University Library

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Yap, T;Celentano, A;Seers, C;McCullough, MJ;Farah, CS

Title:
Molecular diagnostics in oral cancer and oral potentially malignant disorders—A clinician's guide

Date:
2020-01-01

Citation:
Yap, T., Celentano, A., Seers, C., McCullough, M. J. & Farah, C. S. (2020). Molecular diagnostics in oral cancer and oral potentially malignant disorders—A clinician's guide. Journal of Oral Pathology and Medicine, 49 (1), pp.1-8. https://doi.org/10.1111/jop.12920.

Persistent Link:
https://hdl.handle.net/11343/286246

Check for updates

DR TAMI YAP (Orcid ID : 0000-0001-8412-1114)

PROFESSOR CAMILE S FARAH (Orcid ID : 0000-0002-1642-6204)

Article type : Special Issue Article

INVITED REVIEW

Molecular Diagnostics in Oral Cancer and Oral Potentially Malignant Disorders – A Clinician's Guide

Tami Yap[1], Antonio Celentano[1], Christine Seers[1], Michael J McCullough[1] , Camile S Farah[1,2]

[1] Melbourne Dental School, University of Melbourne, Victoria, Australia

[2] Australian Centre for Oral Oncology Research & Education, Western Australia, Australia

Running title: Molecular diagnostics - a clinician's guide

Word count: 3721 (excluding abstract, references, tables)

References: 48

Keywords: oral cancer, potentially malignant oral mucosal disorders, biomarkers, molecular diagnostics, prognosis

Author Manuscript

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/JOP.12920

This article is protected by copyright. All rights reserved

Corresponding Author:

Dr Tami Yap
Melbourne Dental School – University of Melbourne
Level 4, 720 Swanston Street
Carlton VIC 3053
tspyap@unimelb.edu.au

ORCID Numbers:
Tami Yap: ORCID: 0000-0001-8412-1114
Antonio Celentano: ORCID: 0000-0003-4293-251
Christine Seers: ORCID: 0000-0002-7777-443X
Michael McCullough: ORCID: 0000-0001-7829-4684
Camile S Farah: ORCID: 0000-0002-1642-6204

Abstract

Current risk stratification of individuals for the development of oral squamous cell carcinoma (OSCC), including those with oral potentially malignant disorders (OPMD) remains based upon clinical detection of visibly abnormal mucosa and tissue biopsy with histological assessment for the presence of OSCC or oral epithelial dysplasia (OED). In OPMD, the presence of OED remains the only prognostic tool used in standard care for the development of future OSCC, despite its ample limitations. There is assured potential that the analysis of the genome, transcripts and proteome can provide insight into what is occurring at a cellular level preceding the appearance of clinically observable change. The landscape of the role of the genome and its transcriptome on the development of OSCC and relationships with OPMDs is immense with exploration occurring on several fronts. For clinicians involved in the diagnosis and care of patients with OSCC and OPMD, understanding of commonly used molecular diagnostic techniques is imperative to gain useful insight from the expanding literature investigating the development of OSCC and the relationship with the clinical presentations which encompass

Author Manuscript

This article is protected by copyright. All rights reserved

OPMDs. Here we present an introduction to molecular diagnostic methods used in the study of OSCC and OPMD.

Author Manuscript

Conflicts of interest

The authors have no conflicts of interest to declare.

Introduction

Stratifying individuals at risk of oral cancer, including those with oral potentially malignant disorders (OPMD), remains challenging. Early detection saves lives and significant morbidity. Presently, risk stratification remains based upon good clinical judgement for detection of visibly abnormal mucosa, undertaking a biopsy of what is deemed the most appropriate tissue site, with subsequent interpretation of the histological sample for the presence of oral squamous cell carcinoma (OSCC) or oral epithelial dysplasia (OED).

The development of histopathology revolutionized the diagnosis of cancer and is now part of global standard of care. In the modern era, molecular biology facilitates understanding of pathological processes at the level beyond the human eye including the composition, structure and interaction of molecules, such as DNA, RNA and proteins. Molecular diagnostics is the employment of techniques used in molecular biology for the purpose of assessing disease risk, presence and therapeutic efficacy. Understanding of the commonly used terminology and molecular diagnostic techniques is required to gain insight from the expanding literature investigating the development of OSCC and the relationship with the clinical presentations which encompass OPMDs.

