INTRODUCTION

Ion channels are ion-permeable protein complexes in the lipid membranes. They establish and control the voltage gradient across biological membranes by gating the flow of specific small ions down their electrochemical gradient, which play important roles in diverse cellular processes including cardiac, skeletal, and smooth muscle contraction, epithelial transport of nutrients and ions, T-cell activation and pancreatic beta-cell insulin release, hormonal secretion, osmotic regulation of blood pressure and so on [1-5]. Previous studies have shown that ion channels are involved in the development of various human diseases including cancers [6-11]. The roles of ion channels, especially potassium channels, in the growth of tumor cells have been well documented [12-17]. Blocking channel activity affects the growth of certain tumors both in vitro and in vivo [13,18,19].

Among all the voltage-gated ion channels, potassium channels (K+ channels) have attracted most of the attentions in regard to the relationship with human cancers because K+ channels constitute the largest heterogeneous family, such as voltage-gated K+ (Kv) channels, calcium-activated K+ (KCa) channels, inwardly rectifying K+ (Kir) channels, and two-pore K+ (K2p) channels, with more than eighty genes encoding membrane proteins that control membrane potential [20]. More importantly, K+ channels play a pivotal role in the proliferation, migration, and invasion of various tumor cells and have been considered as new targets for designing cancer treatment strategies [20,21]. For example, Kv channels have been found to be involved in the proliferation of uterine and breast cancer cells [22,23]. Similarly, the activation of Kir channels has been shown to drive tumor cell proliferation in
astrocytoma and glioma [24,25]. Actually, increasing evidences indicate that, instead of a single subfamily of K⁺ channel being present in tumor cells from tissues such as colon, prostate, and breast, a variety of K⁺ channels are found in these tissues [17]. For example, the expression of Kᵥ, KᵥCa, and Kᵥp channels has been reported in prostate cancer cells [26-28].

Lung cancer is the most common cause of cancer-related death in human worldwide [29,30]. Several molecular signatures have been developed to predict survival in lung cancer [31-36]. However, little is known regarding the role of K⁺ channels in promoting tumor growth and its influence on the prognosis of lung cancer compared with other cancers. In this study, we analyzed high-throughput gene expression data to identify K⁺ channel genes implicated in lung cancer. A K⁺ channel gene signature was sought that differentiates lung tumor tissues from normal ones. This molecular signature predicted clinical prognosis in human lung cancer effectively, which is independent of, but cooperative with, standard clinical and pathological prognostic factors including patient age, lymph node involvement, tumor size, and tumor grade. We also indicated that this gene signature potentially may be used to predict outcome in human breast and colon cancers.

METHODS
K⁺ channel genes

The definition of ion channel gene for human was obtained from IUPHAR-DB [37]. In total, we collected eighty-eight K⁺ channel genes, including 47 Kᵥ channels, 12 KᵥCa channels, 15 Kᵣ channels, and 14 Kᵥp channels (Supplementary Table 1).

Gene expression data

Two independent gene expression datasets on lung cancer (GSE10072 and GSE19804) were downloaded from the Gene Expression Omnibus (GEO) database. These two datasets were chosen based on the availability of both normal and tumor tissue specimens, as well as having a large number of samples available. We used these two datasets to filter out the differentially expressed genes between normal and tumor tissues.

Besides the two GEO datasets we mentioned above, we also downloaded the gene expression datasets representing human lung, breast, and colon cancers from publicly available repositories. These datasets were chosen based on the large number of samples, the availability of clinical outcome data, and the diversity of tumor types. For each tumor type, training and validation cohorts were constructed. For lung cancer, we obtained three datasets (n = 359) which were available from a single study [31]. Two datasets were combined as training cohort (n = 257) and the other one was used as validation cohort (n = 102). The dataset for breast cancer (n = 295) was available from previous study [38]. Lastly, for colon cancer, we downloaded the dataset (n = 232) from GEO database (GSE17538). In breast and colon cancer, the tumor samples were randomly separated into two parts (1/2 for training and 1/2 for validation). Clinical characteristics of the patients used for training and validation are listed in Supplementary Table 2.

Statistical analysis

For the datasets GSE10072 and GSE19804, two-tailed t-test was used to identify the genes that differentially expressed between normal and tumor tissues. Genes with adjusted p-value < 0.01 after Bonferroni procedure were deemed dysregulated. The common differentially expressed genes shared by the two datasets were saved for further analysis.

