HHS Public Access

Author manuscript

Nanomedicine. Author manuscript; available in PMC 2016 February 08.

Published in final edited form as:

Nanomedicine. 2012 July ; 8(5): 580–589. doi:10.1016/j.nano.2011.10.001.

Classification of lung cancer histology by gold nanoparticle sensors

Orna Barash, MSc[a,1], **Nir Peled, MD, PhD**[b,c,1], **Ulrike Tisch, PhD**[a], **Paul A. Bunn Jr, MD**[b], **Fred R. Hirsch, MD, PhD**[b], and **Hossam Haick, PhD**[a,*]

[a]The Department of Chemical Engineering and Russell Berrie Nanotechnology Institute, Technion–Israel Institute of Technology, Haifa, Israel

[b]University of Colorado Cancer Center, Division of Medical Oncology, University of Colorado–Denver, Aurora, Colorado, USA

[c]The Lung Cancer Unit, Sheba Medical Center, Tel Aviv University, Tel Aviv, Israel

Abstract

We propose a nanomedical device for the classification of lung cancer (LC) histology. The device profiles volatile organic compounds (VOCs) in the headspace of (subtypes of) LC cells, using gold nanoparticle (GNP) sensors that are suitable for detecting LC-specific patterns of VOC profiles, as determined by gas chromatography–mass spectrometry analysis. Analyzing the GNP sensing signals by support vector machine allowed significant discrimination between (i) LC and healthy cells; (ii) small cell LC and non–small cell LC; and between (iii) two subtypes of non–small cell LC: adenocarcinoma and squamous cell carcinoma. The discriminative power of the GNP sensors was then linked with the chemical nature and composition of the headspace VOCs of each LC state. These proof-of-concept findings could totally revolutionize LC screening and diagnosis, and might eventually allow early and differential diagnosis of LC subtypes with detectable or unreachable lung nodules.

Keywords

Sensor; Gold nanoparticle; Lung cancer; Histology; Volatile organic compound

[*]Corresponding author: Department of Chemical Engineering and Russell Berrie Nanotechnology Institute, Technion-Israel Institute of Technology, Haifa 32000, Israel. hhossam@technion.ac.il (H. Haick).

[1]These two authors contributed equally to this work.

The topic of this invited contribution has been presented at the 4th IEEE International Conference on Nano/Molecular Medicine and Engineering (IEEE-NANOMED 2010), December 5–9, 2010, Hong Kong SAR, China.

O.B., N.P., U.T., P.B., and H.H. have no conflict to declare related to the study. F.H is a member of the consultant/advisory boards of AstraZeneca, Roche, Lilly, Pfizer, Boehringer-Ingelheim, Merck Serono, Ventana-Roche, Glaxo Smith Kline, BMS/Imclone, and Syndax.

All authors designed the research; O.B. and N.P. performed the research; U.T., O.B., N.P., and H.H. analyzed data, and U.T. and H.H. wrote the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.nano.2011.10.001.

Lung cancer (LC) is the leading cause of cancer-related mortality, with 159,390 deaths in the United States during 2009 and more than 1 million deaths each year globally.[1] Small cell LC (SCLC) represents approximately 15% of lung cancers and is the fifth leading cause of cancer mortality. SCLC is distinguished from non–small cell LC (NSCLC) by its rapid growth rate and rapid metabolism, early dissemination to regional lymph nodes and distant sites, and sensitivity to chemotherapy and radiation therapy. Both types of LC occur in a similar population of heavy smokers but require different disease management. Surgery is performed on many of the NSCLC patients (stages I, II, and some of IIIA), whereas chemotherapy and radiation therapy are the backbone of SCLC treatment, except in the very early stage (IA).[2] For advanced NSCLC, the recommended treatment approach depends on the subhistology. For example, pemetrexed is favored for the treatment of large cell carcinoma and adenocarcinoma; bevacizumab is avoided when treating squamous cell carcinoma.[3] Currently, the diagnostic steps for lung nodule evaluation include an invasive procedure that yields a tissue specimen. However, such procedures are not always available, and their sensitivity might be limited due to small nodule size and difficult access.[4] Therefore, a simple and highly sensitive test would be extremely valuable for the classification and early screening of lung cancer.

Analysis of volatile organic compounds (VOCs) is a new frontier in noninvasive medical diagnostics.[5–18] This approach is based on the fact that the cell membrane consists primarily of amphipathic phospholipids, carbohydrates, and many integral membrane proteins that are distinct in different cell types.[19,20] Tumor growth is accompanied by gene changes and/or protein changes that may lead to oxidative stress and peroxidation of the cell membrane species, thus leading to the emission of VOCs.[21–23] Some of these VOCs appear in distinctively different mixture compositions, depending on whether a cell is healthy or cancerous.[5,6,8,24] These VOCs can be detected either directly from the headspace of the cancer cells or via the exhaled breath. The rationale in the latter case is that cancer-related changes in blood chemistry are reflected in measurable changes in the breath through exchange via the lung.[5–7]

Recently, we have developed an array of gold nanoparticle (GNP) sensors in conjugation with pattern recognition methods for noninvasive discrimination between healthy and LC states via exhaled breath samples[5,6,17] as well as in vitro samples.[8] In this study we have adapted this GNP sensor technology for in vitro differentiation between subtle differences in the VOC profiles of various LC subtypes. Using the GNP sensors we have demonstrated the ability to discriminate between LC and control states, between SCLC and NSCLC, and between subtypes of NSCLC. The chemical nature and composition of the associated VOCs was determined by gas chromatography–mass spectrometry (GC-MS) analysis, allowing the tracking of the metabolic pathways of the cells.