This article is protected by copyright. All rights reserved

Clinical change begins at the genome level

Understanding and interpreting molecular findings with relevance to oral cancer begins with appreciation of the immense complexity of cellular machinery and quality control mechanisms which allow constantly renewing epithelial cells to maintain their integrity. OSCC, as with all other cancer types, occurs when the rate of cellular division becomes uncontrolled, leading to spread and continual growth of these cells into surrounding tissues, near and far. Clinically observable change in oral mucosa of OPMDs and OSCC is a result of alteration in replication and maturation of oral keratinocytes. This alteration in replication and differentiation is a result of changes within cells, directed by processes beginning within the genome that ultimately lead to changes in expression of translated proteins.

DNA contains the genetic material of the cell, housed within the nucleus, providing an archive of "blueprints" that are transcribed into RNA. Some of this RNA is translated into protein, whereas other types of RNA are involved in the complex symphony of protein production and cellular regulation (**Figure 1A**). DNA is a double stranded ladder that consists of four types of paired deoxynucleoside triphosphates or dNTPs: adenine with thymine and guanine with cytosine. The transcribed DNA code includes sections referred to as intervening sequences, or introns, which are spliced out with the mature messenger RNA (mRNA) consisting only of exons, ready to code proteins **(Figure 1B)** [1].

Gene alterations

Gene alterations which may be associated with any disease, not just cancer, can be divided into four categories. A *single nucleotide variant* (SNV) or *point mutation* is when a single nucleotide is substituted. This may result in a different protein amino acid sequence (missense mutation) or a premature truncation of the encoded protein (nonsense mutation). A single nucleotide polymorphism (SNP) is a SNV which is shared amongst a defined population. *Insertions* or *deletions* of one or a few consecutive nucleotides, can be "in-frame" resulting in the addition or subtraction of amino acids in a protein, or "frameshift" which can affect the intron regulatory signal or protein truncation. Changes in *exon copy number* by deletions or duplications can impact the functional domains of the encoded protein. Finally, *structural variants (SVs)* are large

Author Manuscript

This article is protected by copyright. All rights reserved

structural anomalies of genetic material between locations on chromosomes or between multiple chromosomes [2].

The DNA content (DNA ploidy) of a cell gives a rough measurement of chromosomal instability and DNA aberration. DNA aneuploidy refers to abnormal quantities of DNA in a nucleus. Chromosomal instability often leads to imbalanced DNA content and the generation of near-diploid or aneuploid clones [3]. Aneuploidy is observed in 14–45% of OPMDs [4] and has been found in 33% of progressive OED compared to 11.3% of dysplastic lesions which did not progress to malignancy [5]. Loss of heterozygosity (LOH) is defined as a loss of genomic material (from a few thousand nucleotides to a whole chromosome) in a pair of chromosomes.

Chemical changes to DNA or changes to DNA accessibility are known as epigenetic modifications. *Epigenetic modification*s can result in the heritable silencing of genes without a change in the sequence [6]. Examples of processes which can initiate and sustain epigenetic change include DNA methylation, RNA-associated silencing and histone modification [7]. There can also be alterations in the transcription of DNA to RNA and the translation of RNA into protein. It is also now known that the majority of RNA does not code for protein but is rather involved in the assembly and control of protein translation.

Other terminology is used in the discussion of gene studies. A *locus* refers to a specific location of a gene on a chromosome (e.g. 9q31.3). There are 23 different pairs of chromosomes. Chromosomes have a long and short arm: “p” for *petit* refers to the short arm and “q” (*queue* or tail) refers to the long arm. Subsequent numbers refer to the position on the arm: region, band, sub-band, sub-sub-bands. The bands are numbered from 1 starting closest to the centromere. A variant of a given locus is called an *allele*. *Wild type* refers to the typical phenotype of an allele which occurs in nature, as opposed to a non-standard *mutant* allele.

The cell cycle

Oral epithelium, as a self-renewing tissue, is constantly undergoing active proliferation. Pro- and anti- cell cycle progression proteins and their related genes are often of interest in the study of oral cancer. Cells undergo division through a cycle with four phases: The S phase, during which DNA is replicated, an M phase, during which the cell divides into two, and two gap phases, G_1

Author Manuscript

This article is protected by copyright. All rights reserved

and G_2, following M phase and S phase, respectively[8]. Cells can also be withdrawn from the cell cycle, reversibly into a G_0 phase, or irreversibly into terminal differentiation or senescence and eventual apoptosis (**Figure 1A).**

Progression through the cell cycle is driven by proteins aptly called cyclins and cyclin-dependent kinase (CDK) family of serine/threonine kinases [9]. Transition through cell cycle phases is closely monitored by sensor mechanisms or checkpoints, which are able to detect abnormal cell cycle events and signal pathways leading to cell cycle arrest [10]. Cell cycle inhibitors can block specific cyclins and CDKs and fall into two classes: the INK4 proteins including p16 (INK4a) and the Cip/Kip family, and can block specific cyclins and CDKs leading to cell cycle arrest [11]. Another well-known part of the cell cycle control machinery is the tumour suppressor protein p53.

