Kobe University Repository : Kernel

PDF issue: 2024-06-09

A novel gas chromatography mass spectrometry-based serum screening method for oral squamous cell carcinoma

Kimoto, Akira

(Degree)
博士（医学）
(Date of Degree)
2013-09-25
(Date of Publication)
2014-09-01
(Resource Type)
doctoral thesis
(Report Number)
甲第5951号
(URL)
https://hdl.handle.net/20.500.14094/D1005951

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。

A novel gas chromatography mass spectrometry-based serum screening method for oral squamous cell carcinoma

口腔扁平上皮癌に対するガスクロマトグラフ質量分析計を用いた新規血清スクリーニング手法

木本　明,　西海　信,　小林　隆,　寺島　禎彦,　鈴木　泰明,　竹内　純一郎,
東　健,　古森　孝英,　吉田　優

神戸大学大学院医学研究科医科学専攻
口腔外科学
（指導教員：古森孝英教授）

木本　明

Key words : Oral squamous cell carcinoma, Oral cancer, Biomarker, Gas chromatography/mass spectrometry, Metabolomics

Title:

A novel gas chromatography mass spectrometry-based serum screening method for oral squamous cell carcinoma

Author names and affiliations:

Akira Kimoto [1], Shin Nishiumi [2], Takashi Kobayashi [2], Yoshihiko Terashima [2], Hiroaki Suzuki [1], Junichiro Takeuchi [1], Takeshi Azuma [2], Takahide Komori [1], Masaru Yoshida [2,3,4]

[1]Department of Oral and Maxillofacial Surgery, Kobe University Graduate School of Medicine, Kobe, Japan

[2]Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kobe, Japan

[3]The Integrated Center for Mass Spectrometry, Kobe University Graduate School of Medicine, Kobe, Japan

[4]Division of Metabolomics Research, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kobe, Japan

Corresponding author:

Masaru Yoshida

Division of Gastroenterology, The Integrated Center for Mass Spectrometry, Division of Metabolomics Research, Kobe University Graduate School of Medicine, Kusunoki-cho 7-5-1, Chuo-ku, Kobe, Hyogo 650-0017, Japan

E-mail: myoshida@med.kobe-u.ac.jp

Abstract

Oral squamous cell carcinoma (OSCC) is the eleventh most common cancer in the world. The five-year survival rate of patients who are diagnosed with OSCC at an early stage is about 80-90%. On the other hand, at an advanced stage, the five-year survival rate is less than 30%. Therefore, it is important to detect OSCC at an early stage. In this study, we subjected OSCC patients' serum samples to gas chromatography/mass spectrometry (GC/MS)-based metabolomic analysis. The serum samples of 18 OSCC patients were used as a training set. Their preoperative serum metabolite levels were compared with those of healthy volunteers. In addition, the pre- and postoperative serum metabolite levels of these patients were also compared. The serum samples collected from another 17 OSCC patients and healthy volunteers were used as a validation set. In addition, 12 patients with oral diseases other than OSCC were also used to evaluate the reliability of the extracted serum metabolite biomarker candidates. The metabolites that displayed significant differences were subjected to multiple logistic regression analysis, and it was found that models based on candidate biomarker pairs possessed greater accuracy than single biomarker candidates. GC/MS-based metabolomic analysis is a promising method for detecting early-stage OSCC.

Keywords

Oral squamous cell carcinoma; Oral cancer; Biomarker; Gas chromatography/mass spectrometry; Metabolomics

Introduction

Oral squamous cell carcinoma (OSCC) is the eleventh most common cancer in the world. In 2008, there were about 263,000 new oral cancer cases and about 127,000 deaths from the disease [1]. The five-year survival rate of patients that are diagnosed with OSCC at an early stage is about 80-90%. On the other hand, when OSCC is diagnosed at an advanced stage the five-year survival rate is less than 30%. Overall, only about 50% of OSCC patients survive for five years. The treatment of OSCC at an advanced stage can have adverse effects on the patient's appearance and can also lead to a variety of problems with eating and speaking. In addition, advanced-stage OSCC causes a greater reduction in quality of life than other cancers. Thus, it is important to detect OSCC at an early stage; however, this is currently difficult. The standard initial diagnostic approaches to OSCC are mainly dependent on inspection and manipulation. Conventional tumor markers, e.g., SCC antigen, and diagnostic imaging, such as computed tomography and biopsy, are not suitable for OSCC screening because conventional tumor markers have low sensitivity for the disease, diagnostic imaging is expensive, and biopsy is invasive. Therefore, new methods for detecting OSCC at an early stage are required.

