CrossMark

ORIGINAL ARTICLE

Investigation of urinary volatomic alterations in head and neck cancer: a non-invasive approach towards diagnosis and prognosis

Ravindra Taware[1] · Khushman Taunk[1] · Jorge A. M. Pereira[2] · Rahul Dhakne[3] ·
Narayanan Kannan[4] · Dharmesh Soneji[4] · José S. Câmara[2,5] · H. A. Nagarajaram[3,6] ·
Srikanth Rapole[1]

Abstract

Introduction   Head and neck cancer (HNC), like many other forms of cancer, is usually detected in advanced stages, causing poor survival outcomes. Lack of specific and sensitive screening markers for early detection of HNC has worsened the scenario for the patients as well as the clinicians. Therefore, identification of efficient, noninvasive and affordable screening marker/methodology with high specificity and sensitivity is imminent need of situation.
Objectives   This study aims to identify and characterize urinary volatomic alterations specific to HNC.

Methods   Volatomic analysis of urine samples collected from HNC patients ($n = 29$) and healthy controls ($n = 31$) was performed using headspace solid phase microextraction coupled to gas chromatography mass spectrometry (GC–MS). Both univariate and multivariate statistical approaches were used to investigate HNC specific volatomic alterations.
Results   Statistical analysis revealed a total of 28 metabolites with highest contribution towards discrimination of HNC patients from healthy controls (VIP >1, $p < 0.05$, Log$_2$ FC $\geq 0.58/\leq -0.57$). The discrimination efficiency and accuracy of urinary VOCs was ascertained by ROC curve analysis that allowed the identification of four metabolites viz. 2,6-dimethyl-7-octen-2-ol, 1-butanol, *p*-xylene and 4-methyl-2-heptanone with highest sensitivity and specificity to discriminate HNC patients from healthy controls. Further, the metabolic pathway analysis identified several dysregulated pathways in HNC patients and their detailed investigations could unravel novel mechanistic insights into the disease pathophysiology.
Conclusion   Overall, this study provides valuable fingerprint of the volatile profile of HNC patients, which in turn, might help in improving the current understanding of this form of cancer and lead to the development of non-invasive approaches for HNC diagnosis.

Keywords   Head and neck cancer · Volatomics · VOMs · GC–MS · Urinary markers

✉   Srikanth Rapole
    rsrikanth@nccs.res.in

1   Proteomics Lab, National Centre for Cell Science, Ganeshkhind, Pune 411007, India

2   CQM/UMa – Centro de Química da Madeira, Universidade da Madeira, Campus Universitário da Penteada, 9000-390 Funchal, Portugal

3   Laboratory of Computational Biology, Centre for DNA fingerprinting & Diagnostics (CDFD), Nampally, Hyderabad 500001, India

4   Malignant Disease Treatment Centre, Military Hospital (Cardio Thoracic Centre), Armed Forces Medical College, Pune 411040, India

5   Faculdade de Ciências Exatas e da Engenharia, Universidade da Madeira, Campus Universitário da Penteada, 9000-390 Funchal, Portugal

6   Department of Biotechnology & Bioinformatics, School of Life Sciences, University of Hyderabad, Hyderabad 500 046, India

1 Introduction

According to the recent world health organization report, cancer is the leading cause of fatalities all over the world and was responsible for 8.2 million deaths in 2012. This report also points that 14 million new cases were diagnosed in 2012

and this number is expected to reach 22 million in the next couple of decades (Stewart and Wild 2014). Particularly in India, GLOBOCAN 2012 indicated that 0.68 million people lost their lives due to various forms of cancer, representing 8.3% of total global cancer casualties (Ferlay et al. 2014). Head and neck cancer (HNC) is a heterogeneous oncological disease comprising tumors from the mucosal lining of the mouth, salivary glands, oropharynx, nasopharynx, and hypopharynx (Rezende et al. 2010). HNC is the third most common cancer across the globe and collectively contribute to substantial mortality, with estimated 526,481 annual incidences worldwide (Ferlay et al. 2014). In India, however, HNC is the most common malignancy found in males, a fact that is correlated to the high prevalence of human papillomavirus (HPV) infection, tobacco and alcohol abuse (Wyss et al. 2013).

Among the HNCs, mouth and tongue malignancies are the most common in the Indian subcontinent (Mishra and Meherotra 2014). For Epstein-Barr virus (EBV) and HPV positive patients, reliable prognostic biomarkers along with positron emission tomography imaging technique are routinely used in clinical practices. However, for HPV negative cases, lack of suitable screening biomarkers results in a disease diagnosis at advanced stages which ultimately leads to high mortality due to the narrow window for therapeutic intervention (Yonezawa et al. 2013). Squamous cell carcinoma antigen, the most reliable marker for the head and neck squamous cell carcinoma has only 43% sensitivity (Inal et al. 2004). Other putative markers such as p53, Ki67, Collagen XVII and EGFR with a sensitivity of 63, 54, 78.5 and 82.6% respectively are also not reliable for diagnosis use in the clinical practice (Tamás et al. 2011). Beyond the absence of reliable biomarkers, the late diagnosis, invasive as well as expensive surgical procedures and the limited access to modern medical facilities are important reasons for the high prevalence of HNC in India (Mallath et al. 2014).

Therefore, the development of reliable, affordable and noninvasive methods/markers for the diagnosis and screening of high risk population in the clinical environment is the utmost need of the current scenario. An exciting alternative to overcome the referred constraints of conventional diagnosis is the detection of discriminant sets of volatile organic metabolites (VOMs) for HNCs in biofluids such as urine (Amann et al. 2014; Downes et al. 2007; Shen et al. 2015; Silva et al. 2011, 2012). The principle involved in this approach is supported by the fact that disease and normal condition can be distinguished from each other by their changed physiology and metabolic rates, as well as differential protein expression patterns, which ultimately leads to the production of disease specific alterations in VOMs (Burke et al. 1983). In turn, these putative volatile markers can be easily transferred to different sensor technologies like eNOSE, facilitating their use (Wilson 2015). This approach is noninvasive, highly sensitive and comprehensive and could be readily applied to the clinical environment to screen high-risk population within short span of time without requiring trained medical staff with well-equipped hospitals. In the past, volatomic analysis has been applied toward diagnosis of several cancers, including lung (Liu et al. 2013; Phillips et al. 1999), breast (Silva et al. 2012), gastric (Chen et al. 2010), prostate (Roberts et al. 2011), thyroid (Guo et al. 2015) and colorectal (Arasaradnam et al. 2014).

To the best of our knowledge, here, we report for the first time, the characterization of the urinary VOMs from HNC patients and healthy controls in a cohort of 60 samples using headspace solid phase microextraction (HS-SPME) in combination with gas chromatography mass spectrometry (GC–MS). As a result, we identified a set of four VOMs viz. 2,6-dimethyl-7-octen-2-ol, 1-butanol, *p*-xylene and 4-methyl-2-heptanone with promising discrimination ability for HNC diagnosis.