Methods

Collection of the headspace samples

The headspace samples were collected from commercially available cell lines. Fourteen NSCLC cell lines, subcategorized into 10 adenocarcinoma cell lines and 4 squamous cell carcinoma cell lines, as well as four SCLC cell lines (Table 1), were obtained from the

Colorado cell bank registry. The cell lines were grown in 100-mm cell culture dishes from seeding (~2 × 10^6 cells) up to 95% confluency (7 × 10^6 cells), using a two-dimensional medium (medium 1) under standard conditions (RPMI 1640 medium + 10% fetal bovine serum;5%CO_2 environment). Several of the cell lines were grown in two or three replicates, as indicated in Table 1. Seven replicates of an immortal bronchial epithelium (IBE) cell line (Table 1) were grown likewise in 100-mm cell culture dishes from seeding (~2 × 10^6 cells) up to 95% confluency (7 × 10^6 cells), using another two-dimensional medium (medium 2) under standard conditions (BEBM 1640 medium + 10% fetal bovine serum; 5% CO_2 environment). Seven replicates of medium 1 served as a baseline control for the LC cell lines, and five replicates of medium 2 served as a baseline control for the IBE cell lines (same incubation time and conditions, but without the cells). Each cell culture was placed in a 150-mm dish. Two Ultra II SKC badges with Tenax TA as a sorbent (265 mg; SKC Inc. Eighty Four, Pennsylvania) were placed above the dish for absorbing the headspace atmosphere during the total growth time (median time 68 hours; range 60–72 hours), as shown in Figure 1, *A*. Two headspace samples per cell line were obtained for identifying the headspace VOCs. For analyzing the samples, the LC histology– specific VOC patterns were obtained by heating the Tenax sorbent material in stainless-steel thermal desorption chambers (350 mL and 750 mL, respectively) that were preheated to 270°C and kept at that temperature for 10 minutes.

Headspace analysis with GC-MS

The headspace VOCs were identified using GC-MS (GC HP 6890; MS-5973; Agilent Technologies, Santa Clara, California); H5-5MS capillary column (5% phenyl methyl siloxane; 30 m in length, 0.25 mm internal diameter, 0.25 mm in thickness, column pressure 8.22 psi, column flow rate 1.0 mL/min); splitless mode; oven profile as described previously.[8] The GC-MS analysis was preceded by solid-phase microextraction (SPME) for preconcentrating the headspace VOCs (Figure 1, *D*). A manual SPME holder with a divinylbenzene/carboxen/polydimethylsilox-ane-coated extraction fiber (Sigma-Aldrich, Rehovot, Israel) was inserted for 30 minutes into the thermal desorption device containing the headspace sample before being delivered to the GC-MS (Figure 1, *E*). The fiber was then inserted into the GC injector (direct mode) for thermal desorption at 270°C. The molecular structures were determined by a spectral library match using the Automated Mass Spectral Deconvolution and Identification System (ADMIS) software (Gaithersburg, Maryland). The data were processed using the open-source XCMS (V. 1.22.1) package (http://metlin.scripps.edu/download/), which provides m/z and retention times. Statistical analysis was carried out using SAS JMP, V. 8.0 (SAS Institute, Cary, North Carolina, 1989– 2005) for Wilcoxon/Kruskal-Wallis tests.

Headspace analysis with the GNP sensors

The headspace samples were analyzed using cross-reactive chemiresistors that were based on spherical GNPs (3- to 4-nm core diameter) coated with organic ligands, where the organic ligands provided the broadly cross-selective adsorption sites for the breath VOCs.[5,6,18] The GNPs were synthesized as described elsewhere.[6,18,25–27] and dispersed in chloroform. Supplementary Figure S1 (available online at http://www.nanomedjournal.com) shows a representative transmission electron microscopy image of the GNPs in solution. The

metallic cores, appearing as dark dots in Supplementary Figure S1, are separated from each other by their capping organic ligands, which appear as a bright medium between the adjacent dark dots. Macroscopically continuous chemiresistive layers were formed by drop-casting the solution onto semicircular microelectronic transducers (Figure 1, *C*, inset). The baseline resistance of the devices ranged from 0.1 MΩ to 24 MΩ with typical values of several MΩ (Supplementary Table S1). The device was dried for 2 hours at 23–25°C and then baked overnight at 50°C in a vacuum oven. The microelectronic transducers consisted of 10 pairs of circular interdigitated gold electrodes that were deposited by an electron-beam evaporator TFDS-870 (Vacuum Systems & Technologies, Petah Tikva, Israel) on a piece of silicon wafer capped with 1 μm thermal oxide (Silicon Quest International, Reno, Nevada). The outer diameter of the circular electrode area was 3 mm (Figure 1, *C*, inset), and the gap between two adjacent electrodes and the width of each electrode were both 20 μm.

Eighteen GNP sensors with different organic functionalities (Supplementary Table S1) were mounted on a custom polytetrafluoroethylene circuit board inside a stainless-steel test chamber with a volume of 100 cm^3, as shown in Figure 1, *B, C*. The sampling system delivered pulses of the headspace sample from the thermal desorption device to the sensors. The chamber was evacuated between exposures. An Agilent multifunction switch 34980 was used to measure the resistance of all 18 sensors simultaneously as a function of time. Typically, the sensors' responses were recorded for 5 minutes in vacuum, followed by 5 minutes under breath exposure, followed by another 5 minutes in vacuum. The cycles were repeated two or three times to test reproducibility. The most suitable sensors were selected for each studied problem from the reservoir of 18 available sensors as described in the Results section under "Identification of LC histology–specific pattern using GNP sensors" (Table 2).

Statistical analysis

LC histology–specific patterns were determined from the collective response of the GNP sensors by applying support vector machine (SVM) analysis as a statistical pattern recognition algorithm.[15,16] SVM analysis is a supervised learning method that finds the best separating line (or plane) between two data sets, through computerized analysis of the sensing signals and automatic choice of the most suitable set of sensing features. It can be used as a heuristic to select the most suitable sensing features from a multidimensional data set for data classification and pattern establishment, and does not require normal distribution of the data points around the average value.[28] The subpopulations were compared by building a multiclass classifier based on a linear nu-SVC SVM classifier.[29] In this study, SVM was also used as a heuristic to select from the chemical compounds identified by GC-MS those compounds that contribute most to the distinction between the compared groups.

Cross-validation was used to evaluate the classification success in terms of specificity, sensitivity, and accuracy by randomly dividing each subpopulation into two sets, which were then used as the training set and the test set. Cross-validation has a training stage followed by a test stage. A new test set that is blinded against the model was created each time before the training stage. The model was built based on the remaining samples (i.e., the training set). Thus, the method is less biased toward any group that is initially selected as a

test set. Cross-validation tries to remove that bias by generating all possible test sets, giving a stronger statistical result and making it superior, in the case of a relatively limited sample size, to the choice of fixed training and test sets. All possible combinations of divisions into test sets and training sets were tested, and the results were averaged. The results were stable against changing the number of folds in the cross-validation.

