

Salivary protein markers: a noninvasive protein profile-based method for the early diagnosis of oral premalignancy and malignancy

Ajeetkumar Patil
Khoobram S. Choudhari
Vayakkara Kolaprath Unnikrishnan
Nandita Shenoy
Ravikiran Ongole
Keerthilatha M. Pai
Vasudevan Bhaskaran Kartha
Santhosh Chidangil

Salivary protein markers: a noninvasive protein profile-based method for the early diagnosis of oral premalignancy and malignancy

Ajeetkumar Patil,^a Khoobram S. Choudhari,^a Vayakkara Kolaprath Unnikrishnan,^a Nandita Shenoy,^b Ravikiran Ongole,^b Keerthilatha M. Pai,^c Vasudevan Bhaskaran Kartha,^d and Santhosh Chidangil^a

^aManipal University, Centre for Atomic and Molecular Physics, Manipal 576104, India

^bManipal College of Dental Sciences, Department of Oral Medicine and Radiology, Attavar, Mangalore 575001, India

^cManipal College of Dental Sciences, Department of Oral Medicine and Radiology, Manipal 576104, India

^dUdayagiri, Deonar, Sion Mumbai 400088

Abstract. An ultra-sensitive hyphenated technique, high-performance liquid chromatography-laser-induced fluorescence detection protein profiling of saliva, is evaluated for early detection and diagnosis of oral premalignancy and malignancy. Calibration sets of protein profiles of unstimulated whole saliva are collected from clinically/pathologically normal, premalignant, and malignant subjects and used as standards. Three parameters—scores of factors, sum of squared residuals, and Mahalanobis distance—derived from principal component analysis of protein profiles of the standard calibration sets, and blind samples are used for “match/no-match” diagnosis of the blind samples. Analyses of the results show that the method is capable of differentiating normal, premalignant, and malignant conditions with the sensitivity and specificity of 79% and 78%, respectively. The technique provides a fast, highly objective (free from personal judgment and statistically defined), and noninvasive diagnostic method for screening and early detection of oral cancer. © 2013 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.18.10.101317]

Keywords: salivary proteins; oral cancer; saliva; protein profiling; high-performance liquid chromatography-laser-induced fluorescence detection; principal component analysis.

Paper 130113SSRR received Mar. 1, 2013; revised manuscript received Jul. 3, 2013; accepted for publication Jul. 3, 2013; published online Aug. 19, 2013.

1 Introduction

Oral cancer is the sixth most common cancer among men in the world.¹ Nowadays, in the developing countries, a higher rate of oral cancer incidence is observed in younger people, presumably due to the increased use of tobacco products. Though the five-year mortality rate of 42% and high morbidity rate are lower when compared with cervical, breast, and colorectal cancers,^{2–8} oral cancer is among leading cause of cancer-related deaths in the male population.⁹ In India, 320,000 new cancer cases are diagnosed annually in males, and 19% of this is oral cancer; in females out of 350,000 new cancer cases, 7% is oral cancer.¹⁰ Annually 640,000 oral cancer cases are diagnosed worldwide.¹¹ In India, oral cancer stands in third position after breast and cervical cancers.¹⁰ One third of the world's oral cancer cases are reported in India.¹² Incidence of oral cancer is greater in males compared to females with a ratio of 2:1.^{11,13–15}

A morphologically altered tissue, where cancer is more prone to occur than the normal, is considered as a precancerous lesion.¹⁶ Leukoplakia, erythroplakia, oral submucous fibrosis (OSMF), and lichen planus are the main oral precancerous lesions. The malignant transformation of leukoplakia, erythroplakia, and OSMF is well described at present.^{17,18} Leukoplakia is reported to have a malignant transformation rate from 0.13% to 19.8%.^{19–22} Erythroplakia is said to have more than 50% malignant transformation rate.^{23–27} Only the erosive form of lichen planus

has a potential to turn into malignant form with a transformation rate of below 1%.²⁵ The risk of malignant transformation of dysplasia ranges from 10% to 14%.²⁸ Malignant transformation rate of 4.5% to 7.6% was reported for OSMF.²⁹ Areca nut chewing is reported to be the causative agent for OSMF.²⁹

In spite of the easy accessibility of the oral cavity, in the majority of cases the disease is detected only in the advanced stages contributing to higher morbidity and mortality rates. Early detection of the cancer is the most effective means of controlling the death rate and the cost of treatment.³⁰ The failure in early detection by dentists and physicians is attributed to the fact that “physicians do not routinely inspect their patients for suspicious oral lesions, and dentists are also remiss in the early diagnosis and referral for oral cancer.”³¹ We have shown in our earlier work that pattern analysis of protein profiles of saliva, serum,^{32–35} and other clinical samples^{36–39} may provide a sensitive method for early detection of oral (and other) premalignant and malignant conditions. The suggestion that “multiparametric biomarker analysis with artificial neural networks and pattern recognition will likely represent one of the promising methodologies for diagnosing and monitoring cancer” has also been made by other workers.⁴⁰ Recently, other groups also have proposed the use of markers in saliva^{41–43} for possible early diagnosis of different forms of malignancies. It is now established⁴¹ that more than 1000 proteins are present in saliva, and the level of many of them changes⁴² in malignancy. In addition, many salivary proteins are degraded under normal conditions of

