (continued).

Chr	PC1			PC2			PC3		
	Karyotype	aCGH		Karyotype	aCGH		Karyotype	aCGH	
	Alteration	Gains	Losses	Alteration	Gains	Losses	Alteration	Gains	Losses
17			q21.33 - q22	-17	p13.3 - p11.1			q11.1 - q25.3	
					q11.2 - q25.3				
18	-18		p11.32 - p11.21	-18		p11.32 - p11.21	-18	p11.32 - p11.21	
			q11.2 - q23			q11.1 - q23			q11.1 - q23
19				add(19)(q13?)			-19		p13.3 - p12
								q11 - q13.43	
20	i(20)(q10)	q11.21 - q13.33				p13 - p12.2	-20	p13 - p11.1	
	+20					p12.1 - p11.21		q11.21 - q13.33	
					q11.21 - q13.33				
21	add(21)(p11.1?)		q11.2 - q22.3	-21		q11.2 - q22.3	add(21)(p11.2)		q11.2 - q22.3
22				-22		q11.22 - q13.33	-22		
X	-X		p22.33 - p11.1		p22.33 - p11.23		der(X;17)(p22.1;q12)		p22.33 - p11.21
	der(X)t(X;13)(q10;q10)		q11.1 - q28			p11.23 - p11.22		q11.1 - q28	
					q11.1 - q28				
Y				-Y		p11.31 - p11.2	-Y		p11.32 - p11.2
						q11.1 - q12			q11.1 - q12
	+mar1			der(?)t(?;?)			+mar2		

Through the integration of these techniques we found several chromosomal aberrations and genes previous linked in other studies with clinical outcome of the HNSCC patients (Table 3), which could be crucial to choose the most appropriate in vitro models for each experimental study.

Chapter 3.5.2 - Table 3. Chromosomal aberrations and genes identified in the primary cultures and BICR 10 cell line that are associated with disease outcome.

Chr.	Chromosomal region	Genes	Alteration	Clinical association	References
1	1q34		Gain	LRP-free survival	[Bauer et al. 2008]
3	3p		Loss	Early Event	[Hogg et al. 2002]
	3q25.31	*CCNL1*	Amplification	Loco-regional metastases	[Sticht et al. 2005]
		PIK3CA	Amplification	Poor prognosis	[Suda et al. 2012]
4	4q35.1	*ING2*	LOH	Advanced stage	[Borkosky et al. 2009]
7	7p11.2	*EGFR*	Amplification or protein overexpression	Poor prognosis; poor clinical outcome	[Sheu et al. 2009; Temam et al. 2007]
8	8p23		Loss	Poor prognosis	[Bockmuhl et al. 2001]
	8p21.2	*NEFL*	LOH	Increased mortality risk	[Coon et al. 2004]
		MYC	Amplification	Advanced primary tumors/ late event in the tumorigenesis	[Rodrigo et al. 1996]
9	9p21-22		LOH	Early Event	[van der Riet et al. 1994]
		CDKN2A	Inactivation	High frequency of recurrences	[Danahey et al. 1999]
11/18	+11q13 and -18q			Poor prognosis	[Pathare et al. 2011]
16	16q23-24	*FANCA*	Gain	LRP-free survival	[Bauer et al. 2008]
18	18q		LOH	Poor patient survival	[Pearlstein et al. 1998]

Discussion

Human tumor-derived cell lines are important preclinical models to understand the molecular bases of carcinogenesis, identify therapeutic targets and drugs screening. We established three primary cultures from tongue and retromolar trigone carcinomas. A comprehensive (cyto)genomic and epigenetic characterization of theses cell cultures and also the BICR 10 cell line were performed. These cells with a detailed molecular characterization are attractive resources to further studies of pharmacogenomics, to investigate potential candidate genes playing a role in radiation sensitivity and resistance and to explore the candidate genes associated to tumor stage, location and clinical outcome, with respect to their clinical-pathological and prognostic properties.

The karyotype of BICR 10 and primary cultures is highly complex and contains several structural and numeric aberrations. It was possible to identify several

isochromosomes and translocations, which is described as underlying mechanisms of carcinogenesis and oncogenes activation by gene fusion [Haluska et al. 1987]. In our study, we frequently identified der(1), der(9), der(11), isochromosomes of 1p, 3q, 5p, 8q, 9p and 20q, which is common rearrangements of HNSCC lines [Singh et al. 2001a; Squire et al. 2002; Van Dyke et al. 1994]. Furthermore, banding and molecular cytogenetic analyses allowed to characterize the overall genome-wide alterations occurring in these cells, in which the molecular genetic composition mirrors that of primary tumors, namely, gains at 3q, 7p, 8q, 20q and losses at 3p, 8p, 9p, 18q, 21q, previously reported in various HNSCC studies [Freier et al. 2010; Susanne M. Gollin 2001; C. Jin et al. 2006; S. C. Lin et al. 2002; Parikh et al. 2007; Ribeiro et al. 2014b; Ribeiro et al. 2014a]. All cells in study presented *PAX5*, *WT1* and *GATA5* gene promoter methylated, which seem to be interesting candidate genes that may play a role in tumor development and/or disease progression as we previous reported [Ribeiro et al. 2016]. Thus, the numerical and structural rearrangements identified in these cells could suggest that aberrations involving these chromosomes may occur non-randomly and bear a relationship with HNSCC carcinogenesis, therapy response and clinical outcome.

The integration of the techniques applied in this study introduced advantages and an improvement in the molecular characterization of the BICR 10 cell line and the new established primary cell cultures. In all characterized cells, we observed aneuploidy and several structural rearrangements, identifying the majority of the chromosomes involved in each rearrangement. This comprehensive characterization of individual tumor cells is crucial due to the high percentage of intratumoral heterogeneity. The aCGH technique demonstrated the specific chromosomal bands with copy number gains and losses, helping in the delimitation of the size of the imbalances and the establishment of specific breakpoints. Additionally, MS-MLPA technique identified a methylation signature of these cells.

Conclusions

The here reported (cyto)genomic and epigenetic landscape of BICR 10 cell line and the three new established HNSCC-derived cells provided a new opportunity to understand the molecular basis of HNSCC, since we have identified multiple alterations underlying the most frequent oncogenic events in this neoplasm. Our integrated study allowed to deeper insights into specific numerical and chromosomal aberrations, defining the size and breakpoints of the imbalances. These well molecular characterized

cells constitute an attractive cellular model system to identify novel actionable molecular targets, which could facilitate the pre-clinical evaluation of emerging therapeutic modalities regarding their effectiveness according each specific genomic signature exhibited by different clinical-pathological HNSCC tumors.

The present comprehensive and complementary characterization will be useful to accurate design basic and translation studies in future and maybe to interpret better previously undertaken ones.

Acknowledgments

Ribeiro I.P. is a recipient of a PhD fellowship (SFRH/BD/52290/2013) from the Portuguese Foundation for Science and Technology. This work was in part supported by CIMAGO (Center of Investigation on Environment Genetics and Oncobiology - Faculty of Medicine, University of Coimbra).

Conflict of interest

The authors have no conflict of interest to disclose.

Chapter 3.5.2 - Supplementary Table 1. STR DNA fingerprinting of BICR 10 cell line. The asterisk means that these STRs are not described in the European Collection of Authenticated Cell Cultures (ECACC).

STR Marker	DNA profile	Fragment size (pb)
Amelogenin X:	X	107
Amelogenin Y:	NA	NA
TPOX	(8;8)	(230;230)
D2S1338*	(25;25)	(348;348)
Vwa	(17;17)	(179;179)
D3S1358*	(17;17)	(132;132)
FGA*	(22;22)	(235;235)
CSF1PO	(12;14)	(330;338)
D5S818	(11;13)	(151;160)
D7S820	(11;12)	(275;279)
D8S1179*	(14;15)	(149;153)
TH01	(9.3)	(187;187)
D13S317	(8;14)	(217;241)
D16S539	(9;13)	(268;284)
D18S51*	(15;20)	(295;317)
D19S433*	(13.2;15)	(120; 126)
D21S11*	(30.2;32.2)	(210;218)

4. General Discussion and Conclusions

General Discussion

Tumorigenesis is described as a multistep process involving genetic and epigenetic aberrations that guarantee a selective growth advantage to the tumor cells [Colella et al. 2008]. Malignant cells acquire several capabilities during transformation process that are shared among different tumor types, which constitutes the hallmarks of cancer [Hanahan and Weinberg 2000]. Fragile sites associated with replication stress are important contributors to genome instability in neoplastic lesions [Gaillard et al. 2015; Macheret and Halazonetis 2015]. Incomplete replication and/or damage repair at common fragile sites are correlated with increased chromosomal breakpoints, resulting in chromosomal mis-segregations, CNVs, chromosomal rearrangements and consequently in cancer development [Sarni and Kerem 2016]. Nowadays, the molecular mechanisms underlying the recurrent cancer aberrations are still elusive, and the genotype-phenotype correlations would be vital to predict the clinical course of the neoplasm, to adequate the surveillance programs of individuals at high risk and also to develop gene-based therapies.

The current studies describing the molecular HNSCC profiles have highlighted the great inter and intra tumor heterogeneity [Mroz and Rocco 2016; The Cancer Genome Atlas 2015], but this knowledge had no effect in terms of improvement of survival and early detection. It is necessary to improve the understanding of the molecular basis and genomic instability behind HNSCC initiation and progression and consequently to identify and validate biomarkers that can guide early diagnosis, therapeutic strategies and patients' surveillance. Tumor outcome can differ even among patients whose tumors have similar clinical-pathological parameters and who undergo the same treatments. Besides that, the prognosis of HNSCC patients' remains mostly determined by the tumor stage at the diagnosis, which takes into account the extent of the tumor, the presence of lymph-node metastases and the distant metastases. Lately, HPV in oropharyngeal tumors has also been shown as of significant prognostic importance [Elrefaey et al. 2014]. Since patients with HNSCC in the same anatomic location, with the same tumor stage and treated with a similar protocol may have different outcomes, it is of paramount importance to identify prognostic factors at diagnosis that could be able to predict the patients' risk of recurrence and metastasis during the clinical follow-up as well as the response to therapies. Predictors of recurrence may induce a wider surgical resection, neck dissection or adjuvant therapy.

There are several studies that proposed single biomarkers of outcome for HNSCC; however, these biomarkers kept their status of candidates without translation to clinical practice. Frequently, these biomarkers presented insufficient predictive power in HNSCC patients and even discordant results among studies, which could be due to different anatomic sites of the primary tumors, and their molecular heterogeneity, to different sensitivity of the techniques and platforms used as well as to the quality of the specimens analyzed (frozen or paraffin-embedded) [G. R. Thomas et al. 2005]. The presence of low quantity of tumor DNA in the samples is also a problem to the correct interpretation of molecular abnormalities that could be missing due the mixture of tumor cells and normal cells. Thus, the signal of the aberrant DNA from the tumor cells could be overshadowed by an excess of normal DNA from non-tumor cells. Lack of validation of the candidate biomarkers in different HNSCC cohorts as well as integrative approaches of different molecular data generated for the same patients with the same methodologies are urgently needed in order to find precise diagnosis/prognosis biomarkers and to predict therapeutic responses. Moreover, a non-invasive way to search these biomarkers in high-risk populations and during/after patients' treatment is also vital to improve survival rates and decrease mortality. Considering CNAs detection, it has been reported several chromosomal regions and genes with aberrations in HNSCC and a sequential accumulation of genomic alterations from oral precursor lesions to invasive tumors. Identification of epigenomic biomarkers has so far focused mainly on DNA methylation, as this is a well-known epigenetic phenomenon with an established role in cancer, leading to altered gene function and malignant cellular transformation. Simillarly, gene expression analyses have shown the great heterogeneity of HNSCC and the classification of these tumors into subgroups with apparent different clinical outcomes.

Having this in mind, we divided our study in different phases using different technologies to accomplish a detailed molecular characterization of HNSCC.

1st Phase

In the first phase of this work we evaluated the genetic and epigenetic profile of tumor and non-tumor samples from the same patients with OSCC diagnosis, using probe panels directed to genes related to human carcinogenesis process (Chapter 3.1). In 35 OSCC patients, we analyzed 133 cancer related genes using MLPA technique and

we verified that the most frequently copy number gains were mapped on chromosomal arms 3q, 6p, 8q, 11q, 16p, 16q, 17p, 17q and 19q, whereas the most frequently copy number losses were mapped on chromosomal arms 2q, 3p, 4q, 5q, 8p, 9p, 11q and 18q (Chapter 3.1.1). We observed a pattern of imbalances in several chromosomes, e.g., specific chromosomal regions presented gains and losses in the great majority of our OSCC patients. After analyzing the most frequent altered chromosomes we highlighted specific genes that showed copy number gains or losses, such as: *ERBB4*, *CTNNB1*, *NFKB1*, *IL2*, *IL12B*, *TUSC3*, *CDKN2A* and *CASP1* with loss and *MME*, *BCL6*, *VEGF*, *PTK2*, *PTP4A3*, *RNF139*, *CCND1*, *FGF3*, *CTTN*, *MVP*, *CDH1*, *BRCA1*, *CDKN2D*, and *BAX*, as well as exon 4 of *TP53* with gain. Additionally, we also evaluated macroscopically tumor-free tissue close to the surgical margin, being the majority of these tissues histologically classified as tumor negative. We observed some genetic imbalances common to those observed in the tumor tissue in both histological positive and negative margins, without any morphological manifestation. The presence of tumor cells or cells with genetic imbalances related to the tumorigenesis process within or close to a surgical margin could be indicative of relapse risk, which also happens in patients with histologically tumor-free margins after surgery. Thus, the identification of genetic imbalances that could indicate the risk of relapses is of utmost importance in patients' management, which could include tighter clinical and molecular follow-up periods, using non-invasive approaches. By comparing tumor and matched macroscopically tumor-free tissue, we built a logistic regression model to predict the two types of tissue based on genetic data. In this model, loss of *TUSC3* gene showed statistical significance, which seems to be indicative of its relevance in the development of oral tumors, namely the role of this gene in the transition from normal oral mucosa to potentially malignant oral mucosa. We also identified specific imbalances in the smokers comparatively to non-smokers, namely the presence of losses at 3p (*MLH1*) and 11q (*ATM*), which suggests different molecular signatures according to the risk factors that originated these tumors.

We also tested, for the first time, in 30 OSCC patients of our cohort, a specifically and newly developed HNSCC probe panel (P248 MLPA probe panel), comprising 36 different genes, in order to verify its relevance to genetic characterization of these tumors and also the clinical utility in terms of decision guiding (Chapter 3.1.2). Similarly, to the previous study, we observed a consistent pattern of genetic imbalances distribution in terms of losses and gains of genetic material, which were predominantly

detected in specific chromosomes, such as 3, 8 and 11. Specific genes were highlighted due to the high frequent losses and gains detected, such as *RARB, FHIT, CSMD1, GATA4, MTUS1* and *MCCC1, MYC, WISP1, PTK2, CCND1, FGF4, FADD, CTTN*, respectively. Moreover, reinforcing the presence of different genetic signatures accordingly to the tumor localization, we verified that gain of *MYC* and *WISP1* genes was more frequent in tumors localized in the floor of the mouth.

With both these studies (Chapters 3.1.1 and 3.1.2) we proved the heterogeneity of oral tumors and also the presence of consistent specific chromosomal imbalances in these patients. We reinforced the utility of MLPA technique and this specific MLPA probe panel (P248 - HNSCC) for genetic characterization of oral tumors. However, we also verified that a more comprehensive genomic study of these tumors is needed, using high-throughput technologies in order to do a better selection of chromosomal regions and genes to be included in these probe mixes, since we observed that some of our patients did not exhibited imbalances in the genes selected for the MLPA probe panels.

In order to complement the genetic study, we performed an epigenetic analysis using a MS-MLPA technique that allows to simultaneously evaluate CNAs and methylation status of a specific set of genes (Chapters 3.1.3, 3.1.4). This epigenetic analysis was divided in two studies. In the first study (Chapter 3.1.3) we analyzed tumor samples of 93 OSCC patients using two MS-MLPA probe mixes, where 54 tumor suppressor genes and aberrant methylation patterns in a subset of 35 of these genes were analyzed. Through the correlation of the obtained genetic and epigenetic profiles with the patients' clinical-pathological features we found that *WT1* gene promoter methylation is a predictor of a better prognosis and that *MSH6* and *GATA5* gene promoter methylation serve as predictors of worse prognosis. *GATA5* gene promoter methylation was found to be significantly associated with a shorter survival rate. Additionally, *PAX5* gene promoter methylation was significantly associated with tongue tumors, reinforcing once again the molecular heterogeneity in the different tumor locations of the oral cavity, which could explain the different disease behavior and therapeutic response. We observed CNAs in all 54 analyzed genes. Interestingly, none of the evaluated genes exhibited both copy number loss and methylation, but in contrast, we observed copy number gain and methylation in the *MSH6, CASP8, PAX5, WT1, CHFR, GATA5* and *TIMP*3 genes.

The intrinsic reversibility of epigenetic alterations makes the DNA methylation an attractive target for the development of novel strategies for cancer prevention and

treatment, and several drugs with ability to change the levels of methylation are in clinical trials [Egger et al. 2004].

In the second methylation study (Chapter 3.1.4) we analyzed tumor samples and tissue from surgery resection margin (macroscopically tumor-free tissue) of 49 OSCC patients using MS-MLPA technique. We identified a genomic and epigenetic signature associated with worse OSCC patients' survival that comprises copy number alterations in *ATM, CASR, TP73, CADM1, RARB, CDH13, PAX5* and *RB1* genes and *GATA5, PAX6, CADM1* and *CHFR* promoter methylation. Copy number alterations in *BRCA1, CDKN2A, CHFR, GATA5, PYCARD, STK11, TP53* and *VHL* genes and *GATA5, CADM1, KLNN, MSH6, PAX5* and *WT1* promoter methylation were shown to be associated with development of metastasis/relapses during or after OSCC treatment. A validation in a larger and geographically separate cohort, from TCGA database reinforced the clinical applicability of these signatures to stratify the patients.

Additionally, cells scrapped from the tumor surface of these 49 OSCC patients were also analyzed. We verified good agreement in the status of *CDKN2A* promoter methylation evaluated in both noninvasive and tumor tissue samples; however, further studies regarding the non-invasive collection of samples and the genomic and epigenetic analysis are required.

With these two studies (Chapters 3.1.3, 3.1.4), we identified a specific set of genes that together has the ability to predict the prognosis of OSCC patients.

2nd Phase

In the second phase of the study, we moved forward the high-throughput technologies. We applied a whole genome approach, aCGH, to identify new chromosomal regions and genes related to HNSCC and to correlate these genomic imbalances with clinical-pathologic features of the patients (Chapter 3.2). In order to achieve these goals, we analyzed 104 HNSCC patients (the OSCC patients previously analyzed were included plus new samples from hypopharynx, larynx, supraglottis and epiglottis). Several copy number gains and losses in all chromosomes were detected (Chapter 3.2.1), being the chromosomes 3, 5, 7, 8, 9, 11, 12, 14, 15, 16, 17, 18, 19, 20 and 22 the most frequently altered. Copy number gains were commonly observed at chromosomes 3q, 5p, 7p, 7q, 8q, 11q, 12p, 14, 15, 16, 17, 19, 20, 22 and copy number losses at chromosomes 3p, 8p, 9p, 11qter and 18. By applying these genomic patterns of

CNVs we build a robust genomic model for predicting the development of HNSCC recurrence/metastasis. This predictive model showed a good accuracy (> 80%) and was validated in an independent population from TCGA data portal. This predictive genomic model presented three phases and comprises chromosomal regions from 5p, 6p, 8p, 9p, 11q, 12q, 15q and 17p, where several upstream and downstream members of signaling pathways are mapped, enhancing cell proliferation and invasion, namely components of PI3K/Akt, mTOR, Wnt, Hedgehog, Hippo, Notch, MAP/ERK, regulation of telomerase, cytoskeletal, metabolism and DNA repair. With this genomic predictive model, it is possible to identify newly diagnosed HNSCC patients at risk of developing recurrence/metastasis and therefore, monitor them closely, avoid or perform early diagnosis of the recurrences and even provide more aggressive treatment.

In both studies using either MLPA (Chapter 3.1) or aCGH (Chapter 3.2), we verified that the most common altered genes are mapped in chromosomes 3, 8, 9, 11, 16, 17, 18 and 19. Additionally, as expected, whole-genome aCGH technique revealed new chromosomal regions and genes whose role as players in the HNSCC development and behavior needs to be explored in larger independent cohorts.