Study design

The present study was conducted in three phases. The first phase aimed to discriminate between LC and healthy cells through (i) chemical analysis of the substances in the cell line headspace, and through (ii) LC-specific patterns that were obtained from the collective responses of one to three GNP sensors. Headspace samples were obtained in duplicates for the chemical analysis and the pattern identification. The IBE cell line without other cancer-specific modifications was chosen as a model for the healthy lung cells (Table 1). The majority of the LC headspace samples were collected from NSCLC cell lines. Because NSCLC originates from epithelial cells, and epithelial cells made up the majority of the lung tissue, the IBE cell line is an adequate control for the identification of LC states. The second phase aimed to discriminate the NSCLC and SCLC, which account for the vast majority of LCs. The third phase aimed to discriminate between histologically different subtypes within the NSCLC group (i.e., adenocarcinoma and squamous cell carcinoma).

Results

Chemical analysis of the headspace LC cell lines

Our GC-MS/SPME analysis identified over 700 different VOCs in each headspace sample. Nonparametric Wilcoxon/K-ruskal-Wallis tests could identify several VOCs from the families of aldehydes, alkanes, ketones, alcohols, and benzene derivatives that were on average significantly elevated or reduced in the LC subtypes studied, as compared to the empty medium (Table 3).

Marked differences were observed between the average headspace composition of all LC cell lines and of their simulated healthy controls (i.e., the replicates of the IBE cell line; compare Table 1). The headspace of the IBE cell lines was almost identical to the headspace of medium 2, with only slightly elevated levels of 1,3-dimethyl-benzene. In contrast, the LC cells caused on average more significant changes to the head-space environment, as could be expected from their faster metabolism. For direct comparison between the LC and IBE states, a correction of the VOC concentrations due to the differences in their related growth medium (see “Collection of the headspace samples” in the Methods section) was carried out. For this purpose the average VOC concentrations in the headspace of medium 1 and medium 2 were subtracted from the headspace concentrations of the LC and IBE cell lines, respectively. Under consideration of the experimental error, only decanal showed a significant difference between the LC and the IBE states (Table 3). In this regard, SVM analysis and cross-validation showed that the decanal concentration in the headspace of the LC cells was so strongly decreased with respect to the IBE control cells that LC and IBE states could be completely separated, with an accuracy of 100% (Table 4).

Twelve VOCs showed significant differences in the concentration between SCLC and NSCLC (Table 3). Among these VOCs, SVM analysis identified three substances [decanal, acetophenone, and 1,3-bis(1,1-dimethylethyl)-benzene] as chief contributors to the separation between the two groups, which allowed distinguishing NSCLC from SCLC with 100% sensitivity and 75% specificity.

Nine VOCs (aldehydes, one alkane, two ketones, one alcohol, and three benzene derivatives) showed differences between the subtypes of NSCLC (i.e., between adenocarcinoma and squamous cell carcinoma). Among these VOCs, SVM identified 2-ethyl-1-hexanol, 1,3-dimethyl-benzene, and 1,3-bis(1,1-dimethylethyl)-benzene as key distinguishing VOCs, which are all found at higher concentration in the headspace of adenocarcinoma than in the headspace of squamous cell carcinoma. Using these three VOCs, adenocarcinomas could be distinguished from squamous cell carcinomas with 100% sensitivity, 67% specificity, and 90% accuracy.

Identification of LC histology–specific pattern using GNP sensors

The choice of suitable sensors was the most crucial step for identifying patterns that are specific for the LC subtypes. Each of the 18 GNP sensors of the reservoir responded to all (or to a certain subset) of the VOCs found in the samples, because the organic ligands of the GNPs provided only a moderate chemical selectivity. The ligands of the cross-reactive 18 GNP sensors were selected based on their ability to absorb certain (classes of) VOCs that are typically emitted from cell membranes as metabolic products. So far, GNP sensors developed by Haick and co-workers have been shown to be sensitive to typical headspace VOCs such as aldehydes, alkanes, ketones, alcohols, and benzene derivatives, with typical detection limits for the separate VOCs of 1–5 parts per billion, and device-to-device variations between sensors based on the same type of GNP within ±15%.[6] Some of the GNP sensors showed a very low response to water.[6] This is an important feature, because otherwise a sensor's response to the high background humidity in the headspace of cell lines could easily mask the signal to the much lower concentrations of the VOCs that indicate a specific LC histology.[14]

The histologically different types of LC were characterized by subtle differences in the concentration of a multitude of metabolites. On the other hand, the concentrations of many other metabolites remained unaffected. Some of the GNP sensors were especially sensitive to the classes of LC histology–specific VOCs, thus adding to the discrimination between the subtypes. Nevertheless, the majority of the sensors were more sensitive to the VOCs that were unaffected by the LC histology, and hence added mostly noise. Therefore, the identification of the sensors that contributed most to the separation was crucial for identifying the LC subtype patterns. However, a priori knowledge of the VOC profiles of the different LC subtypes was not necessary for the choice of the sensors, because SVM was used as a heuristic approach to identify the most suitable GNP sensors in each phase of this study (Table 2).

Each sensor underwent a rapid and fully reversible change in electrical resistance upon exposure to the sample, which could be either an increase (i.e., positive resistance change) or a decrease (i.e., negative resistance change) (Figure 1, *F*). For example, the sensors 1 and

3 that were selected for this study showed negative resistance changes, whereas sensors 2 and 4 showed positive resistance changes for all cell types examined. Several sensing features were extracted from the time-dependent resistance responses that related to (i) the net resistance change upon exposure, ΔR, at the start of the exposure signal (ΔR_{start}); (ii) ΔR at the midrange of exposure signal (ΔR_{center}); (iii) ΔR at the end of the exposure signal (ΔR_{end}); (iv) the resistance response normalized with respect to the baseline resistance, $\Delta R/R_o$ at the start of the exposure signal; and/or (v) the area under the resistance-vs.-time response curve, *A*.

LC-specific patterns were obtained from the collective response of the GNP sensors by applying SVM analysis as a statistical pattern recognition algorithm.[5,6,27,30] In this way, sensing features of one single sensor were selected through SVM to distinguish between the LC states and the IBE controls. The number of input parameters was kept low enough to avoid overfitting during the SVM analysis. The pseudo-3D representation of the three sensing features in Figure 2, *A*, shows that the two states could clearly be distinguished. Table 5, lists the classification success that was determined through cross-validation, in terms of correct and false classifications. From this classification success, we estimated 96% sensitivity, 86% specificity, and 93% accuracy for the identification of LC from headspace samples. Note that medium 1 was used to grow the LC cell lines and medium 2 was used to grow the IBE cell lines, so that the sensing features had to be corrected to enable the direct comparison. Because the response signals of many GNP sensors were additive, the sensing features were corrected by subtracting from the measured responses the mean responses to the corresponding measured medium. To verify that the influence of the medium on the headspace composition was effectively excluded, the sensing features that were collected from the headspaces of medium 1 and of medium 2 were compared. The results showed a total overlap of the sensing features, and SVM analysis could not distinguish between the corrected signals of the two media (Supplementary Figure S2 and Table S2).