The metabolome, which is the comprehensive characterization of the small molecular weight metabolite profiles, is beginning to gain attention. Analyzing metabolite profiles allows genotype-phenotype relationships to be evaluated because the metabolome is closely related to physiological functions and pathological characteristics. In metabolomics, nuclear magnetic resonance (NMR) spectroscopy, gas chromatography/mass spectrometry (GC/MS), liquid chromatography/mass

spectrometry (LC/MS), and capillary electrophoresis/mass spectrometry (CE/MS) have been used to measure the levels of a variety of metabolites. In this study, we used GC/MS-based metabolomic analysis to discover new biomarkers and screening methods for OSCC. The main advantages of GC/MS are its simplicity of use, the abundance of databases for serum metabolite analysis, and its high repeatability; therefore, GC/MS is suitable for screening studies aimed at discovering metabolite biomarkers of OSCC. We have performed several metabolomic studies using GC/MS [2] and reported serum metabolite biomarkers for various cancers [3,4,5,6,7]. In this study, we used 18 OSCC patients and 18 healthy volunteers as a training set, and an OSCC prediction model was established via multiple logistic regression analysis. Then, the validity of the prediction model was assessed using a validation set consisting of OSCC patients (N=17) and healthy volunteers (N=17).

Materials and methods

Participants

This study was approved by the ethics committee at Kobe University Graduate School of Medicine. The human samples were used in accordance with the guidelines of Kobe University Hospital, and written informed consent was obtained from all subjects. For the training set, serum samples were collected from 18 patients who underwent resection for OSCC. Serum samples were collected from these patients both before surgery and over 2 weeks after the operation. In addition, serum samples were also obtained from healthy volunteers in the morning after overnight fasting. For the validation set, serum samples were collected prior to surgery from 17 patients who were scheduled to undergo resection for OSCC as well as from 17 healthy volunteers. Serum samples were also obtained prior to surgery from 12 patients who underwent operations for oral diseases other than oral cancer, and further samples were collected from 10 of these patients over 2 weeks after the operation. The serum samples from the OSCC and oral disease patients were collected at Kobe University Hospital between April 2011 and March 2012. The patients were diagnosed by microscopy, biopsy, or surgical resection and classified using the TNM classification of the Union for International Cancer Control (UICC). The serum samples from the healthy volunteers were obtained at another facility. The characteristics of all subjects are summarized in Table 1. The serum was separated by centrifugation at 3,000 x g for 10 min at 4°C, before being transferred to a clean tube and stored at -80°C until use.

Experimental procedures

The extraction of low molecular weight metabolites was performed according to the

method described in our previous report [8], and 0.5 mg/ml of 2-isopropylmalic acid (Sigma-Aldrich, Tokyo, Japan) dissolved in distilled water was used as an internal standard. Then, oximation and the subsequent derivatization for GC/MS measurement were carried out according to a previously described method [8].

GC/MS analysis was performed using a GCMS-QP2010 Ultra spectrometer (Shimadzu Co., Kyoto, Japan) with a fused silica capillary column (CP-SIL 8 CB low bleed/MS; inner diameter, 30 mm × 0.25 mm; film thickness, 0.25 μm; Agilent Co., Palo Alto, CA), according to a previously described method [9]. The front inlet temperature was 230°C, and the flow rate of helium gas through the column was 39.0 cm/sec. The column temperature was held at 80°C for 2 min and then raised by 15°C/min to 330°C and held there for 6 min. The transfer line and ion-source temperatures were 250°C and 200°C, respectively. Twenty scans per second were recorded over the mass range 85-500 m/z using the Advanced Scanning Speed Protocol (ASSP, Shimadzu Co.).

Data processing was performed according to the methods described in previous reports [9,10]. Briefly, the MS data were exported in netCDF format. The peak detection and alignment were performed using the MetAlign software (Wageningen UR, The Netherlands). The resultant data were exported in CSV format and then analyzed with in-house analytical software (AIoutput). This software enables peak identification and semi-quantification using an in-house metabolite library. For semi-quantification, the peak height of each ion was calculated and normalized to the peak height of 2-isopropylmalic acid as an internal standard. Names were assigned to each metabolite peak based on the method described in a previous report [10].

Statistical analysis

The patients (N=35) were allocated to the training or validation sets (Table 1). Regarding the healthy volunteers used for the training set, age- and sex-controlled samples were prepared (N=18). In the training set study, the OSCC patients' preoperative serum metabolite levels were compared with those of the healthy volunteers using the Wilcoxon rank sum test, and comparisons between the OSCC patients' pre- and postoperative serum metabolite levels were performed using the Wilcoxon signed-rank test. The metabolites that displayed significant differences ($p<0.05$) in their levels in both of these comparisons were selected for the validation set study. In the validation set study, the preoperative levels of the selected metabolites in the OSCC patients were compared with those of the healthy volunteers using the Wilcoxon rank sum test. Comparisons between the oral disease patients and OSCC patients, and between the pre- and postoperative serum metabolite levels of the oral disease patients were also performed. Based on these univariate analyses, tumor biomarker candidates were selected. In the training set, multiple logistic regression models were used to predict OSCC positivity and negativity. Receiver operating characteristic (ROC) analysis was carried out using JMP10 (SAS Institute Inc.), and the optimal cut-off, AUC, specificity, and sensitivity values were calculated. In the validation set and the training + validation set studies, the predictive models were evaluated using different samples, and the specificity and sensitivity of the models were examined using the cut-off values obtained from the training set. p values of less than 0.05 were considered to indicate a significant difference.