Since we observed a pattern of imbalances using aCGH, e.g., frequent CNAs observed in the same chromosomal regions in several of our 104 HNSCC patients, we raised the hypothesis that the imbalances observed were caused by non-random breaks, which could be due to the presence of low copy repeat DNA sequences localized at or in close proximity to the identified breakpoint sites (Chapter 3.2.2). In our cohort, we observed frequent non-random break events clustered in chromosomes 12p, 8p, 3q, 14q, 6p, 4q, Xq and 8q. Chromosomes 6, 14, 3, 8 and X exhibited a higher susceptibility to breaks in different sites. This pattern of highly localized rearrangements in these clinical heterogeneous HNSCC patients seems to be non-random since it targets some chromosomes more than other, leading to genomic instability and eventually to tumor progression. We used the Repeated Masker track and joined fragments of interrupted repeats of the UCSC genome Browser to determine repeat DNA elements in the breakpoint site, 100 bp, 300 bp, 500 bp, 100 kb, 500 kb, 1 Mb and 10 Mb upstream and downstream flanking each breakpoint. The probability to find repeat DNA elements showed a maximum 100 kb downstream and 100 kb to 10 Mb of breakpoint. LINEs, SINEs and LTRs were the most frequent repeat elements identified. However, less than 14% of breakpoint sites have DNA repeat elements, which seems to show that the

presence of DNA repeats is not the only reason for the non-random pattern of break events observed in HNSCC.

3rd Phase

Genomic studies do not represent all aspects of cellular biology, having integrative models higher predictive power comparatively with each individual dataset [Goodspeed et al. 2016]. Thus, in pursuit of a multi-biomarker strategy for HNSCC patients and with the belief that only with omic integration it will be possible to face the high mortality and morbidity of this neoplasm, we move to a third phase of this study - a transcriptome approach (Chapter 3.3), in order to complement the genomic and epigenetic data of our patients. In order to address this transcriptome signature, we performed gene expression microarray analyses in 48 OSCC patients (34 tongue tumors and 14 floor of the mouth tumors), which were previously genomically and epigenetically characterized. We identified a gene expression signature with capability to discriminate between tumor sites (tongue vs. floor of the mouth). A set of 10 differently expressed genes in both tumor locations seems to be predictor of outcome in OSCC and may be useful for the development of novel diagnostic markers and therapeutic modalities. Thus, among the differently expression of these 10 genes, we highlighted *HIF3A* and *HOXA10* down-regulation and *LOX* up-regulation in floor of the mouth tumors comparatively to tongue tumors, which seems to be related with a highly aggressive behavior of tongue tumors, opening the door to differently treat these tumors accordingly to their anatomic locations and consequently gene expression profile.

The development of genome-wide technologies has opened the possibility of identifying, simultaneously, several alterations at genetic, epigenetic and gene expression level of the cancer cells. The major question that rose with this study was which combination of omics alterations at the different analyzed levels could be integrated to build a robust and reliable set of prognosis biomarkers, helping in the routine clinical management of the patients. Further studies are needed to conclude this integration of the omic data and to validate in different cohort the identified biomarkers.

Whole genome technologies and computational interpretation and correlation of molecular data with clinical-pathological features of the patients might become revolutionary to the clinical management of HNSCC patients.

Initially, genetic data did not drive diagnosis but had a confirmatory role, being now a great challenge to convert pathogenic genetic data into a diagnostic tool that, in combination with clinical observation, could help and guide the clinical decisions and long-term management of the patients [Katsanis and Katsanis 2013].

4th Phase

In the last phase of this work, we applied genomic and epigenetic technologies to study specific selected paradigmatic patients (Chapter 3.4.1, 3.4.2, 3.4.3) in order to molecularly characterize these patients and individually explore the improvement that this molecular knowledge could bring to routine clinical practice and therapeutic decision-making.

In the first case report (Chapter 3.4.1) we analyzed two patients with tongue squamous cell carcinoma, being the first patient also diagnosed with a simultaneous leukoplakia while the second patient developed an erythroleukoplakia following the primary tumor treatment. In both tumor and potentially malignant lesions, we identified common alterations in chromosomal regions and genes, namely *FBXL5*, *UGT2B15*, *UGT2B28*, *KANSL1*, *GSTT1* and *DUSP22*, being some of these typical aberrations described in oral cancer and others linked to chemoradioresistance. Additionally, we also highlighted some putative genes related to hallmarks of malignancy that could have an important role in predicting the progression of leukoplakia and erythroleukoplakia to squamous cell carcinoma, namely gains in *BNIPL*, *MCL1*, *STAG2*, *CSPP1* and *ZNRF3* genes.

Several oral cancers and their related complications could be preventable through early detection, unfortunately, oral potentially malignant lesions are usually misdiagnosed, which leads to delay of treatment and consequently malignant transformation [Mortazavi et al. 2014]. At present, there are no reliable clinical, pathological or molecular biomarkers to predict malignant transformation that can be used in clinical practice to prevent this event in individual patients [van der Waal 2014]. Thus, since these lesions present a risk of malignancy, it is crucial to identify molecular biomarkers with the ability to predict their behavior over the years and in this sense to make early diagnosis and treatment a reality for every patient. With genomic data of our two reported patients, we attempt to give a step forward in the identification of several

putative biomarkers of malignant transformation and contribute to the creation of a clinical and molecular follow-up program for these lesions.

In the second case report (Chapter 3.4.2) we analyzed an 88-year-old woman diagnosed with synchronous bilateral tongue carcinoma and without the traditional risk factors related to oral cancer - alcohol, tobacco or presence of HPV. In both tumors from left and right side of the tongue, we identified several molecular traits similar to those described for tumors of the oral cavity, such as gains at 3q, 5p, 7p, and 8q and losses at 3p, 8p, 9p, 11q, 18q and 21q. Moreover, specific genomic and epigenetic signatures for each of these two tumors were also observed allowing its molecular discrimination, which could eventually be associated with the laterality of these tumors, the synchronism and the tongue site-specific anatomic site. Several CNAs and methylation patterns identified in the tumors of this patient have already been associated with clinical outcome, such as *CREM*, *GATA5*, *MCCC1*, *MGMT*, *MYC*, *NEFL*, *PAX5*, *PTK2*, *SEMA3F* and *WT1*, which reinforces the importance of accurately establish diagnostic, prognostic and, ultimately patient selection biomarkers for the choice of treatment options and general clinical management of the patients. In this patient, we showed the molecular heterogeneity of oral cavity tumors even in the same patient and anatomic site, which could be the key to explain the different outcomes of oral tumor patients and consequently the need of different therapeutic options.

In the third case report (Chapter 3.4.3) we analyzed a 49-year-old man with a primary left side tongue tumor, a local recurrence after 49-months of follow-up and following three months, a second primary tumor at the pharyngoesophageal region. The three tumors of this patient present a clonal origin since they shared several chromosomal imbalances, such as losses at 1p, 2p, 3p, 4q, 5q, 6q, 7q, 8p, 10p, 11q, 12p, 12q, 13q, 15q, 16p, 16q, 17p, 17q, 18q, 19p, 19q, 21q and Xp and gains at 3q, 7q, 14q and 15q, where genes related to important functional mechanisms of tumorigenesis are mapped. Moreover, we identified some putative candidate biomarkers of prognosis, relapse risk and treatment response that could have a great impact in the clinical practice of HNSCC patients.

The study of these four patients opens new directions of research, addressing the role of the highlighted genes in oral cancer development and progression, reinforcing the molecular and clinical heterogeneity of these tumors that should be treated as different entities, where the genomic profile seems to have a great potential to help in the early detection, prognosis and selection of treatment. We applied array-CGH to all

these patients, since this technology offers the potential for high throughput and high-resolution genome analysis, identifying the most relevant chromosomal regions and genes with imbalances. In this way, a first whole genome screening of imbalances related to HNSCC was made; opening the door to the selection of directed and specific probe panels with clinically useful genes that could be used in routine clinical practice, in a cost- and time-effective way.

We finalized our work by performing two studies using two commercial OSCC cell lines and three primary cultures established in our laboratory from OSCC patients. A comprehensive (cyto)genomic and epigenetic characterization of HSC-3 and BIRC 10 cell lines, as well as of three newly established cell cultures from tongue and retromolar trigone primary tumors were performed (Chapters 3.5.1, 3.5.2).

In cancer research, tumor-derived cell lines are frequently used as models since they carry imbalances that arose in the tumor from which they were derived [Goodspeed et al. 2016]. Cancer cell lines are used to study several biological processes and have been widely used in pharmacogenomics studies. Tumor-derived cell lines have been used since long time ago as drug discovery tools, but only recently have the researchers realized the presence of genomic heterogeneity across the human cancer patient population, and therefore across tumor-derived cell lines, which plays a significant role in the response to treatment [S. V. Sharma et al. 2010b]. Cancer cell lines genomically well characterized allows linking genomic profiles to therapeutic response. With the improvement of high-throughput data generation techniques, we could "omic profile" the human cancer cell lines and consequently choose the appropriate *in vitro* genomic model for studying a primary tumor or specific aspects of tumor biology [Goodspeed et al. 2016].

Initially, chromosomal imbalances such as aneuploidies and segmental aneusomies had to be larger than 3–5 Mb in size to be detected by conventional cytogenetics techniques [Shaffer and Lupski 2000]. However, significant improvements were made in chemistry and microscopy in order to increase the resolution of cytogenetics through the development of multi-probe FISH and chromosomal CGH [Katsanis and Katsanis 2013]. Nowadays we are in the era of genomic technologies that are able to detect genomic variations in patients with high accuracy and at reduced cost, offering the promise of personalized medicine [Katsanis and Katsanis 2013]. Additionally, it is expected that, in the future, researchers will be working with cell lines that were routinely molecularly characterized in order to increase the quality of the conclusions and the validity of the

studies. It is important to note that when the genomes of cell lines are unstable, such as in HeLa, it is often needed to update their characterization, guaranteeing the integrity of the cell lines and the quality of the biological insights derived from them [Landry et al. 2013].

With the characterization of the two cell lines and the three primary cell cultures (Chapters 3.5.1, 3.5.2), we identified the most frequent numerical and structural chromosomal rearrangements in these cells as well as the genes with CNAs and methylation, thus enriching the resources available for oral cancer research, especially for testing therapeutic agents. Overall, these cells revealed complex karyotypes with multiple structural rearrangements involving the great majority of chromosomes. The most common observed rearrangements in HSC-3 and BICR 10 cell lines were translocations in centromeric/near-centromeric regions. In primary cell cultures the most common rearrangements observed were isochromosomes and derivative chromosomes from translocations.

After this comprehensive characterization, these cells became a more attractive and helpful *in vitro* resource to understand the evolution of OSCC and to study the molecular mechanisms behind its progression.

Conclusions

With this work, we provide new insights in HNSCC field, namely the identification of different molecular signatures associated to patients' prognosis and correlated to different clinical-pathological features, such as anatomic site, tumor stage and risk factors. The molecular characterization of HNSCC recurring to an integrative omic approach in pursuit of accurate and robust prognostic biomarkers was the golden target of this project that gives a step forward through the analysis of different omics results for the same HNSCC patients. A validation of these identified biomarkers using online available data of geographical different HNSCC patient from the TCGA portal proved the predictive power of the identified biomarkers as well as the utility of their translation to clinical practice.

We performed a comprehensive and robust molecular characterization of HNSCC, identifying new prognostic biomarkers with potential to lead to the creation of new screening methods to early detection, diagnosis and ultimately to develop new therapeutic strategies. Our main conclusions are:

- The genetic, epigenetic and transcriptomic characterization of our cohort showed different molecular signatures related to tumor stage, anatomic site and tobacco consumption (Chapters 3.1, 3.3);
- Specific genomic and epigenetic signatures associated with HNSCC patients' prognosis and survival were identified (Chapters 3.1, 3.2);
- Genomic and epigenetic signatures with capability to predict recurrences and/or metastasis (Chapters 3.1, 3.2) were identified, presenting potentially great impact in the clinical management of patients;
- Genomic and epigenetic imbalances in non-tumor samples (from surgical margins) and scrapped cells (Chapters 3.1.1, 3.1.4) were observed, which reinforce the importance of molecularly analyze the high-risk patients even before the visible morphological changes and also the suspicious lesions in order to early diagnose these tumors and their recurrences.

The identification of these biomarkers/signatures of prognosis and survival opens the door of personalized medicine for these clinically and molecularly heterogeneous HNSCC tumors. Due to variations in the sample population, small sample size of each anatomic subsites, different experimental designs and analysis methods, different studies for the same disease can exhibit different candidate biomarkers, which reinforces the importance of the integrated analysis of different omics data to identify consistent biomarkers. Thus, considering the molecular integration of the highlighted biomarkers in this study and with further validation in different HNSCC cohorts, it might be possible to develop a probe panel with costs compatible with routine diagnostic to be used in patients' follow-up.

The development of a non-invasive approach to repeatedly search these biomarkers is extremely important. Our attempt using scrapped cells did not present the expected results (Chapter 3.1.4), further studies need to be do with our clinicians. Biofluids are an alternative that could be considered in the future.

We proved the clinical utility of this molecular characterization since we analyzed four paradigmatic HNSCC patients (Chapter 3.4) and we identified several chromosomal regions and genes that could help and guide the clinical management, namely in terms of type or intensification of treatment modalities, since we identified some genomic patterns that seem to be related to radiation and/or chemotherapy

resistance and also several genomic and epigenetic patterns related to patients' prognosis and survival. We took advantage of high-throughput technologies not only to molecularly characterize the HNSCC patients but also to characterize commercial HNSCC cell lines and primary cell cultures established from our patients (Chapter 3.5), which allowed us to understand the ploidy and the complex structural chromosomal rearrangements of these tumors. This comprehensive characterization enables cell models for further studies both in radiation and pharmacogenomics fields, as well as to understand the basic biological mechanisms of HNSCC tumor development and progression.

5. Future Perspectives

Taking into account our results, there are several paths that should be explored in the future, such as:

- The integration of our omic data;
- The study of other omics approaches, namely proteomics and metabolomics;
- A validation of the identified signatures in geographically separate HNSCC populations;
- A validation in large cohorts from each anatomic site, tumor stage and risk factors;
- The selection of directed and specific probe panels with clinically useful genes that could be used in routine clinical practice, in a cost- and time-effective way;
- The identification of the highlighted biomarkers in the biofluids, namely in the circulating tumor DNA - "liquid biopsies", of HNSCC patients, in order to non-invasively search these biomarkers in high-risk populations and during/after patients' treatment;
- The establishment and comprehensive molecular characterization of more primary cell cultures to perform radiation and pharmacological studies using these molecularly well-characterized cell cultures.

It is important to emphasize that the follow-up of patients is crucial, thus our HNSCC cohort should be continually followed up in order to allow the establishment of strong genotype/phenotype correlations.

As a final remark, it is important to mention that nowadays, although there are some studies with full characterization of HNSCC at different biological levels, using different omics approaches, the challenge is to move forward from these huge molecular profiles to clinical practice translation. Integrate and validate the data from different biological levels, namely, genomic, epigenetic and transcriptomic are expected to produce new insights of biological behavior of these tumors, improving their clinical classification and identifying diagnostic and prognostic biomarker signatures. New study designs using bioinformatics algorithms and single cell DNA/RNA sequencing of tumors and biofluids should allow the identification and characterization of distinct intratumor subpopulations, which could indicate the distinct contributions of each of these populations to tumor pathogenesis.

6. References

Ai, L., Vo, Q. N., Zuo, C., Li, L., Ling, W., Suen, J. Y., Hanna, E., Brown, K. D., and Fan, C. Y. (2004), 'Ataxia-telangiectasia-mutated (ATM) gene in head and neck squamous cell carcinoma: promoter hypermethylation with clinical correlation in 100 cases', *Cancer Epidemiol Biomarkers Prev*, 13 (1), 150-6.

Akiyama, Y., Watkins, N., Suzuki, H., Jair, K. W., van Engeland, M., Esteller, M., Sakai, H., Ren, C. Y., Yuasa, Y., Herman, J. G., and Baylin, S. B. (2003), 'GATA-4 and GATA-5 transcription factor genes and potential downstream antitumor target genes are epigenetically silenced in colorectal and gastric cancer', *Mol Cell Biol*, 23 (23), 8429-39.

Aksu, G., Karadeniz, A., Saynak, M., Fayda, M., Kadehci, Z., and Kocaelli, H. (2006), 'Treatment results and prognostic factors in oral tongue cancer: analysis of 80 patients', *Int J Oral Maxillofac Surg*, 35 (6), 506-13.

Ambatipudi, Srikant, Gerstung, Moritz, Gowda, Ravindra, Pai, Prathamesh, Borges, Anita M., Schaffer, Alejandro A., Beerenwinkel, Niko, and Mahimkar, Manoj B. (2011), 'Genomic Profiling of Advanced-Stage Oral Cancers Reveals Chromosome 11q Alterations as Markers of Poor Clinical Outcome', *PLoS One*, 6 (2), e17250.

Ang, K. K., Berkey, B. A., Tu, X., Zhang, H. Z., Katz, R., Hammond, E. H., Fu, K. K., and Milas, L. (2002), 'Impact of epidermal growth factor receptor expression on survival and pattern of relapse in patients with advanced head and neck carcinoma', *Cancer Res*, 62 (24), 7350-6.

Arantes, L. M., de Carvalho, A. C., Melendez, M. E., Centrone, C. C., Gois-Filho, J. F., Toporcov, T. N., Caly, D. N., Tajara, E. H., Goloni-Bertollo, E. M., Carvalho, A. L., and Gencapo (2015), 'Validation of methylation markers for diagnosis of oral cavity cancer', *Eur J Cancer*, 51 (5), 632-41.

Argiris, A., Karamouzis, M. V., Raben, D., and Ferris, R. L. (2008), 'Head and neck cancer', *Lancet*, 371 (9625), 1695-709.

Ashman, J. N., Patmore, H. S., Condon, L. T., Cawkwell, L., Stafford, N. D., and Greenman, J. (2003), 'Prognostic value of genomic alterations in head and neck squamous cell carcinoma detected by comparative genomic hybridisation', *Br J Cancer*, 89 (5), 864-9.

Awan, Kh (2014), 'Oral Cancer: Early Detection is Crucial', *J Int Oral Health*, 6 (5), i-ii.

Ayoub, C., Wasylyk, C., Li, Y., Thomas, E., Marisa, L., Robe, A., Roux, M., Abecassis, J., de Reynies, A., and Wasylyk, B. (2010), 'ANO1 amplification and expression in HNSCC with a high propensity for future distant metastasis and its functions in HNSCC cell lines', *Br J Cancer*, 103 (5), 715-26.

Babic, A. M., Kireeva, M. L., Kolesnikova, T. V., and Lau, L. F. (1998), 'CYR61, a product of a growth factor-inducible immediate early gene, promotes angiogenesis and tumor growth', *Proc Natl Acad Sci U S A*, 95 (11), 6355-60.

Bagan, J. V. and Scully, C. (2008), 'Recent advances in Oral Oncology 2007: epidemiology, aetiopathogenesis, diagnosis and prognostication', *Oral Oncol*, 44 (2), 103-8.

Bahl, A., Kumar, P., Dar, L., Mohanti, B. K., Sharma, A., Thakar, A., Karthikeyan, V., Sikka, K., Singh, C., Poo, K., and Lodha, J. (2013), 'Prevalence and trends of human papillomavirus in oropharyngeal cancer in a predominantly north Indian population', *Head Neck*.

Barnes, N. L., Khavari, S., Boland, G. P., Cramer, A., Knox, W. F., and Bundred, N. J. (2005), 'Absence of HER4 expression predicts recurrence of ductal carcinoma in situ of the breast', *Clin Cancer Res*, 11 (6), 2163-8.

Bauer, V. L., Braselmann, H., Henke, M., Mattern, D., Walch, A., Unger, K., Baudis, M., Lassmann, S., Huber, R., Wienberg, J., Werner, M., and Zitzelsberger, H. F. (2008), 'Chromosomal changes characterize head and neck cancer with poor prognosis', *J Mol Med (Berl)*, 86 (12), 1353-65.

Beck, T. N., Kaczmar, J., Handorf, E., Nikonova, A., Dubyk, C., Peri, S., Lango, M., Ridge, J. A., Serebriiskii, I. G., Burtness, B., Golemis, E. A., and Mehra, R. (2015), 'Phospho-T356RB1 predicts survival in HPV-negative squamous cell carcinoma of the head and neck', *Oncotarget*, 6 (22), 18863-74.

Bei, L., Lu, Y., Bellis, S. L., Zhou, W., Horvath, E., and Eklund, E. A. (2007), 'Identification of a HoxA10 activation domain necessary for transcription of the gene encoding beta3 integrin during myeloid differentiation', *J Biol Chem*, 282 (23), 16846-59.

6. References

Belbin, T. J., Singh, B., Barber, I., Socci, N., Wenig, B., Smith, R., Prystowsky, M. B., and Childs, G. (2002), 'Molecular classification of head and neck squamous cell carcinoma using cDNA microarrays', *Cancer Res,* 62 (4), 1184-90.

Belbin, T. J., Schlecht, N. F., Smith, R. V., Adrien, L. R., Kawachi, N., Brandwein-Gensler, M., Bergman, A., Chen, Q., Childs, G., and Prystowsky, M. B. (2008), 'Site-specific molecular signatures predict aggressive disease in HNSCC', *Head Neck Pathol,* 2 (4), 243-56.