Next, we tested the power of the differentially expressed genes in predicting the clinical outcome in lung, breast, and colon cancers. For each cancer training/validation dataset, we normalized the gene expression level into the scale of [-1, 1] by probability of expression (POE) algorithm [39,40] implemented in the metaArray library of the R Statistical Package (R Foundation for Statistical Computing, Vienna, Austria) [41]. Based on the gene expression and clinical outcome data from the training dataset, we can assign a Wald statistic generated by Cox proportional-hazard regression to each gene as a weight. A risk score was calculated for each patient using a linear combination of weighted gene expression as below:

\[ s = \sum_{i=1}^{n} w_i (e_i - \mu_i) / \tau_i \]

Here, \( s \) is the risk score of the patient; \( n \) is the number of differentially expressed genes; \( w_i \) denotes the weight of gene \( i \); \( e_i \) denotes the expression level of gene \( i \); and \( \mu_i \) and \( \tau_i \) are the mean and standard deviation of the gene expression values for gene \( i \) across all samples, respectively. Patients were then divided into high-score and low-score groups with the median of the risk score as the threshold value. A high score indicated a poor outcome. The weight of each gene was fixed, based on the training groups, and then tested in the validation groups [34].

RESULTS
Differentially expressed K⁺ channel genes in lung cancer

Two independent lung cancer datasets (GSE10072 and GSE19804) were obtained from GEO database. Both datasets contain the gene expression data from both normal and tumor lung tissues. For the dataset GSE10072, all the subjects (59 normal and 58 tumor tissues) were collected in the Lombardy region of Italy [42]. Affymetrix Human Genome U133A Array was used...
Potassium channels as biomarker in lung cancer

to measure the gene expression levels in this dataset. In contrast, the subjects in the dataset GSE19804 (60 normal and 60 tumor tissues) were from Taiwan [43]. All the subjects in this cohort are female. The gene expression levels were assayed by Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, CA, USA).

In the dataset GSE10072, all the $\text{K}^+$ channel genes were represented by 97 probesets. Among these probesets, 7 probe sets representing 6 genes were found to be up-regulated in tumor tissue while 12 probesets coding for 9 genes were down-regulated in tumor tissue, at the specified significance level (Bonferroni-adjusted $p < 0.01$) (Supplementary Table 3). Most of these results can be reproduced in the dataset GSE19804. 5 out of 6 up-regulated $\text{K}^+$ channel genes in the dataset GSE10072 were validated in the dataset GSE19804 (Fig. 1 and Supplementary Fig. 1). Similarly, 5 out of 9 down-regulated genes in the dataset GSE10072 were confirmed by the dataset GSE19804 (Fig. 1 and Supplementary Fig. 1). Therefore, we designated the ten common dysregulated genes (5 up-regulated and 5 down-regulated in tumor tissue) shared by both datasets as K-10 signature (Table 1). This K-10 signature accurately distinguished tumor tissue specimens from normal ones in both Italian and Taiwanese cohorts (Fig. 2).

![Fig. 1. Distribution of expression level for the K-10 signature in the Italian cohort (GSE10072).](image)

The 10 dysregulated genes were verified in the Taiwanese cohort (GSE19804), among which 5 genes (6 probe sets, the upper row) were up-regulated while 5 genes (7 probe sets, the lower row) were down-regulated in tumor tissue. Each panel was labelled with the corresponding gene symbol and probeset ID. Y-axis: log$_2$ transformed expression level. The solid line depicts the mean of expression value for normal or tumor tissue. The dot line indicates the overall mean of expression value across all individual samples.

![Table 1. K-10 signature](table)

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene title</th>
<th>Regulation in lung tumor tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>KCNK1</td>
<td>potassium channel, subfamily K, member 1</td>
<td>Up</td>
</tr>
<tr>
<td>KCNK12</td>
<td>potassium channel, subfamily K, member 12</td>
<td>Up</td>
</tr>
<tr>
<td>KCNK5</td>
<td>potassium channel, subfamily K, member 5</td>
<td>Up</td>
</tr>
<tr>
<td>KCNMB3</td>
<td>potassium large conductance calcium-activated channel, subfamily M beta member 3</td>
<td>Up</td>
</tr>
<tr>
<td>KCNN4</td>
<td>potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4</td>
<td>Up</td>
</tr>
<tr>
<td>KCNA5</td>
<td>potassium voltage-gated channel, shaker-related subfamily, member 5</td>
<td>Down</td>
</tr>
<tr>
<td>KCNAB1</td>
<td>potassium voltage-gated channel, shaker-related subfamily, beta member 1</td>
<td>Down</td>
</tr>
<tr>
<td>KCNAB2</td>
<td>potassium voltage-gated channel, shaker-related subfamily, beta member 2</td>
<td>Down</td>
</tr>
<tr>
<td>KCNJ8</td>
<td>potassium inwardly-rectifying channel, subfamily J, member 8</td>
<td>Down</td>
</tr>
<tr>
<td>KCNK3</td>
<td>potassium channel, subfamily K, member 3</td>
<td>Down</td>
</tr>
</tbody>
</table>
We hypothesized that K-10 signature is predictive of tumor outcome in lung cancer patients. We constructed a risk scoring system that combined gene expression of K-10 with risk for death in the training dataset. High-score patients were defined as those having a risk score greater than or equal to the group median score. In independent validation cohorts, we tested the ability of the risk score to stratify patients into prognostic groups. We performed Kaplan–Meier survival analysis comparing the high-score and low-score groups and determined statistical significance by log-rank tests. The K-10 signature was able to identify patients with poor overall survival in lung cancer in both the training (p < 0.001) and testing (p = 0.033) cohorts (Fig. 3). This association between K-10 risk score and survival was also confirmed by univariate Cox proportional hazard analysis of overall survival. High-score patients had an increased risk for death of 1.78-fold in the training cohort (p < 0.001) and 2.08-fold in the testing cohort (p = 0.036).