The study of human samples

Molecular analysis techniques can be applied to fixed or frozen human tissue specimens, cell culture products or biofluids. The techniques include in-situ analysis or extraction of material from clinical samples for subsequent analysis.

Tissue collected from the patient for analysis for standard histopathological assessment undergoes fixation in formalin and is subsequently embedded in paraffin. The embedded tissue is then thinly sliced and mounted on a glass slide in preparation for light microscopy. The tissue can then be submitted for a variety of commonly used stains, the most common being hematoxylin and eosin [12]. Techniques can also be employed to study specific molecules in these tissues with the advantage that the molecules are in their normal topographical surroundings, also known as "in-situ". Proteins can be assessed using antibodies (immunohistochemistry) or nucleic acids using nucleotide probes (in-situ hybridization).

Tissue as well as biofluids, such as blood or saliva, can also be analyzed for other molecular components as preparation for cell-level studies, or following isolation or 'extraction' of targeted molecules. Analyzing extracted material allows for more sensitive and specific techniques to be employed as will be described later.

Selecting the target and technique for molecular analysis should consider the clinical question as well as the sample type being utilized **(Table 1)**. Techniques can be 'low throughput', which

Author Manuscript

This article is protected by copyright. All rights reserved

means that few targets in few samples can be analyzed at a time, or 'high throughput' which means that high number of targets or a high number of samples can be analyzed at one time. The concept of multiplexing refers to the analysis of several targets of interest in one sample at the same time or in parallel, utilizing techniques to identify the results of several targets. For example various molecular "labels", such as a fluorophores which emit light of different colours, are used to indicate which target is being visualized.

Common techniques used in molecular diagnostics

Presented below are common techniques employed in molecular diagnostics of solid tissue neoplasia including OSCC and OPMD.

Immunohistochemistry

Immunohistochemistry (IHC) analysis is a method for demonstrating the presence and location of proteins in tissue sections. Sections of 3- to 4-μm thin fixed tissue are mounted on glass slides, after which they undergo a process such as enzyme digestion by a protease to "unmask" antigens. Protein targets can be labelled using a 'direct' method, where an antibody directed against a desired target is conjugated to a molecule to facilitate detection. This for example could be a fluorophore or an enzyme that reacts with a provided substrate to deposit a coloured compound. Enzymes commonly used are horseradish peroxidase or calf intestine alkaline phosphatase.

The 'indirect' method uses an unlabeled primary antibody to bind to the target molecule. Labelled secondary antibodies that bind the primary antibody are then applied for detection purposes. This is useful if labelled primary antibodies are not available and can give better detection sensitivity [13]. IHC can also be used with cytology specimens. Fluorescent probes can also be utilized to visualize binding of target antibodies. Multiple coloured probes allow for multiplexing where spatially-patterned visualization of several antigens can be observed [14].

This article is protected by copyright. All rights reserved

In-situ hybridization

In-situ hybridization (ISH) is used to examine DNA or RNA in its topographical tissue location or in cellular preparations [15], and utilizes a similar concept to immunohistochemistry. ISH uses the formation of a hybrid molecule consisting of an endogenous single-stranded cellular RNA or DNA and a complementary single-stranded RNA or DNA probe. ISH can be performed on fixed or frozen tissue and hundreds of different hybridizations can be performed on the same tissue. The tissue is treated with proteases, usually proteinase K, to allow access to the nucleic acid of interest. A probe can be of variable length, but is typically 30-300 nucleic acid bases [16]. There are two basic ways to visualize RNA and DNA target molecules in tissue sections using ISH: fluorescence (FISH) and chromogenic (CISH) detection [17]. Target molecules in-situ can be limited in number which has been a limitation of using ISH. Methods to overcome this challenge are the use of amplification of the target DNA or RNA in the tissue preparation, known as in-situ polymerase chain reaction (in-situ PCR) or amplification of the reporter signal, such as catalyzed reporter deposition or branched DNA technology [18-20].