Bello, I. O., Soini, Y., and Salo, T. (2010), 'Prognostic evaluation of oral tongue cancer: means, markers and perspectives (II)', *Oral Oncol,* 46 (9), 636-43.

Bensaad, K. and Vousden, K. H. (2007), 'p53: new roles in metabolism', *Trends Cell Biol,* 17 (6), 286-91.

Bezabeh, T., Odlum, O., Nason, R., Kerr, P., Sutherland, D., Patel, R., and Smith, I. C. (2005), 'Prediction of treatment response in head and neck cancer by magnetic resonance spectroscopy', *AJNR Am J Neuroradiol,* 26 (8), 2108-13.

Bitu, C. C., Destro, M. F., Carrera, M., da Silva, S. D., Graner, E., Kowalski, L. P., Soares, F. A., and Coletta, R. D. (2012), 'HOXA1 is overexpressed in oral squamous cell carcinomas and its expression is correlated with poor prognosis', *BMC Cancer,* 12, 146.

Bloching, M., Reich, W., Schubert, J., Grummt, T., and Sandner, A. (2007), 'The influence of oral hygiene on salivary quality in the Ames Test, as a marker for genotoxic effects', *Oral Oncol,* 43 (9), 933-9.

Bockmuhl, U., Ishwad, C. S., Ferrell, R. E., and Gollin, S. M. (2001), 'Association of 8p23 deletions with poor survival in head and neck cancer', *Otolaryngol Head Neck Surg,* 124 (4), 451-5.

Bockmuhl, U., Schluns, K., Kuchler, I., Petersen, S., and Petersen, I. (2000), 'Genetic imbalances with impact on survival in head and neck cancer patients', *Am J Pathol,* 157 (2), 369-75.

Bockmuhl, U., Schluns, K., Schmidt, S., Matthias, S., and Petersen, I. (2002), 'Chromosomal alterations during metastasis formation of head and neck squamous cell carcinoma', *Genes Chromosomes Cancer,* 33 (1), 29-35.

Bockmuhl, U., Petersen, S., Schmidt, S., Wolf, G., Jahnke, V., Dietel, M., and Petersen, I. (1997), 'Patterns of chromosomal alterations in metastasizing and nonmetastasizing primary head and neck carcinomas', *Cancer Res,* 57 (23), 5213-6.

Bonner, J. A., Harari, P. M., Giralt, J., Azarnia, N., Shin, D. M., Cohen, R. B., Jones, C. U., Sur, R., Raben, D., Jassem, J., Ove, R., Kies, M. S., Baselga, J., Youssoufian, H., Amellal, N., Rowinsky, E. K., and Ang, K. K. (2006), 'Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck', *N Engl J Med,* 354 (6), 567-78.

Borkosky, S. S., Gunduz, M., Nagatsuka, H., Beder, L. B., Gunduz, E., Ali, M. A., Rodriguez, A. P., Cilek, M. Z., Tominaga, S., Yamanaka, N., Shimizu, K., and Nagai, N. (2009), 'Frequent deletion of ING2 locus at 4q35.1 associates with advanced tumor stage in head and neck squamous cell carcinoma', *J Cancer Res Clin Oncol,* 135 (5), 703-13.

Bornstein, S., White, R., Malkoski, S., Oka, M., Han, G., Cleaver, T., Reh, D., Andersen, P., Gross, N., Olson, S., Deng, C., Lu, S. L., and Wang, X. J. (2009), 'Smad4 loss in mice causes spontaneous head and neck cancer with increased genomic instability and inflammation', *J Clin Invest,* 119 (11), 3408-19.

Bose, P., Brockton, N. T., and Dort, J. C. (2013), 'Head and neck cancer: from anatomy to biology', *Int J Cancer,* 133 (9), 2013-23.

Bossi, P., Locati, L., and Licitra, L. (2013), 'Emerging tyrosine kinase inhibitors for head and neck cancer', *Expert Opin Emerg Drugs,* 18 (4), 445-59.

Bouquot, J. E., Weiland, L. H., and Kurland, L. T. (1988), 'Leukoplakia and carcinoma in situ synchronously associated with invasive oral/oropharyngeal carcinoma in Rochester, Minn., 1935-1984', *Oral Surg Oral Med Oral Pathol,* 65 (2), 199-207.

Bova, R. J., Quinn, D. I., Nankervis, J. S., Cole, I. E., Sheridan, B. F., Jensen, M. J., Morgan, G. J., Hughes, C. J., and Sutherland, R. L. (1999), 'Cyclin D1 and p16INK4A expression predict reduced survival in carcinoma of the anterior tongue', *Clin Cancer Res,* 5 (10), 2810-9.

Braakhuis, B. J., Brakenhoff, R. H., and Leemans, C. R. (2005), 'Second field tumors: a new opportunity for cancer prevention?', *Oncologist,* 10 (7), 493-500.

Braakhuis, Boudewijn J. M., Bloemena, Elisabeth, Leemans, C. RenÃ©, and Brakenhoff, Ruud H. (2010), 'Molecular analysis of surgical margins in head and neck cancer: More than a marginal issue', *Oral Oncol*, 46 (7), 485-91.

Brauswetter, D., Danos, K., Gurbi, B., Felegyhazi, E. F., Birtalan, E., Meggyeshazi, N., Krenacs, T., Tamas, L., and Petak, I. (2016), 'Copy number gain of PIK3CA and MET is associated with poor prognosis in head and neck squamous cell carcinoma', *Virchows Arch*, 468 (5), 579-87.

Brekken, R. A., Overholser, J. P., Stastny, V. A., Waltenberger, J., Minna, J. D., and Thorpe, P. E. (2000), 'Selective inhibition of vascular endothelial growth factor (VEGF) receptor 2 (KDR/Flk-1) activity by a monoclonal anti-VEGF antibody blocks tumor growth in mice', *Cancer Res*, 60 (18), 5117-24.

Bremmer, J. F., Graveland, A. P., Brink, A., Braakhuis, B. J., Kuik, D. J., Leemans, C. R., Bloemena, E., van der Waal, I., and Brakenhoff, R. H. (2009), 'Screening for oral precancer with noninvasive genetic cytology', *Cancer Prev Res (Phila)*, 2 (2), 128-33.

Bremmer, J. F., Braakhuis, B. J., Ruijter-Schippers, H. J., Brink, A., Duarte, H. M., Kuik, D. J., Bloemena, E., Leemans, C. R., van der Waal, I., and Brakenhoff, R. H. (2005), 'A noninvasive genetic screening test to detect oral preneoplastic lesions', *Lab Invest*, 85 (12), 1481-8.

Burke, J. R., Hura, G. L., and Rubin, S. M. (2012), 'Structures of inactive retinoblastoma protein reveal multiple mechanisms for cell cycle control', *Genes Dev*, 26 (11), 1156-66.

Bussink, J., van der Kogel, A. J., and Kaanders, J. H. (2008), 'Activation of the PI3-K/AKT pathway and implications for radioresistance mechanisms in head and neck cancer', *Lancet Oncol*, 9 (3), 288-96.

Cabanillas, R. and Llorente, J. L. (2009), 'The Stem Cell Network model: clinical implications in cancer', *Eur Arch Otorhinolaryngol*, 266 (2), 161-70.

Cabanillas, R., Astudillo, A., Valle, M., de la Rosa, J., Alvarez, R., Duran, N. S., and Cadinanos, J. (2013), 'Novel germline CDKN2A mutation associated with head and neck squamous cell carcinomas and melanomas', *Head Neck*, 35 (3), E80-4.

Cabelguenne, A., Blons, H., de Waziers, I., Carnot, F., Houllier, A. M., Soussi, T., Brasnu, D., Beaune, P., Laccourreye, O., and Laurent-Puig, P. (2000), 'p53 alterations predict tumor response to neoadjuvant chemotherapy in head and neck squamous cell carcinoma: a prospective series', *J Clin Oncol*, 18 (7), 1465-73.

Califano, J., Westra, W. H., Meininger, G., Corio, R., Koch, W. M., and Sidransky, D. (2000), 'Genetic progression and clonal relationship of recurrent premalignant head and neck lesions', *Clin Cancer Res*, 6 (2), 347-52.

Califano, J., van der Riet, P., Westra, W., Nawroz, H., Clayman, G., Piantadosi, S., Corio, R., Lee, D., Greenberg, B., Koch, W., and Sidransky, D. (1996), 'Genetic progression model for head and neck cancer: implications for field cancerization', *Cancer Res*, 56 (11), 2488-92.

Campbell, B. H., Spinelli, K., Marbella, A. M., Myers, K. B., Kuhn, J. C., and Layde, P. M. (2004), 'Aspiration, weight loss, and quality of life in head and neck cancer survivors', *Arch Otolaryngol Head Neck Surg*, 130 (9), 1100-3.

Cancer Genome Atlas, Network (2015), 'Comprehensive genomic characterization of head and neck squamous cell carcinomas', *Nature*, 517 (7536), 576-82.

Canel, M., Secades, P., Rodrigo, J. P., Cabanillas, R., Herrero, A., Suarez, C., and Chiara, M. D. (2006), 'Overexpression of focal adhesion kinase in head and neck squamous cell carcinoma is independent of fak gene copy number', *Clin Cancer Res*, 12 (11 Pt 1), 3272-9.

Capaccio, P., Pruneri, G., Carboni, N., Pagliari, A. V., Quatela, M., Cesana, B. M., and Pignataro, L. (2000), 'Cyclin D1 expression is predictive of occult metastases in head and neck cancer patients with clinically negative cervical lymph nodes', *Head Neck*, 22 (3), 234-40.

Cardoso, A. R., Oliveira, M., Amorim, A., and Azevedo, L. (2016), 'Major influence of repetitive elements on disease-associated copy number variants (CNVs)', *Hum Genomics*, 10 (1), 30.

Carneiro, A., Isinger, A., Karlsson, A., Johansson, J., Jonsson, G., Bendahl, P. O., Falkenback, D., Halvarsson, B., and Nilbert, M. (2008), 'Prognostic impact of array-based genomic profiles in esophageal squamous cell cancer', *BMC Cancer*, 8, 98.

6. References

Carrera, M., Bitu, C. C., de Oliveira, C. E., Cervigne, N. K., Graner, E., Manninen, A., Salo, T., and Coletta, R. D. (2015), 'HOXA10 controls proliferation, migration and invasion in oral squamous cell carcinoma', *Int J Clin Exp Pathol,* 8 (4), 3613-23.

Cha, J. D., Kim, H. J., and Cha, I. H. (2011), 'Genetic alterations in oral squamous cell carcinoma progression detected by combining array-based comparative genomic hybridization and multiplex ligation-dependent probe amplification', *Oral Surg Oral Med Oral Pathol Oral Radiol Endod,* 111 (5), 594-607.

Chang, S. S. and Califano, J. (2008), 'Current status of biomarkers in head and neck cancer', *J Surg Oncol,* 97 (8), 640-3.

Chen, A. J., Zhou, G., Juan, T., Colicos, S. M., Cannon, J. P., Cabriera-Hansen, M., Meyer, C. F., Jurecic, R., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Fletcher, F., Tan, T. H., and Belmont, J. W. (2002), 'The dual specificity JKAP specifically activates the c-Jun N-terminal kinase pathway', *J Biol Chem,* 277 (39), 36592-601.

Chen, H. C., Yang, C. M., Cheng, J. T., Tsai, K. W., Fu, T. Y., Liou, H. H., Tseng, H. H., Lee, J. H., Li, G. C., Wang, J. S., Hou, Y. Y., Weng, T. J., and Ger, L. P. (2015a), 'Global DNA hypomethylation is associated with the development and poor prognosis of tongue squamous cell carcinoma', *J Oral Pathol Med*.

Chen, K., Sawhney, R., Khan, M., Benninger, M. S., Hou, Z., Sethi, S., Stephen, J. K., and Worsham, M. J. (2007), 'Methylation of multiple genes as diagnostic and therapeutic markers in primary head and neck squamous cell carcinoma', *Arch Otolaryngol Head Neck Surg,* 133 (11), 1131-8.

Chen, L., Zhou, W., Zhang, L., and Zhang, F. (2014), 'Genome architecture and its roles in human copy number variation', *Genomics Inform,* 12 (4), 136-44.

Chen, X., Liu, L., Mims, J., Punska, E. C., Williams, K. E., Zhao, W., Arcaro, K. F., Tsang, A. W., Zhou, X., and Furdui, C. M. (2015b), 'Analysis of DNA methylation and gene expression in radiation-resistant head and neck tumors', *Epigenetics,* 10 (6), 545-61.

Chen, Y. J., Lin, S. C., Kao, T., Chang, C. S., Hong, P. S., Shieh, T. M., and Chang, K. W. (2004), 'Genome-wide profiling of oral squamous cell carcinoma', *J Pathol,* 204 (3), 326-32.

Chen, Z. G. (2009), 'The cancer stem cell concept in progression of head and neck cancer', *J Oncol*, 2009, 894064.

Chinn, S. B. and Myers, J. N. (2015), 'Oral Cavity Carcinoma: Current Management, Controversies, and Future Directions', *J Clin Oncol,* 33 (29), 3269-76.

Chmelarova, M., Dvorakova, E., Spacek, J., Laco, J., Mzik, M., and Palicka, V. (2013), 'Promoter methylation of GATA4, WIF1, NTRK1 and other selected tumour suppressor genes in ovarian cancer', *Folia Biol (Praha),* 59 (2), 87-92.

Choi, S. and Myers, J. N. (2008), 'Molecular pathogenesis of oral squamous cell carcinoma: implications for therapy', *J Dent Res,* 87 (1), 14-32.

Chuang, S. C., Jenab, M., Heck, J. E., Bosetti, C., Talamini, R., Matsuo, K., Castellsague, X., Franceschi, S., Herrero, R., Winn, D. M., La Vecchia, C., Morgenstern, H., Zhang, Z. F., Levi, F., Dal Maso, L., Kelsey, K., McClean, M. D., Vaughan, T., Lazarus, P., Muscat, J., Ramroth, H., Chen, C., Schwartz, S. M., Eluf-Neto, J., Hayes, R. B., Purdue, M., Boccia, S., Cadoni, G., Zaridze, D., Koifman, S., Curado, M. P., Ahrens, W., Benhamou, S., Matos, E., Lagiou, P., Szeszenia-Dabrowska, N., Olshan, A. F., Fernandez, L., Menezes, A., Agudo, A., Daudt, A. W., Merletti, F., Macfarlane, G. J., Kjaerheim, K., Mates, D., Holcatova, I., Schantz, S., Yu, G. P., Simonato, L., Brenner, H., Mueller, H., Conway, D. I., Thomson, P., Fabianova, E., Znaor, A., Rudnai, P., Healy, C. M., Ferro, G., Brennan, P., Boffetta, P., and Hashibe, M. (2012), 'Diet and the risk of head and neck cancer: a pooled analysis in the INHANCE consortium', *Cancer Causes Control,* 23 (1), 69-88.

Chung, C. H., Parker, J. S., Karaca, G., Wu, J., Funkhouser, W. K., Moore, D., Butterfoss, D., Xiang, D., Zanation, A., Yin, X., Shockley, W. W., Weissler, M. C., Dressler, L. G., Shores, C. G., Yarbrough, W. G., and Perou, C. M. (2004), 'Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression', *Cancer Cell,* 5 (5), 489-500.

Citri, A. and Yarden, Y. (2006), 'EGF-ERBB signalling: towards the systems level', *Nat Rev Mol Cell Biol,* 7 (7), 505-16.

Coffa, Jordy and Berg, Joost van den (2011), 'Analysis of MLPA Data Using Novel Software Coffalyser.NET by MRC-Holland', in Ahmed Badr Eldin (ed.), *Modern Approaches to Quality Control* (Rijeka, Croatia InTech), 125-50.

Cognetti, D. M., Weber, R. S., and Lai, S. Y. (2008), 'Head and neck cancer: an evolving treatment paradigm', *Cancer*, 113 (7 Suppl), 1911-32.

Colella, S., Richards, K. L., Bachinski, L. L., Baggerly, K. A., Tsavachidis, S., Lang, J. C., Schuller, D. E., and Krahe, R. (2008), 'Molecular signatures of metastasis in head and neck cancer', *Head Neck*, 30 (10), 1273-83.

Colevas, A. D. (2006), 'Chemotherapy options for patients with metastatic or recurrent squamous cell carcinoma of the head and neck', *J Clin Oncol*, 24 (17), 2644-52.

Conway, D. I., McMahon, A. D., Smith, K., Black, R., Robertson, G., Devine, J., and McKinney, P. A. (2010), 'Components of socioeconomic risk associated with head and neck cancer: a population-based case-control study in Scotland', *Br J Oral Maxillofac Surg*, 48 (1), 11-7.

Cooke, S. L., Pole, J. C., Chin, S. F., Ellis, I. O., Caldas, C., and Edwards, P. A. (2008), 'High-resolution array CGH clarifies events occurring on 8p in carcinogenesis', *BMC Cancer*, 8, 288.

Coon, S. W., Savera, A. T., Zarbo, R. J., Benninger, M. S., Chase, G. A., Rybicki, B. A., and Van Dyke, D. L. (2004), 'Prognostic implications of loss of heterozygosity at 8p21 and 9p21 in head and neck squamous cell carcinoma', *Int J Cancer*, 111 (2), 206-12.

Cooper, J. S., Pajak, T. F., Forastiere, A. A., Jacobs, J., Campbell, B. H., Saxman, S. B., Kish, J. A., Kim, H. E., Cmelak, A. J., Rotman, M., Machtay, M., Ensley, J. F., Chao, K. S., Schultz, C. J., Lee, N., Fu, K. K., and Radiation Therapy Oncology Group, Intergroup (2004), 'Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck', *N Engl J Med*, 350 (19), 1937-44.

COSMIC (2016), 'https://cansar.icr.ac.uk/cansar/cell-lines/HSC-3/', June, 26th

Costello, J. F., Berger, M. S., Huang, H. S., and Cavenee, W. K. (1996), 'Silencing of p16/CDKN2 expression in human gliomas by methylation and chromatin condensation', *Cancer Res*, 56 (10), 2405-10.

Cowpe, J. G., Longmore, R. B., and Green, M. W. (1988), 'Quantitative exfoliative cytology of abnormal oral mucosal smears', *J R Soc Med*, 81 (9), 509-13.

D'Souza, G., Kreimer, A. R., Viscidi, R., Pawlita, M., Fakhry, C., Koch, W. M., Westra, W. H., and Gillison, M. L. (2007), 'Case-control study of human papillomavirus and oropharyngeal cancer', *N Engl J Med*, 356 (19), 1944-56.

da Silva Veiga, L. C., Bergamo, N. A., dos Reis, P. P., Kowalski, L. P., and Rogatto, S. R. (2003), 'DNA gains at 8q23.2: a potential early marker in head and neck carcinomas', *Cancer Genet Cytogenet*, 146 (2), 110-5.

Danahey, D. G., Tobin, E. J., Schuller, D. E., Bier-Laning, C. M., Weghorst, C. M., and Lang, J. C. (1999), 'p16 mutation frequency and clinical correlation in head and neck cancer', *Acta Otolaryngol*, 119 (2), 285-8.

Dancey, J. E. (2006), 'Therapeutic targets: MTOR and related pathways', *Cancer Biol Ther*, 5 (9), 1065-73.

de Carvalho, A. C., Kowalski, L. P., Campos, A. H., Soares, F. A., Carvalho, A. L., and Vettore, A. L. (2012), 'Clinical significance of molecular alterations in histologically negative surgical margins of head and neck cancer patients', *Oral Oncol*, 48 (3), 240-8.

Di Cristofano, A. and Pandolfi, P. P. (2000), 'The multiple roles of PTEN in tumor suppression', *Cell*, 100 (4), 387-90.

Dionne, K. R., Warnakulasuriya, S., Zain, R. B., and Cheong, S. C. (2015), 'Potentially malignant disorders of the oral cavity: current practice and future directions in the clinic and laboratory', *Int J Cancer*, 136 (3), 503-15.

Doci, C. L., Mikelis, C. M., Lionakis, M. S., Molinolo, A. A., and Gutkind, J. S. (2015), 'Genetic Identification of SEMA3F as an Antilymphangiogenic Metastasis Suppressor Gene in Head and Neck Squamous Carcinoma', *Cancer Res*, 75 (14), 2937-48.

Domcke, S., Sinha, R., Levine, D. A., Sander, C., and Schultz, N. (2013), 'Evaluating cell lines as tumour models by comparison of genomic profiles', *Nat Commun*, 4, 2126.

6. References

Douglas, E. J., Fiegler, H., Rowan, A., Halford, S., Bicknell, D. C., Bodmer, W., Tomlinson, I. P., and Carter, N. P. (2004), 'Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas', *Cancer Res*, 64 (14), 4817-25.

Duvvuri, U., Shiwarski, D. J., Xiao, D., Bertrand, C., Huang, X., Edinger, R. S., Rock, J. R., Harfe, B. D., Henson, B. J., Kunzelmann, K., Schreiber, R., Seethala, R. S., Egloff, A. M., Chen, X., Lui, V. W., Grandis, J. R., and Gollin, S. M. (2012), 'TMEM16A induces MAPK and contributes directly to tumorigenesis and cancer progression', *Cancer Res*, 72 (13), 3270-81.