Independence of K-10 from other clinicopathologic factors

We investigated the performance of the K-10 signature in comparison with clinicopathologic variables associated with prognosis in lung cancer. A multivariate Cox analysis of overall survival indicated that K-10 status remained a significant covariate in relation to the standard clinicopathologic factors in lung cancer, including patient age, lymph node status, tumor size, and tumor grade (Table 2). We next stratified the patients according
to the factors significant on multivariate analysis, i.e., age, lymph node status, and tumor size, respectively. For patients with age < 65 and ≥ 65, the high-score patients had a significantly increased risk for death of 1.31-fold (p = 0.013) and 1.30-fold (p = 0.006), respectively. For patients with and without lymph node involvement, the high-score patients had a significantly increased risk for death of 1.31-fold (p = 0.005) and 1.31-fold (p = 0.014), respectively. For patients with tumor size < T3 and ≥ T3, the high-score patients had a significantly increased risk for death of 1.25-fold (p = 0.005) and 1.41-fold (p = 0.066), respectively. Kaplan–Meier survival analysis also demonstrated a significantly reduced overall survival for high-score patients in each subset grouped by age, lymph node status, and tumor size (Fig. 4). Taken together, these results suggest that the expression of K-10 signature is associated with clinical outcomes and is an independent prognostic factor.

Predicting survival in other cancers by K-10 signature

We further tested the performance of K-10 in predicting clinical outcome in breast and colon cancers. A risk score based on the expression of K-10 signature was assigned to each patient with either breast or colon cancer. Kaplan–Meier survival analysis indicated that the K-10 signature was also able to identify patients with poor overall survival in breast and colon cancers in both the training (p = 0.004 for breast cancer and p = 0.016 for colon cancer) and testing (p = 0.047 in breast cancer and p = 0.007 in colon cancer) cohorts (Fig. 5). Univariate Cox proportional hazard regression on overall survival demonstrated that, in the training cohorts, high-score patients had an increased risk for death of 2.59-fold in breast cancer (p = 0.005) and 2.29-fold in colon cancer (p = 0.019). In the validation cohorts, high-score patients had an increased risk for death of 1.97-fold in breast cancer (p = 0.052) and 2.05-fold in colon cancer (p = 0.009).

We also conducted multivariate analyses using standard clinical and pathological factors available for the breast and colon cancer datasets. Cox proportional hazards regression indicated that K-10 status was still a significant covariate in colon cancer but not in breast cancer (Supplementary Table 4). In breast cancer with poor prognosis (age < 40, tumor size ≥ T2, tumor grade ≥ 2, or ER-negative), we further stratified the patients by using the K-10 risk score. Kaplan–Meier survival analysis demonstrated a significantly reduced overall survival for the high-score patients when compared to the low-score patients except for the ER-negative group (Supplementary Fig. 2). A similar analysis for the colon cancer patients with poor prognosis (age ≥ 60 or with stage 3, 4) also demonstrated an increased risk for death in the high-score patients (Supplementary Fig. 3).

Table 2. Cox proportional hazard analysis of overall survival for the lung cancer patients

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Hazard ratio (95% confidence interval)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-10 high-score vs. low-score</td>
<td>1.29 (1.11, 1.48)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Age (per year)</td>
<td>1.02 (1.00, 1.04)</td>
<td>0.011</td>
</tr>
<tr>
<td>Lymph node status (+) vs. (–)</td>
<td>2.72 (2.02, 3.65)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Tumor size ≥ T3 vs. &lt; T3</td>
<td>2.72 (1.85, 4.01)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Tumor grade 2, 3 vs. 1</td>
<td>0.86 (0.54, 1.36)</td>
<td>0.509</td>
</tr>
</tbody>
</table>

Fig. 4. Kaplan–Meier survival curves of the lung cancer patients grouped by clinicopathologic factors. K-10 is independent from other clinicopathologic factors in lung cancer. (A) Patients were stratified by age; (B) Patients were stratified by lymph node status; and (C) Patients were stratified by tumor size. High-score patients were defined as those having a K-10 risk score greater than or equal to the group median score. p-value were calculated by log-rank tests for the differences in overall survival between the high-score and low-score groups.
DISCUSSION