Gel electrophoresis and blotting

Gel electrophoresis is a technique which utilizes an electric current to push charged molecules, such as proteins, DNA or RNA, through a gel for a set amount of time before comparison to a reference guide. The electrophoresis separates fragments according to physical properties as they move away from the negatively charged cathode in a process called fractionation. Smaller fragments travel farther in the gel whereas larger pieces migrate a shorter distance [21].

Southern blotting was developed by Edwin Southern in 1975 and is used to study DNA [22]. In this method DNA fragments separated using electrophoresis are denatured whilst in the gel to produce single stranded DNA that is then transferred to a membrane support, retaining the pattern of separation. The blot is then incubated with radioactive or enzyme-bound labeled complementary DNA or RNA probes followed by detection of the label on an x-ray film by autoradiography or by chemiluminescence. A similar methodology for RNA is called Northern blotting. Protein detection uses antibodies to locate targets and is called Western blotting [23].

Author Manuscript

This article is protected by copyright. All rights reserved

Blotting is commonly used in combination with other molecular techniques to confirm the presence of targets of interest.

Flow cytometry

Flow cytometry (FC) has several applications including cell counting, analysis of DNA content (DNA ploidy), immuno-phenotyping and fluorescence-based cellular and particle studies. FC is based upon analysis of single cells or particles as they file past an interrogation point where light sourced from a laser is scattered according to the properties of the cell. The forward and side light scatter or fluorescence emission is collected by perpendicular optic-sensors. When utilized in fluorescence studies, the light emitted from the laser is at a wavelength which allows visualization of the fluorophore-tagged target in the cell. FC can analyze more than 10,000 cells or particles a second [24]. FC can now be employed utilizing target antibodies and nucleic acid primers labelled with fluorescent probes to study the presence of DNA or RNA within cells. It can be utilized in biofluids, cell suspensions, as well as fresh or FFPE tissue samples although fixation artefacts may arise [25].

Mass spectrometry

Mass spectrometry (MS) utilizes the unique mass to ionic charge ratio of molecules to identify them. For example, as each protein or peptide has a unique amino acid sequence, the mass to ionic charge ratio produces a 'signature' which can be a reference. Software and databases exist for reference of identified proteins.

The main components common to all mass spectrometers are an ion source, a mass analyzer, an ion collection/detection system, a vacuum pump and computer for data receipt [26]. The prepared extracted sample is ionized or vaporized by an ion source. The ion source can come in different forms, such as thermal, chemical, electron bombardment, electrospray ionization (ESI) or Matrix assisted laser desorption/ionization (MALDI) [27]. Examples of common mass analyzers include ion trap, magnetic sector, quadrupole, time-of-flight (TOF), and Fourier transform-ion cyclotron resonance analyzers [28,29]. The ions within the set range of the spectrometer then hit a detection plate, and the resultant electrical current is measured.

Author Manuscript

This article is protected by copyright. All rights reserved

MS has been used in the study of protein structure since the 1990s [30] and is now used in the detection and identification of protein, DNA, RNA and epigenetic studies [31,32]. The output MS data can be compared to databased protein signatures based on their amino acid sequences *in silico.* Tandem MS utilizes a second ion acceleration following the first before detection and can be utilized to identify unknown proteins within a sample [33]. Proteins can be analyzed intact, known as “top-down” proteomics, or from peptides released from proteins by proteolysis, called the “bottom-up” or “shotgun” approach [34].

Polymerase chain reaction

Polymerase chain reaction (PCR) was introduced in 1983. PCR is a single tube reaction customized to amplify a specific genomic sequence. PCR amplification can produce approximately 100 billion copies of one molecule of DNA in a few hours [35]. A specific target sequence of genomic DNA or complementary DNA (cDNA) made from RNA, forms the “template DNA”. The standard PCR reaction mixture consists of template DNA, a large excess of specific oligonucleotide primer pairs, the four deoxynucleotide triphosphates (dNTPs), reaction buffer, and thermostable DNA polymerase. A DNA-dependent DNA polymerase, usually from *Thermus aquaticus* (Taq) that can withstand heating to 95°C with optimal function around 70°C is used [35]. Sometimes, high-fidelity polymerases with proofreading ability, such as those derived from the archaebacterium *Pyrococcus furiosus*, have error rates that are approximately 10 times lower than Taq and are used for applications that mandate more stringent DNA synthesis [36].

The PCR typically consists of 30-40 cycles involving 3 steps: Firstly, denaturing involves heating to 95°C, to allow the DNA template to denature into single strands. Secondly, *annealing* involves cooling the mixture to a temperature to allow oligonucleotide primer binding, followed by binding of DNA polymerase to the 3’ end of the primers. Lastly, *elongation* raises the temperature to around 70-72°C to allow the polymerase activity catalyzing complementary base pairs to be added.