2 Materials and methods

2.1 Subject selection

HNC urine samples were collected from the Malignant Disease Treatment Centre (MDTC), Unit of the Military Hospital-Cardio Thoracic Centre (MH-CTC), Armed Forces Medical College (AFMC), Pune, India. The samples from healthy controls were obtained through the health check-up camp organized by the hospital. The sample cohort for this study comprised of patients with HNC (n = 29) as malignant group while the age and gender matched healthy volunteers (n = 31) served as the control group. The inclusion criteria for this study comprised of minimum 18 years old subjects with histological confirmation of HNC without anticancer therapeutic interventions. Age and gender matched healthy individuals devoid of hypertension, diabetes and not having any medication during last 3 months were selected as healthy controls. Moreover, healthy controls and HNC patients were matched for smoking habit.

2.2 Sample collection

Early morning first urine samples from each subject were collected in sterile 50 mL falcon tubes after maintaining the subjects in overnight fasting condition. Samples were labelled, centrifuged at 5000 g for 10 min at 4 °C, filtered through 0.45 μm filters and stored at −80 °C until further analysis within 2 h of collection. This study was approved by the institutional ethics committee of AFMC and National Centre for Cell Science (NCCS). All the participants in this study were informed about the investigation and prior informed consent was obtained. The demographic and

clinical information of the subjects employed in this study is summarized in Table S1.

2.3 Sample processing

The VOMs from the urine samples obtained from healthy controls and HNC patients were extracted using the HS-SPME sampling approach as described previously (Silva et al. 2011). Briefly, 4 mL of urine sample was transferred into 8 mL headspace glass vial (Thermo Fisher, USA), acidified with 0.5 mL of 5 M hydrochloric acid (Merck, Germany) and added 0.8 g of sodium chloride (Merck, Germany). The processed sample was then closed with a Teflon (PTFE) septum using a screw cap after adding a magnetic stirring bar (0.5×0.1 mm) to the vial. The sample vial was then incubated at 50 °C with continuous stirring at 800 rpm. The carboxen/polydimethylsiloxane (CAR/PDMS) SPME fibre 75 μm (Supelco, USA) was immediately exposed into the headspace of the vial for 1 h in order to allow efficient adsorption of the VOMs. After incubation, the SPME fibre was retracted back into the safety needle and inserted into the inlet port of the GC–MS (250 °C) system for 6 min for thermal desorption of the VOMs.

2.4 Volatomic analysis and data preprocessing

The headspace VOMs were analysed using an Agilent 7890B gas chromatograph (Palo Alto, USA) equipped with BP-20 (SGE, Germany) fused silica capillary column (60 m $\times 0.25$ mm $\times 0.25$ μm) and interfaced with Agilent 5977A quadrupole inert mass selective detector. The oven temperature was programmed as follows: 45 °C held for 5 min, then increased at 2 °C min^{-1} upto 150 °C followed by 10 min hold, then again ramped at 15 °C min^{-1} up to 220 °C and finally held for 15 min with the total GC run time of 87 min. Helium with a purity of 99.999% (Prama Enterprises, India) was used as the carrier gas at 1 mL min^{-1} flow rate. The injections were performed in splitless mode with the inlet port temperature at 250 °C. The transfer line, quadrupole and ionization source temperatures were maintained at 250, 150 and 230 °C respectively. Data acquisitions were carried out in full scan mode in the range of 30 to 300 m/z and the electron impact mass spectra were recorded at a 70 eV. Each sample was analysed in duplicate. Metabolite identification was achieved by manual inspection of chromatograms and spectra by Agilent ChemStation data analysis software (Palo Alto, USA) linked with the NIST11 mass spectral library, using match score of ≥75%. Integration of the chromatograms was also performed by the ChemStation data analysis software. VOMs showing missing values >50% were removed from the analysis.

2.5 Statistical analysis

Peak area of the all identified VOMs of HNC and healthy control groups were subjected to statistical analysis by MetaboAnalyst 3.0 web based tool (Xia et al. 2015). Features with >50% missing value were removed while features with <50% missing values were replaced by half of minimum positive value in the original data. Two control samples were detected as outliers and subsequently removed from the final analysis. The data was median normalised, transformed by cubic root and scaled by auto scaling approaches. The normalised data was further subjected to univariate and multivariate statistical analysis. The univariate statistical analysis such as Wilcoxon rank-sum test ($p < 0.05$) and log$_2$ fold change ($\geq 0.58 / \leq -0.57$) was performed to evaluate the significant differences among urinary VOMs of healthy controls and HNC patients. Positive log$_2$ fold change is treated as upregulated metabolite expression while negative log$_2$ fold change reflects its downregulation. The multivariate supervised statistical classification models, namely partial least squares-discriminant analysis (PLS-DA) and random forest (RF) classification were used to visualise the distribution of variables among HNC and control groups and to identify important VOMs that discriminate between the two groups. PLSDA variable important in projection (VIP) scores allowed us to rank each of the individual features according to its ability to segregate different classes of experimental data. The VIP score is a weighted sum of square of the PLS loadings. The weights correspond to the percentage variation explained by the PLS-DA component in the model. The features with VIP >1 were considered to be important for discrimination (Xia and Wishart 2011). Hierarchical cluster analysis (HCA) was performed using metabolite profiles of healthy and HNC urine samples to identify inherent clustering patterns. Pearson distance analysis is used as similarity measure. It shows the linear relationship between two quantifiable, continuous variables as well as it is a measure of strength of the association between the two variables (Xia and Wishart 2011). To further assess the predictive value of identified urinary VOMs, we performed receiver operator characteristic (ROC) curve analysis to identify potential diagnostic metabolites that can discriminate HNC patients from healthy controls with high specificity and sensitivity. Also, the metabolic profiles of HNC and healthy samples were utilized to get a holistic view of the dysregulated pathways in HNC by using MetPA tool in Metaboanalyst 3.0 web application (Xia et al. 2011). All the VOMs discussed here are identified as per the guidelines laid down by metabolomics standard initiative (Fiehn et al. 2007). Metadata compliant with metabolomics standard initiative (MSI) for the VOMs identified is reported in supporting information 2 (Griffin et al. 2007).

3 Results

## 3.1 Identification of the altered urinary VOMs by GC–MS

To identify the urinary volatomic profile of HNC patients and the healthy volunteers, 29 samples of HNC patients and 31 samples of healthy control group were extracted using HS-SPME and further analysed by GC–MS. Through GC–MS analysis of three chromatograms from each group, 110 VOMs (>75% match score) were identified through the NIST11 mass spectral library and were further used in the data integration method for the remaining chromatograms (85% occurrence). Furthermore, only the VOMs having ≥85% occurrence of frequency among all the samples were considered for following statistical treatment. The representative chromatograms of HNC and healthy control are shown in Fig. S1. The VOMs identified belongs to different chemical families, such as sulphur and nitrogen containing compounds, furanic compounds, ketones, alcohols, benzene derivatives, alkanes, terpenic compounds, organic acids, aldehydes, phenolic compounds etc. Representative chromatograms showing differential regulation of some of the VOMs are depicted in Fig. 1a, b.