The major aim of this work is to identify the differentially expressed K⁺ channel genes in lung cancer. Leveraging microarray expression data in two independent cohorts, we identified a gene signature comprised of K-10 for differentiating lung tumor tissue from normal ones. K-10 signature shows a considerable power to predict clinical outcome in human lung cancer. Multivariate analysis indicates that K-10 is independent of, but cooperative with, standard clinical and pathological prognostic factors including patient age, lymph node involvement, tumor size, and tumor grade. Further analysis suggests that K-10 is also potential to serve as a novel biomarker to assess the risk in breast and colon cancers.

The differentially expressed K⁺ channel genes were first derived from the Italian lung cancer cohort. At this step, we in total identified 15 dysregulated genes between normal and tumor tissue. We next validated these genes in the Taiwanese female cohort. Despite racial and gender difference, we found 10 common dysregulated genes shared by the two groups. Heatmap of expression level of these genes demonstrates that K-10 signature clearly separates the tumor samples from the normal ones in both datasets, which suggests that K-10 signature is not likely to be racial and gender specific.

These findings suggest that K⁺ channel is a mediator of tumor growth and progression. K⁺ channels participate in the regulation of membrane potential and electrochemical driving force for Ca²⁺ entry and both mechanisms are associated with mitotic cell cycling, proliferation, cell volume and development of cancer [44-46]. Previous studies have described that alteration in expression and function of diverse types of K⁺ channels, including Kᵣ, Kᵥ, and K₂p, in colon, breast, and pancreatic cancers [47-49]. Kᵣ channels are widely implicated in the development of different tumors [50,51]. Specifically, we found that KCNA5, KCNA9B1, and KCNAB2 were downregulated in lung tumor tissue. Kᵥ channels are gated by calcium and contribute to cell proliferation by change of intracellular calcium during tumor cell migration [48,52]. Based on our K-10 signature, both KCNMB3 and KCNN4 among Kᵥ channels were upregulated in lung tumor tissue. K₂p channels are background channels that open at rest and affect membrane potential, which also play a critical role in both cell apoptosis and tumorigenesis [49,53]. We also found that altered expression of K₂p channels in lung tumor tissue, in particular, KCNK1, KCNK12, and KCNK5, were overexpressed in tumor tissue, while KCNK3 was down-regulated.

The mutation of channels is also critical factor for functional dysregulation of K⁺ channels. For example, insertion/deletion of 12 nucleotides in the 5’ untranslated region (5’UTR) of KCNK1 was identified in long QT syndrome patients [54]. The missense variants in KCNK5 have been observed among patients with Balkan-onset nephropathy [55]. Lorenz et al. [56] reported an allele association between the KCNMB3 delA750- mutant and idiopathic systemic epilepsy syndrome. Genomic DNA analysis revealed a nonsense mutation in KCNA5 in patients with idiopathic atrial fibrillation, and two mutations, G182R and E211D, localized to the KCNA5, were identified in idiopathic pulmonary arterial hypertension (IPAH) patients [57,58]. Moreover, homozygous mutations of KCNK3 have been identified in PAH patients [59]. The finding in our study, however, the mutations of K⁺ channels are not expected to correspond to the K-10 signature genes.

Molecular signatures discovered through gene expression profiling have been shown to add prognostic value to clinical and pathological findings in several human cancers [31-36]. There is also accumulated evidence indicating that differentially expression of K⁺ channels have correlated with tumorigenesis in lung, breast and colon cancer pathology. Previous reports suggesting that Kᵥ channels may be novel molecular-markers with differential expression related to breast cancer. For example, KCNK5, KCNK9, and KCNK12 were highly expressed, and KCNK6 and KCNK15 were down-regulated in breast cancer [60]. The dysregulation of KCNA5 has also identified in colorectal cancer [51] and breast cancer [61]. Similar to our results, increased KCNN4 expression was identified with aggressive features of lung carcinoma.
Potassium channels as biomarker in lung cancer

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (no. NRF-2017R1D1B06035273, NRF-2017R1A2B4002052), and this work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through Animal Disease Management Technology, Development Program, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (119052-3).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

SUPPLEMENTARY MATERIALS

Supplementary data including three figures and four tables can be found with this article online at http://pdf.medrang.co.kr/paper/pdf/Kjpp/Kjpp2019-23-6-11-s001.pdf.

REFERENCES

10. Hanna MG, Wood NW, Kullmann DM. Ion channels and neurological disease: DNA based diagnosis is now possible, and ion channels may be important in common paroxysmal disorders. J Neurol Neurosurg Psychiatry. 1998;65:427-431.


