Metabolomics approaches in oral Leukoplakia: A mini review

Shalini Basu, Vertika Rai, Surajit Bose, Rashmi Mukherjee and Chandan Chakraborty

Abstract

Objectives: In this review paper, we studied the Oral Leukoplakia (OLK) disease and its diagnosis and explored the application of Omics approaches in the study of OLK focussing mainly on understanding how omics approaches can lead to identification of novel biomarker molecules or molecular signatures with enough potential to be applied clinical practice.

Methods: “Omics” approaches [FTIR, GC-MS, LCMS, NMR] were applied for differentiating OLK from normal cases. Electronic databases (PubMed, Springer and Google Scholar) were searched for a detailed study of the researches.

Results: Analysis of the various diagnostic methods for OLK detection was studied and the early stage metabolomic profiling with the help of various “omics” approaches were explored.

Conclusions: It may be concluded that despite various advancements in OC therapy, few OC can show symptoms only at an advanced stage like in few OLK cases. Thus there arises a need of early detection of this potential neo plasticity by metabolomic profiling by biomarkers. Modern “Omics” strategies can potentially serve this need by giving major contribution to metabolomics profiling.

Keywords: head and neck cancer; biomarker; oral leukoplakia; metabolomics

1. Introduction

Oral cancer occurrence ranges from 2 to 18% worldwide and 0.1-13.5% in India. The development of oral leukoplakia by malignant transformation ranges from between 1.1%-11.7% [1]. The term “leukoplakia” refers to a “white lesion” that cannot be easily scraped down from the surface on which it occurs [2]. “Oral Leukoplakia” (OLK) is defined as a premalignant, persistent lesion of white color on the oral mucosa. Latest research suggest that aberrant and uncoordinated cellular proliferation reflected in tissue and serum also [3-5], which leads to loss in uniformity of individual cells as well as in their architectural orientation resulting in delayed cell growth and maturation. It can be considered as an important predictive factor for malignant transformation [6]. Although non-dysplastic lesions can also be malignant and all dysplastic lesions may become nonmalignant.

The provisional diagnosis of OLK is established when a persistent white lesion is detected that cannot be categorized as any other disease of the oral mucosa. A definitive diagnosis is then made following the provisional diagnosis for identification and elimination of possible etiological factors [7]. If the lesions are seen to persist, histopathological examinations are performed [2]. Those are Hyperplastic candidiasis versus Candida-associated leukoplakia, Hairy leukoplakia (“Greenspan lesions”), Tobacco-induced white lesions, and Idiopathic leukoplakia [8]. A 10 year study was carried out in India, where large random samples were taken out from different geographical areas with different kinds of tobacco usage and the occurrence rates of leukoplakia per 1000 population per year was found to vary from 1.1 to 2.4% among men and 0.2 to 1.3% among women; prevalence varying from 0.2-4.9%. According to statistics from National Cancer Research Institute, the estimation of a number of cases in Brazil for the year 2016 was found to comprise of 11,140 new cases in men and 4,350 New cases in women [9]. The general occurrence age of OLK is 30-50 years with the peak incidence age being above 50 years. In India, this occurs predominantly in males but in the western world, the ratio of occurrence of this disease has a ratio of almost 1:1 [10].
As proven in a lot of studies tobacco serves as a main aetiological factor for OLK. Although tobacco serves as a main factor for the development of OLK, there are few others that contribute to it. Over alcohol consumption plays a supportive role in causing OLK, although its role as an independent aetiological factor is still questionable.

The C. albicans one of the possible aetiological factor in OLK. Also in recent years, the possible contribution of viral agents in the pathogenesis of oral leukoplakia has also been observed and analyzed, particularly with regard to exophytic, verrucous leukoplakia. Also a significant deficiency of various vitamins like vitamin A, B12 and C are seen in patients with OLK [3].