3.2 Statistical analysis of the VOMs identified in HNC patients and healthy control group

In total, out of 110, 101 VOMs were retained after the missing value estimation and were considered for further statistical analysis. To reduce the complexity of data and make features more comparable, data normalisation was carried out with a combination of row wise normalisation methods which includes normalisation to sample median, data transformation by cubic root and data scaling by autoscaling (Fig. S2). The normalised data obtained was further subjected to univariate statistical analysis tools such as $\log_2$ fold change (≥0.58/≤−0.57) and Wilcoxon rank-sum test ($p < 0.05$). The aim of fold change analysis is to compare the mean value changes between HNC and healthy control group and identify the metabolites showing differential regulation. $\log_2$ fold change analysis revealed that 19 VOMs were up regulated while 24 VOMs were downregulated (Table S2). The highest fold change was observed for 5-methyl-3-hexanone. Wilcoxon rank-sum test was carried out to evaluate significant statistical differences among concentrations of urinary VOMs of HNC and healthy control groups and identified 2,6-dimethyl-7-octen-2-ol as the most significantly altered metabolite in HNC patients. Important features identified by both $\log_2$ fold change analysis and Wilcoxon rank-sum test are summarised in supplementary Table S2.

To decide if there was an altered VOM signature in the urine of the HNC patients and healthy controls studied, we first examined the separation of study population based on

Fig. 1 Representative merged chromatograms of HNC patients and healthy controls depicting **a** upregulation of 4-heptanone, and **b** downregulation of 1,6-dimethyl-4-(1-methylethyl)-napthalene
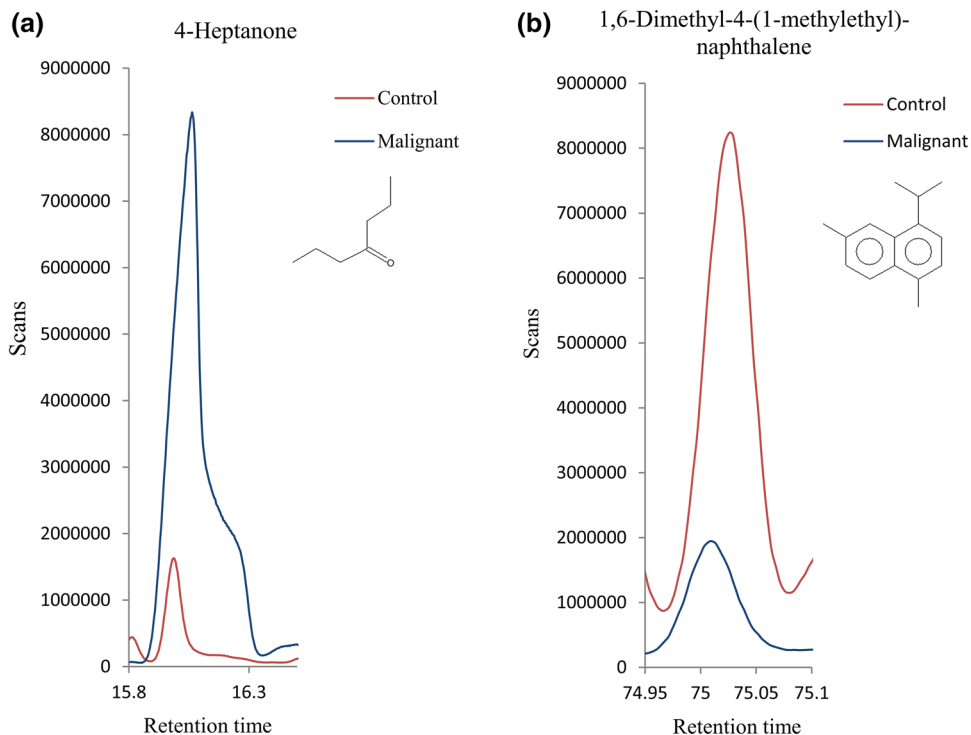
urinary VOMs using supervised partial least square discriminant analysis (PLS-DA) models (Boccard and Rudaz 2013). Decent group separation was achieved between HNC patients and healthy controls, suggesting the presence of characteristic metabolic alterations among each group (Fig. 2a). The permutation test (100 no.) based on prediction accuracy was carried out for PLS-DA model validation. The *p* value based on permutation obtained was *p < 0.01* (0/100) which clearly indicate that data is not over fitted (Fig. 2b). The model was further cross validated using tenfold cross validation method (max components = 5; performance indicator = Q2) (Table S4a). The top 28 differentially expressed metabolites were selected with variable importance in projection (VIP) score of >1, $p < 0.05$, $\log_2$ FC $\geq 0.58/\leq -0.57$ and regarded as critical for the discrimination among study groups in the score plot. The statistically significant VOMs identified by both univariate ($p < 0.05$ and $\log_2$ FC $\geq 0.58/\leq -0.57$) and multivariate (VIP >1) analyses are summarised in Table 1. Additionally, random forest (RF) analysis was carried out to determine the ability of urinary VOMs to accurately classify the subjects into their respective groups, as well as identify the metabolites that have strong correlation with the disease, and hence were critical for class prediction. RF analysis classified HNC patients

and healthy control group with accuracy of 100 and 90% respectively, with out of bag (OOB) classification error rate of 0.0517 (Fig. S3a). Major contributory metabolites for classification are shown in Fig. S3b. RF classification performance is presented in Table S4b. HCA was performed using metabolite profiles of both study groups to identify the inherent clustering patterns. The heat map of metabolites ($p < 0.01$) was generated by average algorithm and Pearson distance analysis, which clearly depict distinct clustering patterns among the studied groups (Fig. 2c).