Figure 2, *B* shows that NSCLC could be well distinguished from SCLC, using one feature each from the sensors 1–3 (Table 2). No correction of the sensing features was necessary in this case, because all LC cell lines were grown on medium 1. SVM and cross-validation yielded 100% sensitivity, 75% specificity, and 96% accuracy. The headspace atmospheres of the NSCLC subtypes (adenocarcinoma and squamous cell carcinoma) were distinguished using three features from sensors 1, 2, and 4 (Figure 2, *C*). SVM and cross-validation analysis yielded 86% sensitivity, 100% specificity, and 90% accuracy.

Discussion

The design of this study followed the hierarchical order of a possible future test for the screening and subsequent differential diagnosis of LC. In the first phase of such a test, a wide population would be screened for LC, using a test that can distinguish between LC and healthy states. In the second phase, the histological LC type would be determined in the LC-positive subjects, using a test that can distinguish between the two most prevalent histological types: NSCLC (80.4%) and SCLC (16.8%).[31] In the third phase the subhistology in the NSCLC-positive population would be identified, using a test that can

distinguish between the two most prevalent types of NSCLC: adenocarcinoma (44%) and squamous cell carcinoma (38%).[32]

Studying the metabolic activity of isolated in vitro cancer cells by analyzing their headspace VOCs avoids the organism's confounding factors (e.g., variations in patient age, gender, lifestyle, medication, and other chronic diseases). Cell lines provide an almost unlimited supply of cells with similar genotypes and phenotypes, avoiding variation between individuals and bypassing ethical issues associated with animal and human experiments. However, in vitro studies may fail to replicate the precise cellular conditions of an organism, because they disregard the synergetic effect of cancer on the whole organism (immune system, hepatic, etc.). Furthermore, with time the cell lines might undergo additional mutations and might no longer reflect the properties of the cell from which they were derived.

Eighteen different SCLC and NSCLC cell lines having different subhistologies were used (Table 1), rather than replicates of a single cell line, to simulate the natural diversity of LC. An IBE cell line served as a model for the healthy state, because epithelial cells make up most of the lung tissue. GC-MS was chosen as a powerful tool for detecting the LC-marker VOCs and their concentrations. However, the GC-MS analysis required expensive equipment and as well as high levels of expertise required to operate it; moreover, the analysis of the results was time-consuming and required significant expertise. In addition, the samples must be *dehumidified and preconcentrated by* SPME, to achieve the necessary sensitivity. To this end, using GNP sensors offers several advantages over GC-MS: it is faster and easier to perform, it does not require any *pretreatment for the samples*, and the technology is potentially more cost-effective.

Chemical composition of the LC cell headspace

Previous in vitro studies[33–36] of LC cell headspace samples identified a number of LC-characteristic VOCs, using various MS techniques. In contrast, the natural diversity of the different SCLC and NSCLC cell lines studied here has reduced the variety of the LC-specific VOCs and singled out decanal for the majority of LC states, compared to the IBE samples. A possible cause for the decrease of decanal in the LC headspace samples could be an impairment of the oxidation phosphorylation process due to mitochondrial defects in LC cells. This mitochondrial defect could result in a decrease in reactive oxygen species in the microenvironment of the cells and consequently, to a decrease in lipid peroxidation products such as aldehydes and ketones.[37] On the other hand, exposure to carcinogens could cause an upregulation of aldehyde dehydrogenase in the lung, which could also contribute to the observed decrease in decanal. Studies have shown that this enzyme is overexpressed in NSCLC cell lines, especially in adenocarcinoma and squamous cell carcinoma, from which the majority of the LC headspace samples studied were taken.[38]

SCLC is distinguished from NSCLC by its rapid growth rate and rapid metabolism (Table 3). A greater quantity and variety of metabolites released by the SCLC would not be surprising. This is because SCLC cells are rapidly dividing cells that require more adenosine triphosphate, nucleotides, fatty acids, membrane lipids, and proteins.[39] The observed difference of the headspace atmosphere could be taken as an indication of these metabolic

differences. Note, however, that the relatively large number of substances is more probably due to the small sample size of the SCLC group. SVM analysis identified decanal, acetophenone, and 1,3-bis(1,1-dimethylethyl)-benzene as the key VOCs distinguishing between the SCLC and NSCLC groups. The marked decrease of the decanal concentration in the headspace of the NSCLC cell lines can be understood in terms of the increased aldehyde dehydrogenase expression in NSCLC cells, compared to SCLC.[38] Acetophenone and benzene derivatives occur in tobacco smoke and/or can be found among environmental toxins, but no connection to LC subhistology has yet been established.[40] SCLC occurs almost exclusively in smokers, whereas NCCLC affects both smokers and nonsmokers.[32,41] Therefore, the observed increased concentration of these substances in the headspace of the SCLC cells might result from different smoking habits of the donors of the tumor tissue from which the cell lines were derived. However, a larger study would be necessary to establish reliable concentration profiles.

Two benzene derivatives [1,3-dimethyl-benzene and 1,3-bis(1,1-dimethylethyl)-benzene] and 2-ethyl-1-hexanol were identified by SVM as the key VOCs for distinguishing between adenocarcinoma and squamous cell carcinoma, from the group of seven compounds that showed significant differences between the NCLC subtypes. These three dominant VOCs were found in elevated concentrations in the headspace of adenocarcinoma, compared to squamous cell carcinoma. An increase of 2-ethyl-1-hexanol and in 1,3-bis(1,1-dimethylethyl)-benzene has been reported in the literature in the headspace of NSCLC adenocarcinoma cells,[8,36] but the biochemical cause for the different VOCs levels in adenocarcinoma and squamous cell carcinoma is still unclear.