Results

In our GC/MS-based metabolomic analysis system, which mainly targeted water-soluble metabolites, 101 metabolites were detected in the subjects' sera. Among the 101 metabolites, 1 metabolite; i.e., 2-isopropylmalic acid, was used as an internal standard. Therefore, this metabolite was excluded from the subsequent analyses, and a total of 100 metabolites were subjected to a comparative evaluation between OSCC patients and healthy volunteers.

In the training set study, the preoperative serum metabolite levels of the OSCC patients were compared with those of the healthy volunteers using the Wilcoxon rank sum test. The training set was composed of OSCC patients (N=18) and age- and sex-matched healthy volunteers (N=18). Thirty-eight metabolites displayed significant differences between the pre-operative levels of the OSCC patients and those of the healthy volunteers ($p<0.05$) (Table 2). Then, a comparison of the OSCC patients' pre- and postoperative metabolite levels was performed with the training set using the Wilcoxon signed-rank test, and 32 of the metabolites displayed significant alterations (Table 3). Thirteen metabolites; i.e., benzoic acid, glyceric acid, serine (3TMS), acetylsalicylic acid, lauric acid, N-acetyl-L-aspartic acid_1, asparagine, ornithine, glucuronate_1, heptadecanoate, phosphate, O-phosphoethanolamine and cysteine+cystine, were extracted from both of the abovementioned comparisons. The 13 selected metabolites were re-evaluated using the validation set, which was composed of OSCC patients (N=17) and healthy volunteers (N=17). As a result, 7 metabolites were selected as biomarker candidates (Table 4) because significant and similar alterations in their serum levels were observed in both the training and validation sets. Among the oral disease

patients, 6 of the metabolites (not including ornithine) displayed significantly different levels compared with healthy volunteers (Table 5). None of the 7 metabolites exhibited significant differences between their preoperative and postoperative levels in the oral disease patients (Table 6). In Figure 1, the differences in the levels of the 7 metabolites among tumor stages are shown. The 7 biomarker candidates displayed higher values in the OSCC patients than in the healthy volunteers at each stage, and the more advanced tumor stages tended to exhibit higher levels of each metabolite than the earlier tumor stages. ROC curves were prepared using the data for these 7 metabolites, and the cut-off, sensitivity, and specificity values of each metabolite were calculated (Table 7). In the training set, glyceric acid (77.7%), lauric acid (100%), N-acetyl-L-aspartic acid (94.4%), ornithine (88.8%), and heptadecanoate (83.3%) displayed relatively high sensitivity values (shown in brackets). Regarding specificity, lauric acid exhibited 100% specificity, and glyceric acid (77.7%), serine (3TMS) (94.4%), N-acetyl-L-aspartic acid (94.4%), asparagine (88.8%), and ornithine (94.4%) also demonstrated relatively high specificity values (shown in brackets). Lauric acid, N-acetyl-L-aspartic acid, and ornithine all displayed AUC values of greater than 80%. We then performed evaluations using multiple biomarkers to improve the reliability of our predictive models. Multiple logistic regression models involving pairs of the 7 selected metabolite biomarkers metabolites were employed as a new evaluation method (Table 8). As a result, we found that 4 pairs of biomarker candidates; i.e., ornithine+asparagine, ornithine+glyceric acid, ornithine+N-acetyl-L-aspartic acid, and ornithine+serine (3TMS), displayed sensitivity values of over 80% and specificity values of over 70% in the training set, validation set, and training+validation set.

Discussion

The five-year survival rate of patients that are diagnosed with OSCC at an early stage is about 80-90%, but when it is diagnosed at an advanced stage the five-year survival rate is less than 30%, suggesting that it is important to detect OSCC at an early stage. The SCC antigen is used as a tumor marker of OSCC; however, it is estimated that it displays a sensitivity of about 30% during the early stages of the disease. Therefore, the SCC antigen is unsuitable for the early detection of OSCC, although it is partially useful for assessing the effects of treatment and the early detection of recurrence. In this study, it was suggested that serum metabolome analysis might be useful for diagnosing early-stage OSCC, and 7 metabolite biomarker candidates; i.e., glyceric acid, lauric acid, N-acetyl-L-aspartic acid, ornithine, heptadecanoate, serine, and asparagine, were found. In particular, a diagnostic system for OSCC based on combining several metabolite biomarker candidates exhibited high sensitivity and specificity. During the early stages of the disease; i.e., stages 1 and 2, the sensitivity values of the 7 metabolite biomarker candidates for OSCC were 77.7% for glyceric acid, 83.3% for lauric acid, 66.6% for N-acetyl-L-aspartic acid, 88.8% for ornithine, 66.6% for heptadecanoate, 38.8% for serine, and 50.0% for asparagine. In addition, their specificity values were 88.5% for glyceric acid, 100% for lauric acid, 97.1% for N-acetyl-L-aspartic acid, 85.7% for ornithine, 80.0% for heptadecanoate, 97.1% for serine, and 91.4% for asparagine. Furthermore, the sensitivity and specificity values of ornithine+asparagine, ornithine+glyceric acid, ornithine+N-acetyl-L-aspartic acid, and ornithine+serine were 94.4% and 82.8%, 88.8% and 85.7%, 88.8% and 97.1%, and 88.8% and 82.8%, respectively. Thus, our results suggest that various biomarker pairs might be useful for the early detection of OSCC.