There are two clinical variants of OLK observed – Homogeneous and non-homogeneous forms. Homogeneous leukoplakia is defined as a persistent white lesion that has a uniform flat, thin appearance that might contain shallow cracks and has a smooth, wrinkled or corrugated surface with a constant texture throughout. It has a homogeneous whitish color along with flat, thin and rather smooth surface. Non – homogenous ones are white and red with the irregularly flat white lesion and nodular or exophytic surface. Oral cancer if detected at an early stage can be treated properly and even cured. However, the majority of the patients come forward for treatment only when the disease is in a well-advanced stage. If we go by history it can be seen that maximum cancers have a pre-cancer development stage, and it is known that intervention at this stage may result in regression of the lesion. Hence the manuscripts reviews the latest diagnosis techniques studied for early detection which can be used in future for diagnosis.

2. Materials and Methods

Extensive comprehensive searches of related literature were made based on mainly the keywords like Omics, OLK, Metabolomics, oral squamous cell carcinoma, omics technology used in OLK, FTIR, GC-MS, LCMS, NMR by using various search engines like google scholar, Springer and PubMed from 1997 to 2018. The papers having relevant information focusing on OLK diagnosis were selected. We have summarized the information. We have gone through a total of 45 articles, and only 32 papers matched our article.

![Pie chart of type of sample used in papers](image)

Fig 1: It depicts the type of sample used in the papers studied

3. Why do we need early biomarkers for OC detection?

In spite of new innovations and developments in the field of surgery, radiation and chemotherapy the survival rate of OC and OLK in last 5 years have increased to 40%. Hence there is an immediate need for development of alternative, sensitive and non-invasive test [11]. It has been proved that an early detection of cancer increases the survival rates of the patients suffering from OLK. There is an urgent need to develop new and more reliable markers to improve early detection of the disease, allow disease stratification and increase survival rates [12]. Toluidine Blue (TB) is a very sensitive staining method used in the diagnosis of OC for decades and can be used in the diagnosis of OLK specifically [13]. It provides information about lesion margins and identification of biopsy sites. The sensitivity of detection by this method varies from 78-100% and it specificity varies from 31-100% [14]. Detection can be done by two other methods- Oral CDx and brush biopsies. Besides sensitivity and specificity they also lack accuracy which are 52-100% and 29-100% respectively.

4. Molecular Approaches for OLK

The various molecular biology methods such as: blotting techniques (Western, Northern and Southern for protein, RNA and DNA respectively); electrophoresis; gene silencing and RNA interference; enzyme-linked immunosorbent assay (ELISA); gene cloning; conventional and RT-qPCR (real-time qualitative polymerase chain reaction (PCR); Comparative genomic hybridization (CGH); karyotyping & fluorescence in situ hybridization (FISH) and chromosomal/cytogenetic analysis; have to a large extent helped in the oral pathological studies. This is by the identification of biomarkers based on cytogenic level alterations such as copy number variation which is pure, indicative of premalignancy of the oral lesions. Along with this various molecular alteration such as microsatellite instability (MSI), Loss of heterozygosity (LOS), abnormal mismatch repair protein (MMR) etc. have been identified to be indicative of oral malignancy [15].

Previously DNA sequencing based analytical procedures used techniques like Polymerase chain reaction (PCR) etc to sequence small fragments of DNA, but after the completion of human genome project new techniques like NGS which is a massive parallel sequencing also known as Next Generation Sequencing was developed to study the genetic basis of diseases to a deeper extent. Along with this development of
array technologies, Genomics, Transcriptomics, Genome-wide association studies (GWAS), Whole exome sequencing, proteomics, epigenomics, lipidomics, and metabolomics have proved to be promising applications of the Omics-based approach in disease analytics which includes the diagnosis of diseases like OLK. This may be an effective and non-invasive early-stage cancer screening tool and may be potentially applied to the detection of OLK.