Classification of the LC histology with GNP sensors

We found that the organic ligands of the GNP sensors that were selected using a statistical algorithm were closely correlated to the key separating VOCs identified by complementary GC-MS. The superior discrimination of sensor 1 between LC and IBE cell lines can be understood in terms of the structural similarity between the decanethiol coating of the GNPs and the decanal that appears as the key separating VOC in the headspace of the LC cell lines. It is reasonable to assume that decanal would form van der Waals interactions with the decanethiol ligand, because similar chemical structures tend to stabilize each other. The separation between NSCLC and SCLC could not be achieved with sensor 1 alone but required an additional GNP sensor with hexanethiol ligands (sensor 2) and a GNP sensor with butanethiol ligands (sensor 3). This can be understood in terms of the larger variety of VOCs that contribute to the separation of the LC subtypes. Decanal, which would have a particular affinity to sensor 1, was found among the key separating compounds, but two other VOCs [benzene-1,3-bis(1,1-dimethylethyl) and acetophenone] have important roles as well. The steric hindrance between these two VOCs and the decanethiol ligands of sensor 1 could impede the effective sensor-analyte interaction. It can be assumed that benzene-1,3-bis(1,1-dimethylethyl) and acetophenone are more likely to be absorbed on hexanethiol and butanethiol, which are less steric molecules and less hydrophobic than decanethiol. Effective separation between adenocarcinoma and squamous cell carcinoma was achieved by replacing the chemiresistive butanethiol-GNPs (sensor 3) with 2-mercaptobenzoxazole-GNPs (sensor 4). This could be related to the observed increase of benzene derivatives

among the key separating VOCs. The 2-mercaptobenzoxazole ligands of sensor 4 would have a particular affinity to the two benzene derivatives among the dominant headspace VOCs due to their structural similarity. Moreover, GNPs coated with benzene derivatives have many vacant metal sites on the surface due to strong steric effects of the large ligand molecules. These metal surface sites are more likely to absorb polar analytes such as the third separating VOC (i.e., 2-ethyl-1-hexanol), which would have a strong effect on the conductivity of the GNP films.[42] 2-Ethyl-1-hexanol is also easily absorbed by the structurally similar hexanethiol ligands of sensor 2, because they share a similar hexane chain, and to a lesser extent, to the longer decanethiol ligands of sensor 1. Finally, the nonpolar alkane ligands of sensors 1 and 2 are generally likely to absorb nonpolar VOCs such as the dominant benzene derivatives in the studied headspace samples.

In summary, analyzing the VOCs in the headspace of LC cells with GNP sensors allowed significant discrimination between LC and IBE cells, as well as between two major LC subtypes, NSCLC and SCLC, and between two subtypes of NSCLC that both occur in similar populations of smokers. The results presented here should be considered a proof of concept. A wider study would be necessary to confirm the criteria for distinguishing the different lung cancer subtypes. It is reasonable to expect that the clusters for a larger sample size would be less defined, and some overlap could occur. In turn, cluster separation could be improved again by further refining the GNP sensors. The iterative improvement of the sensors while expanding the sample size would eventually yield an optimized method. The use of organically functionalized GNPs allows tailoring the properties of the constituent sensors to tune their sensitivity to the particular classes of molecules that are of interest here —that is, the LC histology–specific VOCs. Arguably, measuring the volatile metabolites could be potentially useful for detection the full spectrum of the cancerous metabolic transformation, as each cancer type has its own metabolic specificity.[43] This potential might totally revolutionize LC screening and diagnosis, and might eventually make early, differential diagnosis of LC subtypes with undetectable or unreachable lung nodules possible. Although similar classifications could conceptually be achieved by spectrometry methods, these techniques are impeded by the need for expensive equipment and the considerable expertise required to operate such instruments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

O.B. and U.T. acknowledge helpful discussions with Dr. Maya Ilouze.

The research leading to these results has received funding from the FP7-Health Program under the LCAOS (grant 258868; H.H. and N.P.) and FP7′s ERC grant under DIAG-CANCER (grant 256639; H.H.), the NIH/Lung SPORE (F.H., P.B.), and the International Association for the Study of Lung Cancer (N.P.).

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin. 2009; 59:225–249. [PubMed: 19474385]

2. Vallieres E, Shepherd FA, Crowley J, Van Houtte P, Postmus PE, Carney D, et al. The IASLC Lung Cancer Staging Project: proposals regarding the relevance of TNM in the pathologic staging of small cell lung cancer in the forthcoming (seventh) edition of the TNM classification for lung cancer. J Thorac Oncol. 2009; 4:1049–1059. [PubMed: 19652623]
3. Peled N, Yoshida K, Wynes MW, Hirsch FR. Predictive and prognostic markers for epidermal growth factor receptor inhibitor therapy in non-small cell lung cancer. Ther Adv Med Oncol. 2009; 1:137–144. [PubMed: 21789118]
4. Holty JE, Kuschner WG, Gould MK. Accuracy of transbronchial needle aspiration for mediastinal staging of non-small cell lung cancer: a meta-analysis. Thorax. 2005; 60:949–955. [PubMed: 15994251]
5. Peng G, Hakim M, Broza YY, Billan S, Abdah-Bortnyak R, Kuten A, et al. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. Br J Cancer. 2010; 103:542–551. [PubMed: 20648015]
6. Peng G, Tisch U, Adams O, Hakim M, Shehada N, Broza YY, et al. Diagnosing lung cancer in exhaled breath using gold nanoparticles. Nat Nanotechnol. 2009; 4:669–673. [PubMed: 19809459]
7. Chen X, Xu F, Wang Y, Pan Y, Lu D, Wang P, et al. A study of the volatile organic compounds exhaled by lung cancer cells in vitro for breath diagnosis. Cancer. 2007; 110:835–844. [PubMed: 17599760]
8. Barash O, Peled N, Hirsch FR, Haick H. Sniffing the unique qodor printq of non-small-cell lung cancer with gold nanoparticles. Small. 2009; 5:2618–2624. [PubMed: 19705367]
9. Wehinger A, Schmid A, Mechtcheriakov S, Ledochowski M, Grabmer C, Gastl GA, et al. Lung cancer detection by proton transfer reaction mass-spectrometric analysis of human breath gas. Int J Mass Spectrom. 2007; 265:49–59.
10. Phillips M, Gleeson K, Hughes JMB, Greenberg J, Cataneo RN, Baker L, et al. Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. Lancet. 1999; 353:1930–1933. [PubMed: 10371572]
11. Gordon SM, Szidon JP, Krotoszynski BK, Gibbons RD, Oneill HJ. Volatile organic-compounds in exhaled air from patients with lungcancer. Clin Chem. 1985; 31:1278–1282. [PubMed: 4017231]
12. Preti G, Labows JN, Kostelc JG, Aldinger S, Daniele R. Analysis of lung air from patients with bronchogenic carcinoma and controls using gas chromatography-mass spectrometry. J Chromatogr. 1988; 432:1–11. [PubMed: 3220881]
13. O'Neill HJ, Gordon SM, O'Neill MH, Gibbons RD, Szidon JP. A computerized classification technique for screening for the presence of breath biomarkers in lung cancer. Clin Chem. 1988; 34:1613–1618. [PubMed: 3042190]
14. Tisch U, Haick H. Arrays of chemisensitive monolayer-capped metallic nanoparticles for diagnostic breath testing. Rev Chem Eng. 2011; 26:171–179.
15. Shuster G, Gallimidi Z, Heyman-Reiss A, Dovgolevsky E, Billan S, Abdah-Bortnyak R, et al. Classification of breast cancer precursors through exhaled breath. Breast Cancer Res Treat. 2011; 126:791–796. [PubMed: 21190078]
16. Hakim M, Billan S, Tisch U, Peng G, Dvrokind I, Abdah-Bortnyak R, et al. Diagnosis of head-and-neck-cancer from exhaled breath. Br J Cancer. 2011; 104:1649–1655. [PubMed: 21505455]
17. Tisch U, Haick H. Nanomaterials for cross-reactive sensor arrays. MRS Bull. 2010; 35:797.
18. Tisch, U.; Haick, H. Sensors based on monolayer-capped metal nanoparticles. In: Korotcenkov, G., editor. Chemical sensors. Vol. 2: Nanstructured materials. New York: Momentum Press, LLC; 2010. p. 141-202.
19. Alberts, B.; Johnson, A.; Lewis, J. Molecular biology of the cell. 4th. New York: Garland Publishing; 2002.
20. Singer SJ, Nicolson GL. The fluid mosaic model of the structure of cell membranes. Science. 1972; 175:720–731. [PubMed: 4333397]
21. Kneepkens CM, Lepage G, Roy CC. The potential of the hydrocarbon breath test as a measure of lipid peroxidation. Free Radic Biol Med. 1994; 17:127–160. [PubMed: 7959173]
22. Vousden KH, Ryan KM. p53 and metabolism. Nat Rev Cancer. 2009; 9:691–700. [PubMed: 19759539]