Many studies of OSCC biomarkers have been performed by various research groups, but most of them did not intend to find biomarkers for OSCC screening, e.g., they examined the utility of SCC antigen. The sensitivity of SCC antigen for early-stage OSCC is about 30%, although it was found to be useful for aiding the early diagnosis of recurrence [11]. In another paper, it was demonstrated that SCC antigen is a prognostic indicator of OSCC but is not useful for OSCC screening [12]. A previous study suggested that metabolomics analysis of plasma samples using nuclear magnetic resonance spectroscopy might be useful for OSCC screening [13], and the plasma levels of ornithine and asparagine were found to be higher in OSCC patients than in healthy volunteers [14], which was consistent with our findings. In addition, it was reported that most essential and non-essential amino acids displayed lower levels in OSCC patients than in healthy volunteers [15,16]; however, these results were partly caused by dysphagia and insufficient food intake. However, all of the patients in the study by Stefano Tiziani et al. [14] were provided with sufficient nutrition; therefore, the observed alterations in the levels of ornithine and asparagine might not have been due to the patients' nutrition status. Among the 7 metabolite biomarker candidates extracted in the present study, lauric acid and heptadecanoate are lipid-soluble compounds that are found in food, suggesting that the effects of nutrition in OSCC and/or other oral diseases might partially contribute to these variations. In addition, the metabolite profile of the blood represents the sum of metabolic events in multiple organs, and it is difficult to understand the reasons for changing serum metabolite levels in OSCC patients from serum metabolite profiling. However, in this study we performed comparisons not only between OSCC patients and healthy volunteers but also between the pre- and

postoperative values of OSCC patients. In addition, the preoperative levels of patients with oral diseases other than OSCC were investigated to diminish the influence of oral surgery on our results and enhance the reliability of the OSCC biomarker candidates we selected. Therefore, our findings are considered to closely reflect the differences between the presence and absence of OSCC. In particular, the level of ornithine did not differ significantly between the healthy volunteers and oral disease patients (Table 5); therefore, this metabolite might be heavily involved in OSCC. Heavy smoking and alcohol consumption are primary causes of OSCC. Moreover, a correlation between human papillomavirus (HPV) positivity and OSCC has also been found recently. However, the relationships between these risk factors and the metabolome have not been elucidated. To establish a metabolomics-based method for detecting early-stage OSCC, it will be necessary to investigate whether these risk factors influence the serum metabolite profile.

In conclusion, the present study could represent a first step towards a novel serum metabolomics-based screening method for OSCC. The extracted serum metabolites were demonstrated to be more effective at detecting OSCC than conventional tumor biomarkers; i.e., they displayed high sensitivity. Moreover, based on comparisons of the pre- and postoperative values of OSCC patients, this novel approach might also be useful for diagnosing recurrence. This study was limited to analyses of serum samples obtained from OSCC patients, and comparative evaluations between OSCC and other forms of cancer should be performed in future. Thus, it will be necessary to perform a further study involving many serum samples from individuals with a variety of conditions. Taken together, our findings highlight the importance of large-scale studies

aimed at developing metabolomics-based methods for detecting early-stage OSCC.

Acknowledgements

This study was supported, in part, by grants for the Global COE Program, Global Center of Excellence for Education and Research on Signal Transduction Medicine in the Coming Generation from The Ministry of Education, Culture, Sports, Science, and Technology [T.A., M.Y.] and a Grant-in-Aid for Scientific Research (C) from the Japan Society for the Promotion of Science [T.K.].

References

1, Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int. J. Cancer 2010; 127: 2893-2917.

2, Yoshida M, Hatano N, Nishiumi S, Irino Y, Izumi Y, Takenawa T, Azuma T. Diagnosis of gastroenterological diseases by metabolome analysis using gas chromatography-mass spectrometry. Journal of Gastroenterology 2012; 47: 9-20. Review.

3,Kobayashi T, Nishiumi S, Ikeda A, Yoshie T, Sakai A, Matsubara A, et al. A novel serum metabolomics-based diagnostic approach to pancreatic cancer. Cancer Epidemiol Biomarkers Prev. 2013 Apr; 22: 571-9.

4, Hori S, Nishiumi S, Kobayashi K, Shinohara M, Hatakeyama Y, Kotani Y, Hatano N, Maniwa Y, Nishio W, Bamba T, Fukusaki E, Azuma T, Takenawa T, Nishimura Y, Yoshida M. A metabolomic approach to lung cancer. Lung Cancer 2011; 74: 284-292.

5, Kondo Y, Nishiumi S, Shinohara M, Hatano N, Ikeda A, Yoshie T, Kobayashi T, Shiomi Y, Irino Y, Takenawa T, Azuma T, Yoshida M. Serum fatty acid profiling of colorectal cancer by gas chromatography/mass spectrometry. Biomarkers in Medicine 2011; 5: 451-460.