5. Omics technologies in cancer diagnosis

“OMICS” is short for metabolomics. Metabolism is defined as a collection of processes that generate energy and various cellular level building blocks by the cell utilizing the molecules, substances, and nutrients accumulated from the surroundings. These products formed along with the various cellular intermediates formed during these processes constitute the metabolites. Metabolomics is the detailed analysis and technical study of these substances with low molecular weights called metabolites present in cellular level, tissue level or of whole organisms and are dependent upon can be manipulated by various factors. It enables the analysis of a wide range of exogenous and endogenous metabolites which includes substances including lipids, peptides, amino acids, nucleic acids, vitamins, organic acids, carbohydrates and thiols. Although it is not limited to only these low molecular weight structures. The analyses of these metabolites are important to enable proper treatment on that particular therapeutic metabolite.

Metabolomics is a very advantageous method because of two facts that help this technique to be highly sensitive and an effective way to measure system phenotype. Firstly the final downstream product of transcription and translational activities is the metabolite and hence it is closest to the phenotype and secondly, the primary metabolism dynamics operate in the timescale of seconds. These analyses are typically divided into two categories- targeted and untargeted methods. Targeted methods deals with identification and quantification of selected metabolites, enzymes, substrates, and known pathways. Untargeted methods deal with unknown untargeted metabolites that can lead to the generation of new hypothesis from further tests. The common analytical tools used in metabolomics are nuclear magnetic resistance (NMR), Liquid chromatography-mass spectrometry (LC-MS), Gas chromatography-mass spectrometry (GC-MS), and Fourier transform infrared spectroscopy (FTIR).

Fourier transform infrared spectroscopy (FTIR) is working on the principle that when an infrared radiation is passed through the sample, the vibrational atoms in the sample absorb radiations in different intensities in different atomic structures in the sample. This creates a distinct pattern for a sample based on the pattern of their infrared spectrum absorption. This pattern for a sample is then compared to the normal pattern and also the malignant and cancerous pattern so that it can be analyzed that the leukoplakia from where the sample is isolated is cancerous or not and whether or not it is premalignant or potentially malignant. It is not possible to determine, the premalignant or malignant areas of leukoplakia based only on their clinical appearance. It is the case of innocuous clinical lesions might show malignant transformation based on microscopic evidence. Leukoplakia is characterized by a hyperkeratotic thickening of the prickle cell layer of the epithelium, a certain amount of acanthosis, corium infiltration by the plasma cells and cellular atypia. These characteristics are better studied and determined by FTIR. Thus FTIR has proven to be an important tool used in metabolomics for the analysis of OLK.

LC-MS is liquid chromatography coupled to mass spectroscopy. This coupling is important due to its greater sensitivity and specificity as compared to other chromatographic detectors. LCMS is a modified form of HPLC (high-performance liquid chromatography) and slightly different from it. Liquid chromatography functions to separate the components of the mixture which will further be quantified and identified. The analyte molecules are converted to charged ion state by the action of mass spectrometers, following which ions and fragment ions produced during the ionization process undergo subsequent analysis based on their charge to mass ratio. The major parts of MS include an ion source, mass analyzer, and detector. The mass analyzer resolves the ions before they are measured by the detector. These ions come from the transformation of sample molecules into ions.

Gas chromatography-mass spectroscopy (GCMS) is used to separate and consequently analyze the different components of a mixture so that they can be studied and their malignant transformation can be further determined from that. The special feature of this is that a large number of substances can be identified with sensitivity in a single analysis. For example, up to 300 compounds can be distinguished by GCMS in one single go after deconvolution of overlapping peaks that are formed by the collective analysis of these compounds. This deconvolution prior to data analysis can be avoided by comparing non processed MS data files rapidly. These involve baseline correction, alternating regression, time window determinations, alignment, PLS-DA and identification of retention time windows in the chromatograms that explain the differences between the samples.