23. Okunieff P, Fenton B, Chen Y. Past, present, and future of oxygen in cancer research. Adv Exp Med Biol. 2005; 566:213–222. [PubMed: 16594155]
24. Hakim M, Billan S, Tisch U, Peng G, Dvrokind I, Marom O, et al. Diagnosis of head-and-neck cancer from exhaled breath. Br J Cancer. 2011; 104:1649–1655. [PubMed: 21505455]
25. Dovgolevsky E, Haick H. Direct observation of the transition point between quasi-spherical and cubic nanoparticles in two-step seed-mediated growth method. Small. 2008; 4:2059–2066. [PubMed: 18932188]
26. Dovgolevsky E, Tisch U, Haick H. Chemically sensitive resistors based on monolayer-capped cubic nanoparticles: towards configurable nanoporous sensors. Small. 2009; 5:1158–1161. [PubMed: 19274647]
27. Dovgolevsky E, Convolina G, Tisch U, Haick H. Monolayer-capped cubic platinum nanoparticles for sensing nonpolar analytes in highly humid atmospheres. J Phys Chem C. 2010; 114:14042–14049.
28. Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995; 30:273–297.
29. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. SIGKDD Explorations. 2009; 11:10–18.
30. Haick H. Chemical sensors based on molecularly modified metallic nanoparticles. J Phys D. 2007; 40:7173–7186.
31. Travis WD, Travis LB, Devesa SS. Lung cancer. Cancer. 1995; 75(Suppl 1):191–202. [PubMed: 8000996]
32. Bryant A, Cerfolio RJ. Differences in epidemiology, histology, and survival between cigarette smokers and never-smokers who develop non-small cell lung cancer. Chest. 2007; 132:185–192. [PubMed: 17573517]
33. Filipiak W, Sponring A, Filipiak A, Ager C, Schubert J, Miekisch W, et al. TD-GC-MS analysis of volatile metabolites of human lung cancer and normal cells in vitro. Cancer Epidemiol Biomark Prev. 2010; 19:182–195.
34. Filipiak W, Sponring A, Mikoviny T, Ager C, Schubert J, Miekisch W, et al. Release of volatile organic compounds (VOCs) from the lung cancer cell line CALU-1 in vitro. Cancer Cell Int. 2008; 8:1–11. [PubMed: 18177496]
35. Sponring A, Filipiak W, Ager C, Schubert JK, Miekisch W, Amann A, et al. Analysis of volatile organic compounds (VOCs) in the headspace of NCI-H1666 lung cancer cells. Cancer Biomark. 2010; 7:153–161. [PubMed: 21263191]
36. Sponring A, Filipiak W, Mikoviny T, Ager C, Schubert JK, Miekisch W, et al. Release of volatile organic compounds from the lung cancer cell line NCI-H2087 in vitro. Anticancer Res. 2009; 29:419–426. [PubMed: 19331181]
37. Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? Nat Rev Cancer. 2004; 4:891–899. [PubMed: 15516961]
38. Patel M, Lu L, Zander DS, Sreerama L, Coco D, Moreb JS. ALDH1A1 and ALDH3A1 expression in lung cancers: correlation with histologic type and potential precursors. Lung Cancer. 2008; 59:340–349. [PubMed: 17920722]
39. Hsu PP, Sabatini DM. Cancer cell metabolism: Warburg and beyond. Cell. 2008; 134:703–707. [PubMed: 18775299]
40. Wallace L, Pellizzari E, Hartwell TD, Perritt R, Ziegenfus R. Exposures to benzene and other volatile compounds from active and passive smoking. Arch Environ Health. 1987; 42:272–279. [PubMed: 3452294]
41. Barbone F, Bovenzi M, Cavallieri F, Stanta G. Cigarette smoking and histologic type of lung cancer in men. Chest. 1997; 112:1474–1479. [PubMed: 9404741]
42. Joseph Y, Guse B, Yasuda A, Vossmeyer T. Chemiresistor coatings from Pt- and Au-nanoparticle/nonanedithiol films: sensitivity to gases and solvent vapors. Sens Actuat B. 2004; B98:188–195.
43. Kroemer G, Pouyssegur J. Tumor cell metabolism: cancer's Achilles' heel. Cancer Cell. 2008; 13:472–482. [PubMed: 18538731]