Address all correspondence to: Santhosh Chidangil, Manipal University, Centre for Atomic and Molecular Physics, Manipal 576104, India. Tel: +91-820-2925071; Fax: 0820-2570062; E-mail: santhosh.cls@manipal.edu

handling, storage, processing, etc., and these reactions themselves can be different under normal and disease conditions due to presence of altered enzymes and reactive proteins. The changes originally arising from disease conditions as well as those taking place after procurement have to be properly identified and quantified for diagnostic applications. Minimally invasive, inexpensive collection, simple handling and transportation, storage, and no risk in sampling make saliva an ideal biofluid for disease diagnosis.^{44–46} The simple collection method reduces the anxiety and discomfort, thereby making it user friendly for patients. Scientists often call saliva the “mirror of the body” and the “window on health status”.⁴⁶ In view of the serious limitations of present oral cancer diagnostic methods,⁴⁷ it is thus very essential to establish the feasibility of the use of saliva for early diagnosis of oral (or any other) malignancy, and to standardize procedures for protein profile analysis of saliva samples for early detection of oral cancer.

In the present work we have used our method of High-performance liquid chromatography-laser-induced fluorescence (HPLC-LIF) protein profile analysis of saliva to systematically investigate the suitability of the technique for early detection of oral premalignancy and malignancy.

2 Methods and Materials

Whole unstimulated saliva was collected from a group of 37 healthy, 40 premalignant, and 23 malignant subjects (sample information given in Table 1). The 40 premalignant subjects consist of 15 oral leukoplakia, 19 OSMF, and 6 oral lichen planus cases. The patients' age groups were: premalignant 23 to 58 years and malignant 25 to 75 years. Ethical clearance for the current study has been obtained from the Kasturba Medical College, Manipal University Ethical Committee. Samples were collected with informed consent. Samples were collected after clinical/pathological diagnosis from the Department of Oral Medicine and Radiology, Manipal College of Dental Sciences, Mangalore. The clinical/pathological results of validation set samples (blind) were conveyed to the protein profile analysis group only after their data processing and diagnosis of test samples. Classification of samples was done by comparing with a control group of 37 healthy subjects from both genders (age range of 20 to 36 years) with no oral or systemic diseases. Only one malignant and two premalignant patients had no history of smoking, tobacco chewing, or alcohol intake.

2.1 Sample Collection, Processing, and Storage

Whole saliva was collected from the subjects after clinical examination. Subject's mouth was cleaned with distilled water prior to sample collection. Volunteers were asked to generate saliva and spit into a wide mouth sterile sample collection bottle

for about 5 to 10 min. Approximately 3 to 5 ml of whole saliva was collected from subjects. Bottles were immediately kept in ice buckets. Samples were transferred immediately to Eppendorf tubes and centrifuged twice at 5000 rpm for 10 min to remove any debris released in spite of the cleaning and also to eliminate exfoliated cells from the sample. Though exfoliated cells are useful in the diagnosis of malignancy,⁴⁸ the major aim of the present studies was to identify “true” changes in salivary proteins in premalignant and malignant states. More importantly, it is well known⁴⁹ that in early stages of malignancy/dysplasia, proliferation is confined to the basal layer. The cells move outward slowly to get sloughed off from the surface mostly only in advanced stages. The protein profiles of exfoliated cells that may appear randomly in any early stage may thus be not very useful for early detection and diagnosis. In the present studies, therefore, only the supernatant after centrifugation was collected and used. The supernatant was distributed into 3 to 5 aliquots, stored at -80°C for short-term storage, and used as soon as the system was ready for the next run.

2.2 Experimental Setup

HPLC-LIF Setup: The HPLC-LIF experimental setup (Fig. 1) is described in detail elsewhere.³⁹ It consists of an HPLC system (Agilent 1200 series) with solvent reservoirs, a degasser (G1322A), a quaternary pump (G1311A), and a manual injector (7725i, Rheodyne, Rohnert Park, CA) with a 20 μl sample

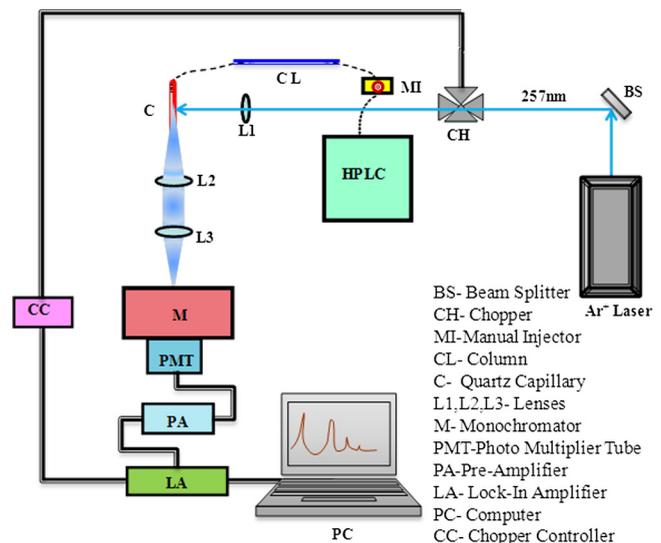


Fig. 1 Schematic diagram of high-performance liquid chromatography (HPLC)-laser-induced fluorescence (LIF) setup.

Table 1 Sample information.