3.3 Receiver operator characteristic (ROC) curve analysis

To further characterize the predictive value of the identified metabolites to discriminate HNC patients from healthy controls, we carried out ROC curve analysis. The ROC curve analysis is used for the classification of true positives and false positives (Obuchowski et al. 2005).The predictive utility of ROC curve is measured using area under the curve (AUC). The AUC for a metabolite equal to 0.5 indicate random prediction and <0.5 is worse than random prediction and therefore less significant to distinguish two sample groups. In turn, a value close to 1 is considered as a perfect
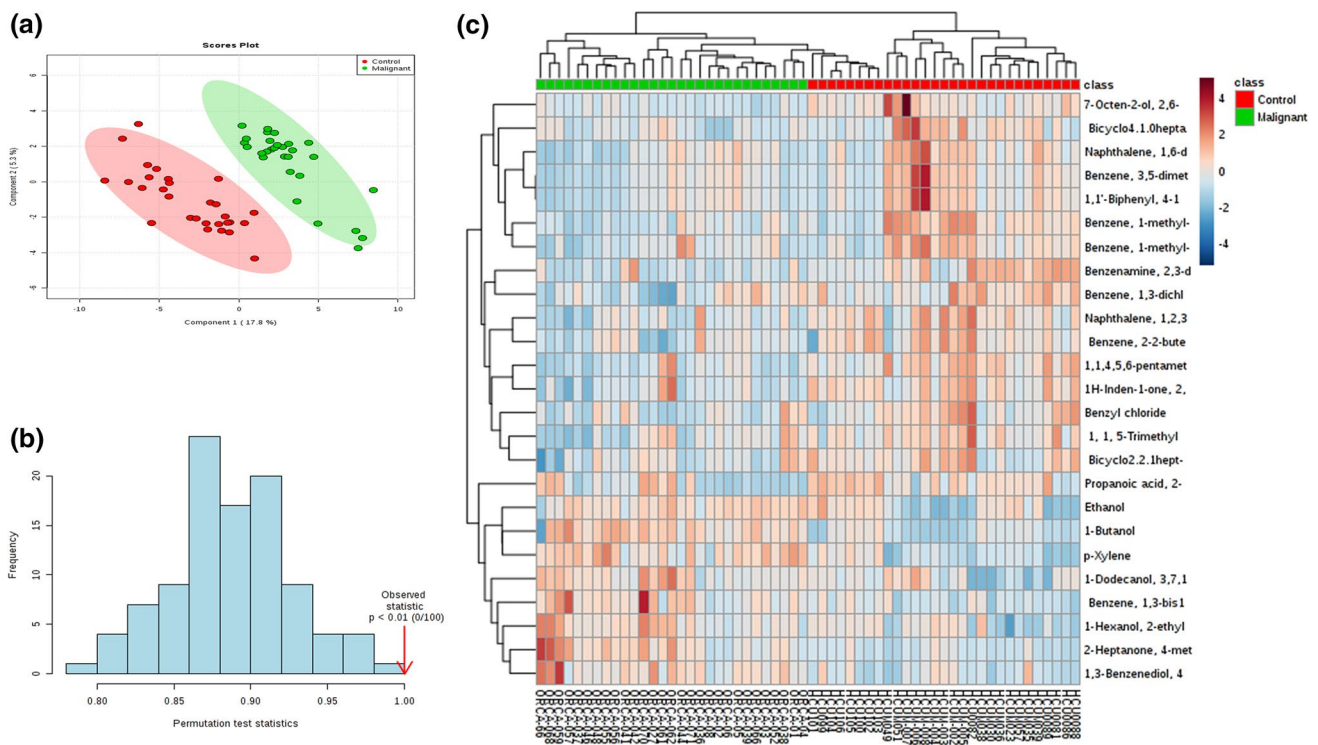
Fig. 2 Multivariate statistical analysis of urinary VOMs of HNC patients and healthy controls. **a** Partial least square discrimination analysis of urinary VOMs among malignant and healthy groups, **b** PLS-DA model validation by permutation test based on prediction accuracy, **c** hierarchical cluster analysis performed using metabolite profiles of healthy and malignant urine samples. The heat map was generated using 25 most significant metabolites (p value < 0.01) by Average algorithm and Pearson distance analysis

Table 1 Important discriminatory metabolites identified by statistical analysis

Sr. No.	Compounds	VIP	p value	FDR	FC (log$_2$)
1	1-Butanol	2.1025	2.12E−09	1.07E−07	3.07
2	*p*-Xylene	2.044	8.19E−08	2.32E−06	1.87
3	4-Methyl-2-heptanone	1.8761	9.18E−08	2.32E−06	3.31
4	1,1,4,5,6-Pentamethyl-2,3-dihydro-1H-indene	1.8502	2.22E−07	4.48E−06	−1.67
5	2,3-Dichloro-benzenamine	1.7188	8.58E−06	8.66E−05	−2.37
6	1,3-Bis-(1,1-dimethylethyl)-benzene	1.6602	5.54E−06	7.00E−05	1.37
7	1-Methyl-4-2-methylpropyl-benzene	1.6537	8.89E−05	0.000748	−1.78
8	1,2,3,4-Tetrahydro-1,1,6 trimethylnaphthalene	1.6533	7.21E−06	8.10E−05	−1.58
9	4-Ethyl1,3-benzenediol	1.6533	5.07E−06	7.00E−05	1.58
10	2,6-Dimethyl-7-octen-2-ol	1.6524	8.97E−10	9.06E−08	−4.52
11	2,3-Dihydro-3,3,5,6-tetramethyl-1H-inden-1-one	1.6231	1.53E−06	2.57E−05	−1.17
12	2-Ethyl-1-hexanol	1.4303	0.002822	0.010557	1.82
13	2-(2-Butenyl)-1,3,5-trimethylbenzene	1.426	7.43E−05	0.000682	−1.95
14	1,6-Dimethyl-4-(1-methylethyl) naphthalene	1.4026	0.00095	0.005049	−1.31
15	3,5-Dimethyl-1-(Phenylmethyl) benzene	1.3797	0.000421	0.003025	−1.73
16	4-(1-Methylethyl)-1,1′-biphenyl	1.3794	0.000449	0.003025	−1.73
17	3,7,11-Trimethyl-1-dodecanol	1.3348	0.004547	0.014351	1.13
18	1-Methyl-4-(1-methylethenyl)-benzene	1.2916	0.000894	0.005017	−1.20
19	1,7,7-Trimethyl-bicyclo[2.2.1]hepten-2-ene	1.2683	0.002143	0.009408	−1.06
20	Benzyl chloride	1.2645	0.001915	0.009209	−1.24
21	2-Methyl-5-(methylthio)furan	1.1811	0.008277	0.021999	−1.22
22	2-Pentanone	1.1599	0.002673	0.010557	1.68
23	Trans-calamenene	1.1424	0.004098	0.013352	−1.08
24	3-Hepten-2-one	1.126	0.005577	0.016095	1.01
25	1-Decen-3-one	1.1164	0.002822	0.010557	1.06
26	Dodecane	1.0679	0.000894	0.005017	1.63
27	2-Heptanone	1.0372	0.016538	0.037961	3.14
28	1-Ethenyl-4-(2-methylpropyl)-benzene	1.0015	0.015133	0.035545	5.10

VIP variable importance in projection, *FDR* false discovery rate, *FC* fold change

discrimination test where all positives are true positives and all negatives are true negatives (Broadhurst and Kell 2006). The top four metabolites with highest discriminative ability of HNC against the healthy controls includes 2,6-dimethyl-7-octen-2-ol (AUC = 0.925), 1-butanol (AUC = 0.920), 4-methyl-2-heptanone (AUC = 0.885) and *p*-xylene (AUC = 0.885) (Fig. 3). The ROC curve analysis, including specificity, sensitivity and respective fold change between HNC and healthy control group is summarized in Table 2.