6, Ikeda A, Nishiumi S, Shinohara M, Yoshie T, Hatano N, Okuno T, Bamba T, Fukusaki E, Takenawa T, Azuma T, Yoshida M. Serum metabolomics as a novel diagnostic approach for gastrointestinal cancer. Biomedical Chromatography 2012; 26: 548-558.

7, Nishiumi S, Kobayashi T, Ikeda A, Yoshie T, Kibi M, Izumi Y, Okuno T, Hayashi N, Kawano S, Takenawa T, Azuma T, Yoshida M. A novel serum metabolomics-based

diagnostic approach for colorectal cancer. PLoS ONE 2012; 7: e40459.

8, Nishiumi S, Shinohara M, Ikeda A, Yoshie T, Hatano N, Kakuyama S, et al. Serum metabolomics as a novel diagnostic approach for pancreatic cancer. Metabolomics 2010; 6: 518−28.

9, Tsugawa H, Bamba T, Shinohara M, Nishiumi S, Yoshida M, Fukusaki E. Practical non-targeted gas chromatography/mass spectrometry-based metabolomics platform for metabolic phenotype analysis. Journal of Bioscience and Bioengineering 2011; 112: 292−8.

10, Tsugawa H, Tsujimoto Y, Arita M, Bamba T, Fukusaki E. GC/MS based metabolomics: development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA). BMC Bioinformatics 2011; 12:131.

11, Krimmel M, Hoffmann J, Krimmel C, Cornelius CP, Schwenzer N. Relevance of SCC-Ag, CEA, CA 19.9 and CA 125 for diagnosis and follow-up in oral cancer. Journal of Cranio-Maxillofacial Surgery 1998; 26: 243- 248.

12, Lin WH, Chen IH, Wei FC, Huang JJ, Kang CJ, Hsieh LL, Wang HM, Huang SF. Clinical Significance of Preoperative Squamous Cell Carcinoma Antigen in Oral-Cavity Squamous Cell Carcinoma. Laryngoscope 2011; 121(5): 971-977.

13, Zhou J, Xu B, Huang J, Jia X, Xue J, Shi X, Xiao L, Li W. ^{1}H NMR-based metabonomic and pattern recognition analysis for detection of oral squamous cell carcinoma. Clin Chim Acta 2009; 401(1-2): 8-13.

14, Tiziani S, Lopes V, Günther UL. Early stage diagnosis of oral cancer using ^{1}H NMR-based metabolomics. Neoplasia 2009; 11(3): 269-76.

15, Kubota A, Meguid MM, and Hitch DC. Amino-acid profiles correlate diagnostically

with organ site in 3 kinds of malignant-tumors. Cancer 1992; 69: 2343–2348.

16, Ching N, Grossi C, Jham G, Angers J, Zurawinsky H, Ching CY, and Nealon TF. Plasma amino-acid and serum unesterified fatty-acid deficits and the effect of nutritional support in chemotherapy treatment. Surgery 1984; 95: 730–738.

TABLE LEGENDS

Table 1. Subject information for the training and validation sets

The cancer patients in this study all had squamous cell carcinoma. There was no significant difference ($p>0.05$) in any of the background characteristics between the OSCC group and healthy volunteer group according to the Wilcoxon rank sum test.

Table 2. A list of metabolites whose serum levels differed significantly ($p<0.05$) between the OSCC patients and healthy volunteers in the training set

The values are shown as fold-induction values (the peak intensity of the OSCC patients (n=18) divided by that of the healthy volunteers (n=18)). P values were calculated using the Wilcoxon rank sum test.

Table 3. A list of metabolites that displayed significantly different ($p<0.05$) pre- and postoperative serum levels in the OSCC patients

P values were calculated using the Wilcoxon signed-rank test.

Table 4. A list of metabolites (from among the 13 selected metabolites) whose serum levels differed significantly ($p<0.05$) between the OSCC patients and healthy volunteers in the validation set

Values are represented as fold-induction values (the peak intensity of the OSCC patients (n=17) divided by that of the healthy volunteers (n=17)). P values were calculated using the Wilcoxon rank sum test. The 13 selected metabolites, whose serum levels differed significantly between the preoperative OSCC patients and the healthy volunteers as well as between the pre- and postoperative OSCC patients in the training set ($p<0.05$), were

evaluated, and then the listed metabolites were selected.

Table 5. A list of metabolites whose serum levels displayed significant differences (p<0.05) between the oral disease patients and healthy volunteers

Values are represented as fold-induction values (the peak intensity of the oral disease patients (n=12) divided by that of the healthy volunteers (n=35; data are from the training set + validation set)). P values were calculated using the Wilcoxon rank sum test. The oral disease patients had the following conditions: jaw deformities (n=9), a bone graft (n=1), a fractured jaw (n=1), or a benign tumor (n=1). The mean age of the oral disease patients was 28.9, and the male to female ratio was 5:7.