Sample information sample no.	Habits	Sex	Age	Clinical/histopathological diagnosis
1 to 37	33 none, 2 alcohol, 2 alcohol + tobacco chewing	16 Male, 21 female	20 to 42	Healthy
38 to 60	19 tobacco chewing, 3 alcohol, 7 smoking	20 Male, 3 female	32 to 71	All carcinoma
61 to 100	3 none, 27 tobacco chewing, 14 alcohol, 18 smoking	35 Male, 5 female	22 to 69	15 OSMF, 17 leukoplakia, 8 lichen planus

injector loop. A reversed phase narrow bore biphenyl column (Sl. No 219TP52, Grace Vydac, Columbia, Maryland) was used for the separation of proteins. The solvents used were: A-Water + 0.1% v/v trifluoroacetic acid (TFA), B-Acetonitrile + 0.1% v/v TFA, C-Methanol, and D-Isopropanol. A and B were used as mobile phases for protein elution. The flow rate was set at 200 $\mu\text{l}/\text{min}$. The column was connected to a UV-grade quartz capillary (75 μm ID, 360 μm OD, FS-175, Upchurch Scientific, Oak Harbor, Washington) flow cell through PEEK tubings using finger tight fittings (5063-6591, Agilent, Santa Clara, CA). The rigidly fixed capillary flow cell was designed and fabricated in our laboratory. For optimized excitation and collection of fluorescence from the sample, the flow cell was accurately mounted on precision mounts. The sample in the flow cell was excited efficiently by illuminating with Ar^+ laser beam (continuous wave, frequency doubled, Innova 90C FreD, Coherent, Santa Clara, CA). The excitation beam was chopped continuously using optical chopper (651, Signal Recovery, Oak Ridge, TN) and controlled by chopper controller, which provides reference frequency to the lock-in amplifier (7265, Signal Recovery, Oak Ridge, TN). The focusing and collection optics were mounted on precision translation stages. The fluorescence was collected and focused onto the slit of the monochromator (Micro HR, Horiba Jobin Yvon, Edison, NJ) set at 340 nm. The signal was detected by a Photomultiplier tube (R 928P, Hamamatsu Inc., Hamamatsu, Japan) and sent to the preamplifier (5113, Signal Recovery, Oak Ridge, TN). Output of the preamplifier was subsequently fed to the lock-in amplifier (7265, EG & G, United Kingdom). The lock-in amplifier was interfaced with the computer for control and storage of the data.

Saliva samples were diluted 20 to 50 times with prefiltered HPLC grade water. Hundred microliter freshly diluted samples were injected into the HPLC system with the 20 μl loop. Proteins were eluted under gradient run with A 70% to 50%, B 30% to 50% in 0 to 20 min, A 50% to 40%, B 50% to 60% in 25 min followed by B 60% to 100% in 26 to 30 min. After each sample run, the column was regenerated with water for 15 min. Prior to the injection of a sample, a blank run was recorded to ensure that the column was clean.

2.3 Data Processing

Data processing involves preprocessing of the data, principal component analysis (PCA), calculation of suitable parameters, and carrying out a match/no-match test of the parameters of test sample against parameters of standard calibration sets.

The duration of each chromatogram is 30 min with the signal collection every 2 s; hence, each chromatogram has 900 data points. Preprocessing includes interpolation, smoothing, baseline correction, normalization, and alignment along the time axis. All data processing was done using PLSplus/IQ module of the GRAMS software (Thermo Scientific Inc., Rockford, IL).

We have shown in our earlier studies on Raman and fluorescence spectra of tissues⁵⁰⁻⁵⁵ and protein profile studies of serum for ovarian, cervical, and breast malignancy conditions^{37,38} that the technique of match/no-match with statistical parameters of standard calibration sets from PCA is a very good method for discrimination between normal, premalignant, and malignant conditions. The method is described in detail in our earlier work. In brief, the method, like any other analytical method, uses standard calibration sets of each category, against which a test sample is matched using parameters derived from PCA. In the present work, standard calibration sets of protein profiles

are formed using 10 to 15 saliva samples collected from subjects who have been judged to be normal/premalignant/malignant clinically and pathologically. They are subjected to PCA, and three parameters—scores of factors, squared sum of residuals (observed—simulated chromatogram)², and M. distance^{34,36-39,52,53}—are derived for the standard sets. Any test sample is then added to each calibration set, and the parameters of the test sample are compared with those of the calibration set. The sample is diagnosed as belonging to that set with which it matches within the range for that set for the decision-making threshold.

3 Results and Discussions

Figure 2(a) shows typical preprocessed normal, malignant, and premalignant saliva chromatograms. Figure 2(b) contains all the three chromatograms expanded to five times showing several minor peaks. A closer look at these plots reveals that the noise level is such that a signal-to-noise ratio of 2 may correspond to a peak height of about 0.004 V. The peak around 1200 s is from α -amylase. The signal for α -amylase in this run is about 1 V. It was observed that in most of the chromatograms, this peak often goes out of the scale with the signal amplification. Higher values of amplification, PMT voltage, and laser power were used in all our runs to analyze and acquire information from smaller peaks in the chromatograms. Hence, the portion of the chromatogram belonging to α -amylase was excluded for the PCA. It may be mentioned here, however, that the α -amylase levels in saliva are also sensitive to disease conditions. The concentration of α -amylase in saliva is about 19 to 308 U/ml.⁵⁶ This shows