Nuclear magnetic resonance (NMR) spectroscopy an analytical method commonly used to analyze the composition of a small molecule called metabolome. It measures the metabolite concentration in the sample being studied so as to reflect any changes in the metabolism of those organisms by determining the biochemical status of organisms hence indicative of any disease or response to chemical treatment. Metabolic signatures in OLK samples has been identified by H and P magnetic resonance spectroscopy for differentiation in benign and normal tissue cells and identification of the formation of oral squamous cell carcinoma (OSCC). For example over expression of glucose transporters especially of Glut-1 is indicative of OSCC.

6. Metabolomics in OLK

A specific technique was developed by Naurecka et.al, 2017 that can be used to develop an understanding of fundamental biochemistry, tissue physiology, and pathology. It is a very sensitive molecular specific technique. This method has the potential to reduce the morbidity of leukoplakia and oral cancers by a detailed examination of preneoplastic and neoplastic tissues. The data that was obtained inferred that these infrared techniques find its application in biomedical and clinical diagnostics.

Imaging using FTIR microscope allows analysis of biochemical compounds and is a technique well suited for individual cells and tissue analysis. It provides information about the biochemical nature of tissue or cell samples and has been applied in many different areas of medical research and testing. IR absorption spectra of abnormal tissues and normal tissues are compared by lipid (2800 - 3000 cm−1), protein (1500 - 1700 cm−1), and nucleic acids (1000 - 1250 cm−1).
regions. OLK is one of the mucous membrane lesions of the mouth is. This change has a “fingerprint region” in the range of 900 - 1800 cm⁻¹ [6].

Rekha Pachaiappan et al. used ATR-FTIR data the advanced statistical model (LDA-ANN) in the diagnosis of oral cancer from normal with better accuracy. The ATR-FTIR Jasco spectrophotometer at 4 cm⁻¹ resolution, was used to acquire infrared spectra by using 30 scans in the 1800-900 cm⁻¹ spectral range. The number of spectra recorded from each blood plasma sample was 5. For the evaluation of the statistical efficacy, the spectral data were routed through the multilayer perception of artificial neural. Among the spectral data, it was found that lipid (1526 cm⁻¹) and amide II (1486 cm⁻¹) affords about 90% in the discrimination between groups using LDA. These results led to a preliminary conclusion that ATR-FTIR can be used to differentiate the normal subject from oral cancer patients using blood plasma [28].

The spectra were obtained using a JASCO FTIR – 6600 spectrometers (Japan), in the range of 4000-400 cm⁻¹, equipped with ATR single reflection diamond and at a resolution of 4 cm⁻¹. The reference spectrum was collected from the diamond crystal as a back group under the same experimental condition for each sample. For correction background absorption spectrum was recorded. Then 1 µl of blood plasma was placed onto the ATR-FTIR diamond crystal and spectra were acquired at the in experimental condition for each sample. For correction background absorption spectrum was recorded. Then 1 µl of blood plasma was placed onto the ATR-FTIR diamond crystal and spectra were acquired at the in experimental condition for each sample. Absolute ethanol was used to clean the blood plasma spot on the diamond crystal and the spectra were repeated multiple times, average spectra were used for the statistical analysis [28].

In a book named “Interface Oral Health Science 2011”, K. Sazaki et al. published that complex data containing thousands of signals generated from the metabolomics experiments of NMR- and MS can be used to represent the metabolite profile for a characteristic biological state. As NMR is a computational tool, it is easy to explain structural composition, spectral data, and pathway diagrams. It can provide the best comprehensive data because it remains connected to the databases. A database can be considered good if it can be easily accessed online through a web interface, uses standardized data formatting to describe, user-friendly interface to retrieval and visualization, and convenient exchange to another database. The 1H NMR spectra obtained from plasma of OSCC patients, OLK patients, and healthy patients, showed great complexity and significant information of the bio fluid. A good model to detect the NMR data for differentiating the OSCC patients from the OLK patients and the controls by using a test set is the result of the PLS-DA analysis.