Egger, G., Liang, G., Aparicio, A., and Jones, P. A. (2004), 'Epigenetics in human disease and prospects for epigenetic therapy', *Nature*, 429 (6990), 457-63.

Eisenberg, E. and Levanon, E. Y. (2013), 'Human housekeeping genes, revisited', *Trends Genet*, 29 (10), 569-74.

Elrefaey, S., Massaro, M. A., Chiocca, S., Chiesa, F., and Ansarin, M. (2014), 'HPV in oropharyngeal cancer: the basics to know in clinical practice', *Acta Otorhinolaryngol Ital*, 34 (5), 299-309.

Erdem, N. F., Carlson, E. R., Gerard, D. A., and Ichiki, A. T. (2007), 'Characterization of 3 oral squamous cell carcinoma cell lines with different invasion and/or metastatic potentials', *J Oral Maxillofac Surg*, 65 (9), 1725-33.

Erler, J. T., Bennewith, K. L., Nicolau, M., Dornhofer, N., Kong, C., Le, Q. T., Chi, J. T., Jeffrey, S. S., and Giaccia, A. J. (2006), 'Lysyl oxidase is essential for hypoxia-induced metastasis', *Nature*, 440 (7088), 1222-6.

Esteller, M., Silva, J. M., Dominguez, G., Bonilla, F., Matias-Guiu, X., Lerma, E., Bussaglia, E., Prat, J., Harkes, I. C., Repasky, E. A., Gabrielson, E., Schutte, M., Baylin, S. B., and Herman, J. G. (2000), 'Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors', *J Natl Cancer Inst*, 92 (7), 564-9.

Fang, Q. G., Shi, S., Liu, F. Y., and Sun, C. F. (2014), 'Tongue squamous cell carcinoma as a possible distinct entity in patients under 40 years old', *Oncol Lett*, 7 (6), 2099-102.

Federico, M. and Bagella, L. (2011), 'Histone deacetylase inhibitors in the treatment of hematological malignancies and solid tumors', *J Biomed Biotechnol*, 2011, 475641.

Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C., and Parkin, D. M. (2010), 'Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008', *Int J Cancer*, 127 (12), 2893-917.

Ferlay, J., Steliarova-Foucher, E., Lortet-Tieulent, J., Rosso, S., Coebergh, J. W., Comber, H., Forman, D., and Bray, F. (2013), 'Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012', *Eur J Cancer*, 49 (6), 1374-403.

Ferris, R. L. (2015), 'Immunology and Immunotherapy of Head and Neck Cancer', *J Clin Oncol*, 33 (29), 3293-304.

Forastiere, A., Koch, W., Trotti, A., and Sidransky, D. (2001), 'Head and neck cancer', *N Engl J Med*, 345 (26), 1890-900.

Fortin, A., Wang, C. S., and Vigneault, E. (2009), 'Influence of smoking and alcohol drinking behaviors on treatment outcomes of patients with squamous cell carcinomas of the head and neck', *Int J Radiat Oncol Biol Phys*, 74 (4), 1062-9.

Fracchiolla, Nicola S., Pruneri, Giancarlo, Pignataro, Lorenzo, Carboni, Nadia, Capaccio, Pasquale, Boletini, Alketa, Buffa, Roberto, and Neri, Antonino (1997), 'Molecular and immunohistochemical analysis of thebcl-1/cyclin D1 gene in laryngeal squamous cell carcinomas', *Cancer*, 79, 1114-21.

Frampton, J. E. (2010), 'Cetuximab: a review of its use in squamous cell carcinoma of the head and neck', *Drugs*, 70 (15), 1987-2010.

Freier, K., Hofele, C., Knoepfle, K., Gross, M., Devens, F., Dyckhoff, G., Plinkert, P., Lichter, P., and Herold-Mende, C. (2010), 'Cytogenetic characterization of head and neck squamous cell carcinoma cell lines as model systems for the functional analyses of tumor-associated genes', *J Oral Pathol Med*, 39 (5), 382-9.

Freier, K., Sticht, C., Hofele, C., Flechtenmacher, C., Stange, D., Puccio, L., Toedt, G., Radlwimmer, B., Lichter, P., and Joos, S. (2006), 'Recurrent coamplification of cytoskeleton-associated genes EMS1 and SHANK2 with CCND1 in oral squamous cell carcinoma', *Genes Chromosomes Cancer*, 45 (2), 118-25.

Freudlsperger, C., Burnett, J. R., Friedman, J. A., Kannabiran, V. R., Chen, Z., and Van Waes, C. (2011), 'EGFR-PI3K-AKT-mTOR signaling in head and neck squamous cell carcinomas: attractive targets for molecular-oriented therapy', *Expert Opin Ther Targets,* 15 (1), 63-74.

Gaillard, H., Garcia-Muse, T., and Aguilera, A. (2015), 'Replication stress and cancer', *Nat Rev Cancer,* 15 (5), 276-89.

Galbiatti, A. L., Padovani-Junior, J. A., Maniglia, J. V., Rodrigues, C. D., Pavarino, E. C., and Goloni-Bertollo, E. M. (2013), 'Head and neck cancer: causes, prevention and treatment', *Braz J Otorhinolaryngol,* 79 (2), 239-47.

Galeazzi, E., Olivero, M., Gervasio, F. C., De Stefani, A., Valente, G., Comoglio, P. M., Di Renzo, M. F., and Cortesina, G. (1997), 'Detection of MET oncogene/hepatocyte growth factor receptor in lymph node metastases from head and neck squamous cell carcinomas', *Eur Arch Otorhinolaryngol,* 254 Suppl 1, S138-43.

Gao, W., Li, J. Z., Chan, J. Y., Ho, W. K., and Wong, T. S. (2012), 'mTOR Pathway and mTOR Inhibitors in Head and Neck Cancer', *ISRN Otolaryngol,* 2012, 953089.

Garnis, C., Baldwin, C., Zhang, L., Rosin, M. P., and Lam, W. L. (2003), 'Use of complete coverage array comparative genomic hybridization to define copy number alterations on chromosome 3p in oral squamous cell carcinomas', *Cancer Res,* 63 (24), 8582-5.

Garnis, C., Coe, B. P., Ishkanian, A., Zhang, L., Rosin, M. P., and Lam, W. L. (2004a), 'Novel regions of amplification on 8q distinct from the MYC locus and frequently altered in oral dysplasia and cancer', *Genes Chromosomes Cancer,* 39 (1), 93-8.

Garnis, Cathie, Coe, Bradley P., Zhang, Lewei, Rosin, Miriam P., and Lam, Wan L. (2004b), 'Overexpression of LRP12, a gene contained within an 8q22 amplicon identified by high-resolution array CGH analysis of oral squamous cell carcinomas', *Oncogene,* 23 (14), 2582-86.

Gazdar, A. F., Kurvari, V., Virmani, A., Gollahon, L., Sakaguchi, M., Westerfield, M., Kodagoda, D., Stasny, V., Cunningham, H. T., Wistuba, II, Tomlinson, G., Tonk, V., Ashfaq, R., Leitch, A. M., Minna, J. D., and Shay, J. W. (1998), 'Characterization of paired tumor and non-tumor cell lines established from patients with breast cancer', *Int J Cancer,* 78 (6), 766-74.

Ghosh, A., Ghosh, S., Maiti, G. P., Sabbir, M. G., Zabarovsky, E. R., Roy, A., Roychoudhury, S., and Panda, C. K. (2010), 'Frequent alterations of the candidate genes hMLH1, ITGA9 and RBSP3 in early dysplastic lesions of head and neck: clinical and prognostic significance', *Cancer Sci,* 101 (6), 1511-20.

Gibcus, J. H., Menkema, L., Mastik, M. F., Hermsen, M. A., de Bock, G. H., van Velthuysen, M. L., Takes, R. P., Kok, K., Alvarez Marcos, C. A., van der Laan, B. F., van den Brekel, M. W., Langendijk, J. A., Kluin, P. M., van der Wal, J. E., and Schuuring, E. (2007), 'Amplicon mapping and expression profiling identify the Fas-associated death domain gene as a new driver in the 11q13.3 amplicon in laryngeal/pharyngeal cancer', *Clin Cancer Res,* 13 (21), 6257-66.

Gillison, M. L. (2004), 'Human papillomavirus-associated head and neck cancer is a distinct epidemiologic, clinical, and molecular entity', *Semin Oncol,* 31 (6), 744-54.

Goerner, M., Seiwert, T. Y., and Sudhoff, H. (2010), 'Molecular targeted therapies in head and neck cancer--an update of recent developments', *Head Neck Oncol,* 2, 8.

Gollin, S. M. (2014), 'Cytogenetic alterations and their molecular genetic correlates in head and neck squamous cell carcinoma: a next generation window to the biology of disease', *Genes Chromosomes Cancer,* 53 (12), 972-90.

Gollin, Susanne M. (2001), 'Chromosomal alterations in squamous cell carcinomas of the head and neck: Window to the biology of disease', *Head & Neck,* 23 (3), 238-53.

Gonzalez-Garcia, R., Naval-Gias, L., Roman-Romero, L., Sastre-Perez, J., and Rodriguez-Campo, F. J. (2009), 'Local recurrences and second primary tumors from squamous cell carcinoma of the oral cavity: a retrospective analytic study of 500 patients', *Head Neck,* 31 (9), 1168-80.

Goodspeed, A., Heiser, L. M., Gray, J. W., and Costello, J. C. (2016), 'Tumor-Derived Cell Lines as Molecular Models of Cancer Pharmacogenomics', *Mol Cancer Res,* 14 (1), 3-13.

Greshock, J., Nathanson, K., Martin, A. M., Zhang, L., Coukos, G., Weber, B. L., and Zaks, T. Z. (2007), 'Cancer cell lines as genetic models of their parent histology: analyses based on array comparative genomic hybridization', *Cancer Res,* 67 (8), 3594-600.

6. References

Guerra, E. N., Rego, D. F., Elias, S. T., Coletta, R. D., Mezzomo, L. A., Gozal, D., and De Luca Canto, G. (2016), 'Diagnostic accuracy of serum biomarkers for head and neck cancer: A systematic review and meta-analysis', *Crit Rev Oncol Hematol*, 101, 93-118.

Guerrero-Preston, R., Michailidi, C., Marchionni, L., Pickering, C. R., Frederick, M. J., Myers, J. N., Yegnasubramanian, S., Hadar, T., Noordhuis, M. G., Zizkova, V., Fertig, E., Agrawal, N., Westra, W., Koch, W., Califano, J., Velculescu, V. E., and Sidransky, D. (2014), 'Key tumor suppressor genes inactivated by "greater promoter" methylation and somatic mutations in head and neck cancer', *Epigenetics*, 9 (7), 1031-46.

Guervos, M. A., Marcos, C. A., Hermsen, M., Nuno, A. S., Suarez, C., and Llorente, J. L. (2007), 'Deletions of N33, STK11 and TP53 are involved in the development of lymph node metastasis in larynx and pharynx carcinomas', *Cell Oncol*, 29 (4), 327-34.

Gupta, P. C., Mehta, F. S., Daftary, D. K., Pindborg, J. J., Bhonsle, R. B., Jalnawalla, P. N., Sinor, P. N., Pitkar, V. K., Murti, P. R., Irani, R. R., Shah, H. T., Kadam, P. M., Iyer, K. S., Iyer, H. M., Hegde, A. K., Chandrashekar, G. K., Shiroff, B. C., Sahiar, B. E., and Mehta, M. N. (1980), '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*, 8 (6), 283-333.

Gutierrez, V. F., Marcos, C. A., Llorente, J. L., Guervos, M. A., Iglesias, F. D., Tamargo, L. A., and Hermsen, M. (2012), 'Genetic profile of second primary tumors and recurrences in head and neck squamous cell carcinomas', *Head Neck*, 34 (6), 830-9.

Ha, P. K. and Califano, J. A. (2003), 'The molecular biology of mucosal field cancerization of the head and neck', *Crit Rev Oral Biol Med*, 14 (5), 363-9.

Haddad, R. I. and Shin, D. M. (2008), 'Recent advances in head and neck cancer', *N Engl J Med*, 359 (11), 1143-54.

Haluska, F. G., Tsujimoto, Y., and Croce, C. M. (1987), 'Oncogene activation by chromosome translocation in human malignancy', *Annu Rev Genet*, 21, 321-45.

Hanahan, D. and Weinberg, R. A. (2000), 'The hallmarks of cancer', *Cell*, 100 (1), 57-70.

Hanahan, D. and Weinberg, R. A. (2011), 'Hallmarks of cancer: the next generation', *Cell*, 144 (5), 646-74.

Hardisson, D., Alvarez-Marcos, C., Salas-Bustamante, A., Alonso-Guervos, M., Sastre, N., and Sampedro, A. (2004), 'Numerical aberrations of chromosomes 8, 9, 11, and 17 in squamous cell carcinoma of the pharynx and larynx: a fluorescence in situ hybridization and DNA flow cytometric analysis of 50 cases', *Oral Oncol*, 40 (4), 409-17.

Hashibe, M., Brennan, P., Benhamou, S., Castellsague, X., Chen, C., Curado, M. P., Dal Maso, L., Daudt, A. W., Fabianova, E., Fernandez, L., Wunsch-Filho, V., Franceschi, S., Hayes, R. B., Herrero, R., Koifman, S., La Vecchia, C., Lazarus, P., Levi, F., Mates, D., Matos, E., Menezes, A., Muscat, J., Eluf-Neto, J., Olshan, A. F., Rudnai, P., Schwartz, S. M., Smith, E., Sturgis, E. M., Szeszenia-Dabrowska, N., Talamini, R., Wei, Q., Winn, D. M., Zaridze, D., Zatonski, W., Zhang, Z. F., Berthiller, J., and Boffetta, P. (2007), 'Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium', *J Natl Cancer Inst*, 99 (10), 777-89.

Hashimoto, Y., Shindo-Okada, N., Tani, M., Takeuchi, K., Toma, H., and Yokota, J. (1996), 'Identification of genes differentially expressed in association with metastatic potential of K-1735 murine melanoma by messenger RNA differential display', *Cancer Res*, 56 (22), 5266-71.

Hausser, J. and Strimmer, K. (2009), 'Entropy inference and the James-Stein estimator, with application to nonlinear gene association networks', *J. Mach. Learn. Res.*, 10, 1469-84.

Hayashi, M., Guerrero-Preston, R., Sidransky, D., and Koch, W. M. (2015), 'Paired box 5 methylation detection by droplet digital PCR for ultra-sensitive deep surgical margins analysis of head and neck squamous cell carcinoma', *Cancer Prev Res (Phila)*, 8 (11), 1017-26.

Hellebrekers, D. M., Lentjes, M. H., van den Bosch, S. M., Melotte, V., Wouters, K. A., Daenen, K. L., Smits, K. M., Akiyama, Y., Yuasa, Y., Sanduleanu, S., Khalid-de Bakker, C. A., Jonkers, D., Weijenberg, M. P., Louwagie, J., van Criekinge, W., Carvalho, B., Meijer, G. A., Baylin, S. B., Herman, J. G., de Bruine, A. P., and van Engeland, M. (2009), 'GATA4 and GATA5 are

potential tumor suppressors and biomarkers in colorectal cancer', *Clin Cancer Res*, 15 (12), 3990-7.

Heng, H. H., Regan, S. M., Liu, G., and Ye, C. J. (2016), 'Why it is crucial to analyze non clonal chromosome aberrations or NCCAs?', *Mol Cytogenet*, 9, 15.

Heo, D. S., Snyderman, C., Gollin, S. M., Pan, S., Walker, E., Deka, R., Barnes, E. L., Johnson, J. T., Herberman, R. B., and Whiteside, T. L. (1989), 'Biology, cytogenetics, and sensitivity to immunological effector cells of new head and neck squamous cell carcinoma lines', *Cancer Res*, 49 (18), 5167-75.

Herman, J. G., Latif, F., Weng, Y., Lerman, M. I., Zbar, B., Liu, S., Samid, D., Duan, D. S., Gnarra, J. R., Linehan, W. M., and et al. (1994), 'Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma', *Proc Natl Acad Sci U S A*, 91 (21), 9700-4.

Hermsen, M., Guervos, M. A., Meijer, G., Baak, J., van Diest, P., Marcos, C. A., and Sampedro, A. (2001), 'New chromosomal regions with high-level amplifications in squamous cell carcinomas of the larynx and pharynx, identified by comparative genomic hybridization', *J Pathol*, 194 (2), 177-82.

Hinsch, H. and Hannenhalli, S. (2006), 'Recurring genomic breaks in independent lineages support genomic fragility', *BMC Evol Biol*, 6, 90.

Ho, A. S., Turcan, S., and Chan, T. A. (2013), 'Epigenetic therapy: use of agents targeting deacetylation and methylation in cancer management', *Onco Targets Ther*, 6, 223-32.

Hogg, R. P., Honorio, S., Martinez, A., Agathanggelou, A., Dallol, A., Fullwood, P., Weichselbaum, R., Kuo, M. J., Maher, E. R., and Latif, F. (2002), 'Frequent 3p allele loss and epigenetic inactivation of the RASSF1A tumour suppressor gene from region 3p21.3 in head and neck squamous cell carcinoma', *Eur J Cancer*, 38 (12), 1585-92.

Holloway, R. L., Hellewell, J. L., Marbella, A. M., Layde, P. M., Myers, K. B., and Campbell, B. H. (2005), 'Psychosocial effects in long-term head and neck cancer survivors', *Head Neck*, 27 (4), 281-8.

Homig-Holzel, C. and Savola, S. (2012), 'Multiplex ligation-dependent probe amplification (MLPA) in tumor diagnostics and prognostics', *Diagn Mol Pathol*, 21 (4), 189-206.

Huang, Qiang, Yu, Guo Pei, McCormick, Steven A., Mo, Juan, Datta, Bhakti, Mahimkar, Manoj, Lazarus, Philip, Schäffer, Alejandro A., Desper, Richard, and Schantz, Stimson P. (2002), 'Genetic differences detected by comparative genomic hybridization in head and neck squamous cell carcinomas from different tumor sites: construction of oncogenetic trees for tumor progression', *Genes, Chromosomes and Cancer*, 34 (2), 224-33.

Hussain, S. P. and Harris, C. C. (2006), 'p53 biological network: at the crossroads of the cellular-stress response pathway and molecular carcinogenesis', *J Nippon Med Sch*, 73 (2), 54-64.

Ince, T. A., Sousa, A. D., Jones, M. A., Harrell, J. C., Agoston, E. S., Krohn, M., Selfors, L. M., Liu, W., Chen, K., Yong, M., Buchwald, P., Wang, B., Hale, K. S., Cohick, E., Sergent, P., Witt, A., Kozhekbaeva, Z., Gao, S., Agoston, A. T., Merritt, M. A., Foster, R., Rueda, B. R., Crum, C. P., Brugge, J. S., and Mills, G. B. (2015), 'Characterization of twenty-five ovarian tumour cell lines that phenocopy primary tumours', *Nat Commun*, 6, 7419.

Iskow, R. C., McCabe, M. T., Mills, R. E., Torene, S., Pittard, W. S., Neuwald, A. F., Van Meir, E. G., Vertino, P. M., and Devine, S. E. (2010), 'Natural mutagenesis of human genomes by endogenous retrotransposons', *Cell*, 141 (7), 1253-61.

Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., and Thun, M. J. (2009), 'Cancer statistics, 2009', *CA Cancer J Clin*, 59 (4), 225-49.

Jerjes, W., Upile, T., Petrie, A., Riskalla, A., Hamdoon, Z., Vourvachis, M., Karavidas, K., Jay, A., Sandison, A., Thomas, G. J., Kalavrezos, N., and Hopper, C. (2010), 'Clinicopathological parameters, recurrence, locoregional and distant metastasis in 115 T1-T2 oral squamous cell carcinoma patients', *Head Neck Oncol*, 2, 9.

Jeuken, J., Cornelissen, S., Boots-Sprenger, S., Gijsen, S., and Wesseling, P. (2006), 'Multiplex ligation-dependent probe amplification: a diagnostic tool for simultaneous identification of different genetic markers in glial tumors', *J Mol Diagn*, 8 (4), 433-43.

6. References

Jin, C., Jin, Y., Gisselsson, D., Wennerberg, J., Wah, T. S., Stromback, B., Kwong, Y. L., and Mertens, F. (2006), 'Molecular cytogenetic characterization of the 11q13 amplicon in head and neck squamous cell carcinoma', *Cytogenet Genome Res*, 115 (2), 99-106.

Jin, Y., Mertens, F., Jin, C., Akervall, J., Wennerberg, J., Gorunova, L., Mandahl, N., Heim, S., and Mitelman, F. (1995), 'Nonrandom chromosome abnormalities in short-term cultured primary squamous cell carcinomas of the head and neck', *Cancer Res*, 55 (14), 3204-10.

Kagan, H. M. and Li, W. (2003), 'Lysyl oxidase: properties, specificity, and biological roles inside and outside of the cell', *J Cell Biochem*, 88 (4), 660-72.