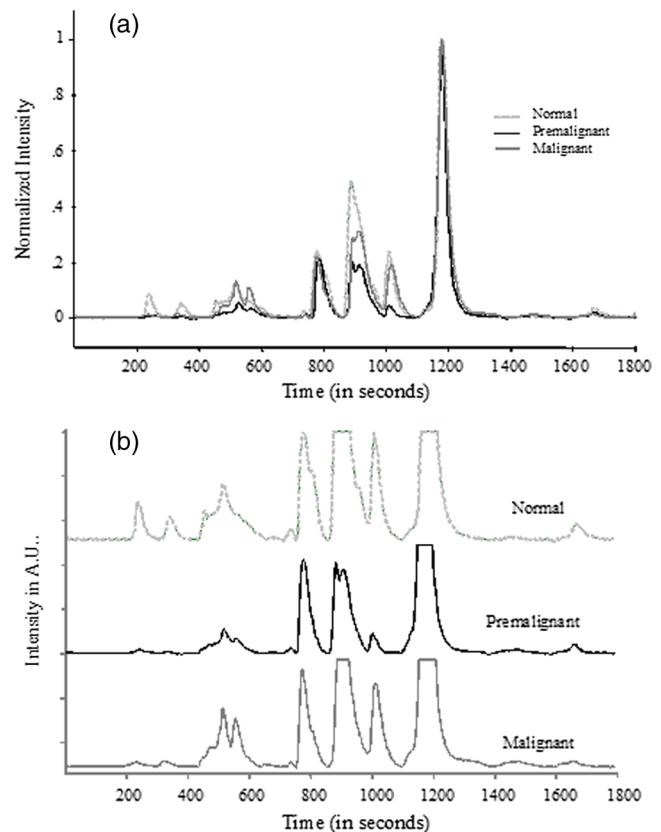


Fig. 2 (a) Typical preprocessed saliva chromatograms of normal, premalignant, and malignant classes (normalized with respect to 1200 s peak). (b) Expanded (five times) chromatograms in (a).

that our system has a detection limit of a few femtomoles of protein under present experimental conditions. This detection limit can be easily reduced by 1 to 2 orders of magnitude by appropriate use of laser power, for example multipassing and more efficient collection of fluorescence.⁵⁷

PCA was done with separate sections of the chromatograms, namely, in the regions of 400 to 650 s, 700 to 1100 s, 1400 to 1500 s, 1550– to 1750 s and also using complete chromatograms excluding α -amylase peak. The choice of the regions were made on the basis of visual inspection by which it was found that these regions may contain the maximum information and give better results. It was observed that the third region above gives good sensitivity and specificity in comparison with the other regions/ the complete chromatogram excluding α -amylase peak. The significant number of factors was decided using the Eigen values, total percent variance, and squared residuals. Squared residual, as mentioned above, is the sum of squares of differences for all data points between observed and simulated chromatograms. The Mahalanobis distance (M. distance) is expressed in units of standard deviation.⁵⁵ Receiver operating characteristic (ROC) curve analysis was done in an earlier study to find out the optimum value for M. distance⁵⁸ for the classification of clinical samples. An M. distance of >2 for a sample indicates the probability of the sample being in the same cluster with which the match was made to be $<5\%$, and a value >3 gives a probability of a value $<0.1\%$. Hence, the samples with M. distance >3 are classified as out of the calibration set.

The PCA and match/no-match test results of region 3 are given in Tables 2–4 and the final results for other regions and complete chromatograms are shown in Table 5. Test runs were made using 10 factors. It was found that the first five factors contributed to about 98% of the total variance. It can be seen from Fig. 3 that first five factors' contribution gives maximum variations, and from sixth factor onward the contribution can be attributed to random variations, which may include variations from day-to-day, noise, etc. The contributions from the higher factors are so small that they are not sensitive to sample changes.

In our method of PCA,^{52–54} the mean of data from all samples is first formed, and the variation of each sample from this mean is calculated. PCA is done with these variations. The variations from the mean for any sample are thus represented by its scores for the different factors. Thus, scores can be used in a first step as a discriminating parameter. Attempts were made to find the classification of samples by plotting the different scores along x - and y -axes and also by plotting scores of factors against the sample number. The plot of scores of factor 1 versus scores of

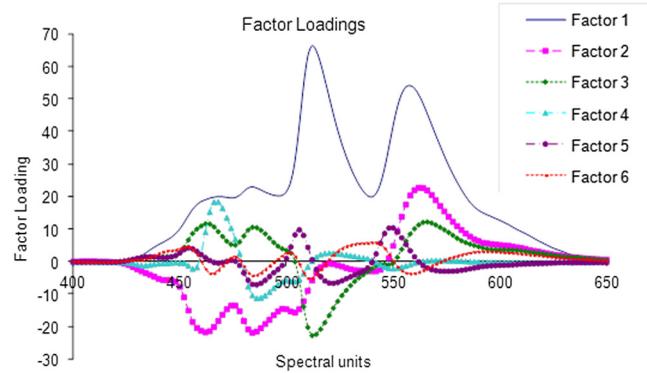


Fig. 3 Factor loadings for different factors.

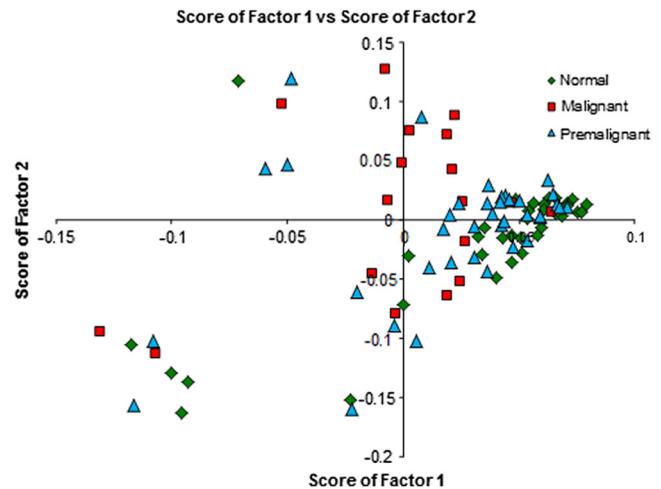


Fig. 4 Plot of scores of factor 1 versus scores of factor 2.

factor 2 is shown in Fig. 4. From the figure, it can be seen that there is discrimination between normal and malignant samples; however, practically no decision can be made regarding premalignant samples simply based on the first two principal component scores.