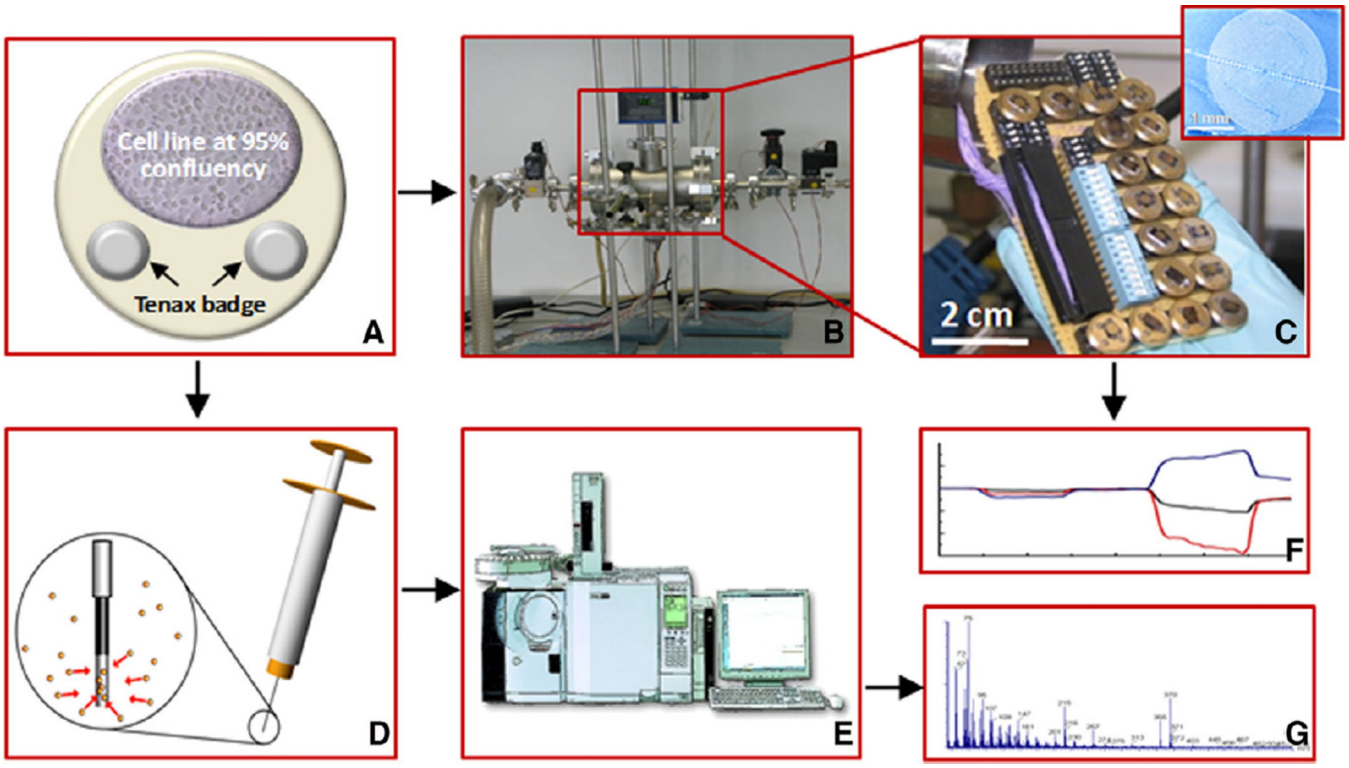

Figure 1.

Illustration of the experimental procedure that was used in the in vitro studies for sampling lung cell line headspace. **(A)** The cell culture dish and two Ultra II SKC badges with Tenax TA (265 mg; SKC) sorbent material were placed together in a bigger cell culture dish for headspace sampling during the growth time from seeding to 95% confluence. When cells reached ~95% confluency, the Tenax TA material was transferred from the badges into sealed vials and analyzed by both the GNP sensors **(B, C, F)** and GC-MS **(D, E, G). (B)** The headspace sample was fed via a gas flow system into an exposure cell. **(C)** GNP sensors used for the diagnosis of the headspace samples. See inset for a scanning electron microscopy image of the sensor's interdigitated microelectrodes. **(D)** SPME for preconcentrating the headspace sample. **(E)** GC-MS for the preconcentrated headspace sample. **(F)** Typical signal output of the GNP sensors and **(G)** typical GC-MS chromatogram.

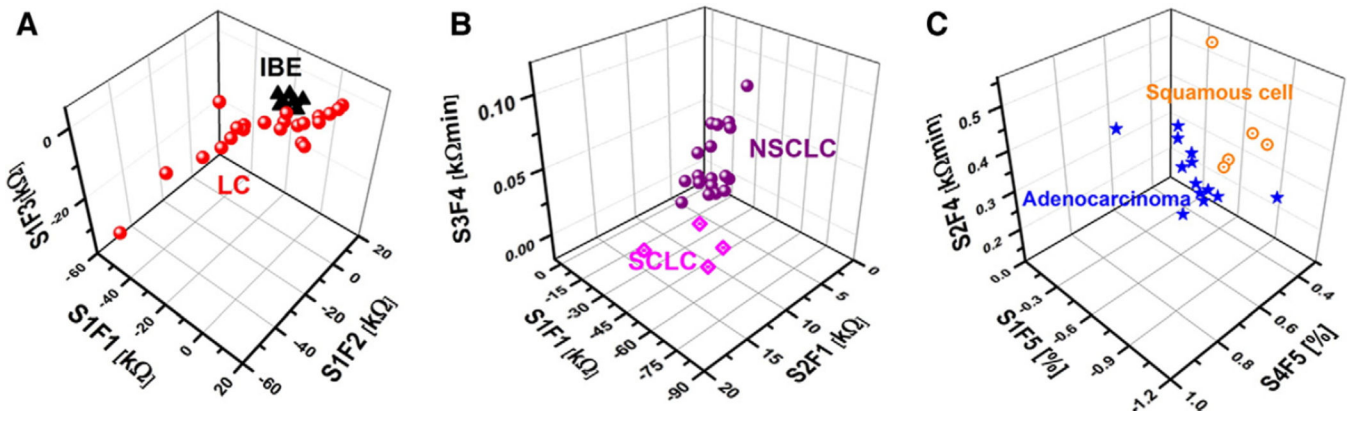

Figure 2.