Traditional PCR detects reaction products at the last phase, thus it is so called “End-point PCR”, can be used for determination of “presence vs absence” of the DNA target, whereas “real time”

Author Manuscript

This article is protected by copyright. All rights reserved

measures the fluorescent reporter molecules or signals at each cycle, thus allowing for calculation of relative quantification from a point of reference (qPCR).

Using real-time PCR, two different types of chemical strategies are used to generate a fluorescent signal which is then measured. One is based on double stranded intercalating dye (e.g. SYBR-Green) and the other on different dye-labelled probe-reporter molecule systems (i.e., exonuclease-based double-labelled dye oligo-deoxynucleotide molecular beacons). The reporter molecule is assessed at each PCR cycle and collected into specialized software for analysis. The earlier the cycle that the report molecule is detected, the more abundant the target is. Relative abundance using qPCR compares the cycle at which a target molecule is detected with the detection of a reference molecule.

Microarrays

Microarrays miniaturize thousands of assays on one small solid surface. Microarrays can be used to analyze DNA, RNA or proteins which have been isolated or extracted from a tissue specimen. There are different types of nucleic acid microarrays including printed double-stranded DNA and oligonucleotide arrays, in situ-synthesized arrays, high-density bead arrays, electronic microarrays, and suspension bead arrays [37]. Protein microarrays include analytical, functional protein and reverse-phase protein microarrays [38]. They share the concept of a chip, bead or solid surface with thousands of variants of a particular polymer of interest [39]. Each variant has a probe or dot containing many copies of the same sequence.

The extracted test sample material is labelled with an identifying fluorophore. A control or reference sample, labelled with a different fluorescent emission fluorophore, is used for comparison. The test and control sample materials are then washed across the solid surface with the opportunity to hybridize to specific spots on the glass slide containing a complementary sequence [40]. The spots in the hybridized microarray are then excited by a laser of appropriate wavelength to visualize the fluorescence. Where the dots bind their target, the fluorescence is detected and imaged. The signal intensity of each probe indicates the relative quantity of the target present [41]. The digital image information is processed by specialized software to convert the image of each spot to a numerical reading [42]. The resultant data is then interpreted

Author Manuscript

This article is protected by copyright. All rights reserved

statistically depending on the research question and data available, and may include clustering to classify data into groups of genes or samples with similar characteristics [40].

Sanger Sequencing

The concept of sequencing a genome lay in the desire to overcome the bias of pre-selecting genes of interest, and to instead, discover the sequence of all genetic material present. The Sanger method [43], or the chain-terminator method, was utilized by the Human Genome Project. Automated Sanger sequencing is now considered the "first-generation" of DNA sequencing technologies. Sanger sequencing reactions are still used routinely for sequencing small amounts of DNA fragments. It remains the gold-standard for clinical cytogenetics due to the >99.99% accuracy reported for most genes sequenced [44].

The original Sanger sequencing method takes place in four separate reactions, each containing a small amount of one of the four modified versions of dNTP, called a dideoxynucleotides (ddNTP) which lack a 3' hydroxyl group and do not allow a subsequent base to be added by the DNA polymerase. The four ddNTPs are called ddATP, ddCTP, ddGTP and ddTTP. The four reactions also contain the four dNTPs (dATP, dCTP, dGTP and dTTP) for building DNA. However, whenever a ddNTP instead of a dNTP is recruited, the replication is truncated at that base. Thus, the result is various DNA fragments of different lengths ending with the same known base [45]. Subsequent gel electrophoresis of these fragments is used to measure the length of the fragments, with one separate lane of gel for each of the four reactions, thus the sequence is slowly recorded. This protocol was later modified to use fluorescently labelled ddNTPs, which permits one reaction containing all four chain terminating bases, rather than four separate reactions [41].

Next Generation Sequencing

Next generation sequencing (Next Gen Sequencing, NGS, high throughput sequencing, deep sequencing, second generation sequencing, massively parallel sequencing) uses sequencing of multiple DNA fragments, performed in parallel. The speed and data acquired in NGS is exponentially greater compared to Sanger sequencing, accompanied by significantly reduced

This article is protected by copyright. All rights reserved

Author Manuscript

time and costs. NGS can be utilized in coding and non-coding RNA studies by creating complementary DNA strands to sample RNA. Like other methods, NGS is restricted by target abundance, dictating the "depth" to which a sample can be sequenced.