Much better discrimination can be achieved by using our match/no-match technique in which scores of factors, spectral residual, and M. distance are used as matching parameters. A standard calibration set of each class is formed using chromatograms from clinically/pathologically certified samples of that class. Parameters from these are used for the diagnosis of

Table 2 Prediction report using normal calibration test in the region 1400 to 1500 s.

Class	Sample nos.	Match	M. distance range	S. residual range
Normal	1 to 21, 23, 24, 28, 29, 31 to 35, 37	Yes	0.32 to 2.83	0 to 0.026
	22, 25 to 27, 30, 36	No	3.12 to 107.34	0.026 to 0.69
Premalignant	61, 65, 68, 70, 74, 76, 77, 80 to 100	No	3.02 to 2718.41	0.027 to 17.02
	62 to 64, 66, 67, 69, 71 to 73, 75, 78, 79	Yes	0.56 to 2.22	0.01 to 0.02
Malignant	38 to 51	No	3.11 to 758.72	0.027 to 47.38
	52	Yes	1.04	0.014

Table 3 Prediction report using malignant calibration test in the region 1400 to 1500 s.

Class	Sample nos.	Match	M. distance range	S. residual range
Malignant	38, 40 to 43, 46, 47, 50 to 52, 54, 55	Yes	0.42 to 2.66	0.002 to 0.28
	39, 44, 45, 48, 49, 53, 56 to 60	No	3.46 to 681.53	0.32 to 50.54
Normal	3, 5, 10, 11, 13, 14, 16, 19, 21, 23, 25 to 27, 30, 34 to 37	No	3.14 to 53.10	0.30 to 3.97
	1, 2, 4, 6 to 9, 12, 15, 17, 18, 20, 22, 24, 28, 29, 31 to 33	Yes	0.94 to 2.98	0.15 to 0.33
Premalignant	65, 80 to 83, 85 to 91, 94, 95, 97 to 100	No	3.73 to 1190.39	0.38 to 88.60
	61 to 64, 66 to 79, 84, 92, 93, 96	Yes	0.71 to 2.57	0.058 to 0.29

test samples for match/no-match test. If the parameters of a sample match with the parameters of the calibration set, the sample is considered to be of the same class as that of the test calibration set. If a sample does not belong to the class of the calibration set, the spectral residual and M. distance of the sample under test will be high compared with the ones that belong to the same class. Thus, by matching a given sample to each of the standard calibration sets, a more reliable classification can be achieved with good sensitivity and specificity. Table 2 shows the prediction report of the test made with the normal calibration set. It can be seen that almost all the normal samples lie in this range with the interference of only one malignant and a few premalignant samples representing a clear classification. ROC curve of two independent classes (pre-malignant and malignant) and the same for the combined diagnosis are shown in Fig. 5. Area under the curve calculated for pre-malignant, malignant, and combined (pre-malignant + malignant) curve are 0.69, 0.89, and 0.74, respectively. From the plot, it is clear that using an M. distance value of 2 will give the best possible discrimination between the classes.

Table 2 shows that when tested with the normal calibration set, the majority of the normal samples match all malignant samples (except one), and a large number of pre-malignant samples do not match, giving much better discrimination between normal, pre-malignant, and malignant classes compared with the use of scores alone.

All validation samples were subjected to the match/no-match test also with the malignant and pre-malignant standard sets. The results of tests made with malignant and pre-malignant calibration sets are shown in Tables 3 and 4, respectively. Table 5 lists

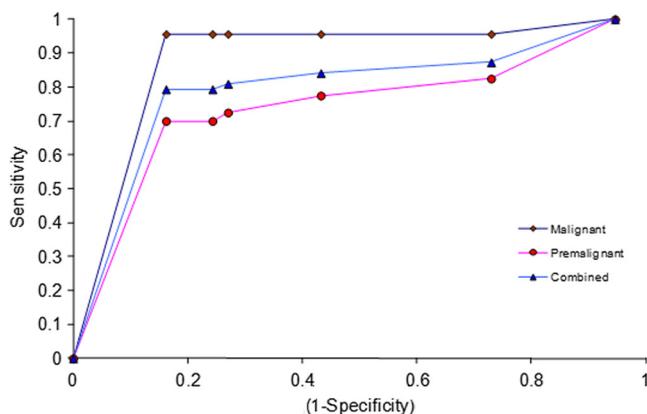


Fig. 5 Results of receiver operating characteristic (ROC) analysis with normal calibration set for the region 1400 to 1500 s.

the results of all the regions and the complete chromatograms for all the three classes. Only about 10 clinically/pathologically certified samples from each class were used to form the calibration sets. In the match/no-match test, every member of any standard calibration set is rotated out of the set and matched against the rest. Since the experimental system and mathematical data processing in the computer do not depend on the sample origin, the method is totally blind as to which sample (standard or validation) is involved. In other words, each member of the calibration set can also be considered as a validation sample. The sensitivity and specificity in the present studies, as seen from Table 5 are 79.37% and 77.38%, respectively. The positive and negative predictive values are 89.29% and 61.76%, respectively.