In a study by Zing Ling Zhou et al., the preliminary results show that the 1H NMR-based metabolomics analysis on the human plasma can distinguish the OSCC patients from the OLK patients and the healthy controls, but, the plasma varies greatly among different individuals, and is very complicated in the physicochemical terms of lipoprotein, protein, small organic molecule, and so on. A great amount of useful metabolic information and organic reactions of a tumor is present in the H-NMR spectrum. This metabolomics study can be used in the study of the molecular properties of the abnormal lipoproteins from the plasma and is considered to be a great step in the blood screening for an early detection of OSCC. This kind of the technique is expected to provide a useful tool in the diagnosis of oral cancer in the future.

Another researcher Shalini Gupta et al., 2014 through some NMR experiments revealed various important features for the results of the NMR derived metabolomics approaches of human serum samples. First, it was observed that NMR-derived fingerprint of various serum metabolic profile can precisely differentiate OLK, OSCC and HC cohorts. Second, the OPLS-DA approach can help to divulge buried and potential biomarkers from the complex NMR spectra. Third, critical appraisal (manual integration of metabolites approach) and external validation (double-blind study) offer the precision of the results. Fourth, ROC analysis of important biomarkers (2 step procedure) revealed the clinical significance of the identified metabolites. Fifth, it concludes that the definite permutations of eight important metabolites are adequate to precisely discriminate OLK, OSCC and HC cohorts.

Therefore this technique accurately differentiate oral cancer (OLK + OSCC) and HC subjects but also the OLK and OSCC stages of oral cancer. They are east invasive and at par with routine clinical examination. Discovery of the correlation between the metabolic fingerprint of serum and histopathological events of oral cancer in the clinic not only opened new avenues in the probing and determining of oral cancer but also complement present explanatory modalities with the goal of advancing the treatment mode and constructing follow-up protocols.

In a study by Guo X. Xie et al., urine samples of patients with oral squamous cell carcinoma (OSCC, n = 37), oral leukoplakia (OLK, n = 32) and healthy subjects (n = 34) were analyzed by gas chromatography-mass spectrometry (GC–MS). These urinary metabolite profiles of OSCC, OLK and healthy control samples can be differentiated using multivariate statistical analysis and a panel of differentially expressed metabolites was obtained. Metabolites, valine and 6-hydroxynicotinic acid, in combination yielded an accuracy of 98.9%, sensitivity of 94.4%, specificity of 91.4%, and positive predictive value of 91.9% in distinguishing OSCC from the controls. The combination of three differential metabolites, 6-hydroxynicotinic acid, cysteine, and tyrosine, was able to discriminate between OSCC and OLK with an accuracy of 92.7%, sensitivity of 85.0%, specificity of 89.7%, and positive predictive value of 91.9% [21].

The data obtained from the GC–MS analysis were converted to NetCDF format via the data analysis interface of the PE Instrument (PerkinElmer Inc., USA). Each file was extracted subsequently by using custom scripts in the MATLAB 7.0 (The Math Works, Inc., USA) for data pre-treatment procedures such as baseline correction, de-noising, smoothing and alignment, time-window splitting, and peak deconvolution (based on multivariate curve resolution algorithm) (Ni et al. 2007). The OPLS-DA models of this GC–MS metabolomics analysis study demonstrated difference among OSCC and OLK patients and healthy controls, highlighting the diagnostic potential of this non-invasive analytical approach. By comparing the mass fragments with NIST 05 standard mass spectral databases and available reference compounds, different metabolites were obtained in OSCC and OLK samples.