3.4 Metabolic pathway analysis

Metabolic pathway analysis of the identified urinary VOMs revealed several metabolic pathways that are significantly altered in HNC patients (Fig. 4). The important metabolic pathways that are excessively active in HNC were butanoate metabolism, glycolysis or gluconeogenesis, propanoate metabolism, pyruvate metabolism, sulfur metabolism, taurine and hypotaurine metabolism, selenoamino acid metabolism and synthesis and degradation of ketone bodies. The pathways showing suppressed activity in HNC were related with cysteine and methionine metabolism (Table S3).

4 Discussion

Otto Warburg showed in 1920 that cancer cells are highly glycolytic compared to normal cells, a condition which is known as the Warburg effect (Warburg et al. 1927; Potter et al. 2016). Since then, it has become more evident that malignant cells differ significantly in their metabolism from their normal counterpart. The altered metabolism is essential for survival and to sustain the proliferative nature of cancer cells by excessively diverting glycolysis intermediates to replenish biosynthesis of nucleotides, proteins and fatty acids, essential to maintain malignant characteristics (Lunt and Vander Heiden 2011). Specifically, capturing these altered metabolic profiles could help us to identify clinically relevant biosignatures that may be helpful in noninvasive
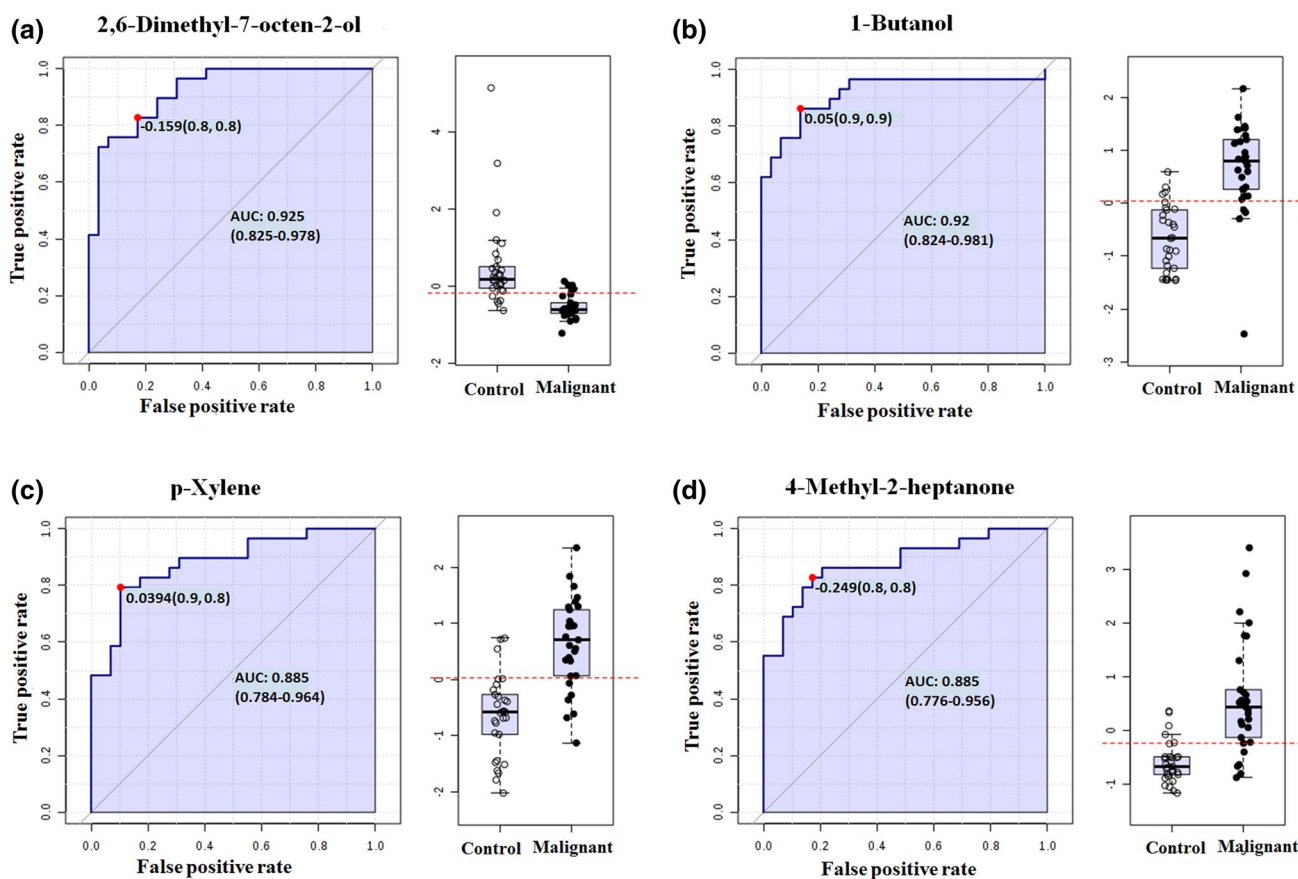
Fig. 3 Receiver operator characteristic curve of **a** 2,6-dimethyl-7-octen-2-ol, **b** 1-butanol, **c** 4-methyl-2-heptanone and **d** *p*-xylene

Table 2 The AUC values including sensitivity, specificity, p value and fold change of top four differentially expressed metabolites with highest discriminative ability of HNC against the healthy controls obtained by ROC curve analysis

Compound	AUC	Sensitivity	Specificity	p value	FDR	FC ($\log_2$)
2,6-Dimethyl-7-octen-2-ol	0.92	0.8	0.8	8.97E−10	9.06E−08	−4.52
1-Butanol	0.92	0.9	0.9	2.12E−09	1.07E−07	3.07
p-Xylene	0.88	0.9	0.8	8.19E−08	2.32E−06	1.87
4-Methyl-2-heptanone	0.88	0.8	0.8	9.18E−08	2.32E−06	3.31

AUC area under curve, *FDR* false discovery rate, *FC* fold change

detection of cancer, improve diagnosis and prognosis as well as aid in finding better drug targets for enhanced therapeutic outcome. Different tools, techniques and biospecimens such as saliva, blood, urine, cells and tissue were routinely employed by several investigators to capture metabolomic profiles of cancer, including HNC. In the present study, we demonstrated for the first time that urinary volatomic signature of HNC exists and can be detected by HS-SPME coupled to GC–MS analysis. The identified metabolites were grouped under various chemical families based on their chemical nature and could be of microbial, endogenous or exogenous in origin. Organic acids and alcohols are produced mainly due to carbohydrate degradation by anaerobic bacteria in the intestine, while alkanes are derived from lipid peroxidation of polyunsaturated fatty acids found in the cell membrane (Cummings 1981; Haick et al. 2014). In turn, exogenous metabolites can be accumulated through feeding, air inhalation or skin absorption (Ulanowska et al. 2010).