Table 6. A list of metabolites whose pre- and postoperative serum levels in oral disease patients were significantly different (p<0.05)

P values were calculated using the Wilcoxon rank sum test. The 12 preoperative samples were obtained from patients with the following oral conditions: jaw deformities (n=9), a bone graft (n=1), a fractured jaw (n=1), or a benign tumor (n=1). The 10 postoperative samples were collected from patients with the following conditions: jaw deformities (n=7), a bone graft (n=1), a fractured jaw (n=1), or a benign tumor (n=1). The mean age and male to female ratio of the oral disease patients were 28.9 and 5:7 for the preoperative samples and 29.7 and 1:1 for the postoperative samples, respectively.

Table 7. Sensitivity, specificity, and AUC values of the 7 biomarker candidates

Sensitivity, specificity, and AUC values were calculated in the training set, validation set, and training + validation set using ROC analysis.

Table 8. Sensitivity, specificity, and AUC values of metabolite pairs selected from among the 7 biomarker candidates

The metabolite pairs were selected from among the 7 biomarker candidates and had their sensitivity, specificity, and AUC values evaluated by logistic regression analysis.

FIGURE LEGENDS

Figure 1. The serum levels of the 7 selected biomarker candidates according to clinical stage

The 7 selected biomarker candidates included glyceric acid, serine, lauric acid, N-acetyl-L-aspartic acid, asparagine, ornithine, and heptadecanoate. The letter 'a' indicates a significant difference according to the Steel-Dwass test ($p<0.05$). The data were obtained from the training + validation set.

Table 1. Subject information for the training and validation sets

		Training set		Validation set	
		OSCC patients	Healthy volunteers	OSCC patients	Healthy volunteers
N		18	18	17	17
	Male	13	13	11	12
	Female	5	5	6	5
Age	(years)				
	Mean	66.5	66.6	65.8	64.8
	Range	38-88	59-75	23-87	57-75
TNM stage	I	4	-	4	-
	II	6	-	4	-
	III	2	-	4	-
	IV	6	-	5	-
Cancer location	Tongue	7	-	6	-
	Gingiva	7	-	5	-
	Cheek	2	-	4	-
	Floor of the mouth	2	-	1	-
	Lip	0	-	1	-

The cancer patients in this study all had squamous cell carcinoma. There was no significant difference ($p>0.05$) in any of the background characteristics between the OSCC group and healthy volunteer group according to the Wilcoxon rank sum test.

Table 2. A list of metabolites whose serum levels differed significantly ($p<0.05$) between the OSCC patients and healthy volunteers in the training set

Compound name	Fold induction (cancer/healthy)	p value
Lactic acid	1.57	<0.0001
Oxalate	1.91	0.0062
Malonic acid	0.10	0.0003
Urea	1.67	0.0007
Benzoic acid	1.40	0.0155
2-Aminoethanol	1.48	0.009
Phosphate	0.81	0.0279
Glycine (3TMS)	1.34	0.0009
Succinic acid (or aldehyde)	1.24	0.009
Glyceric acid	1.48	0.0025
Serine (3TMS)	1.36	0.0382
Thymine	0.24	0.0025
Acetylsalicylic acid	2.69	0.0237
Aspartic acid	1.66	0.0005
Pyroglutamic acid	1.31	0.0068
Glutamic acid	1.82	0.0005
Phenylalanine	1.50	0.0008
p-Hydroxybenzoic acid	10.25	<0.0001
4-Hydroxyphenylacetic acid	2.59	0.0068
Lauric acid	2.54	<0.0001
N-acetyl-L-aspartic acid_1	2.35	<0.0001
Ribose	0.49	<0.0001
Asparagine	1.25	0.0119
Xylitol	1.64	0.0013
Arabitol	2.28	<0.0001

Orotic acid	3.57	<0.0001
O-Phosphoethanolamine	0.77	0.0257
Ornithine	1.58	<0.0001
Glucose_1	1.31	0.0006
Sebacic acid	0.58	0.0042
Galactosamine_1	0.38	0.0002
Glucuronate_1	0.995	0.0354
Inositol	0.86	0.0042
Heptadecanoate	1.38	0.0142
Kynurenine	1.75	0.048
Tryptophan	0.69	0.0038
Spermidine	2.79	0.0445
Cysteine+Cystine	0.80	0.0082

The values are shown as fold-induction values (the peak intensity of the OSCC patients (n=18) divided by that of the healthy volunteers (n=18)). P values were calculated using the Wilcoxon rank sum test.