Sensing features selected by SVM analysis from the GNPs' multidimensional data output for optimal separation between headspace samples from: **(A)** LC and IBE cells; **(B)** SCLC and NSCLC; and from **(C)** subcategories of NSCLC (i.e., adenocarcinoma and squamous cell carcinoma cells). The pseudo-3D representation of the selected features illustrates the separation between histologically different LC subtypes. F1 is the ΔR_{start}, F2 is the ΔR_{mid}, F3 is the ΔR_{end}, F4 is the $\Delta R/R_0$, and F5 is the are under the normalized response signal (NOTE: ΔR is the net resistance change upon exposure and R_0 is the baseline resistance). S1-S4 refers to the sensors defined in Table 2.

Nanomedicine. Author manuscript; available in PMC 2016 February 08.

Table 1

Characteristics of the headspace samples analyzed using the GNP sensors and/or GC-MS

Histology			Cell line†	Growth medium	GNP sensors		GC-MS	
Lung cancer	**NSCLC**	**Adenocarcinoma**	H1650	Medium 1	1	14	1	14
			H820		1		1	
			A549		1		1	
			H1975		1		1	
			H4006		3*		3*	
			H1435		3*		3*	
			Calu-3		1		1	
			A549		1		1	
			H2009		1		1	
			Calu-3		1		1	
		Squamous cell carcinoma	HCC95		2*	5	2*	6
			HCC15		2*		2*	
			H226		1		1	
			NE18		0		1	
	SCLC		H774		1	4	1	4
			H69		1		1	
			H187		1		1	
			H526		1		1	
IBE			Minna 3KT	Medium 2	7*	7	7*	7

GC-MS, gas chromatography–mass spectrometry; GNP, gold nanoparticle; IBE, immortal bronchial epithelium; NSCLC, non–small cell lung cancer; SCLC, small cell lung cancer.

Headspace samples were collected in duplicates with one exception (no sample for the GNP sensors was collected from NE18).

*Samples from cell line replicates.

†All the cell lines used in this study were commercially available.

Table 2

Organic ligands of the GNP sensors

Sensor no.	Organic ligand	LC–IBE	SCLC–NSLC	Adenocarcinoma–squamous cell carcinoma
1	Decanethiol	X	X	X
2	Hexanethiol		X	X
3	Butanethiol		X	
4	2-Mercaptobenzoxazole			X

IBE, immortal bronchial epithelium; LC, lung cancer; NSCLC, non–small cell lung cancer; SCLC, small cell lung cancer.

Table 3

Trends of VOC profiles in the headspace samples that characterize LC and its subhistology as determined by GC-MS/SPME

Group	VOC	Medium 1→LC	Medium 2→IBE	IBE†→LC†	Medium 1→NSCLC	Medium 1→SCLC	SCLC→NSCLC	Medium 1→ adenocarcinoma	Medium 1→squamous cell carcinoma	Squamous cell carcinoma→ Adenocarcinoma
Aldehydes	Benzaldehyde						↓**		↓**	↑*
	Nonanal					↓**	↑**		↓**	↑**
	Decanal	↓*		↓***	↓**	↑**	↓**	↓**	↓**	
Alkanes	Telradccanc					↓**	↑**			
	5-Melhyl-lridccanc					↓**	↑**		↓**	↑**
Ketones	6-Mehyl-5-hcplene-2-one	↑*			↑*	↓**	↑**	↑*		↑*
	Acetophcnone					↑**	↓**			↑*
Alcohols	2,4-Bis(1,1-dimethylethyl)-phenol	↑**			↑**		↓*	↑*	↑*	
	2-Ethyl-1-hexanol				↑*	↓**	↑**	↑*	↓*	↑**
Benzene Derivatives	1,3-Bis(1,1-dimcthylcthyl)-benzene	↑*			↑*	↓**	↑**	↑**		↑**
	1,3-Dimethyl-benzene		↑*			↓**	↑**	↑**	↓**	↑***
	Styrene	↓*			↑*	↑*	↓*	↑*		↑**

GC-MS, gas chromatography–mass spectrometry; IBE, immortal bronchial epithelium; LC, lung cancer; NSCLC, non–small cell lung cancer; SCLC, small cell lung cancer; SPME, solid-phase microextraction; VOC, volatile organic compound.

↑, Increase in concentration; ↓, decrease in concentration;

* $P < 0.05$;

** $P < 0.01$;

*** $P < 0.0001$.

†The abundance of these VOCs was corrected by subtracting the mean value in the corresponding medium.

Table 4

Classification success of the GC-MS chemical analysis, as expressed by number of correct and incorrect sample classifications estimated by support vector machine (SVM) and cross-validation

(A) Lung cancer (LC) and immortal bronchial epithelium (IBE) cell lines

	Classified as LC	Classified as IBE
LC*	24	0
IBE*	0	7

(B) Non–small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) cell lines

	Classified as NSCLC	Classified as SCLC
NSCLC	20	0
SCLC	1	3

(C) Adenocarcinoma and the squamous cell carcinoma cell lines

	Classified as adenocarcinoma	Classified as squamous cell carcinoma
Adenocarcinoma	14	0
Squamous cell carcinoma	2	4

A: Using only the abundance of decanal. Note that different media were used to grow the LC and IBE cell lines, so that the sensing features had to be corrected to enable the direct comparison.

B: Using the abundance of decanal, 1,3-bis(1,1-dimethylethyl)-benzene, and acetophenone.

C: Using the abundance of 1,3-dimethyl-benzene, 1,3-bis(1,1-dimethy-lethyl)-benzene, and 2-ethyl-1-hexanol.

*The abundance of decanal was corrected by subtracting the mean value in the corresponding medium.

Table 5

Classification success of the GNP sensors, as expressed by the number of correct and incorrect sample classifications estimated by supportive vector machine (SVM) and cross-validation

(A) Lung cancer (LC) and immortal bronchial epithelium (IBE) cell lines

	Classified as LC	Classified as IBE
LC*	22	1
IBE*	1	6

(B) Non–small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) cell lines

	Classified as NSCLC	Classified as SCLC
NSCLC	19	0
SCLC	1	3

(C) Adenocarcinoma and the squamous cell carcinoma cell lines

	Classified as adenocarcinoma	Classified as squamous cell carcinoma
Adenocarcinoma	12	2
Squamous cell carcinoma	0	5

A: Using three sensing from sensor 1. Note that different media were used to grow the LC and IBE cell lines, so that the sensing features had to be corrected to enable the direct comparison.

B: Using one feature each from sensors 1–3.

C: Using one feature each from sensors 1, 2, and 4.

*The sensing features were corrected by subtracting the mean value of the corresponding medium.