Kalyankrishna, S. and Grandis, J. R. (2006), 'Epidermal growth factor receptor biology in head and neck cancer', *J Clin Oncol*, 24 (17), 2666-72.

Kaplan, R. N., Riba, R. D., Zacharoulis, S., Bramley, A. H., Vincent, L., Costa, C., MacDonald, D. D., Jin, D. K., Shido, K., Kerns, S. A., Zhu, Z., Hicklin, D., Wu, Y., Port, J. L., Altorki, N., Port, E. R., Ruggero, D., Shmelkov, S. V., Jensen, K. K., Rafii, S., and Lyden, D. (2005), 'VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche', *Nature*, 438 (7069), 820-7.

Katsanis, S. H. and Katsanis, N. (2013), 'Molecular genetic testing and the future of clinical genomics', *Nat Rev Genet*, 14 (6), 415-26.

Kearney, L. (2006), 'Multiplex-FISH (M-FISH): technique, developments and applications', *Cytogenet Genome Res*, 114 (3-4), 189-98.

Kobayashi, T., Maruyama, S., Abe, T., Cheng, J., Takagi, R., Saito, C., and Saku, T. (2012), 'Keratin 10-positive orthokeratotic dysplasia: a new leucoplakia-type precancerous entity of the oral mucosa', *Histopathology*, 61 (5), 910-20.

Kowalski, L. P. (2002), 'Results of salvage treatment of the neck in patients with oral cancer', *Arch Otolaryngol Head Neck Surg*, 128 (1), 58-62.

Kozaki, Ken-ichi, Imoto, Issei, Pimkhaokham, Atiphan, Hasegawa, Shogo, Tsuda, Hitoshi, Omura, Ken, and Inazawa, Johji (2006), 'PIK3CA mutation is an oncogenic aberration at advanced stages of oral squamous cell carcinoma.', *Cancer science*, 97, 1351-8.

Kozeretska, I. A., Demydov, S. V., and Ostapchenko, L. I. (2011), 'Mobile genetic elements and cancer. From mutations to gene therapy', *Exp Oncol*, 33 (4), 198-205.

Kudo, Y., Kitajima, S., Sato, S., Miyauchi, M., Ogawa, I., and Takata, T. (2001), 'High expression of S-phase kinase-interacting protein 2, human F-box protein, correlates with poor prognosis in oral squamous cell carcinomas', *Cancer Res*, 61 (19), 7044-7.

Kumar, B., Cordell, K. G., Lee, J. S., Worden, F. P., Prince, M. E., Tran, H. H., Wolf, G. T., Urba, S. G., Chepeha, D. B., Teknos, T. N., Eisbruch, A., Tsien, C. I., Taylor, J. M., D'Silva, N. J., Yang, K., Kurnit, D. M., Bauer, J. A., Bradford, C. R., and Carey, T. E. (2008), 'EGFR, p16, HPV Titer, Bcl-xL and p53, sex, and smoking as indicators of response to therapy and survival in oropharyngeal cancer', *J Clin Oncol*, 26 (19), 3128-37.

Lander, E. S. Linton, L. M. Birren, B. Nusbaum, C. Zody, M. C. Baldwin, J. Devon, K. Dewar, K. Doyle, M. FitzHugh, W. Funke, R. Gage, D. Harris, K. Heaford, A. Howland, J. Kann, L. Lehoczky, J. LeVine, R. McEwan, P. McKernan, K. Meldrim, J. Mesirov, J. P. Miranda, C. Morris, W. Naylor, J. Raymond, C. Rosetti, M. Santos, R. Sheridan, A. Sougnez, C. Stange-Thomann, Y. Stojanovic, N. Subramanian, A. Wyman, D. Rogers, J. Sulston, J. Ainscough, R. Beck, S. Bentley, D. Burton, J. Clee, C. Carter, N. Coulson, A. Deadman, R. Deloukas, P. Dunham, A. Dunham, I. Durbin, R. French, L. Grafham, D. Gregory, S. Hubbard, T. Humphray, S. Hunt, A. Jones, M. Lloyd, C. McMurray, A. Matthews, L. Mercer, S. Milne, S. Mullikin, J. C. Mungall, A. Plumb, R. Ross, M. Shownkeen, R. Sims, S. Waterston, R. H. Wilson, R. K. Hillier, L. W. McPherson, J. D. Marra, M. A. Mardis, E. R. Fulton, L. A. Chinwalla, A. T. Pepin, K. H. Gish, W. R. Chissoe, S. L. Wendl, M. C. Delehaunty, K. D. Miner, T. L. Delehaunty, A. Kramer, J. B. Cook, L. L. Fulton, R. S. Johnson, D. L. Minx, P. J. Clifton, S. W. Hawkins, T. Branscomb, E. Predki, P. Richardson, P. Wenning, S. Slezak, T. Doggett, N. Cheng, J. F. Olsen, A. Lucas, S. Elkin, C. Uberbacher, E. Frazier, M. Gibbs, R. A. Muzny, D. M. Scherer, S. E. Bouck, J. B. Sodergren, E. J. Worley, K. C. Rives, C. M. Gorrell, J. H. Metzker, M. L. Naylor, S. L. Kucherlapati, R. S. Nelson, D. L. Weinstock, G. M. Sakaki, Y. Fujiyama, A. Hattori, M. Yada,

T. Toyoda, A. Itoh, T. Kawagoe, C. Watanabe, H. Totoki, Y. Taylor, T. Weissenbach, J. Heilig, R. Saurin, W. Artiguenave, F. Brottier, P. Bruls, T. Pelletier, E. Robert, C. Wincker, P. Smith, D. R. Doucette-Stamm, L. Rubenfield, M. Weinstock, K. Lee, H. M. Dubois, J. Rosenthal, A. Platzer, M. Nyakatura, G. Taudien, S. Rump, A. Yang, H. Yu, J. Wang, J. Huang, G. Gu, J. Hood, L. Rowen, L. Madan, A. Qin, S. Davis, R. W. Federspiel, N. A. Abola, A. P. Proctor, M. J. Myers, R. M. Schmutz, J. Dickson, M. Grimwood, J. Cox, D. R. Olson, M. V. Kaul, R. Raymond, C. Shimizu, N. Kawasaki, K. Minoshima, S. Evans, G. A. Athanasiou, M. Schultz, R. Roe, B. A. Chen, F. Pan, H. Ramser, J. Lehrach, H. Reinhardt, R. McCombie, W. R. de la Bastide, M. Dedhia, N. Blocker, H. Hornischer, K. Nordsiek, G. Agarwala, R. Aravind, L. Bailey, J. A. Bateman, A. Batzoglou, S. Birney, E. Bork, P. Brown, D. G. Burge, C. B. Cerutti, L. Chen, H. C. Church, D. Clamp, M. Copley, R. R. Doerks, T. Eddy, S. R. Eichler, E. E. Furey, T. S. Galagan, J. Gilbert, J. G. Harmon, C. Hayashizaki, Y. Haussler, D. Hermjakob, H. Hokamp, K. Jang, W. Johnson, L. S. Jones, T. A. Kasif, S. Kaspryzk, A. Kennedy, S. Kent, W. J. Kitts, P. Koonin, E. V. Korf, I. Kulp, D. Lancet, D. Lowe, T. M. McLysaght, A. Mikkelsen, T. Moran, J. V. Mulder, N. Pollara, V. J. Ponting, C. P. Schuler, G. Schultz, J. Slater, G. Smit, A. F. Stupka, E. Szustakowki, J. Thierry-Mieg, D. Thierry-Mieg, J. Wagner, L. Wallis, J. Wheeler, R. Williams, A. Wolf, Y. I. Wolfe, K. H. Yang, S. P. Yeh, R. F. Collins, F. Guyer, M. S. Peterson, J. Felsenfeld, A. Wetterstrand, K. A. Patrinos, A. Morgan, M. J. de Jong, P. Catanese, J. J. Osoegawa, K. Shizuya, H. Choi, S. Chen, Y. J. Szustakowki, J. and International Human Genome Sequencing, Consortium (2001), 'Initial sequencing and analysis of the human genome', *Nature*, 409 (6822), 860-921.

Landry, J. J., Pyl, P. T., Rausch, T., Zichner, T., Tekkedil, M. M., Stutz, A. M., Jauch, A., Aiyar, R. S., Pau, G., Delhomme, N., Gagneur, J., Korbel, J. O., Huber, W., and Steinmetz, L. M. (2013), 'The genomic and transcriptomic landscape of a HeLa cell line', *G3 (Bethesda)*, 3 (8), 1213-24.

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L. J., 3rd, Lohr, J. G., Harris, C. C., Ding, L., Wilson, R. K., Wheeler, D. A., Gibbs, R. A., Kucherlapati, R., Lee, C., Kharchenko, P. V., Park, P. J., and Cancer Genome Atlas Research, Network (2012), 'Landscape of somatic retrotransposition in human cancers', *Science*, 337 (6097), 967-71.

Leemans, C. R., Braakhuis, B. J., and Brakenhoff, R. H. (2011), 'The molecular biology of head and neck cancer', *Nat Rev Cancer*, 11 (1), 9-22.

Lese, C. M., Rossie, K. M., Appel, B. N., Reddy, J. K., Johnson, J. T., Myers, E. N., and Gollin, S. M. (1995), 'Visualization of INT2 and HST1 amplification in oral squamous cell carcinomas', *Genes Chromosomes Cancer*, 12 (4), 288-95.

Li, J. P., Fu, Y. N., Chen, Y. R., and Tan, T. H. (2010), 'JNK pathway-associated phosphatase dephosphorylates focal adhesion kinase and suppresses cell migration', *J Biol Chem*, 285 (8), 5472-8.

Li, Y. F., Hsiao, Y. H., Lai, Y. H., Chen, Y. C., Chen, Y. J., Chou, J. L., Chan, M. W., Lin, Y. H., Tsou, Y. A., Tsai, M. H., and Tai, C. K. (2015), 'DNA methylation profiles and biomarkers of oral squamous cell carcinoma', *Epigenetics*, 10 (3), 229-36.

Liang, C., Marsit, C. J., Houseman, E. A., Butler, R., Nelson, H. H., McClean, M. D., and Kelsey, K. T. (2012), 'Gene-environment interactions of novel variants associated with head and neck cancer', *Head Neck*, 34 (8), 1111-8.

Liborio-Kimura, T. N., Jung, H. M., and Chan, E. K. (2015), 'miR-494 represses HOXA10 expression and inhibits cell proliferation in oral cancer', *Oral Oncol*, 51 (2), 151-7.

Liehr, T. (2008), *Fluorescence In Situ Hybridization (FISH) - Application Guide* (Springer Berlin Heidelberg).

Lim, Y., Sun, C. X., Tran, P., and Punyadeera, C. (2016), 'Salivary epigenetic biomarkers in head and neck squamous cell carcinomas', *Biomark Med*, 10 (3), 301-13.

Lin, C. C., Yang, J. S., Chen, J. T., Fan, S., Yu, F. S., Yang, J. L., Lu, C. C., Kao, M. C., Huang, A. C., Lu, H. F., and Chung, J. G. (2007a), 'Berberine induces apoptosis in human HSC-3 oral cancer cells via simultaneous activation of the death receptor-mediated and mitochondrial pathway', *Anticancer Res*, 27 (5A), 3371-8.

6. References

Lin, C. J., Grandis, J. R., Carey, T. E., Gollin, S. M., Whiteside, T. L., Koch, W. M., Ferris, R. L., and Lai, S. Y. (2007b), 'Head and neck squamous cell carcinoma cell lines: established models and rationale for selection', *Head Neck*, 29 (2), 163-88.

Lin, S. C., Chen, Y. J., Kao, S. Y., Hsu, M. T., Lin, C. H., Yang, S. C., Liu, T. Y., and Chang, K. W. (2002), 'Chromosomal changes in betel-associated oral squamous cell carcinomas and their relationship to clinical parameters', *Oral Oncol*, 38 (3), 266-73.

Liu, S. A., Wong, Y. K., Lin, J. C., Poon, C. K., Tung, K. C., and Tsai, W. C. (2007), 'Impact of recurrence interval on survival of oral cavity squamous cell carcinoma patients after local relapse', *Otolaryngol Head Neck Surg*, 136 (1), 112-8.

Llewellyn, C. D., Johnson, N. W., and Warnakulasuriya, K. A. (2001), 'Risk factors for squamous cell carcinoma of the oral cavity in young people--a comprehensive literature review', *Oral Oncol*, 37 (5), 401-18.

Lotan, R., Xu, X. C., Lippman, S. M., Ro, J. Y., Lee, J. S., Lee, J. J., and Hong, W. K. (1995), 'Suppression of retinoic acid receptor-beta in premalignant oral lesions and its up-regulation by isotretinoin', *N Engl J Med*, 332 (21), 1405-10.

Lothaire, P., de Azambuja, E., Dequanter, D., Lalami, Y., Sotiriou, C., Andry, G., Castro, G., Jr., and Awada, A. (2006), 'Molecular markers of head and neck squamous cell carcinoma: promising signs in need of prospective evaluation', *Head Neck*, 28 (3), 256-69.

Luo, M. L., Shen, X. M., Zhang, Y., Wei, F., Xu, X., Cai, Y., Zhang, X., Sun, Y. T., Zhan, Q. M., Wu, M., and Wang, M. R. (2006), 'Amplification and overexpression of CTTN (EMS1) contribute to the metastasis of esophageal squamous cell carcinoma by promoting cell migration and anoikis resistance', *Cancer Res*, 66 (24), 11690-9.

Luryi, A. L., Yarbrough, W. G., Niccolai, L. M., Roser, S., Reed, S. G., Nathan, C. A., Moore, M. G., Day, T., and Judson, B. L. (2014), 'Public awareness of head and neck cancers: a cross-sectional survey', *JAMA Otolaryngol Head Neck Surg*, 140 (7), 639-46.

Ma, C., Quesnelle, K. M., Sparano, A., Rao, S., Park, M. S., Cohen, M. A., Wang, Y., Samanta, M., Kumar, M. S., Aziz, M. U., Naylor, T. L., Weber, B. L., Fakharzadeh, S. S., Weinstein, G. S., Vachani, A., Feldman, M. D., and Brose, M. S. (2009), 'Characterization CSMD1 in a large set of primary lung, head and neck, breast and skin cancer tissues', *Cancer Biol Ther*, 8 (10), 907-16.

Macheret, M. and Halazonetis, T. D. (2015), 'DNA replication stress as a hallmark of cancer', *Annu Rev Pathol*, 10, 425-48.

Mao, E. J., Schwartz, S. M., Daling, J. R., and Beckmann, A. M. (1998), 'Loss of heterozygosity at 5q21-22 (adenomatous polyposis coli gene region) in oral squamous cell carcinoma is common and correlated with advanced disease', *J Oral Pathol Med*, 27 (7), 297-302.

Marcos, C. A., Alonso-Guervos, M., Prado, N. R., Gimeno, T. S., Iglesias, F. D., Hermsen, M., and Llorente, J. L. (2010), 'Genetic model of transformation and neoplastic progression in laryngeal epithelium', *Head Neck*, 33 (2), 216-24.

Markopoulos, A. K. (2012), 'Current aspects on oral squamous cell carcinoma', *Open Dent J*, 6, 126-30.

Marron, M., Boffetta, P., Zhang, Z. F., Zaridze, D., Wunsch-Filho, V., Winn, D. M., Wei, Q., Talamini, R., Szeszenia-Dabrowska, N., Sturgis, E. M., Smith, E., Schwartz, S. M., Rudnai, P., Purdue, M. P., Olshan, A. F., Eluf-Neto, J., Muscat, J., Morgenstern, H., Menezes, A., McClean, M., Matos, E., Mates, I. N., Lissowska, J., Levi, F., Lazarus, P., La Vecchia, C., Koifman, S., Kelsey, K., Herrero, R., Hayes, R. B., Franceschi, S., Fernandez, L., Fabianova, E., Daudt, A. W., Dal Maso, L., Curado, M. P., Cadoni, G., Chen, C., Castellsague, X., Boccia, S., Benhamou, S., Ferro, G., Berthiller, J., Brennan, P., Moller, H., and Hashibe, M. (2010), 'Cessation of alcohol drinking, tobacco smoking and the reversal of head and neck cancer risk', *Int J Epidemiol*, 39 (1), 182-96.

Marta, G. N., Riera, R., Bossi, P., Zhong, L. P., Licitra, L., Macedo, C. R., de Castro Junior, G., Carvalho, A. L., William, W. N., Jr., and Kowalski, L. P. (2015), 'Induction chemotherapy prior to surgery with or without postoperative radiotherapy for oral cavity cancer patients: Systematic review and meta-analysis', *Eur J Cancer*, 51 (17), 2596-603.

Martin, C. L., Reshmi, S. C., Ried, T., Gottberg, W., Wilson, J. W., Reddy, J. K., Khanna, P., Johnson, J. T., Myers, E. N., and Gollin, S. M. (2008), 'Chromosomal imbalances in oral squamous cell carcinoma: examination of 31 cell lines and review of the literature', *Oral Oncol,* 44 (4), 369-82.

Marur, S. and Forastiere, A. A. (2008), 'Head and neck cancer: changing epidemiology, diagnosis, and treatment', *Mayo Clin Proc,* 83 (4), 489-501.

Maruya, S., Issa, J. P., Weber, R. S., Rosenthal, D. I., Haviland, J. C., Lotan, R., and El-Naggar, A. K. (2004), 'Differential methylation status of tumor-associated genes in head and neck squamous carcinoma: incidence and potential implications', *Clin Cancer Res,* 10 (11), 3825-30.

Maynard, M. A., Evans, A. J., Hosomi, T., Hara, S., Jewett, M. A., and Ohh, M. (2005), 'Human HIF-3alpha4 is a dominant-negative regulator of HIF-1 and is down-regulated in renal cell carcinoma', *FASEB J,* 19 (11), 1396-406.

Mehanna, H., Paleri, V., West, C. M., and Nutting, C. (2010), 'Head and neck cancer--Part 1: Epidemiology, presentation, and prevention', *BMJ,* 341, c4684.

Mendez, E., Houck, J. R., Doody, D. R., Fan, W., Lohavanichbutr, P., Rue, T. C., Yueh, B., Futran, N. D., Upton, M. P., Farwell, D. G., Heagerty, P. J., Zhao, L. P., Schwartz, S. M., and Chen, C. (2009), 'A genetic expression profile associated with oral cancer identifies a group of patients at high risk of poor survival', *Clin Cancer Res,* 15 (4), 1353-61.

Menze, B. H., Kelm, B. M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., and Hamprecht, F. A. (2009), 'A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data', *BMC Bioinformatics,* 10, 213.

Messadi, D. V. (2013), 'Diagnostic aids for detection of oral precancerous conditions', *Int J Oral Sci,* 5 (2), 59-65.

Meyer, T. K., Kuhn, J. C., Campbell, B. H., Marbella, A. M., Myers, K. B., and Layde, P. M. (2004), 'Speech intelligibility and quality of life in head and neck cancer survivors', *Laryngoscope,* 114 (11), 1977-81.

Michalides, Rob, Van Veelen, Nicoletta, Hart, Augustinus, Loftus, Barbara, Wientjens, Ellen, and Balm, Alfons (1995), 'Overexpression of Cyclin Dl Correlates with Recurrence in a Group of Forty-seven Operable Squamous Cell Carcinomas of the Head and Neck', *CANCER RESEARCH,* 55, 975-78.

Mineta, H., Miura, K., Takebayashi, S., Ueda, Y., Misawa, K., Harada, H., Wennerberg, J., and Dictor, M. (2000), 'Cyclin D1 overexpression correlates with poor prognosis in patients with tongue squamous cell carcinoma', *Oral Oncol,* 36 (2), 194-8.

Misawa, K., Misawa, Y., Kanazawa, T., Mochizuki, D., Imai, A., Endo, S., Carey, T. E., and Mineta, H. (2016), 'Epigenetic inactivation of galanin and GALR1/2 is associated with early recurrence in head and neck cancer', *Clin Exp Metastasis,* 33 (2), 187-95.

Misawa, K., Ueda, Y., Kanazawa, T., Misawa, Y., Jang, I., Brenner, J. C., Ogawa, T., Takebayashi, S., Grenman, R. A., Herman, J. G., Mineta, H., and Carey, T. E. (2008), 'Epigenetic inactivation of galanin receptor 1 in head and neck cancer', *Clin Cancer Res,* 14 (23), 7604-13.

Mitka, M. (2013), 'Evidence lacking for benefit from oral cancer screening', *JAMA,* 309 (18), 1884.

Mohankumar, K., Pajaniradje, S., Sridharan, S., Singh, V. K., Ronsard, L., Banerjea, A. C., Selvanesan, B. C., Coumar, M. S., Periyasamy, L., and Rajagopalan, R. (2014), 'Apoptosis induction by an analog of curcumin (BDMC-A) in human laryngeal carcinoma cells through intrinsic and extrinsic pathways', *Cell Oncol (Dordr),* 37 (6), 439-54.