There are a number of different NGS platforms, yet each typically has a sample preparation or "library preparation" step in which the DNA or complimentary DNA, that serves as the template, is purified, amplified and fragmented, followed by physical isolation of DNA fragments by attachment to solid surfaces or small beads [44]. Clonal clusters of an original DNA fragment are sequenced in miniaturized chemical reactions, and millions are spatially arranged so that individual reactions are isolated from one another and can be distinctly detected by light and captured digital imaging. For example, in the first commercial NGS platform, Roche 454, the production of light occurred each time a base was incorporated and the amount of light measured was proportional to the number of nucleotides added [46]. The large amount of data generated from NGS presents significant challenges in terms of storage and management, in addition to bioinformatic analysis and interpretation [47].

NGS enables avoidance of the bias of pre-selected gene targets and can be utilized to sequence the whole genome, the coding genome (whole-exome sequencing), or RNA abundance and sequencing (RNA-Seq) but it can also be used to detect a well-established, clinically relevant point mutation in addition to other targeted enquiries [48].

Conclusion

It is hoped that analysis of the genome, transcriptome and proteome will continually build insight into what is occurring at a sub-cellular level preceding the appearance of clinically or histopathologically observable change in OSCC and OPMD. Understanding the molecular processes underpinning the development of OSCC is extremely relevant to clinicians who care for patients in the field of oral mucosal disease. Molecular diagnostic techniques have strengths and limitations. Some have potential to be developed into standard-of-care tests, whereas others afford their most crucial insight through use of a translational study-based environment. The expansion of knowledge afforded by molecular diagnostics continues to expand on several fronts including prevention, early diagnosis and therapeutics. The integration of clinical examination,

Author Manuscript

This article is protected by copyright. All rights reserved

histopathology and molecular diagnostics will advance the ability to risk stratify individuals for the development of OSCC and move towards decreasing the impact of this devastating disease.

Figure Legend

Figure 1A. Oral squamous cell carcinoma is characterized by a loss of cell cycle control leading to an escape from appropriate senescence and apoptosis leading and resultant disordered epithelial growth spreading beyond the basal lamina. The alteration leading to the loss of cell cycle control may take place anywhere from the genome to messenger RNA to the encoded proteins. Major players in this process include microRNAs, effectors at the post-transcription level. **B.** DNA is transcribed to RNA which is then processed. The introns, which may contain transcription regulatory messages, are spliced, leaving a mature messenger RNA consisting of exons only.

Table 1. Comparison of Molecular Diagnostic Techniques utilized in the study of OSCC and OPMD.

References

1. Roy SW, Gilbert W. The evolution of spliceosomal introns: patterns, puzzles and progress. *Nature reviews Genetics.* 2006;7(3):211-221.
2. Vnencak-Jones C, M. Berger, W. Pao. 2015. Types of Molecular Tumor Testing. My Cancer Genome http://www.mycancergenome.org/content/molecular-medicine/types-of-molecular-tumor-testing/ (Updated June 18).
3. Lingen MW, Pinto A, Mendes RA, et al. Genetics/epigenetics of oral premalignancy: current status and future research. *Oral diseases.* 2011;17 Suppl 1:7-22.
4. Pentenero M, Giaretti W, Navone R, et al. DNA aneuploidy and dysplasia in oral potentially malignant disorders: association with cigarette smoking and site. *Oral oncology.* 2009;45(10):887-890.