Univariate statistical analysis revealed that ketonic metabolite 3-Hepten-2-one is significantly elevated in HNC patient's urine with 1.01 $\log_2$ fold change. Furthermore, multivariate statistical analysis confirm that other ketonic metabolites such as 2-pentanone, 2-heptanone, and its methylated derivative 4-methyl-2-heptanone are significantly
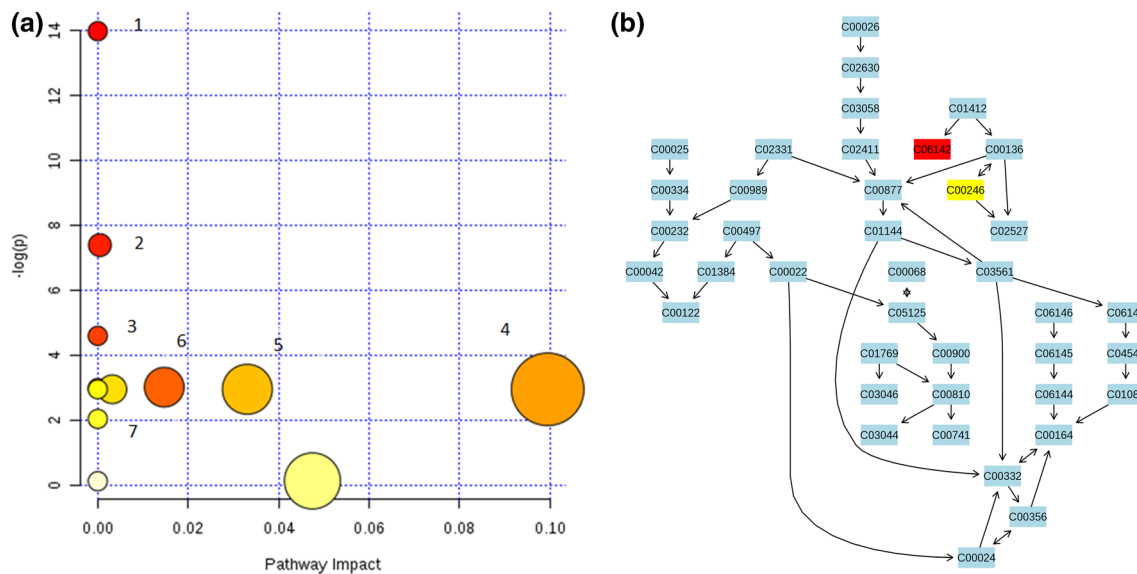
Fig. 4 **a** Altered metabolic pathways among HNC patients as compared to healthy controls. *1* Butanoate pathway, *2* glycolysis or gluconeogenesis, *3* cysteine and methionine metabolism, *4* propanoate metabolism, *5* nicotinate and nicotinamide metabolism, *6* pyruvate metabolism, *7* synthesis and degradation of ketone bodies, **b** the butanoate metabolism pathway. The map was generated using reference map by KEGG. C06142 represent 1-butanol, C00246 represents butyric acid

upregulated in HNC patients as compared to healthy controls and emerged as major discriminatory metabolites between the two study groups. Additionally, pathway analysis also indicated that butanoate metabolism and ketone body metabolism are excessively active in the HNC patients. Interestingly, ketone bodies, a source of high energy fuel that produce more energy compared to other mitochondrial fuel is also preferentially utilized by cancerous cells under the hypoxic condition which allows tumor to grow without proper oxygen supply (Martinez-Outschoorn et al. 2012). Therefore, presence of higher concentration of ketones in HNC patient's urine indicate an important fact that ketone bodies are being excessively produced in the HNC patients as compared to their healthy counterparts. Also, this could be attributed to the fact that food intake in the HNC patients is severely affected and major needs of energy are compensated through oxidation of fatty acids and tissue protein catabolism, which result in weight loss and production of ketone bodies.

It's a well-known fact that cancer cells are highly glycolytic and generate most of the energy through aerobic glycolysis (Warburg 1956). Pyruvic acid, an end product of the glycolytic cycle is subsequently reduced to lactic acid instead utilized in oxidative phosphorylation and may further get converted to acetic acid. Furthermore, acetyl CoA and pyruvic acid, both can directly convert into acetic acid resulting into increased accumulation which reflects upregulated glycolysis and pyruvate metabolism in HNC patients. Interestingly, it should be noticed that chronic alcohol abuse among HNC patients may also lead to the production of

elevated levels of acetic acid by ethanol catabolism catalyzed by alcohol dehydrogense enzyme (Schug et al. 2016).

Dodecane, an acyclic alkane was found in the headspace of the urine samples of both the study groups. The concentration of dodecane was higher in HNC patients in comparison with healthy controls. Alkanes are primarily generated by peroxidation of polyunsaturated fatty acids (PUFA) known as reactive oxygen species (ROS) mediated lipid peroxidation (Frank Kneepkens et al. 1994). The altered redox status and the generation of ROS are common biochemical events that occur in malignant cells (Barrera 2012; Panieri and Santoro 2016). Therefore, detection of higher amounts of lipid peroxidation products such as dodecane in HNC patient's urine was not surprising. Other hydrocarbons, mainly napthalene derivatives are suppressed in HNC patients as compared to healthy volunteers and thought to be derived from degradation of steroids but little is known about their origin and role in cell metabolism.

Alcohols such as 1-butanol, 2-ethyl-1-hexanol and 3,7,11-trimethyl-1-dodecanol are significantly upregulated, while 2,6-dimethyl-7-octen-2-ol considerably down regulated in HNC. Specifically, 2-ethyl-1-hexanol was reported at an increased concentration in saliva of lung cancer patients and NSLC cell lines (Sponring et al. 2009). This could be attributed to the fact that cytochrome P450 enzymes are induced during carcinogenesis and subsequently hydroxylate several VOMs, including alkanes such as butane, octane, hexane and dodecane produced during ROS mediated lipid peroxidation of PUFA and leads to the production of corresponding alcohol (Hakim et al. 2012). Moreover, these

alcohols are converted either into the respective aldehydes by various enzymes, such as alcohol dehydrogenase and cytochrome P450 (CYP2E1) or excreted through breath, urine, saliva, sweat, feces etc.

Hypotaurine is involved in the protection against oxidative stress and ROS induced membrane damage (Brand et al. 1998). In HNC patients, taurine and hypotaurine metabolism is more active, which indicates increased levels of ROS and oxidative stress, a common signature of malignant cells and detected in almost all cancers. The taurine and hypotaurine metabolism is found to be altered in various malignancies including breast, ovarian, colon, lung and renal cancer (Gossai and Lau-Cam 2009; Huang et al. 2016; Pradhan et al. 2013; Roy et al. 2014; Tiruppathi et al. 1992). Cysteine and methionine metabolism is down regulated in the HNC patients and characterized by the presence of methanethiol in reduced quantity as compared to healthy controls. Methanethiol is generated by incomplete metabolism of methionine by transamination pathway and is found significantly down regulated in malignant cells (Silva et al. 2011).