Molin, Y. and Fayette, J. (2011), 'Current chemotherapies for recurrent/metastatic head and neck cancer', *Anticancer Drugs,* 22 (7), 621-5.

Mortazavi, H., Baharvand, M., and Mehdipour, M. (2014), 'Oral potentially malignant disorders: an overview of more than 20 entities', *J Dent Res Dent Clin Dent Prospects,* 8 (1), 6-14.

Mossink, M. H., van Zon, A., Scheper, R. J., Sonneveld, P., and Wiemer, E. A. (2003), 'Vaults: a ribonucleoprotein particle involved in drug resistance?', *Oncogene,* 22 (47), 7458-67.

Mrasek, K., Schoder, C., Teichmann, A. C., Behr, K., Franze, B., Wilhelm, K., Blaurock, N., Claussen, U., Liehr, T., and Weise, A. (2010), 'Global screening and extended nomenclature for 230 aphidicolin-inducible fragile sites, including 61 yet unreported ones', *Int J Oncol,* 36 (4), 929-40.

6. References

Mroz, Edmund A. and Rocco, James W. (2016), 'Intra-tumor heterogeneity in head and neck cancer and its clinical implications', *World Journal of Otorhinolaryngology-Head and Neck Surgery*, 2 (2), 60-67.

Mungall, A. J. Palmer, S. A. Sims, S. K. Edwards, C. A. Ashurst, J. L. Wilming, L. Jones, M. C. Horton, R. Hunt, S. E. Scott, C. E. Gilbert, J. G. Clamp, M. E. Bethel, G. Milne, S. Ainscough, R. Almeida, J. P. Ambrose, K. D. Andrews, T. D. Ashwell, R. I. Babbage, A. K. Bagguley, C. L. Bailey, J. Banerjee, R. Barker, D. J. Barlow, K. F. Bates, K. Beare, D. M. Beasley, H. Beasley, O. Bird, C. P. Blakey, S. Bray-Allen, S. Brook, J. Brown, A. J. Brown, J. Y. Burford, D. C. Burrill, W. Burton, J. Carder, C. Carter, N. P. Chapman, J. C. Clark, S. Y. Clark, G. Clee, C. M. Clegg, S. Cobley, V. Collier, R. E. Collins, J. E. Colman, L. K. Corby, N. R. Coville, G. J. Culley, K. M. Dhami, P. Davies, J. Dunn, M. Earthrowl, M. E. Ellington, A. E. Evans, K. A. Faulkner, L. Francis, M. D. Frankish, A. Frankland, J. French, L. Garner, P. Garnett, J. Ghori, M. J. Gilby, L. M. Gillson, C. J. Glithero, R. J. Grafham, D. V. Grant, M. Gribble, S. Griffiths, C. Griffiths, M. Hall, R. Halls, K. S. Hammond, S. Harley, J. L. Hart, E. A. Heath, P. D. Heathcott, R. Holmes, S. J. Howden, P. J. Howe, K. L. Howell, G. R. Huckle, E. Humphray, S. J. Humphries, M. D. Hunt, A. R. Johnson, C. M. Joy, A. A. Kay, M. Keenan, S. J. Kimberley, A. M. King, A. Laird, G. K. Langford, C. Lawlor, S. Leongamornlert, D. A. Leversha, M. Lloyd, C. R. Lloyd, D. M. Loveland, J. E. Lovell, J. Martin, S. Mashreghi-Mohammadi, M. Maslen, G. L. Matthews, L. McCann, O. T. McLaren, S. J. McLay, K. McMurray, A. Moore, M. J. Mullikin, J. C. Niblett, D. Nickerson, T. Novik, K. L. Oliver, K. Overton-Larty, E. K. Parker, A. Patel, R. Pearce, A. V. Peck, A. I. Phillimore, B. Phillips, S. Plumb, R. W. Porter, K. M. Ramsey, Y. Ranby, S. A. Rice, C. M. Ross, M. T. Searle, S. M. Sehra, H. K. Sheridan, E. Skuce, C. D. Smith, S. Smith, M. Spraggon, L. Squares, S. L. Steward, C. A. Sycamore, N. Tamlyn-Hall, G. Tester, J. Theaker, A. J. Thomas, D. W. Thorpe, A. Tracey, A. Tromans, A. Tubby, B. Wall, M. Wallis, J. M. West, A. P. White, S. S. Whitehead, S. L. Whittaker, H. Wild, A. Willey, D. J. Wilmer, T. E. Wood, J. M. Wray, P. W. Wyatt, J. C. Young, L. Younger, R. M. Bentley, D. R. Coulson, A. Durbin, R. Hubbard, T. Sulston, J. E. Dunham, I. Rogers, J. and Beck, S. (2003), 'The DNA sequence and analysis of human chromosome 6', *Nature*, 425 (6960), 805-11.

Murat, A., Migliavacca, E., Gorlia, T., Lambiv, W. L., Shay, T., Hamou, M. F., de Tribolet, N., Regli, L., Wick, W., Kouwenhoven, M. C., Hainfellner, J. A., Heppner, F. L., Dietrich, P. Y., Zimmer, Y., Cairncross, J. G., Janzer, R. C., Domany, E., Delorenzi, M., Stupp, R., and Hegi, M. E. (2008), 'Stem cell-related "self-renewal" signature and high epidermal growth factor receptor expression associated with resistance to concomitant chemoradiotherapy in glioblastoma', *J Clin Oncol*, 26 (18), 3015-24.

Murphy, W. J., Larkin, D. M., Everts-van der Wind, A., Bourque, G., Tesler, G., Auvil, L., Beever, J. E., Chowdhary, B. P., Galibert, F., Gatzke, L., Hitte, C., Meyers, S. N., Milan, D., Ostrander, E. A., Pape, G., Parker, H. G., Raudsepp, T., Rogatcheva, M. B., Schook, L. B., Skow, L. C., Welge, M., Womack, J. E., O'Brien S, J., Pevzner, P. A., and Lewin, H. A. (2005), 'Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps', *Science*, 309 (5734), 613-7.

Murugan, A. K., Hong, N. T., Fukui, Y., Munirajan, A. K., and Tsuchida, N. (2008), 'Oncogenic mutations of the PIK3CA gene in head and neck squamous cell carcinomas', *Int J Oncol*, 32 (1), 101-11.

Myo, K., Uzawa, N., Miyamoto, R., Sonoda, I., Yuki, Y., and Amagasa, T. (2005), 'Cyclin D1 gene numerical aberration is a predictive marker for occult cervical lymph node metastasis in TNM Stage I and II squamous cell carcinoma of the oral cavity', *Cancer*, 104 (12), 2709-16.

Nagatsuka, H, Ishiwari, Y, Tsujigiwa, H, Nakano, K, and Nagai, N (2001), 'Quantitation of epidermal growth factor receptor gene amplification by competitive polymerase chain reaction in pre-malignant and malignant oral epithelial lesions.', *Oral oncology*, 37, 599-604.

Nagpal, J. K. and Das, B. R. (2003), 'Oral cancer: reviewing the present understanding of its molecular mechanism and exploring the future directions for its effective management', *Oral Oncol*, 39 (3), 213-21.

Nair, D. R., Pruthy, R., Pawar, U., and Chaturvedi, P. (2012), 'Oral cancer: Premalignant conditions and screening--an update', *J Cancer Res Ther,* 8 Suppl 1, S57-66.

Nakaoka, T., Ota, A., Ono, T., Karnan, S., Konishi, H., Furuhashi, A., Ohmura, Y., Yamada, Y., Hosokawa, Y., and Kazaoka, Y. (2014), 'Combined arsenic trioxide-cisplatin treatment enhances apoptosis in oral squamous cell carcinoma cells', *Cell Oncol (Dordr),* 37 (2), 119-29.

Namazie, A., Alavi, S., Olopade, O. I., Pauletti, G., Aghamohammadi, N., Aghamohammadi, M., Gornbein, J. A., Calcaterra, T. C., Slamon, D. J., Wang, M. B., and Srivatsan, E. S. (2002), 'Cyclin D1 amplification and p16(MTS1/CDK4I) deletion correlate with poor prognosis in head and neck tumors', *Laryngoscope,* 112 (3), 472-81.

Nathanson, Arthur, Ågren, Karin, Biörklund, Anders, Lind, Magnus G., Andréason, Lars, Anniko, Matti, Freijd, Anders, Lejdeborn, Lars, Kinman, Staffan, Kumlien, Anders, Köling, Anders, Lindström, Jörgen, Novik, Anders, Olofsson, Jan, and Pettersson, Kari-Lvar (1989), 'Evaluation of some prognostic factors in small squamous cell carcinoma of the mobile tongue: A multicenter study in Sweden', *Head Neck,* 11 (5), 387-92.

Negrini, S., Gorgoulis, V. G., and Halazonetis, T. D. (2010), 'Genomic instability--an evolving hallmark of cancer', *Nat Rev Mol Cell Biol,* 11 (3), 220-8.

Neves, S. S., Sarmento-Ribeiro, A. B., Simoes, S. P., and Pedroso de Lima, M. C. (2006), 'Transfection of oral cancer cells mediated by transferrin-associated lipoplexes: mechanisms of cell death induced by herpes simplex virus thymidine kinase/ganciclovir therapy', *Biochim Biophys Acta,* 1758 (11), 1703-12.

Neville, B. W. and Day, T. A. (2002), 'Oral cancer and precancerous lesions', *CA Cancer J Clin,* 52 (4), 195-215.

Neville, B. W., Dann, D. D., Allen, C. M., and Bouqout, J. E. (2009), 'Patologia epitelial. In: Patologia Oral & Maxilofacial.', *Guanabara Koogan, Brazil,* 388-423.

Nilsson, O. (2013), 'Profiling of ileal carcinoids', *Neuroendocrinology,* 97 (1), 7-18.

Nix, P. A., Greenman, J., Cawkwell, L., and Stafford, N. D. (2004), 'Defining the criteria for radioresistant laryngeal cancer', *Clin Otolaryngol Allied Sci,* 29 (6), 705-8.

Nobre, R. J., Cruz, E., Real, O., de Almeida, L. P., and Martins, T. C. (2010), 'Characterization of common and rare human papillomaviruses in Portuguese women by the polymerase chain reaction, restriction fragment length polymorphism and sequencing', *J Med Virol,* 82 (6), 1024-32.

Noorlag, R., van Kempen, P. M., Stegeman, I., Koole, R., van Es, R. J., and Willems, S. M. (2015), 'The diagnostic value of 11q13 amplification and protein expression in the detection of nodal metastasis from oral squamous cell carcinoma: a systematic review and meta-analysis', *Virchows Arch,* 466 (4), 363-73.

Noutomi, Y., Oga, A., Uchida, K., Okafuji, M., Ita, M., Kawauchi, S., Furuya, T., Ueyama, Y., and Sasaki, K. (2006), 'Comparative genomic hybridization reveals genetic progression of oral squamous cell carcinoma from dysplasia via two different tumourigenic pathways', *J Pathol,* 210 (1), 67-74.

Nygren, A. O., Ameziane, N., Duarte, H. M., Vijzelaar, R. N., Waisfisz, Q., Hess, C. J., Schouten, J. P., and Errami, A. (2005), 'Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences', *Nucleic Acids Res,* 33 (14), e128.

O'Shaughnessy, J. A., Kelloff, G. J., Gordon, G. B., Dannenberg, A. J., Hong, W. K., Fabian, C. J., Sigman, C. C., Bertagnolli, M. M., Stratton, S. P., Lam, S., Nelson, W. G., Meyskens, F. L., Alberts, D. S., Follen, M., Rustgi, A. K., Papadimitrakopoulou, V., Scardino, P. T., Gazdar, A. F., Wattenberg, L. W., Sporn, M. B., Sakr, W. A., Lippman, S. M., and Von Hoff, D. D. (2002), 'Treatment and prevention of intraepithelial neoplasia: an important target for accelerated new agent development', *Clin Cancer Res,* 8 (2), 314-46.

Ohta, M., Inoue, H., Cotticelli, M. G., Kastury, K., Baffa, R., Palazzo, J., Siprashvili, Z., Mori, M., McCue, P., Druck, T., Croce, C. M., and Huebner, K. (1996), 'The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers', *Cell,* 84 (4), 587-97.

6. References

Oji, Y., Yamamoto, H., Nomura, M., Nakano, Y., Ikeba, A., Nakatsuka, S., Abeno, S., Kiyotoh, E., Jomgeow, T., Sekimoto, M., Nezu, R., Yoshikawa, Y., Inoue, Y., Hosen, N., Kawakami, M., Tsuboi, A., Oka, Y., Ogawa, H., Souda, S., Aozasa, K., Monden, M., and Sugiyama, H. (2003), 'Overexpression of the Wilms' tumor gene WT1 in colorectal adenocarcinoma', *Cancer Sci*, 94 (8), 712-7.

Oliveira, L. R. and Ribeiro-Silva, A. (2011), 'Prognostic significance of immunohistochemical biomarkers in oral squamous cell carcinoma', *Int J Oral Maxillofac Surg*, 40 (3), 298-307.

Omar, E. (2015), 'Current concepts and future of noninvasive procedures for diagnosing oral squamous cell carcinoma--a systematic review', *Head Face Med*, 11, 6.

Ottaviano, L., Schaefer, K. L., Gajewski, M., Huckenbeck, W., Baldus, S., Rogel, U., Mackintosh, C., de Alava, E., Myklebost, O., Kresse, S. H., Meza-Zepeda, L. A., Serra, M., Cleton-Jansen, A. M., Hogendoorn, P. C., Buerger, H., Aigner, T., Gabbert, H. E., and Poremba, C. (2010), 'Molecular characterization of commonly used cell lines for bone tumor research: a trans-European EuroBoNet effort', *Genes Chromosomes Cancer*, 49 (1), 40-51.

Owens, Lewis V., Xu, LiHui, Craven, Rolf J., Dent, Georgette A., Weiner, Tim M., Kornberg, Lori, Liu, Edison T., and Cance, William G. (1995), 'Overexpression of the Focal Adhesion Kinase (p125FAK) in Invasive Human Tumors', *Cancer Research*, 55 (13), 2752-55.

Pai, S. I. and Westra, W. H. (2009), 'Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment', *Annu Rev Pathol*, 4, 49-70.

Parikh, R. A., White, J. S., Huang, X., Schoppy, D. W., Baysal, B. E., Baskaran, R., Bakkenist, C. J., Saunders, W. S., Hsu, L. C., Romkes, M., and Gollin, S. M. (2007), 'Loss of distal 11q is associated with DNA repair deficiency and reduced sensitivity to ionizing radiation in head and neck squamous cell carcinoma', *Genes Chromosomes Cancer*, 46 (8), 761-75.

Pathare, S. M., Gerstung, M., Beerenwinkel, N., Schaffer, A. A., Kannan, S., Pai, P., Pathak, K. A., Borges, A. M., and Mahimkar, M. B. (2011), 'Clinicopathological and prognostic implications of genetic alterations in oral cancers', *Oncol Lett*, 2 (3), 445-51.

Pattje, W. J., Melchers, L. J., Slagter-Menkema, L., Mastik, M. F., Schrijvers, M. L., Gibcus, J. H., Kluin, P. M., Hoegen-Chouvalova, O., van der Laan, B. F., Roodenburg, J. L., van der Wal, J. E., Schuuring, E., and Langendijk, J. A. (2013), 'FADD expression is associated with regional and distant metastasis in squamous cell carcinoma of the head and neck', *Histopathology*, 63 (2), 263-70.

Pearlstein, R. P., Benninger, M. S., Carey, T. E., Zarbo, R. J., Torres, F. X., Rybicki, B. A., and Dyke, D. L. (1998), 'Loss of 18q predicts poor survival of patients with squamous cell carcinoma of the head and neck', *Genes Chromosomes Cancer*, 21 (4), 333-9.

Pecina-Slaus, N., Kljaic, M., and Nikuseva-Martic, T. (2005), 'Loss of heterozygosity of APC and CDH1 genes in laryngeal squamous cell carcinoma', *Pathol Res Pract*, 201 (8-9), 557-63.

Peng, Q., Pevzner, P. A., and Tesler, G. (2006), 'The fragile breakage versus random breakage models of chromosome evolution', *PLoS Comput Biol*, 2 (2), e14.

Pershouse, M. A., El-Naggar, A. K., Hurr, K., Lin, H., Yung, W. K., and Steck, P. A. (1997), 'Deletion mapping of chromosome 4 in head and neck squamous cell carcinoma', *Oncogene*, 14 (3), 369-73.

Peters, I., Dubrowinskaja, N., Kogosov, M., Abbas, M., Hennenlotter, J., von Klot, C., Merseburger, A. S., Stenzl, A., Scherer, R., Kuczyk, M. A., and Serth, J. (2014a), 'Decreased GATA5 mRNA expression associates with CpG island methylation and shortened recurrence-free survival in clear cell renal cell carcinoma', *BMC Cancer*, 14, 101.

Peters, I., Gebauer, K., Dubrowinskaja, N., Atschekzei, F., Kramer, M. W., Hennenlotter, J., Tezval, H., Abbas, M., Scherer, R., Merseburger, A. S., Stenzl, A., Kuczyk, M. A., and Serth, J. (2014b), 'GATA5 CpG island hypermethylation is an independent predictor for poor clinical outcome in renal cell carcinoma', *Oncol Rep*, 31 (4), 1523-30.

Petti, S. (2009), 'Lifestyle risk factors for oral cancer', *Oral Oncol*, 45 (4-5), 340-50.

Pickering, C. R., Zhang, J., Neskey, D. M., Zhao, M., Jasser, S. A., Wang, J., Ward, A., Tsai, C. J., Ortega Alves, M. V., Zhou, J. H., Drummond, J., El-Naggar, A. K., Gibbs, R., Weinstein, J. N., Wheeler, D. A., Wang, J., Frederick, M. J., and Myers, J. N. (2014), 'Squamous cell carcinoma

of the oral tongue in young non-smokers is genomically similar to tumors in older smokers', *Clin Cancer Res*, 20 (14), 3842-8.

Pickering, C. R., Zhang, J., Yoo, S. Y., Bengtsson, L., Moorthy, S., Neskey, D. M., Zhao, M., Ortega Alves, M. V., Chang, K., Drummond, J., Cortez, E., Xie, T. X., Zhang, D., Chung, W., Issa, J. P., Zweidler-McKay, P. A., Wu, X., El-Naggar, A. K., Weinstein, J. N., Wang, J., Muzny, D. M., Gibbs, R. A., Wheeler, D. A., Myers, J. N., and Frederick, M. J. (2013), 'Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers', *Cancer Discov*, 3 (7), 770-81.

Pignon, J. P., Bourhis, J., Domenge, C., and Designe, L. (2000), 'Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: three meta-analyses of updated individual data. MACH-NC Collaborative Group. Meta-Analysis of Chemotherapy on Head and Neck Cancer', *Lancet*, 355 (9208), 949-55.

Pinto-Leite, R., Carreira, I., Melo, J., Ferreira, S. I., Ribeiro, I., Ferreira, J., Filipe, M., Bernardo, C., Arantes-Rodrigues, R., Oliveira, P., and Santos, L. (2014), 'Genomic characterization of three urinary bladder cancer cell lines: understanding genomic types of urinary bladder cancer', *Tumour Biol*, 35 (5), 4599-617.

Poh, C. F., Zhu, Y., Chen, E., Berean, K. W., Wu, L., Zhang, L., and Rosin, M. P. (2012), 'Unique FISH patterns associated with cancer progression of oral dysplasia', *J Dent Res*, 91 (1), 52-7.

Poschl, G. and Seitz, H. K. (2004), 'Alcohol and cancer', *Alcohol Alcohol*, 39 (3), 155-65.

Potash, A. E., Karnell, L. H., Christensen, A. J., Vander Weg, M. W., and Funk, G. F. (2010), 'Continued alcohol use in patients with head and neck cancer', *Head Neck*, 32 (7), 905-12.

Prasad, M., Bernardini, M., Tsalenko, A., Marrano, P., Paderova, J., Lee, C. H., Ben-Dor, A., Barrett, M. T., and Squire, J. A. (2008), 'High definition cytogenetics and oligonucleotide aCGH analyses of cisplatin-resistant ovarian cancer cells', *Genes Chromosomes Cancer*, 47 (5), 427-36.

Prince, A., Aguirre-Ghizo, J., Genden, E., Posner, M., and Sikora, A. (2010), 'Head and neck squamous cell carcinoma: new translational therapies', *Mt Sinai J Med*, 77 (6), 684-99.

Pulte, D. and Brenner, H. (2010), 'Changes in survival in head and neck cancers in the late 20th and early 21st century: a period analysis', *Oncologist*, 15 (9), 994-1001.

Qi, X. W., Zhang, F., Wu, H., Liu, J. L., Zong, B. G., Xu, C., and Jiang, J. (2015), 'Wilms' tumor 1 (WT1) expression and prognosis in solid cancer patients: a systematic review and meta-analysis', *Sci Rep*, 5, 8924.

Ragin, C. C., Modugno, F., and Gollin, S. M. (2007), 'The epidemiology and risk factors of head and neck cancer: a focus on human papillomavirus', *J Dent Res*, 86 (2), 104-14.