Other methods like HPLC-mass spectrometry (MS), surface-enhanced laser desorption/ionization-MS, matrix-assisted laser desorption/ionization-MS, and sodium dodecyl sulfate-polyacrylamide gel electrophoresis are available for protein profiling. These techniques have several limitations, discussed elsewhere.^{59,60} The reversed phase HPLC with ultra-sensitive LIF detection has been suggested as the best technique for diagnostic applications of protein profiling due to its superior resolution, high-time efficiency, high-detection capacity, and sensitivity.⁶⁰ The combination of the highly sensitive LIF (which has advantages such as limited stray light effects, low background and accurate focusing capacity of the excitation source on to the micro-bore capillary leading to the excitation of picoliter amounts of samples without any loss of excitation energy etc.) and highly efficient HPLC enables easy detection of the ultra trace biomolecules in complex physiological samples. The other major advantage is the capability of detection of multiple markers in a single run with less time and a small quantity of the sample. Other groups have also studied salivary proteomes.⁶¹⁻⁶³ Yamada et al.⁶¹ have studied salivary proteome, but no oral malignancy samples were investigated. No pre-malignant cases were investigated by Hu et al.⁶² Xie et al.⁶³ have identified proteins in cells separated from saliva from four subjects with oral squamous cell carcinoma of tongue only and have not provided any diagnostic techniques. Almost all studies were done by mass spectroscopic techniques with the inherent disadvantages pointed out by Mohamed et al.⁶⁰ The present study clearly shows that our technique of HPLC-LIF protein profiling-PCA pattern analysis of the profiles can provide a noninvasive, relatively fast, and sensitive method for screening for early detection. The same technique can, obviously, be used for monitoring the effectiveness of therapy and follow up for early detection of any regression after apparent cure. The method is extremely

Table 4 Prediction report using premalignant calibration test in the region 1400 to 1500 s.

Class	Sample nos.	Match	M. distance range	S. residual range
Premalignant	62 to 64, 66 to 69, 71 to 74, 76 to 79, 82, 84, 85, 87, 90, 92, 93, 96(23)	Yes	0.45 to 2.90	0.00 to 0.03
	61, 65, 70, 75, 80, 81, 83, 86, 88, 89, 91, 94, 95, 97 to 100 (17)	No	3.22 to 602.87	0.034 to 5.65
Normal	2, 3, 5, 13, 21, 23, 25 to 27, 30, 32, 36, 37(13)	No	3.29 to 139.58	0.04 to 1.34
	1, 4, 6 to 12, 14 to 20, 22, 24, 28, 29, 31, 33 to 35 (24)	Yes	0.58 to 2.78	0.003 to 0.032
Malignant	38 to 51, 53 to 60 (22)	No	3.87 to 2145.03	0.05 to 20.08
	52 (01)	Yes	0.83	0.01

Table 5 Results of all analyses.

	Region 1 (400 to 650 s)	Region 2 (700 to 1100 s)	Region 3 (1400 to 1500 s)	Region 4 (1550 to 1750 s)	Full region (without α -amylase)
Sensitivity (%)	87.30	85.71	79.37	77.78	87.30
Specificity (%)	25.93	22.22	77.78	51.85	22.22
Positive predictive value (%)	73.34	72.00	89.29	79.03	72.36
Negative predictive value (%)	46.67	40.00	61.76	50.00	46.67
Accuracy (%)	69	66.67	79	70	67.77

sensitive, and sensitivity can be increased further by 1 to 2 orders of magnitude,⁶⁴ if necessary.

4 Conclusion

From the present study, it can be concluded that protein profiling of saliva samples can be used as a screening technique for diagnosis of premalignant and malignant conditions of the oral cavity. As the method is noninvasive (no biopsy) and saliva can be easily obtained from the subject, the subject can be called back any number of times for cross-verification of any anomalies or abnormalities in the results. The technique is based on statistical analysis of instrumental data which can be completely automated and can be operated by a technician. It is thus highly adoptable for developing countries, where the majority of the population resides in areas with very limited hospital/clinical facilities. In the present study, we have used HPLC-LIF combined with PCA as a diagnostic tool for oral cancer screening. Out of the results from match/no-match test, it is observed that good discrimination can be achieved between normal/premalignant/malignant conditions. Using ROC analysis, it can be concluded that the M. distance with a value 2 will give the possible best discrimination of different classes. Though some suspicious premalignant cases may match normal/malignant conditions, any such cases can be called back for periodic examination at regular intervals (say, a few months) with only minimum efforts. Hence, the method can be used as a regular screening technique for the early detection of oral cancer.

Acknowledgments

Authors are thankful to Manipal University for providing the research facility for the work.