Several benzene derivatives were found at varying concentrations in the urine of the both the study groups. For most of the benzene derivatives, their origin and role in metabolism is not clear but assumed as exogenous and derived from various sources such as cigarette smoke, tobacco chewing, alcohol consumption, pollution and exposure to radiation. Since these molecules are highly hydrophobic in nature, they easily pass through membrane; enter the cytoplasm and causes chronic peroxidative damage to the macromolecules which ultimately lead to the development of diseases such as cancer (Halliwell et al. 1992). These absorbed compounds are then functionalized by cytochrome P450 enzyme and further conjugated to a more soluble and excretable form by other enzymes such as glutathione S-transferase, sulfotransferase and N-acetyltransferase (Guengerich et al. 1991) and ultimately removed from the body. Overall, the results obtained in this study are noteworthy and provides valuable insights into the volatile alterations of HNC subjects. However, we believe that further studies in a large cohort of patients are required in order to validate and confirm the utility of the identified urinary VOCs as diagnostic markers.

5 Conclusions

To the best of our knowledge, we have carried out, for the first time, the urinary volatomic profiling of HNC in a cohort of 29 patients along with 31 healthy controls, using HS-SPME and GC–MS approach and identified 101 VOMs. The univariate and multivariate statistical analysis revealed 28 metabolites significantly altered in HNC. From the panel of the metabolites detected,

2,6-dimethyl-7-octen-2-ol, 1-butanol, *p*-xylene and 4-methyl-2-heptanone have emerged as potential volatile signature with highest sensitivity and specificity to discriminate HNC patients from healthy controls. The metabolic pathway analysis indicated that the discriminatory VOMs could be originated from several dysregulated pathways in HNC such as butanoate metabolism and ketone body synthesis and degradation. This further paves a way to utilize these pathways as targets for the development of novel therapeutic interventions.

Author contributions   Conceived the study: JSC, HAN, SR; Designed the study: RT, KT, JSC, HAN, SR; Performed the experiments: RT, KT; Compiled and analyzed data: RT, KT, JAMP, RD, SR; Statistical analysis: RT, KT, RD, HAN, SR; Drafted the manuscript: RT, KT, JAMP, RD, NK, DS, JSC, HAN, SR; Provided clinical samples: NK, DS; Provided chemicals and reagents: SR.

Compliance with ethical standards

References

Amann, A., Costello, B. D. L., Miekisch, W., Schubert, J., Buszewski, B., Pleil, J., et al. (2014). The human volatilome: Volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *Journal of Breath Research, 8*(3), 34001.

Arasaradnam, R. P., McFarlane, M. J., Ryan-Fisher, C., Westenbrink, E., Hodges, P., Thomas, M. G., et al. (2014). Detection of colorectal cancer (CRC) by urinary volatile organic compound analysis. *PloS ONE, 9*(9), e108750.

Barrera, G. (2012). Oxidative stress and lipid peroxidation products in cancer progression and therapy. *ISRN Oncology, 2012*(137289), 21.

Boccard, J., & Rudaz, S. (2013). Mass Spectrometry Metabolomic Data Handling for Biomarker Discovery. In H. J. Issaq & T. D. Veenstra (Eds.), *Proteomic and metabolomic approaches to biomarker discovery* (pp. 425–445). Boston: Academic Press.

Brand, A., Leibfritz, D., Hamprecht, B., & Dringen, R. (1998). Metabolism of cysteine in astroglial cells: Synthesis of hypotaurine and taurine. *Journal of Neurochemistry, 71*(2), 827–832.

Broadhurst, D. I., & &Kell, D. B. (2006).Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics*, 2(4), 171–196.

Burke, D. G., Halpern, B., Malegan, D., McCairns, E., Danks, D., Schlesinger, P., & Wilken, B. (1983). Profiles of urinary volatiles from metabolic disorders characterized by unusual odors. *Clinical Chemistry, 29*(10), 1834–1838.

Chen, J. L., Tang, H. Q., Hu, J. D., Fan, J., Hong, J., & Gu, J. Z. (2010). Metabolomics of gastric cancer metastasis detected by gas chromatography and mass spectrometry. *World Journal of Gastroenterology: WJG, 16*(46), 5874–5880.

Cummings, J. H. (1981). Short chain fatty acids in the human colon. *Gut*, 22(9), 763–779.

Downes, M. R., Byrne, J. C., Pennington, S. R., Dunn, M. J., Fitzpatrick, J. M., & Watson, R. W. G. (2007). Urinary markers for prostate cancer. *BJU International, 99*(2), 263–268.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2014). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer, 136*(5), E359–E386

Fiehn, O., Robertson, D., Griffin, J. L., Van der Werf, M., Nikolau, B., Morrison, N., et al. (2007). The metabolomics standard initiative (MSI). *Metabolomics, 3*(3), 175–178.

Frank Kneepkens, C. M., Lepage, G., & Roy, C. C. (1994). The potential of the hydrocarbon breath test as a measure of lipid peroxidation. *Free Radical Biology and Medicine, 17*(2), 127–160.

Gossai, D., & Lau-Cam, C. A. (2009). The effects of taurine, hypotaurine, and taurine homologs on erythrocyte morphology, membrane fluidity and cytoskeletal spectrin alterations due to diabetes, alcoholism and diabetes-alcoholism in the rat. *Advances in Experimental Medicine and Biology, 643*, 369–379.

Griffin, J. L., Nicholls, A. W., Daykin, C. A., Heald, S., Keun, H. C., Schuppe-Koistinen, I., et al. (2007). Standard reporting requirements for biological samples in metabolomics experiments: Mammalian/in vivo experiments. *Metabolomics, 3*(3), 179–188.

Guengerich, F. P., Kim, D. H., & Iwasaki, M. (1991). Role of human cytochrome P-450 IIE1 in the oxidation of many low molecular weight cancer suspects. *Chemical Research in Toxicology, 4*(2), 168–179.

Guo, L., Wang, C., Chi, C., Wang, X., Liu, S., Zhao, W., et al. (2015). Exhaled breath volatile biomarker analysis for thyroid cancer. *Translational Research, 166*(2), 188–195.

Haick, H., Broza, Y. Y., Mochalski, P., Ruzsanyi, V., & Amann, A. (2014). Assessment, origin, and implementation of breath volatile cancer markers. *Chemical Society Reviews, 43*(5), 1423–1449.

Hakim, M., Broza, Y. Y., Barash, O., Peled, N., Phillips, M., Amann, A., & Haick, H. (2012). Volatile organic compounds of lung cancer and possible biochemical pathways. *Chemical Reviews, 112*(11), 5949–5966.

Halliwell, B., Gutteridge, J. M., & Cross, C. E. (1992). Free radicals, antioxidants, and human disease: Where are we now? *The Journal of Laboratory and Clinical Medicine, 119*(6), 598–620.

Huang, S., Chong, N., Lewis, N. E., Jia, W., Xie, G., & Garmire, L. X. (2016). Novel personalized pathway-based metabolomics models reveal key metabolic pathways for breast cancer diagnosis. *Genome Medicine, 8*(1), 34.