Rate, W. R., Garrett, P., Hamaker, R., Singer, M., Pugh, N., Ross, D., Haerr, R., and Charles, G. (1991), 'Intraoperative radiation therapy for recurrent head and neck cancer', *Cancer*, 67 (11), 2738-40.

Redon, R., Muller, D., Caulee, K., Wanherdrick, K., Abecassis, J., and du Manoir, S. (2001), 'A simple specific pattern of chromosomal aberrations at early stages of head and neck squamous cell carcinomas: PIK3CA but not p63 gene as a likely target of 3q26-qter gains', *Cancer Res*, 61 (10), 4122-9.

Reibel, J. (2003), 'Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics', *Crit Rev Oral Biol Med*, 14 (1), 47-62.

Reis, P. P., Rogatto, S. R., Kowalski, L. P., Nishimoto, I. N., Montovani, J. C., Corpus, G., Squire, J. A., and Kamel-Reid, S. (2002), 'Quantitative real-time PCR identifies a critical region of deletion on 22q13 related to prognosis in oral cancer', *Oncogene*, 21 (42), 6480-7.

Rettori, M. M., de Carvalho, A. C., Longo, A. L., de Oliveira, C. Z., Kowalski, L. P., Carvalho, A. L., and Vettore, A. L. (2013), 'TIMP3 and CCNA1 hypermethylation in HNSCC is associated with an increased incidence of second primary tumors', *J Transl Med*, 11, 316.

Ribeiro, I. P., Marques, F., Caramelo, F., Ferrao, J., Prazeres, H., Juliao, M. J., Rifi, W., Savola, S., de Melo, J. B., Baptista, I. P., and Carreira, I. M. (2014a), 'Genetic imbalances detected by multiplex ligation-dependent probe amplification in a cohort of patients with oral squamous cell carcinoma-the first step towards clinical personalized medicine', *Tumour Biol*.

Ribeiro, I. P., Marques, F., Caramelo, F., Pereira, J., Patricio, M., Prazeres, H., Ferrao, J., Juliao, M. J., Castelo-Branco, M., de Melo, J. B., Baptista, I. P., and Carreira, I. M. (2014b), 'Genetic gains

and losses in oral squamous cell carcinoma: impact on clinical management', *Cell Oncol (Dordr)*.

Ribeiro, I. P., Caramelo, F., Marques, F., Domingues, A., Mesquita, M., Barroso, L., Prazeres, H., Juliao, M. J., Baptista, I. P., Ferreira, A., Melo, J. B., and Carreira, I. M. (2016), 'WT1, MSH6, GATA5 and PAX5 as epigenetic oral squamous cell carcinoma biomarkers - a short report', *Cell Oncol (Dordr)*, 39 (6), 573-82.

Richards, R. I. (2001), 'Fragile and unstable chromosomes in cancer: causes and consequences', *Trends Genet*, 17 (6), 339-45.

Rivera, C. (2015), 'Essentials of oral cancer', *Int J Clin Exp Pathol*, 8 (9), 11884-94.

Robbins, K. T., Clayman, G., Levine, P. A., Medina, J., Sessions, R., Shaha, A., Som, P., Wolf, G. T., American, Head, Neck, Society, American Academy of, Otolaryngology-Head, and Neck, Surgery (2002), 'Neck dissection classification update: revisions proposed by the American Head and Neck Society and the American Academy of Otolaryngology-Head and Neck Surgery', *Arch Otolaryngol Head Neck Surg*, 128 (7), 751-8.

Rocco, J. W. and Sidransky, D. (2001), 'p16(MTS-1/CDKN2/INK4a) in cancer progression', *Exp Cell Res*, 264 (1), 42-55.

Rodrigo, J. P., Lazo, P. S., Ramos, S., Alvarez, I., and Suarez, C. (1996), 'MYC amplification in squamous cell carcinomas of the head and neck', *Arch Otolaryngol Head Neck Surg*, 122 (5), 504-7.

Rodriguez, T., Altieri, A., Chatenoud, L., Gallus, S., Bosetti, C., Negri, E., Franceschi, S., Levi, F., Talamini, R., and La Vecchia, C. (2004), 'Risk factors for oral and pharyngeal cancer in young adults', *Oral Oncol*, 40 (2), 207-13.

Roepman, P., Kemmeren, P., Wessels, L. F., Slootweg, P. J., and Holstege, F. C. (2006), 'Multiple robust signatures for detecting lymph node metastasis in head and neck cancer', *Cancer Res*, 66 (4), 2361-6.

Ross, D. T., Scherf, U., Eisen, M. B., Perou, C. M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S. S., Van de Rijn, M., Waltham, M., Pergamenschikov, A., Lee, J. C., Lashkari, D., Shalon, D., Myers, T. G., Weinstein, J. N., Botstein, D., and Brown, P. O. (2000), 'Systematic variation in gene expression patterns in human cancer cell lines', *Nat Genet*, 24 (3), 227-35.

Rothenberg, S. M. and Ellisen, L. W. (2012), 'The molecular pathogenesis of head and neck squamous cell carcinoma', *J Clin Invest*, 122 (6), 1951-7.

Rothschild, B. L., Shim, A. H., Ammer, A. G., Kelley, L. C., Irby, K. B., Head, J. A., Chen, L., Varella-Garcia, M., Sacks, P. G., Frederick, B., Raben, D., and Weed, S. A. (2006), 'Cortactin overexpression regulates actin-related protein 2/3 complex activity, motility, and invasion in carcinomas with chromosome 11q13 amplification', *Cancer Res*, 66 (16), 8017-25.

Ruiz-Herrera, A. and Robinson, T. J. (2007), 'Chromosomal instability in Afrotheria: fragile sites, evolutionary breakpoints and phylogenetic inference from genome sequence assemblies', *BMC Evol Biol*, 7, 199.

Ruiz-Herrera, A. and Robinson, T. J. (2008), 'Evolutionary plasticity and cancer breakpoints in human chromosome 3', *Bioessays*, 30 (11-12), 1126-37.

Ruiz, C., Martins, J. R., Rudin, F., Schneider, S., Dietsche, T., Fischer, C. A., Tornillo, L., Terracciano, L. M., Schreiber, R., Bubendorf, L., and Kunzelmann, K. (2012), 'Enhanced expression of ANO1 in head and neck squamous cell carcinoma causes cell migration and correlates with poor prognosis', *PLoS One*, 7 (8), e43265.

Sabbir, M. G., Roy, A., Mandal, S., Dam, A., Roychoudhury, S., and Panda, C. K. (2006), 'Deletion mapping of chromosome 13q in head and neck squamous cell carcinoma in Indian patients: correlation with prognosis of the tumour', *Int J Exp Pathol*, 87 (2), 151-61.

Salazar, C., Nagadia, R., Pandit, P., Cooper-White, J., Banerjee, N., Dimitrova, N., Coman, W. B., and Punyadeera, C. (2014), 'A novel saliva-based microRNA biomarker panel to detect head and neck cancers', *Cell Oncol (Dordr)*, 37 (5), 331-8.

Saldivar, J. C., Bene, J., Hosseini, S. A., Miuma, S., Horton, S., Heerema, N. A., and Huebner, K. (2013), 'Characterization of the role of Fhit in suppression of DNA damage', *Adv Biol Regul*, 53 (1), 77-85.

Saldivar, J. C., Miuma, S., Bene, J., Hosseini, S. A., Shibata, H., Sun, J., Wheeler, L. J., Mathews, C. K., and Huebner, K. (2012), 'Initiation of genome instability and preneoplastic processes through loss of Fhit expression', *PLoS Genet*, 8 (11), e1003077.

Sanderson, R. J. and Ironside, J. A. (2002), 'Squamous cell carcinomas of the head and neck', *BMJ*, 325 (7368), 822-7.

Sano, D. and Myers, J. N. (2007), 'Metastasis of squamous cell carcinoma of the oral tongue', *Cancer Metastasis Rev*, 26 (3-4), 645-62.

Saranath, D., Panchal, R. G., Nair, R., Mehta, A. R., Sanghavi, V., Sumegi, J., Klein, G., and Deo, M. G. (1989), 'Oncogene amplification in squamous cell carcinoma of the oral cavity', *Jpn J Cancer Res*, 80 (5), 430-7.

Sarni, D. and Kerem, B. (2016), 'The complex nature of fragile site plasticity and its importance in cancer', *Curr Opin Cell Biol*, 40, 131-6.

Saunders, W. S., Shuster, M., Huang, X., Gharaibeh, B., Enyenihi, A. H., Petersen, I., and Gollin, S. M. (2000), 'Chromosomal instability and cytoskeletal defects in oral cancer cells', *Proc Natl Acad Sci U S A*, 97 (1), 303-8.

Scherer, S. E. Muzny, D. M. Buhay, C. J. Chen, R. Cree, A. Ding, Y. Dugan-Rocha, S. Gill, R. Gunaratne, P. Harris, R. A. Hawes, A. C. Hernandez, J. Hodgson, A. V. Hume, J. Jackson, A. Khan, Z. M. Kovar-Smith, C. Lewis, L. R. Lozado, R. J. Metzker, M. L. Milosavljevic, A. Miner, G. R. Montgomery, K. T. Morgan, M. B. Nazareth, L. V. Scott, G. Sodergren, E. Song, X. Z. Steffen, D. Lovering, R. C. Wheeler, D. A. Worley, K. C. Yuan, Y. Zhang, Z. Adams, C. Q. Ansari-Lari, M. A. Ayele, M. Brown, M. J. Chen, G. Chen, Z. Clerc-Blankenburg, K. P. Davis, C. Delgado, O. Dinh, H. H. Draper, H. Gonzalez-Garay, M. L. Havlak, P. Jackson, L. R. Jacob, L. S. Kelly, S. H. Li, L. Li, Z. Liu, J. Liu, W. Lu, J. Maheshwari, M. Nguyen, B. V. Okwuonu, G. O. Pasternak, S. Perez, L. M. Plopper, F. J. Santibanez, J. Shen, H. Tabor, P. E. Verduzco, D. Waldron, L. Wang, Q. Williams, G. A. Zhang, J. Zhou, J. Allen, C. C. Amin, A. G. Anyalebechi, V. Bailey, M. Barbaria, J. A. Bimage, K. E. Bryant, N. P. Burch, P. E. Burkett, C. E. Burrell, K. L. Calderon, E. Cardenas, V. Carter, K. Casias, K. Cavazos, I. Cavazos, S. R. Ceasar, H. Chacko, J. Chan, S. N. Chavez, D. Christopoulos, C. Chu, J. Cockrell, R. Cox, C. D. Dang, M. Dathorne, S. R. David, R. Davis, C. M. Davy-Carroll, L. Deshazo, D. R. Donlin, J. E. D'Souza, L. Eaves, K. A. Egan, A. Emery-Cohen, A. J. Escotto, M. Flagg, N. Forbes, L. D. Gabisi, A. M. Garza, M. Hamilton, C. Henderson, N. Hernez, O. Hines, S. Hogues, M. E. Huang, M. Idlebird, D. G. Johnson, R. Jolivet, A. Jones, S. Kagan, R. King, L. M. Leal, B. Lebow, H. Lee, S. LeVan, J. M. Lewis, L. C. London, P. Lorensuhewa, L. M. Loulseged, H. Lovett, D. A. Lucier, A. Lucier, R. L. Ma, J. Madu, R. C. Mapua, P. Martindale, A. D. Martinez, E. Massey, E. Mawhiney, S. Meador, M. G. Mendez, S. Mercado, C. Mercado, I. C. Merritt, C. E. Miner, Z. L. Minja, E. Mitchell, T. Mohabbat, F. Mohabbat, K. Montgomery, B. Moore, N. Morris, S. Munidasa, M. Ngo, R. N. Nguyen, N. B. Nickerson, E. Nwaokelemeh, O. O. Nwokenkwo, S. Obregon, M. Oguh, M. Oragunye, N. Oviedo, R. J. Parish, B. J. Parker, D. N. Parrish, J. Parks, K. L. Paul, H. A. Payton, B. A. Perez, A. Perrin, W. Pickens, A. Primus, E. L. Pu, L. L. Puazo, M. Quiles, M. M. Quiroz, J. B. Rabata, D. Reeves, K. Ruiz, S. J. Shao, H. Sisson, I. Sonaike, T. Sorelle, R. P. Sutton, A. E. Svatek, A. F. Svetz, L. A. Tamerisa, K. S. Taylor, T. R. Teague, B. Thomas, N. Thorn, R. D. Trejos, Z. Y. Trevino, B. K. Ukegbu, O. N. Urban, J. B. Vasquez, L. I. Vera, V. A. Villasana, D. M. Wang, L. Ward-Moore, S. Warren, J. T. Wei, X. White, F. Williamson, A. L. Wleczyk, R. Wooden, H. S. Wooden, S. H. Yen, J. Yoon, L. Yoon, V. Zorrilla, S. E. Nelson, D. Kucherlapati, R. Weinstock, G. Gibbs, R. A. Baylor, College of Medicine Human Genome Sequencing Center Sequence Production, Team (2006), 'The finished DNA sequence of human chromosome 12', *Nature*, 440 (7082), 346-51.

Schlessinger, J. (2002), 'Ligand-induced, receptor-mediated dimerization and activation of EGF receptor', *Cell*, 110 (6), 669-72.

Schouten, J. P., McElgunn, C. J., Waaijer, R., Zwijnenburg, D., Diepvens, F., and Pals, G. (2002), 'Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification', *Nucleic Acids Res*, 30 (12), e57.

6. References

Schrock, A., Bode, M., Goke, F. J., Bareiss, P. M., Schairer, R., Wang, H., Weichert, W., Franzen, A., Kirsten, R., van Bremen, T., Queisser, A., Kristiansen, G., Heasley, L., Bootz, F., Lengerke, C., and Perner, S. (2014), 'Expression and role of the embryonic protein SOX2 in head and neck squamous cell carcinoma', *Carcinogenesis*, 35 (7), 1636-42.

Schutt, C., Bumm, K., Mirandola, L., Bernardini, G., D'Cunha, N., Tijani, L., Nguyen, D., Cordero, J., Jenkins, M. R., Cobos, E., Kast, W. M., and Chiriva-Internati, M. (2012), 'Immunological treatment options for locoregionally advanced head and neck squamous cell carcinoma', *Int Rev Immunol*, 31 (1), 22-42.

Schuuring, E. (1995), 'The involvement of the chromosome 11q13 region in human malignancies: cyclin D1 and EMS1 are two new candidate oncogenes--a review', *Gene*, 159 (1), 83-96.

Scully, C., Field, J. K., and Tanzawa, H. (2000), 'Genetic aberrations in oral or head and neck squamous cell carcinoma (SCCHN): 1. Carcinogen metabolism, DNA repair and cell cycle control', *Oral Oncol*, 36 (3), 256-63.

Seibold, S., Rudroff, C., Weber, M., Galle, J., Wanner, C., and Marx, M. (2003), 'Identification of a new tumor suppressor gene located at chromosome 8p21.3-22', *FASEB J*, 17 (9), 1180-2.

Sever, R. and Brugge, J. S. (2015), 'Signal transduction in cancer', *Cold Spring Harb Perspect Med*, 5 (4).

Severino, P., Alvares, A. M., Michaluart, P., Jr., Okamoto, O. K., Nunes, F. D., Moreira-Filho, C. A., Tajara, E. H., Head, and Neck Genome Project, Gencapo (2008), 'Global gene expression profiling of oral cavity cancers suggests molecular heterogeneity within anatomic subsites', *BMC Res Notes*, 1, 113.

Shaffer, L. G. and Lupski, J. R. (2000), 'Molecular mechanisms for constitutional chromosomal rearrangements in humans', *Annu Rev Genet*, 34, 297-329.

Shah, F. D., Begum, R., Vajaria, B. N., Patel, K. R., Patel, J. B., Shukla, S. N., and Patel, P. S. (2011), 'A review on salivary genomics and proteomics biomarkers in oral cancer', *Indian J Clin Biochem*, 26 (4), 326-34.

Shamji, A. F., Nghiem, P., and Schreiber, S. L. (2003), 'Integration of growth factor and nutrient signaling: implications for cancer biology', *Mol Cell*, 12 (2), 271-80.

Sharma, S., Kelly, T. K., and Jones, P. A. (2010a), 'Epigenetics in cancer', *Carcinogenesis*, 31 (1), 27-36.

Sharma, S. V., Haber, D. A., and Settleman, J. (2010b), 'Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents', *Nat Rev Cancer*, 10 (4), 241-53.

Shaw, R. and Beasley, N. (2016), 'Aetiology and risk factors for head and neck cancer: United Kingdom National Multidisciplinary Guidelines', *J Laryngol Otol*, 130 (S2), S9-S12.

Sheu, J. J., Hua, C. H., Wan, L., Lin, Y. J., Lai, M. T., Tseng, H. C., Jinawath, N., Tsai, M. H., Chang, N. W., Lin, C. F., Lin, C. C., Hsieh, L. J., Wang, T. L., Shih Ie, M., and Tsai, F. J. (2009), 'Functional genomic analysis identified epidermal growth factor receptor activation as the most common genetic event in oral squamous cell carcinoma', *Cancer Res*, 69 (6), 2568-76.

Shipitsyna, Elena, Savicheva, Alevtina, Sokolovskiy, Evgenij, Ballard, Ronald C., Domeika, Marius, Unemo, Magnus, and Jensen, J??rgen S. (2010), 'Guidelines for the laboratory diagnosis of Mycoplasma genitalium infections in East European countries', *Acta Dermato-Venereologica*, 90, 461-67.

Singchat, W., Hitakomate, E., Rerkarmnuaychoke, B., Suntronpong, A., Fu, B., Bodhisuwan, W., Peyachoknagul, S., Yang, F., Koontongkaew, S., and Srikulnath, K. (2016), 'Genomic Alteration in Head and Neck Squamous Cell Carcinoma (HNSCC) Cell Lines Inferred from Karyotyping, Molecular Cytogenetics, and Array Comparative Genomic Hybridization', *PLoS One*, 11 (8), e0160901.

Singh, B., Gogineni, S., Goberdhan, A., Sacks, P., Shaha, A., Shah, J., and Rao, P. (2001a), 'Spectral karyotyping analysis of head and neck squamous cell carcinoma', *Laryngoscope*, 111 (9), 1545-50.

Singh, B., Gogineni, S. K., Sacks, P. G., Shaha, A. R., Shah, J. P., Stoffel, A., and Rao, P. H. (2001b), 'Molecular cytogenetic characterization of head and neck squamous cell carcinoma and refinement of 3q amplification', *Cancer Res*, 61 (11), 4506-13.

Skinner, H. D., Giri, U., Yang, L., Woo, S. H., Story, M. D., Pickering, C. R., Byers, L. A., Williams, M. D., El-Naggar, A., Wang, J., Diao, L., Shen, L., Fan, Y. H., Molkentine, D. P., Beadle, B. M.,

Meyn, R. E., Myers, J. N., and Heymach, J. V. (2016), 'Proteomic Profiling Identifies PTK2/FAK as a Driver of Radioresistance in HPV-negative Head and Neck Cancer', *Clin Cancer Res*, 22 (18), 4643-50.

Slaughter, D. P., Southwick, H. W., and Smejkal, W. (1953), 'Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin', *Cancer*, 6 (5), 963-8.

Soon, L. L., Yie, T. A., Shvarts, A., Levine, A. J., Su, F., and Tchou-Wong, K. M. (2003), 'Overexpression of WISP-1 down-regulated motility and invasion of lung cancer cells through inhibition of Rac activation', *J Biol Chem*, 278 (13), 11465-70.

Sparano, A., Quesnelle, K. M., Kumar, M. S., Wang, Y., Sylvester, A. J., Feldman, M., Sewell, D. A., Weinstein, G. S., and Brose, M. S. (2006), 'Genome-wide profiling of oral squamous cell carcinoma by array-based comparative genomic hybridization', *Laryngoscope*, 116 (5), 735-41.

Specenier, P. and Vermorken, J. B. (2013), 'Cetuximab: its unique place in head and neck cancer treatment', *Biologics*, 7, 77-90.

Specenier, P. M. and Vermorken, J. B. (2008), 'Recurrent head and neck cancer: current treatment and future prospects', *Expert Rev Anticancer Ther*, 8 (3), 375-91.

Squire, J. A., Bayani, J., Luk, C., Unwin, L., Tokunaga, J., MacMillan, C., Irish, J., Brown, D., Gullane, P., and Kamel-Reid, S. (2002), 'Molecular cytogenetic analysis of head and neck squamous cell carcinoma: By comparative genomic hybridization, spectral karyotyping, and expression array analysis', *Head Neck*, 24 (9), 874-87.

Stadler, M. E., Patel, M. R., Couch, M. E., and Hayes, D. N. (2008), 'Molecular biology of head and neck cancer: risks and pathways', *Hematol Oncol Clin North Am*, 22 (6), 1099-124, vii.