In a study by Jiamin Zheng et al., for human salivary metabolome analysis, an isotope labeling LC-MS method has been used and has been proven to be very essential in OLK analysis. For this analysis, acetone protein precipitation is used to process 5μl of the sample, followed by dansyl chloride labeling, and then UV measurement of the total concentration of the labeled metabolites. A relative
quantification can only be performed with the help of a dilution curve of a labeled saliva sample, like a labeled pooled sample because the absolute concentration of the total metabolites in a saliva sample could not be determined due to the lack of a proper standard for calibration. This method of relative quantification developed a way to normalize the individual sample concentration by taking varying volumes of samples for labeling and mixing so that it is ensured that the same amount of sample from each individual took part in metabolome comparison. Besides this, sample injection amount for LC–FTICR-MS analysis can be optimized by UV measurement values to maximize the number of metabolites detected. In an LC-MS approach, the concentrations of individual metabolites present in 12C-labeled individual samples were compared to those in a 13C-labeled pooled sample. It is a differential isotope labeling LC-MS approach with very good reproducibility of both sample processing and LC-MS measurement are obtained with CVs of less than 7% in terms of total concentration of metabolites and the number of peak pairs detected. In a mixture of 12C-labeled individual sample and 13C-labeled control, the number of peak pairs detected ranged from 1052 to 1067, with an average of 1058. <100 metabolites detectable by the reported LC–MS3, 22–25 or CE–MS methods were much lesser than this number. With the help of this method, the changes in metabolome profile due to the effect of saliva sample storage was investigated, and it was found that no significant alteration to the metabolome profile was caused at room-temperature sample storage, compared to the use of an ~80 °C freezer for sample storage. Finally, by this method, the metabolome comparison of two different groups of individuals: normal healthy older adults versus older adults with MCI disease were compared. With the help of OPLS-DA, several discriminant metabolites that contributed most to the separation were discovered by clear observation of the two groups. Taurine was one of the metabolites discovered with lower concentrations in individuals with MCI, compared to the normal old adults [21]. This is a technique of metabolic profiling of serum samples of OLK patients in which the various metabolites in the sample were separated using LCMS and they were histopathologically separated in different groups using MPP Software. After the separation into groups, principal component analysis (PCA) was done for the groups to get a separate scoring for each of the groups following which a statistical analysis through ANOVA was done and compared to the normal value of the metabolites hence detecting which of the metabolites have been upregulated or down regulated in OLK thus identifying the actual biomarkers of this disease [29].

7. Conclusion

The “omics” or “metabolomics” technologies are amongst the few most effective hypothesis-generating tools available today that have provided valuable new information in fields of both basic and clinical research of OLK by doing a complete study of the serum samples in patients showing persistent white lesions termed as leukoplakia. Multiple biomarkers have been studied being generated from the comprehensive analyses of genetic information and their expression in the form of DNA, RNA, proteins, and metabolites that can easily differentiate clinically relevant differences within histologically identical tumors [30]. To completely understand the molecular deregulation of OLK in vitro and in vivo, metabolic and proteomic finger-prints, as well as the bioinformatic analysis and modeling of all the omics data, are needed. Exciting opportunities in biomarker discovery and cancer diagnostics have been offered due to the testing and analysis of the biological samples by these technologies. Nevertheless, important questions still remain unanswered like tailoring patient therapies, integration of biochemical, genetic, clinical and various Omics data for obtaining the holistic molecular view of pathogenic processes and developing sophisticated technologies for translating Omics-based diagnostics into clinical reality [31].

Metabolic profiling produces vast dimensional data. There are various reasons for the need to reduce the number of features to a sufficient minimum. The most important and obvious one being computational complexity. Another main reason is that classification model, developing too much features might lead to “curse of dimensionality”. Furthermore, the enormous amount of raw data contains a large degree of unnecessary information, which would lead to poor performance of model and should be removed. Therefore, further extraction and refinement from raw data is necessary [28].

Decoding of Omics data has put forth a more comprehensive molecular profile of OLK and has opened a new era of research with cure or substantial development in the field of prognosis and prediction of therapeutic response [20]. The application of Omics technologies in clinics for treating oral cancer which is the post stage of OLK will be based on the appropriate design and thorough validation studies to ensure that these metabolomics approaches with Omics dimensional analysis are truly worthy of implementation in clinical field.

Additional Information

Competing financial interests: The authors declare that they have no conflict of interests

8. References