References

1. A. Jemal et al., "Cancer statistics," *CA Cancer J. Clin.* **61**(2), 69–90 (2011).
2. A. Jemal et al., "Global patterns of cancer incidence and mortality rates and trends," *Cancer Epidemiol. Biomarkers Prev.* **19**, 1893–1907 (2010).
3. R. W. Hinerman et al., "Post operative irradiation for squamous cell carcinoma of oral cavity: 35 year experience," *Head Neck* **26**(11), 984–994 (2004).
4. S. Kantola, M. Parikka, and K. Jokinen, "Prognostic factors in tongue cancer relative importance of demographic, clinical and histopathological factors," *Br. J. Cancer* **83**(5), 614–619 (2000).
5. J. N. Myers et al., "Squamous cell carcinoma of tongue in adults: increasing incidence and factor that predict treatment outcomes," *Otolaryng. Head Neck Surg.* **122**(1), 44–51 (2000).
6. N. Ramanujam, "Fluorescence spectroscopy of neoplastic and non neoplastic tissues," *Neoplasia* **2**(1–2), 89–117 (2000).
7. K. C. Ribeiro, L. P. Kowalski, and M. R. Latorre, "Impact of comorbidity, symptoms and patients characteristics on prognosis of oral carcinomas," *Otolaryng. Head Neck Surg.* **126**(9), 1079–1085 (2000).
8. A. Sparano et al., "Multivariate predictors of occult neck metastasis in early oral tongue cancer," *Otolaryng. Head Neck Surg.* **131**(4), 472–476 (2004).
9. C. Center, R. Siegel, and A. Jemal, *Global Cancer Facts & Figures*, 2nd ed., American Cancer Society, Atlanta (2011).
10. I. Ali, W. A. Wani, and K. Saleem, "Cancer scenario in India with future perspectives," *Cancer Ther.* **8**, 56–70 (2011).
11. Oral Cancer Foundation, www.oralcancerfoundation.org (10 May 2013).
12. S. K. Naik et al., "Optical screening of oral cancer: technology for emerging markets," in *Conf. Proc. IEEE Eng Med Biol Soc. 2007*, pp. 2807–2810 (2007).
13. B. W. Neville et al., *Oral and Maxillofacial Pathology*, 2nd ed., pp. 337–369, WB Saunders, Philadelphia (2002).
14. L. A. G. Ries et al., *Cancer Statistics Review 1973–1988*, Vol. 91, no. 2789, National Cancer Institute, Bethesda, Maryland (1991).