Table 3. A list of metabolites that displayed significantly different ($p<0.05$) pre- and postoperative serum levels in the OSCC patients

Compound name	p value
Glycolic acid	0.0182
Sarcosine	0.0023
Ketoisoleucine_1	0.0182
Valine (2TMS)	0.0208
Benzoic acid	0.0432
Phosphate	0.0208
Isoleucine	0.0034
Threonine (2TMS)	0.0047
Glyceric acid	0.0237
Serine (3TMS)	0.0483
Nonanoic acid (C9)	0.0056
Homoserine	0.0483
Acetylsalicylic acid	0.0268
Methionine	0.0023
Pyrogallol	0.004
Lauric acid	<0.0001
N-acetyl-L-aspartic acid_1	<0.0001
Asparagine	0.001
1,6-Anhydroglucose	0.0001
Glutamine	0.0007
4-Hydroxymandelate	0.0047
O-Phosphoethanolamine	0.0432
Theanine_2	0.0047
Ornithine	0.0237
Cadaverine	0.0268
Tagatose_2 (or Psicose_2)	0.0432

α-Sorbopyranose_1 (or Fructose_1)	0.0342
Histidine	0.0034
Glucuronate_1	0.0483
1-Hexadecanol	0.0008
Heptadecanoate	<0.0001
Cysteine+Cystine	0.0003

P values were calculated using the Wilcoxon signed-rank test.

Table 4. A list of metabolites (from among the 13 selected metabolites) whose serum levels differed significantly ($p<0.05$) between the OSCC patients and healthy volunteers in the validation set

Compound name	Fold induction (cancer/control)	p value
Glyceric acid	2.26	<0.0001
Serine (3TMS)	2.68	<0.0001
Lauric acid	3.71	<0.0001
N-acetyl-L-aspartic acid_1	3.06	<0.0001
Asparagine	2.55	<0.0001
Ornithine	1.56	0.0003
Heptadecanoate	2.30	<0.0001

Values are represented as fold-induction values (the peak intensity of the OSCC patients (n=17) divided by that of the healthy volunteers (n=17)). P values were calculated using the Wilcoxon rank sum test. The 13 selected metabolites, whose serum levels differed significantly between the preoperative OSCC patients and the healthy volunteers as well as between the pre- and postoperative OSCC patients in the training set ($p<0.05$), were evaluated, and then the listed metabolites were selected.

Table 5. A list of metabolites whose serum levels displayed significant differences ($p<0.05$) between the oral disease patients and healthy volunteers

Compound name	Fold induction (oral disease /healthy)	p value
Lactic acid	1.28	0.0108
Glycolic acid	1.74	<0.0001
Hydroxybutyrate	1.72	0.0009
Oxalate	2.00	0.0013
2-Aminobutyric acid	1.28	0.0441
Ketoisoleucine_1	1.28	0.024
Malonic acid	0.09	0.0308
Valine (2TMS)	1.32	0.0348
2-Aminoethanol	1.42	0.0163
Phosphate	1.26	0.0101
Threonine (2TMS)	1.42	0.0052
Glycine (3TMS)	1.23	0.0036
Succinic acid (or aldehyde)	1.62	0.0101
Glyceric acid	2.44	<0.0001
Uracil	0.94	0.0116
Serine (3TMS)	1.74	0.0009
Nonanoic acid (C9)	1.55	0.0022
Thymine	0.19	0.0441
Homoserine	1.39	0.037
Malic acid	1.30	0.0308
Homoserine lactone	1.17	0.0045
Acetylsalicylic acid	2.44	<0.0001
Aspartic acid	1.79	<0.0001
Methionine	1.29	0.0495

trans-4-hydroxy-L-proline	1.47	0.0174
Pyroglutamic acid	1.23	0.0124
Glutamic acid	2.17	0.0001
Phenylalanine	1.61	0.0006
p-Hydroxybenzoic acid	15.70	<0.0001
4-Hydroxyphenylacetic acid	1.22	0.0081
Arabinose	0.74	0.0007
Lauric acid	3.76	<0.0001
N-acetyl-L-aspartic acid_1	2.94	<0.0001
Ribulose	1.27	0.0045
Ribose	0.44	<0.0001
Asparagine	1.47	0.0036
Xylitol	1.34	0.0163
1,6-Anhydroglucose	2.21	0.0416
Arabitol	1.60	0.037
Theanine_2	0.48	0.0441
Hypoxanthine	1.67	<0.0001
Lysine (3TMS)	0.60	0.0328
α-Sorbopyranose_1 (or Fructose_1)	2.19	0.0002
Mannose_1	0.83	0.0308
Sebacic acid	0.42	0.0005
Gulcono-1,4-lactone	0.49	0.0041
Galactosamine_1	0.30	0.0416
Glucuronate_1	0.54	0.0001
1-Hexadecanol	0.75	0.0108
Inositol	0.85	0.0108

N-α-acetyl-L-lysine_2	3.59	0.0015
Heptadecanoate	1.59	0.0017
Kynurenine	1.46	0.0441
Spermidine	2.17	0.003
Cysteine+Cystine	0.73	0.0348

Values are represented as fold-induction values (the peak intensity of the oral disease patients (n=12) divided by that of the healthy volunteers (n=35; data are from the training set + validation set)). P values were calculated using the Wilcoxon rank sum test. The oral disease patients had the following conditions: jaw deformities (n=9), a bone graft (n=1), a fractured jaw (n=1), or a benign tumor (n=1). The mean age of the oral disease patients was 28.9, and the male to female ratio was 5:7.