This article is protected by copyright. All rights reserved

Author Manuscript

5. Torres-Rendon A, Stewart R, Craig GT, Wells M, Speight PM. DNA ploidy analysis by image cytometry helps to identify oral epithelial dysplasias with a high risk of malignant progression. *Oral Oncol.* 2009;45(6):468-473.
6. Gasche JA, Goel A. Epigenetic mechanisms in oral carcinogenesis. *Future oncology (London, England).* 2012;8(11):1407-1425.
7. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature.* 2004;429(6990):457-463.
8. Murray A. *The cell cycle: an introduction.* 1993.
9. Hydbring P, Malumbres M, Sicinski P. Non-canonical functions of cell cycle cyclins and cyclin-dependent kinases. *Nature reviews Molecular cell biology.* 2016;17(5):280-292.
10. Williams GH, Stoeber K. The cell cycle and cancer. *The Journal of pathology.* 2012;226(2):352-364.
11. Wiman KG, Zhivotovsky B. Understanding cell cycle and cell death regulation provides novel weapons against human diseases. *Journal of Internal Medicine.* 2017;281(5):483-495.
12. Titford M. A Short History of Histopathology Technique. *Journal of Histotechnology.* 2006;29(2):99-110.
13. Schacht V, Kern JS. Basics of Immunohistochemistry. *The Journal of investigative dermatology.* 2015;135(3):e30.
14. Dixon AR, Bathany C, Tsuei M, White J, Barald KF, Takayama S. Recent developments in multiplexing techniques for immunohistochemistry. *Expert review of molecular diagnostics.* 2015;15(9):1171-1186.
15. Gall JG. The origin of in situ hybridization - A personal history. *Methods (San Diego, Calif).* 2016;98:4-9.
16. McNicol AM, Farquharson MA. In Situ hybridization and its diagnostic applications in pathology. *The Journal of pathology.* 1997;182(3):250-261.
17. van Ommen GJ, Breuning MH, Raap AK. FISH in genome research and molecular diagnostics. *Current opinion in genetics & development.* 1995;5(3):304-308.
18. Bagasra O. Protocols for the in situ PCR-amplification and detection of mRNA and DNA sequences. *Nat Protoc.* 2007;2(11):2782-2795.
19. Kerstens HM, Poddighe PJ, Hanselaar AG. A novel in situ hybridization signal amplification method based on the deposition of biotinylated tyramine. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society.* 1995;43(4):347-352.

Author Manuscript

This article is protected by copyright. All rights reserved

20. Urdea MS. Branched DNA signal amplification. *Bio/technology (Nature Publishing Company).* 1994;12(9):926-928.

21. Lee PY, Costumbrado J, Hsu C-Y, Kim YH. Agarose Gel Electrophoresis for the Separation of DNA Fragments. *Journal of Visualized Experiments : JoVE.* 2012(62):3923.

22. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of molecular biology.* 1975;98(3):503-517.

23. Burnette WN. "Western Blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Analytical biochemistry.* 1981;112(2):195-203.

24. Picot J, Guerin CL, Le Van Kim C, Boulanger CM. Flow cytometry: retrospective, fundamentals and recent instrumentation. *Cytotechnology.* 2012;64(2):109-130.

25. Corver WE, ter Haar NT. High-resolution multiparameter DNA flow cytometry for the detection and sorting of tumor and stromal subpopulations from paraffin-embedded tissues. *Curr Protoc Cytom.* 2011;Chapter 7:Unit 7 37.

26. Finehout EJ, Lee KH. An introduction to mass spectrometry applications in biological research. *Biochemistry and Molecular Biology Education.* 2004;32(2):93-100.

27. Meher AK, Chen YC. Electrospray Modifications for Advancing Mass Spectrometric Analysis. *Mass spectrometry (Tokyo, Japan).* 2017;6(Spec Iss):S0057.

28. Nagornov KO, Gorshkov MV, Kozhinov AN, Tsybin YO. High-resolution Fourier transform ion cyclotron resonance mass spectrometry with increased throughput for biomolecular analysis. *Analytical chemistry.* 2014;86(18):9020-9028.

29. Cho YT, Su H, Wu WJ, et al. Biomarker Characterization by MALDI-TOF/MS. *Advances in clinical chemistry.* 2015;69:209-254.

30. Chowdhury SK, Katta V, Chait BT. Electrospray ionization mass spectrometric peptide mapping: a rapid, sensitive technique for protein structure analysis. *Biochem Biophys Res Commun.* 1990;167(2):686-692.

31. Giessing AMB, Kirpekar F. Mass spectrometry in the biology of RNA and its modifications. *J Proteomics.* 2012;75(12):3434-3449.

32. McLean C, Gluckman P, Sheppard A. Phenotypic diversity and epigenomic variation – The utility of mass spectrometric analysis of DNA methylation. *J Proteomics.* 2012;75(12):3400-3409.

33. Edwards NJ. Protein Identification from Tandem Mass Spectra by Database Searching. *Methods in molecular biology (Clifton, NJ).* 2017;1558:357-380.

Author Manuscript

This article is protected by copyright. All rights reserved

34. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR, 3rd. Protein analysis by shotgun/bottom-up proteomics. *Chemical reviews.* 2013;113(4):2343-2394.

35. Templeton NS. The polymerase chain reaction. History, methods, and applications. *Diagnostic molecular pathology : the American journal of surgical pathology, part B.* 1992;1(1):58-72.

36. Eckert KA, Kunkel TA. DNA polymerase fidelity and the polymerase chain reaction. *PCR methods and applications.* 1991;1(1):17-24.