Inal, E., Lacin, M., Asal, K., Ceylan, A., Koybasioglu, A., Ileri, F., & Uslu, S. S. (2004). The significance of ferritin, lipid-associated sialic acid, CEA, squamous cell carcinoma (SCC) antigen, and CYFRA 21 – 1 levels in SCC of the head and neck. *Kulak Burun Bogaz Ihtis Derg, 12*(1–2), 23–30.

Liu, H., Wang, H., Li, C., Wang, L., Pan, Z., & Wang, L. (2013). Investigation of volatile organic metabolites in lung cancer pleural effusions by solid-phase microextraction and gas chromatography/mass spectrometry. *Journal of Chromatography B, Analytical Technologies in the Biomedical and Life Sciences, 945–946*, 53–59.

Lunt, S. Y., & Vander Heiden, M. G. (2011). Aerobic glycolysis: Meeting the metabolic requirements of cell proliferation. *Annual Review of Cell and Developmental Biology, 27*, 441–464.

Mallath, M. K., Taylor, D. G., Badwe, R. A., Rath, G. K., Shanta, V., Pramesh, C. S., et al. (2014). The growing burden of cancer in India: Epidemiology and social context. *The Lancet Oncology, 15*(6), e205–e212.

Martinez-Outschoorn, U. E., Lin, Z., Whitaker-Menezes, D., Howell, A., Lisanti, M. P., & Sotgia, F. (2012). Ketone bodies and two-compartment tumor metabolism: Stromal ketone production fuels mitochondrial biogenesis in epithelial cancer cells. *Cell Cycle, 11*(21), 3956–3963.

Mishra, A., & Meherotra, R. (2014). Head and neck cancer: Global burden and regional trends in India. *Asian Pacific Journal of Cancer Prevention, 15*(2), 537–550.

Obuchowski, N. A., Blackmore, C. C., Karlik, S., & Reinhold, C. (2005). Fundamentals of clinical research for radiologists. *American Journal of Roentgenology, 184*(2), 364–372.

Panieri, E., & Santoro, M. (2016). ROS homeostasis and metabolism: A dangerous liason in cancer cells. *Cell Death and Disease, 7*(6), e2253.

Phillips, M., Gleeson, K., Hughes, J. M. B., Greenberg, J., Cataneo, R. N., Baker, L., & McVay, W. P. (1999). Volatile organic compounds in breath as markers of lung cancer: A cross-sectional study. *Lancet, 353*(9168), 1930–1933.

Potter, M., Newport, E., & Morten, K. J. (2016). The Warburg effect: 80 years on. *Biochemical Society Transactions, 44*(5), 1499–1505.

Pradhan, M. P., Desai, A., & Palakal, M. J. (2013).Systems biology approach to stage-wise characterization of epigenetic genes in lung adenocarcinoma. *BMC Systems Biology, 7*, 141.

Rezende, T. M. B., Freire, M. d. S., & Franco, O. L. (2010). Head and neck cancer. *Cancer, 116*(21), 4914–4925.

Roberts, M. J., Schirra, H. J., Lavin, M. F., & Gardiner, R. A. (2011). Metabolomics: A novel approach to early and noninvasive prostate cancer detection. *Korean Journal of Urology, 52*(2), 79–89.

Roy, D., Mondal, S., Wang, C., He, X., Khurana, A., Giri, S., et al. (2014). Loss of HSulf-1 promotes altered lipid metabolism in ovarian cancer. *Cancer & Metabolism, 2*, 13.

Schug, Z. T., Voorde, J., Vande, & Gottlieb, E. (2016). The metabolic fate of acetate in cancer. *Nature Reviews Cancer, 16*(11), 708–717.

Shen, C., Sun, Z., Chen, D., Su, X., Jiang, J., Li, G., et al. (2015). Developing urinary metabolomic signatures as early bladder cancer diagnostic markers. *Omics: A Journal of Integrative Biology, 19*(1), 1–11.

Silva, C. L., Passos, M., & &Câmara, J. S. (2011). Investigation of urinary volatile organic metabolites as potential cancer biomarkers by solid-phase microextraction in combination with gas chromatography-mass spectrometry. *British Journal of Cancer, 105*(12), 1894–1904.

Silva, C. L., Passos, M., & &Câmara, J. S. (2012). Solid phase microextraction, mass spectrometry and metabolomic approaches for detection of potential urinary cancer biomarkers - A powerful strategy for breast cancer diagnosis. *Talanta, 89*, 360–368.

Sponring, A., Filipiak, W., Mikoviny, T., Ager, C., Schubert, J., Miekisch, W., et al. (2009). Release of volatile organic compounds

from the lung cancer cell line NCI-H2087 in vitro. *Anticancer Research, 29*(1), 419–426.

Stewart, B. W., & Wild, C. P. (2014). *World Cancer Report 2014.* Geneva: World Health Organization.

Tamás, L., Szentkúti, G., Eros, M., Dános, K., Brauswetter, D., Szende, B., et al. (2011). Differential biomarker expression in head and neck cancer correlates with anatomical localization. *Pathology and Oncology Research, 17*(3), 721–727.

Tiruppathi, C., Brandsch, M., Miyamoto, Y., Ganapathy, V., & Leibach, F. H. (1992). Constitutive expression of the taurine transporter in a human colon carcinoma cell line. *American Journal of Physiology, 263*(5), G625–G631.

Ulanowska, A., Ligor, T., Michel, M., & Buszewski, B. (2010). Hyphenated and unconventional methods for searching volatile cancer biomarkers. *Ecological Chemistry and Engineering S-Chemia I InzynieriaEkologiczna S. 17*, 9–23

Warburg, O. (1956). On the origin of cancer cells. *Science, 123*(3191), 309–314.

Warburg, O., Ding, Y., Sharma, J., Paetz, M. B., Meixensberger, J., Gaunitz, F., et al. (1927). The metabolism of tumors in the body. *The Journal of General Physiology, 8*(6), 519–530.

Wilson, A. D. (2015). Advances in electronic-nose technologies for the detection of volatile biomarker metabolites in the human breath. *Metabolites, 5*(1), 140–163.

Wyss, A., Hashibe, M., Chuang, S. C., Lee, Y. C. A., Zhang, Z. F., Yu, G. P., et al. (2013). Cigarette, cigar, and pipe smoking and the risk of head and neck cancers: Pooled analysis in the international head and neck cancer epidemiology consortium. *American Journal of Epidemiology, 178*(5), 679–690.

Xia, J., Sinelnikov, I. V., Han, B., & Wishart, D. S. (2015). MetaboAnalyst 3.0-making metabolomics more meaningful. *Nucleic Acids Research, 43*(W1), W251–W257.

Xia, J., & Wishart, D. S. (2011). Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nature Protocols, 6*(6), 743–760.

Xia, J., Wishart, D. S., & Valencia, A. (2011). MetPA: A web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics, 27*, 2342–2344.

Yonezawa, K., Nishiumii, S., Kitamoto-Matsuda, J., Fujita, T., Morimoto, K., Yamashita, D., et al. (2013). Serum and tissue metabolomics of head and neck cancer. *Cancer Genomics & Proteomics, 10*, 233–238.