Stephen, Josena K, Chen, Kang Mei, Shah, Veena, Havard, Shaleta, Kapke, Alissa, Lu, Mei, Benninger, Michael S, and Worsham, Maria J (2010), 'DNA hypermethylation markers of poor outcome in laryngeal cancer.', *Clinical epigenetics*, 1, 61-69.

Stephens, P. J., Greenman, C. D., Fu, B., Yang, F., Bignell, G. R., Mudie, L. J., Pleasance, E. D., Lau, K. W., Beare, D., Stebbings, L. A., McLaren, S., Lin, M. L., McBride, D. J., Varela, I., Nik-Zainal, S., Leroy, C., Jia, M., Menzies, A., Butler, A. P., Teague, J. W., Quail, M. A., Burton, J., Swerdlow, H., Carter, N. P., Morsberger, L. A., Iacobuzio-Donahue, C., Follows, G. A., Green, A. R., Flanagan, A. M., Stratton, M. R., Futreal, P. A., and Campbell, P. J. (2011), 'Massive genomic rearrangement acquired in a single catastrophic event during cancer development', *Cell*, 144 (1), 27-40.

Sticht, C., Hofele, C., Flechtenmacher, C., Bosch, F. X., Freier, K., Lichter, P., and Joos, S. (2005), 'Amplification of Cyclin L1 is associated with lymph node metastases in head and neck squamous cell carcinoma (HNSCC)', *Br J Cancer*, 92 (4), 770-4.

Stratton, M. R., Campbell, P. J., and Futreal, P. A. (2009), 'The cancer genome', *Nature*, 458 (7239), 719-24.

Suda, T., Hama, T., Kondo, S., Yuza, Y., Yoshikawa, M., Urashima, M., Kato, T., and Moriyama, H. (2012), 'Copy number amplification of the PIK3CA gene is associated with poor prognosis in non-lymph node metastatic head and neck squamous cell carcinoma', *BMC Cancer*, 12, 416.

Sugiyama, H. (2010), 'WT1 (Wilms' tumor gene 1): biology and cancer immunotherapy', *Jpn J Clin Oncol*, 40 (5), 377-87.

Sun, P. C., Uppaluri, R., Schmidt, A. P., Pashia, M. E., Quant, E. C., Sunwoo, J. B., Gollin, S. M., and Scholnick, S. B. (2001), 'Transcript map of the 8p23 putative tumor suppressor region', *Genomics*, 75 (1-3), 17-25.

Tabor, M. P., Brakenhoff, R. H., Ruijter-Schippers, H. J., Van Der Wal, J. E., Snow, G. B., Leemans, C. R., and Braakhuis, B. J. (2002), 'Multiple head and neck tumors frequently originate from a single preneoplastic lesion', *Am J Pathol*, 161 (3), 1051-60.

Tabor, M. P., Brakenhoff, R. H., van Houten, V. M., Kummer, J. A., Snel, M. H., Snijders, P. J., Snow, G. B., Leemans, C. R., and Braakhuis, B. J. (2001), 'Persistence of genetically altered fields in head and neck cancer patients: biological and clinical implications', *Clin Cancer Res*, 7 (6), 1523-32.

6. References

Tafe, Laura J. (2017), 'The Molecular Pathology of Head and Neck Squamous Cell Carcinoma', in William B. Coleman and Gregory J. Tsongalis (eds.), *The Molecular Basis of Human Cancer* (New York, NY: Springer New York), 589-601.

Takebayashi, S., Hickson, A., Ogawa, T., Jung, K. Y., Mineta, H., Ueda, Y., Grenman, R., Fisher, S. G., and Carey, T. E. (2004), 'Loss of chromosome arm 18q with tumor progression in head and neck squamous cancer', *Genes Chromosomes Cancer*, 41 (2), 145-54.

Tanaka, T., Wiesener, M., Bernhardt, W., Eckardt, K. U., and Warnecke, C. (2009), 'The human HIF (hypoxia-inducible factor)-3alpha gene is a HIF-1 target gene and may modulate hypoxic gene induction', *Biochem J*, 424 (1), 143-51.

Temam, S., Kawaguchi, H., El-Naggar, A. K., Jelinek, J., Tang, H., Liu, D. D., Lang, W., Issa, J. P., Lee, J. J., and Mao, L. (2007), 'Epidermal growth factor receptor copy number alterations correlate with poor clinical outcome in patients with head and neck squamous cancer', *J Clin Oncol*, 25 (16), 2164-70.

Thariat, J., Vignot, S., Lapierre, A., Falk, A. T., Guigay, J., Van Obberghen-Schilling, E., and Milano, G. (2015), 'Integrating genomics in head and neck cancer treatment: Promises and pitfalls', *Crit Rev Oncol Hematol*, 95 (3), 397-406.

The Cancer Genome Atlas, Network (2015), 'Comprehensive genomic characterization of head and neck squamous cell carcinomas', *Nature*, 517 (7536), 576-82.

Thomas, G., Hashibe, M., Jacob, B. J., Ramadas, K., Mathew, B., Sankaranarayanan, R., and Zhang, Z. F. (2003), 'Risk factors for multiple oral premalignant lesions', *Int J Cancer*, 107 (2), 285-91.

Thomas, G. R., Nadiminti, H., and Regalado, J. (2005), 'Molecular predictors of clinical outcome in patients with head and neck squamous cell carcinoma', *Int J Exp Pathol*, 86 (6), 347-63.

Tokumaru, Y., Yamashita, K., Osada, M., Nomoto, S., Sun, D. I., Xiao, Y., Hoque, M. O., Westra, W. H., Califano, J. A., and Sidransky, D. (2004), 'Inverse correlation between cyclin A1 hypermethylation and p53 mutation in head and neck cancer identified by reversal of epigenetic silencing', *Cancer Res*, 64 (17), 5982-7.

Tolosi, L., Theissen, J., Halachev, K., Hero, B., Berthold, F., and Lengauer, T. (2013), 'A method for finding consensus breakpoints in the cancer genome from copy number data', *Bioinformatics*, 29 (14), 1793-800.

Tornesello, M. L., Perri, F., Buonaguro, L., Ionna, F., Buonaguro, F. M., and Caponigro, F. (2014), 'HPV-related oropharyngeal cancers: from pathogenesis to new therapeutic approaches', *Cancer Lett*, 351 (2), 198-205.

Treangen, T. J. and Salzberg, S. L. (2011), 'Repetitive DNA and next-generation sequencing: computational challenges and solutions', *Nat Rev Genet*, 13 (1), 36-46.

Trivedy, C. R., Craig, G., and Warnakulasuriya, S. (2002), 'The oral health consequences of chewing areca nut', *Addict Biol*, 7 (1), 115-25.

Tsao, A. S., Garden, A. S., Kies, M. S., Morrison, W., Feng, L., Lee, J. J., Khuri, F., Zinner, R., Myers, J., Papadimitrakopoulou, V., Lewin, J., Clayman, G. L., Ang, K. K., and Glisson, B. S. (2006), 'Phase I/II study of docetaxel, cisplatin, and concomitant boost radiation for locally advanced squamous cell cancer of the head and neck', *J Clin Oncol*, 24 (25), 4163-9.

Uchida, K., Oga, A., Okafuji, M., Mihara, M., Kawauchi, S., Furuya, T., Chochi, Y., Ueyama, Y., and Sasaki, K. (2006), 'Molecular cytogenetic analysis of oral squamous cell carcinomas by comparative genomic hybridization, spectral karyotyping, and fluorescence in situ hybridization', *Cancer Genet Cytogenet*, 167 (2), 109-16.

Uzawa, N., Akanuma, D., Negishi, A., Iwaki, H., Uzawa, Y., Amagasa, T., and Yoshida, M. A. (2001), 'Homozygous deletions on the short arm of chromosome 3 in human oral squamous cell carcinomas', *Oral Oncol*, 37 (4), 351-6.

van den Berg, R. M., Snijders, P. J., Grunberg, K., Kooi, C., Spreeuwenberg, M. D., Meijer, C. J., Postmus, P. E., Smit, E. F., and Steenbergen, R. D. (2011), 'Comprehensive CADM1 promoter methylation analysis in NSCLC and normal lung specimens', *Lung Cancer*, 72 (3), 316-21.

van den Broek, G. B., Wreesmann, V. B., van den Brekel, M. W., Rasch, C. R., Balm, A. J., and Rao, P. H. (2007), 'Genetic abnormalities associated with chemoradiation resistance of head and neck squamous cell carcinoma', *Clin Cancer Res*, 13 (15 Pt 1), 4386-91.

van der Riet, P., Nawroz, H., Hruban, R. H., Corio, R., Tokino, K., Koch, W., and Sidransky, D. (1994), 'Frequent loss of chromosome 9p21-22 early in head and neck cancer progression', *Cancer Res,* 54 (5), 1156-8.

van der Waal, I. (2014), 'Oral potentially malignant disorders: is malignant transformation predictable and preventable?', *Med Oral Patol Oral Cir Bucal,* 19 (4), e386-90.

van der Waal, I., de Bree, R., Brakenhoff, R., and Coebergh, J. W. (2011), 'Early diagnosis in primary oral cancer: is it possible?', *Med Oral Patol Oral Cir Bucal,* 16 (3), e300-5.

Van Dyke, D. L., Worsham, M. J., Benninger, M. S., Krause, C. J., Baker, S. R., Wolf, G. T., Drumheller, T., Tilley, B. C., and Carey, T. E. (1994), 'Recurrent cytogenetic abnormalities in squamous cell carcinomas of the head and neck region', *Genes Chromosomes Cancer,* 9 (3), 192-206.

van Kempen, P. M., van Bockel, L., Braunius, W. W., Moelans, C. B., van Olst, M., de Jong, R., Stegeman, I., van Diest, P. J., Grolman, W., and Willems, S. M. (2014), 'HPV-positive oropharyngeal squamous cell carcinoma is associated with TIMP3 and CADM1 promoter hypermethylation', *Cancer Med,* 3 (5), 1185-96.

van Kempen, P. M., Noorlag, R., Braunius, W. W., Moelans, C. B., Rifi, W., Savola, S., Koole, R., Grolman, W., van Es, R. J., and Willems, S. M. (2015), 'Clinical relevance of copy number profiling in oral and oropharyngeal squamous cell carcinoma', *Cancer Med,* 4 (10), 1525-35.

van Monsjou, H. S., Wreesmann, V. B., van den Brekel, M. W., and Balm, A. J. (2013), 'Head and neck squamous cell carcinoma in young patients', *Oral Oncol,* 49 (12), 1097-102.

Verma, R., Singh, A., Badni, M., Chandra, A., Gupta, S., and Verma, R. (2015), 'Evaluation of exfoliative cytology in the diagnosis of oral premalignant and malignant lesions: A cytomorphometric analysis', *Dent Res J (Isfahan),* 12 (1), 83-8.

Vermorken, J. B. (2005), 'Medical treatment in head and neck cancer', *Ann Oncol,* 16 Suppl 2, ii258-64.

Vermorken, J. B., Mesia, R., Rivera, F., Remenar, E., Kawecki, A., Rottey, S., Erfan, J., Zabolotnyy, D., Kienzer, H. R., Cupissol, D., Peyrade, F., Benasso, M., Vynnychenko, I., De Raucourt, D., Bokemeyer, C., Schueler, A., Amellal, N., and Hitt, R. (2008), 'Platinum-based chemotherapy plus cetuximab in head and neck cancer', *N Engl J Med,* 359 (11), 1116-27.

Vigneswaran, N. and Williams, M. D. (2014), 'Epidemiologic trends in head and neck cancer and aids in diagnosis', *Oral Maxillofac Surg Clin North Am,* 26 (2), 123-41.

Virgilio, L., Shuster, M., Gollin, S. M., Veronese, M. L., Ohta, M., Huebner, K., and Croce, C. M. (1996), 'FHIT gene alterations in head and neck squamous cell carcinomas', *Proc Natl Acad Sci U S A,* 93 (18), 9770-5.

Vora, H. H., Shah, N. G., Patel, D. D., Trivedi, T. I., and Chikhlikar, P. R. (2003a), 'Prognostic significance of biomarkers in squamous cell carcinoma of the tongue: multivariate analysis', *J Surg Oncol,* 82 (1), 34-50.

Vora, H. H., Shah, N. G., Patel, D. D., Trivedi, T. I., and Choksi, T. J. (2003b), 'BRCA1 expression in leukoplakia and carcinoma of the tongue', *J Surg Oncol,* 83 (4), 232-40.

Walden, M. J. and Aygun, N. (2013), 'Head and neck cancer', *Semin Roentgenol,* 48 (1), 75-86.

Wang, W., Han, S., Yao, Z., Li, X., Huang, P., Zhang, M., Chen, Y., and He, J. (2011), 'A Study of Epidemiologic and Recurrence Factors of Oral Cancer', *J Oral Maxillofac Surg*.

Wang, X. L., Uzawa, K., Imai, F. L., and Tanzawa, H. (1999), 'Localization of a novel tumor suppressor gene associated with human oral cancer on chromosome 4q25', *Oncogene,* 18 (3), 823-5.

Warnakulasuriya, S. (2000), 'Lack of molecular markers to predict malignant potential of oral precancer', *J Pathol,* 190 (4), 407-9.

Warnakulasuriya, S. and Ariyawardana, A. (2016), 'Malignant transformation of oral leukoplakia: a systematic review of observational studies', *J Oral Pathol Med,* 45 (3), 155-66.

Warren S and Gates O (1932), 'Multiple primary malignant tumors. A survey of the literature and a statistical study.', *Am J Cancer,* 16, 1358–414.

Webster, J. A., Beck, A. H., Sharma, M., Espinosa, I., Weigelt, B., Schreuder, M., Montgomery, K. D., Jensen, K. C., van de Rijn, M., and West, R. (2010), 'Variations in stromal signatures in breast and colorectal cancer metastases', *J Pathol,* 222 (2), 158-65.

6. References

Weiss, D., Stockmann, C., Schrodter, K., and Rudack, C. (2013), 'Protein expression and promoter methylation of the candidate biomarker TCF21 in head and neck squamous cell carcinoma', *Cell Oncol (Dordr)*, 36 (3), 213-24.

White, J. S., Weissfeld, J. L., Ragin, C. C., Rossie, K. M., Martin, C. L., Shuster, M., Ishwad, C. S., Law, J. C., Myers, E. N., Johnson, J. T., and Gollin, S. M. (2007), 'The influence of clinical and demographic risk factors on the establishment of head and neck squamous cell carcinoma cell lines', *Oral Oncol*, 43 (7), 701-12.

William, W. N., Jr. (2012), 'Oral premalignant lesions: any progress with systemic therapies?', *Curr Opin Oncol*, 24 (3), 205-10.

Wittekind, C., Greene, F.L., Hutter, R.V.P., Klimpfinger, M., and L.H.Sobin (2003), *TNM atlas. Illustrated guide to the TNM/pTNM classification of malignant tumours*, ed. International Union Against Cancer (UICC) (5th ed edn.; Berlin-Heidelberg-New York: Springer-Verlag).

Wittekindt, C., Wagner, S., Mayer, C. S., and Klussmann, J. P. (2012), 'Basics of tumor development and importance of human papilloma virus (HPV) for head and neck cancer', *GMS Curr Top Otorhinolaryngol Head Neck Surg*, 11, Doc09.

Wong, L. Y., Wei, W. I., Lam, L. K., and Yuen, A. P. (2003), 'Salvage of recurrent head and neck squamous cell carcinoma after primary curative surgery', *Head Neck*, 25 (11), 953-9.

Worsham, M. J. (2011), 'Identifying the risk factors for late-stage head and neck cancer', *Expert Rev Anticancer Ther*, 11 (9), 1321-5.

Worsham, M. J., Ali, H., Dragovic, J., and Schweitzer, V. P. (2012), 'Molecular characterization of head and neck cancer: how close to personalized targeted therapy?', *Mol Diagn Ther*, 16 (4), 209-22.

Worsham, M. J., Wolman, S. R., Carey, T. E., Zarbo, R. J., Benninger, M. S., and Van Dyke, D. L. (1999), 'Chromosomal aberrations identified in culture of squamous carcinomas are confirmed by fluorescence in situ hybridisation', *Mol Pathol*, 52 (1), 42-6.

Worsham, M. J., Pals, G., Schouten, J. P., Van Spaendonk, R. M., Concus, A., Carey, T. E., and Benninger, M. S. (2003), 'Delineating genetic pathways of disease progression in head and neck squamous cell carcinoma', *Arch Otolaryngol Head Neck Surg*, 129 (7), 702-8.

Worsham, M. J., Chen, K. M., Tiwari, N., Pals, G., Schouten, J. P., Sethi, S., and Benninger, M. S. (2006), 'Fine-mapping loss of gene architecture at the CDKN2B (p15INK4b), CDKN2A (p14ARF, p16INK4a), and MTAP genes in head and neck squamous cell carcinoma', *Arch Otolaryngol Head Neck Surg*, 132 (4), 409-15.

Worsham, M. J., Stephen, J. K., Chen, K. M., Havard, S., Shah, V., Gardner, G., and Schweitzer, V. G. (2014), 'Delineating an epigenetic continuum in head and neck cancer', *Cancer Lett*, 342 (2), 178-84.

Wreesmann, V. B., Wang, D., Goberdhan, A., Prasad, M., Ngai, I., Schnaser, E. A., Sacks, P. G., and Singh, B. (2004), 'Genetic abnormalities associated with nodal metastasis in head and neck cancer', *Head Neck*, 26 (1), 10-5.

Yamatoji, M., Kasamatsu, A., Yamano, Y., Sakuma, K., Ogoshi, K., Iyoda, M., Shinozuka, K., Ogawara, K., Takiguchi, Y., Shiiba, M., Tanzawa, H., and Uzawa, K. (2010), 'State of homeobox A10 expression as a putative prognostic marker for oral squamous cell carcinoma', *Oncol Rep*, 23 (1), 61-7.

Yang, L., Han, Y., Suarez Saiz, F., and Minden, M. D. (2007), 'A tumor suppressor and oncogene: the WT1 story', *Leukemia*, 21 (5), 868-76.

Ye, H., Pungpravat, N., Huang, B. L., Muzio, L. L., Mariggio, M. A., Chen, Z., Wong, D. T., and Zhou, X. (2007), 'Genomic assessments of the frequent loss of heterozygosity region on 8p21.3-p22 in head and neck squamous cell carcinoma', *Cancer Genet Cytogenet*, 176 (2), 100-6.

Yip, S., Miao, J., Cahill, D. P., Iafrate, A. J., Aldape, K., Nutt, C. L., and Louis, D. N. (2009), 'MSH6 mutations arise in glioblastomas during temozolomide therapy and mediate temozolomide resistance', *Clin Cancer Res*, 15 (14), 4622-9.

Yoshida, H., Broaddus, R., Cheng, W., Xie, S., and Naora, H. (2006), 'Deregulation of the HOXA10 homeobox gene in endometrial carcinoma: role in epithelial-mesenchymal transition', *Cancer Res*, 66 (2), 889-97.

You, J. S. and Jones, P. A. (2012), 'Cancer genetics and epigenetics: two sides of the same coin?', *Cancer Cell*, 22 (1), 9-20.

Yu, Ziwei, Weinberger, Paul M., Haffty, Bruce G., Sasaki, Clarence, Zerillo, Cynthia, Joe, John, Kowalski, Diane, Dziura, James, Camp, Robert L., Rimm, David L., and Psyrri, Amanda (2005), 'Cyclin D1 Is a Valuable Prognostic Marker in Oropharyngeal Squamous Cell Carcinoma', *Clinical Cancer Research*, 11 (3), 1160-66.

Zabora, J., BrintzenhofeSzoc, K., Curbow, B., Hooker, C., and Piantadosi, S. (2001), 'The prevalence of psychological distress by cancer site', *Psychooncology*, 10 (1), 19-28.

Zheng, Rena and Blobel, Gerd A. (2010), 'GATA Transcription Factors and Cancer', *Genes & Cancer*, 1 (12), 1178-88.

Zhou, X., Temam, S., Oh, M., Pungpravat, N., Huang, B. L., Mao, L., and Wong, D. T. (2006), 'Global expression-based classification of lymph node metastasis and extracapsular spread of oral tongue squamous cell carcinoma', *Neoplasia*, 8 (11), 925-32.

Zitzelsberger, H., Bauer, V., Thomas, G., and Unger, K. (2010), 'Molecular rearrangements in papillary thyroid carcinomas', *Clin Chim Acta*, 411 (5-6), 301-8.

Zou, C. P., Youssef, E. M., Zou, C. C., Carey, T. E., and Lotan, R. (2001), 'Differential effects of chromosome 3p deletion on the expression of the putative tumor suppressor RAR beta and on retinoid resistance in human squamous carcinoma cells', *Oncogene*, 20 (47), 6820-7.

Zuo, C., Ai, L., Ratliff, P., Suen, J. Y., Hanna, E., Brent, T. P., and Fan, C. Y. (2004), 'O6-methylguanine-DNA methyltransferase gene: epigenetic silencing and prognostic value in head and neck squamous cell carcinoma', *Cancer Epidemiol Biomarkers Prev*, 13 (6), 967-75.