37. Miller MB, Tang Y-W. Basic concepts of microarrays and potential applications in clinical microbiology. *Clinical microbiology reviews.* 2009;22(4):611-633.

38. Sutandy FXR, Qian J, Chen C-S, Zhu H. Overview of protein microarrays. *Current protocols in protein science.* 2013;Chapter 27:Unit-27.21.

39. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (New York, NY).* 1995;270(5235):467-470.

40. Babu MM. Introduction to microarray data analysis. *Computational genomics: Theory and application.* 2004;17(6):225-249.

41. Bickel PJ, Brown JB, Huang H, Li Q. *An overview of recent developments in genomics and associated statistical methods.* Vol 3672009.

42. Trevino V, Falciani F, Barrera-Saldaña HA. DNA microarrays: a powerful genomic tool for biomedical and clinical research. *Molecular medicine (Cambridge, Mass).* 2007;13(9-10):527-541.

43. Sanger F, Air GM, Barrell BG, et al. Nucleotide sequence of bacteriophage [phi]X174 DNA. *Nature.* 1977;265(5596):687-695.

44. Rizzo JM, Buck MJ. Key principles and clinical applications of "next-generation" DNA sequencing. *Cancer prevention research (Philadelphia, Pa).* 2012;5(7):887-900.

45. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. *Genomics.* 2016;107(1):1-8.

46. Jessri M, Farah CS. Harnessing massively parallel sequencing in personalized head and neck oncology. *Journal of dental research.* 2014;93(5):437-444.

47. Su Z, Ning B, Fang H, et al. Next-generation sequencing and its applications in molecular diagnostics. *Expert review of molecular diagnostics.* 2011;11(3):333-343.

48. Rizzo G, Black M, Mymryk JS, Barrett JW, Nichols AC. Defining the genomic landscape of head and neck cancers through next-generation sequencing. *Oral Dis.* 2015;21(1):e11-24.

Author Manuscript

This article is protected by copyright. All rights reserved

Table 1. Comparison of Molecular Diagnostic Techniques utilized in the study of OSCC and OPMD.

Technique	Sample Type	Target	Throughput	Input material required	Known Target	Requires Subjective Assessment	Accuracy	Technical Complexity	Cost
Immunohistochemistry	In-situ	Protein	Low	-	Required	Yes	Medium	Low	$
In-Situ Hybridization	In-situ	RNA DNA	Low	-	Required	Yes	Medium	Medium	$$
Electrophoresis & Blotting	Extracted	DNA, RNA, Protein	Low	Low	Required	Yes	Low	Medium	$
PCR	Extracted	DNA, RNA	Low	Low	Required	No	High	Medium	$$
Microarrays	Extracted	DNA, RNA or Protein	High	Medium	Partially Required	No	High	High	$$$

Author Manuscript

This article is protected by copyright. All rights reserved

Sanger Sequencing	Extracted	DNA, RNA	Low	High	No	No	High	High	$$$
Next Generation Sequencing	Extracted	DNA, RNA	High	High	No	No	High	High	$$$

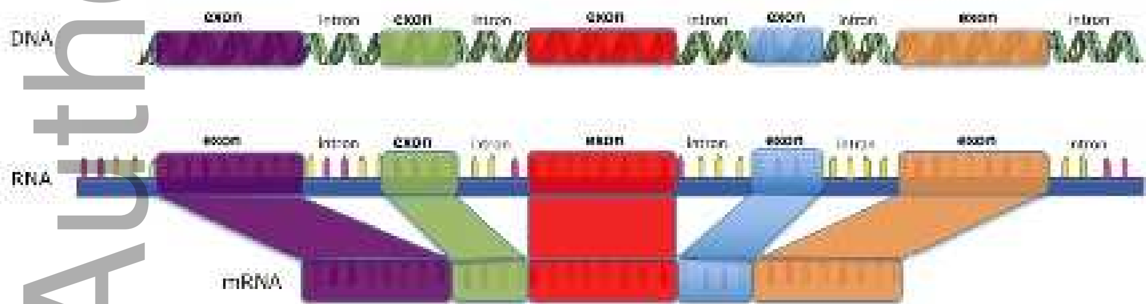
Author Manuscript

This article is protected by copyright. All rights reserved

A

senescence
apoptosis
G0
M
G2
G1
S
Dysregulted
cell division
OSCC
23 pairs
band
mRNA
Ribosome
DNA

B



jop_12920_f1.tiff

Author Manuscript

This article is protected by copyright. All rights reserved