15. P. A. Swango, "Cancers of the oral cavity and pharynx in the United States," *J. Public Health Dent.* **56**(6), 309–318 (1996).
16. K. P. Schepman and I. Van der Waal, "A proposal for classification and staging system for oral leukoplakia: a preliminary study," *Oral Oncol.* **31**(6), 396–398 (1995).
17. J. E. Bouquot and B. S. Whitaker, "Oral leukoplakia—rationale for diagnosis and prognosis of its clinical subtypes or 'phases'," *Quintessence Int.* **25**(2), 133–140 (1994).
18. P. R. Murti et al., "Malignant transformation rate in oral submucous fibrosis over a 17-year period," *Community Dent. Oral Epidemiol.* **13**(6), 340–341 (1985).
19. P. C. Gupta et al., "Incidence rates of oral cancer and natural history of oral precancerous lesions in a 10 year follow up study of Indian villagers," *Community Dent. Oral Epidemiol.* **8**(6), 287–333 (1980).
20. C. Scheifele and P. A. Reichart, "Is there a natural limit of transformation rate of oral leukoplakia?," *Oral Oncol.* **39**(5), 470–475 (2003).
21. M. N. Shiu et al., "Risk factors for leukoplakia and malignant transformation to oral carcinoma: a leukoplakia cohort in Taiwan," *Br. J. Cancer* **82**(11), 1871–1874 (2000).
22. M. N. Shiu and T. H. Chen, "Intervention efficacy and malignant transformation to oral cancer among patients with leukoplakia," *Oncol. Rep.* **10**(6), 1683–1692 (2003).
23. J. E. Bouquot, L. T. Kurland, and L. H. Wieland, "Carcinoma in situ of the upper aerodigestive tract: incidence time trends and follow up in Rochester, Minnesota 1935–1984," *Cancer* **61**(8), 1691–1698 (1988).
24. J. E. Bouquot and H. Ephros, "Erythroplakia: the dangerous red mucosa," *Pract. Periodontics Aesthet. Dent.* **7**(6), 59–68 (1995).
25. B. W. Neville et al., *Oral and Maxillofacial Pathology*, 3rd ed., Elsevier, St. Louis (2009).
26. W. G. Shafer and C. A. Waldron, "Erythroplakia of the oral cavity," *Cancer* **36**(3), 1021–1028 (1975).
27. P. M. Speight and P. R. Morgan, "The natural history and pathology of oral cancer and pre cancer," *Community Dent. Health* **10**(1), 31–41 (1993).
28. S. Silverman, M. Gorsky, and F. Lozada, "Oral leukoplakia and malignant transformation: a follow up study," *Cancer* **53**(3), 563–568 (1984).
29. P. N. Sinol et al., "A case control study of oral submucous fibrosis with special reference to etiological role of areca nut," *J. Oral Pathol. Med.* **19**(2), 94–98 (1990).
30. R. Etzioni et al., "The case of early detection," *Nat. Rev. Cancer* **3**(4), 243–252 (2003).
31. P. W. Smith, "Fluorescence emission-based detection diagnosis of malignancy," *J. Cell. Biochem.* **87**(39), 54–59 (2002).
32. K. Venkatakrishna et al., "LIF-HPLC for detection of biomarkers in malignancy," in *Proc. Natl. Laser Sym.*, A. Mallik, K. N. Srivastava, and S. Pal, Eds., p. 323, Allied Publishers, New Delhi (2000).
33. K. Venkatakrishna et al., "Protein profiles in oral premalignancies: a laser spectroscopy study," in *Proc. Natl. Laser Sym.*, A. Mallik, K. N. Srivastava, and S. Pal, Eds., p. 345, Allied Publishers, New Delhi (2001).
34. K. Venkatakrishna et al., "HPLC-LIF for early detection of oral cancer," *Curr. Sci.* **84**(4), 551–557 (2003).
35. K. Venkatakrishna et al., "Spit and polish for oral cancer test," *News India (Nature)* (5 June 2003).
36. S. Menon et al., "Protein profile study of breast cancer using HPLC-LIF—a pilot study," *Proc. SPIE* **6430**, 64300W (2007).
37. S. K. Singh et al., "Protein profile study of clinical samples of ovarian cancer using HPLC-LIF," *Proc. SPIE* **6092**, 60920N (2006).
38. L. R. Sujatha, V. B. Kartha, and C. Santhosh, "Protein profile study of serum of normal and cervical cancer subjects by high performance liquid chromatography-laser induced fluorescence," *J. Biomed. Opt.* **13**(5), 054062 (2008).
39. S. Bhat et al., "Protein profile analysis of cellular samples from the cervix for the objective diagnosis of cervical cancer using HPLC-LIF," *J. Chromatogr. B* **878**(31), 3225–3230 (2010).
40. E. P. Diamandis, "Proteomic patterns in biological fluids: do they represent the future of cancer diagnostics?," *Clin. Chem.* **49**(8), 1272–1278 (2003).
41. P. Denny et al., "The proteomes of human parotid submandibular / sublingual gland salivas collected as the ductal secretions," *J. Proteome Res.* **7**(5), 1994–2006 (2008).
42. C. F. Streckfus et al., "Breast cancer related proteins are present in saliva and are modulated secondary to ductal carcinoma in situ of the breast," *Cancer Investig.* **26**(2), 159–167 (2008).
43. S. Hu et al., "Salivary proteomics for oral cancer biomarker discovery," *Clin. Cancer Res.* **14**, 6246–6252 (2008).
44. H. P. Lawrence, "Salivary markers of systemic disease: non invasive diagnosis of disease and monitoring of general health," *J. Can. Dent. Assoc.* **68**(3), 170–174 (2002).
45. P. A. Wilmarth et al., "Two dimensional liquid chromatography study of human whole saliva proteome," *J. Proteome Res.* **3**(5), 1017–1023 (2004).
46. D. T. Wong, "Salivary diagnostics: enhancing disease detection and making medicine better," *Eur. J. Dent. Educ.* **12**(1), 22–29 (2008).
47. D. T. Wong, "Towards a simple saliva-based test for the detection of the oral cancer," *Expert Rev. Mol. Diagn.* **6**(3), 267–272 (2006).
48. H. Xie et al., "Proteomics analysis of cells in whole saliva from oral cancer patients via value-added three-dimensional peptide fractionation and tandem mass spectrometry," *Mol. Cell. Proteomics* **7**(3), 486–498 (2007).
49. B. Alberts et al., *Molecular Biology of the Cell*, 4th ed., Chapter 23, p. 1319, Garland Publishing, Inc., New York (2002).
50. R. J. Lakshmi et al., "Tissue Raman spectroscopy for the study of radiation damage: brain irradiation of mice," *Radiat. Res.* **157**(2), 175–182 (2002).
51. C. M. Krishna et al., "Raman spectroscopy studies for diagnosis of cancers in human uterine cervix," *Vib. Spectrosc.* **41**(1), 136–141 (2006).
52. R. Malini et al., "Discrimination of normal, inflammatory, premalignant, and malignant oral tissue: a Raman spectroscopy study," *Biopolymers* **81**(3), 179–193 (2006).
53. B. K. Manjunath et al., "Autofluorescence of oral tissue for optical pathology in oral malignancy," *J. Photochem. Photobiol. B* **73**(1–2), 49–58 (2004).
54. G. S. Nayak et al., "Principal component analysis and artificial neural network analysis of oral tissue fluorescence spectra: classification of normal premalignant and malignant pathological conditions," *Biopolymers* **82**(2), 152–166 (2006).
55. P. C. Mahalanobis, "On the generalised distance in statistics," *Proc. Natl. Inst. Sci. India* **2**(1), 49–55 (1936).
56. D. S. Q. Koh and G. C. H. Koh, "The use of salivary biomarkers in occupational and environmental medicine," *Occup. Environ. Med.* **64**(3), 202–210 (2007).
57. S. Bhat, "Protein profile and laser fluorescence spectroscopic studies of clinical samples for early detection of cervical cancer," Ph.D. Thesis, Manipal University (2009).
58. A. Patil et al., "Evaluation of "high performance liquid chromatography-laser induced fluorescence (HPLC-LIF) for serum protein profiling for the early diagnosis of oral cancer," *J. Biomed. Opt.* **15**(6), 067007 (2010).
59. V. B. Kartha et al., "Diagnosis at the molecular level: analytical laser spectroscopy for clinical applications," in *Photo/Electrochem. Photobiol. Environ. Energy Fuel*, pp. 153–221, Research Signpost, Trivandrum, India (2005).
60. G. S. Mohamed et al., "The importance of protein profiling in the diagnosis and treatment of hematologic malignancies," *Turk. J. Hematol.* **28**, 1–14 (2011).
61. N. Yamada, R. Yuji, and E. Suzuki, "The current status and future prospects of the salivary proteome," *J. Health Sci.* **55**(5), 682–688 (2009).
62. S. Hu1 et al., "Salivary proteomics for oral cancer biomarker discovery," *Clin. Cancer Res.* **14**, 6246–6252 (2008).
63. H. Xie et al., "Proteomics analysis of cells in whole saliva from oral cancer patients via value-added three-dimensional peptide fractionation and tandem mass spectrometry," *Mol. Cell. Proteomics* **7**, 486–498 (2007).
64. A. Patil et al., "Highly sensitive high performance liquid chromatography-laser induced fluorescence for proteomics," *ISRN Spectrosc.* **2012**, 643979 (2012).