Table 6. A list of metabolites whose pre- and postoperative serum levels in oral disease patients were significantly different ($p<0.05$)

Compound name	p value
Malonic acid	0.0272
Nonanoic acid (C9)	0.0378
Homoserine	0.0076
4-Hydroxyphenylacetic acid	0.0027
Arabitol	0.0161
Ribitol	0.0001
Glucarate	0.0378

P values were calculated using the Wilcoxon rank sum test. The 12 preoperative samples were obtained from patients with the following oral conditions: jaw deformities (n=9), a bone graft (n=1), a fractured jaw (n=1), or a benign tumor (n=1). The 10 postoperative samples were collected from patients with the following conditions: jaw deformities (n=7), a bone graft (n=1), a fractured jaw (n=1), or a benign tumor (n=1). The mean age and male to female ratio of the oral disease patients were 28.9 and 5:7 for the preoperative samples and 29.7 and 1:1 for the postoperative samples, respectively.

Table 7. Sensitivity, specificity, and AUC values of the 7 biomarker candidates

	Training set			Validation set			Training + validation set		
Compounds	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
Glyceric acid	77.7	77.7	0.79630	76.4	94.1	0.97232	77.1	88.5	0.88082
Serine (3TMS)	55.5	94.4	0.70370	35.2	100	0.95848	42.8	97.1	0.79755
Lauric acid	100	100	1.00000	70.5	100	0.99308	82.8	100	0.97796
N-acetyl-L-aspartic acid_1	94.4	94.4	0.94136	52.9	100	0.96540	71.4	97.1	0.91755
Asparagine	61.1	88.8	0.74691	52.9	94.1	0.95502	54.2	91.4	0.83592
Ornithine	88.8	94.4	0.94753	88.2	70.5	0.86851	88.5	85.7	0.90204
Heptadecanoate	83.3	61.1	0.74074	58.8	94.1	0.94810	71.4	80.0	0.81061

Sensitivity, specificity, and AUC values were calculated in the training set, validation set, and training + validation set using ROC analysis.

Table 8. Sensitivity, specificity, and AUC values of metabolite pairs selected from among the 7 biomarker candidates

	Training set			Validation set			Training+validation set		
Compounds	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
Asparagine+Heptadecanoate	72.2	94.4	0.87037	41.1	100	0.9654	57.1	97.1	0.86776
Asparagine+Lauric acid	100	100	1	58.8	100	0.91176	80.0	100	0.9551
Asparagine+N-acetyl-L-aspartic acid_1	94.4	94.4	0.93827	82.3	100	0.98962	74.2	97.1	0.92653
Asparagine+Ornithine	94.4	88.8	0.94444	94.1	76.4	0.91003	94.2	82.8	0.92408
Asparagine+Serine (3TMS)	61.1	94.4	0.76235	35.2	94.1	0.95502	48.5	94.2	0.84163
Asparagine+Glyceric acid	77.7	83.3	0.85494	70.5	94.1	0.96886	74.2	88.5	0.89469
Glyceric acid+Heptadecanoate	94.4	83.3	0.93827	64.7	100	0.97924	80	91.4	0.91673
Glyceric acid+Lauric acid	100	100	1	64.7	100	0.94118	82.8	100	0.95102
Glyceric acid+N-acetyl-L-aspartic acid_1	94.4	94.4	0.95062	58.8	100	1	74.2	97.1	0.94041
Glyceric acid+Serine (3TMS)	72.2	94.4	0.83333	64.7	100	0.97578	68.5	97.1	0.88898
Lauric acid+Heptadecanoate	100	100	1	70.5	100	0.91176	85.7	100	0.95224
Lauric acid+N-acetyl-L-aspartic acid_1	100	100	1	64.7	100	0.91176	82.8	100	0.95347
N-acetyl-L-aspartic acid_1+Heptadecanoate	94.4	94.4	0.93827	52.9	100	0.95502	74.2	97.1	0.91429
Ornithine+Glyceric acid	88.8	94.4	0.94444	88.2	76.4	0.89273	88.5	85.7	0.91429
Ornithine+Heptadecanoate	88.8	100	0.97222	64.7	94.1	0.9308	77.1	97.1	0.94694
Ornithine+Lauric acid	100	100	1	58.8	100	0.85294	80.0	100	0.92653
Ornithine+N-acetyl-L-aspartic acid_1	94.4	100	0.95988	88.2	94.1	0.97232	91.4	97.1	0.97061
Ornithine+Serine (3TMS)	94.4	88.8	0.93519	88.2	76.4	0.88235	91.4	82.8	0.90612
Serine (3TMS)+Heptadecanoate	77.7	72.2	0.81481	41.1	100	0.97232	60	85.7	0.84327
Serine (3TMS)+Lauric acid	100	100	1	58.8	100	0.96886	80.0	100	0.97306
Serine (3TMS)+N-acetyl-L-aspartic acid_1	94.4	94.4	0.94136	52.9	100	0.96194	74.2	97.1	0.91837

The metabolite pairs were selected from among the 7 biomarker candidates and had their sensitivity, specificity, and AUC values evaluated by logistic regression analysis.

Figure 1.