Exposure to Water-Pipe Smoking Dysregulates a Set of Genes Associated with Breast Cancer Development and an Unfavorable Outcome

Abstract

Background: Water-pipe smoking (WPS), a predominant method of tobacco consumption, is common amongst young females in the Middle East. WPS smoke consists of toxins analogous to the ones that exist in cigarette smoke and frequently correlates with the onset of several types of human cancers including breast. However, the potential target genes and their underlying mechanisms in the initiation and/or progression of human cancers, especially breast, due to WPS exposure are still unknown. Materials and Methods: In this investigation, we explored the effect of WPS chronic exposure on human normal mammary epithelial cells and analyzed alterations in the differentially ex-pressed gene (DEG) targets using the NanoString nCounter PanCancer Pathways Panel consisting of 770 gene transcripts and a quantitative real-time polymerase chain reaction (PCR) analysis. Results: Our NanoString analysis identified 13 genes dysregulated under the effect of WPS exposure involved in regulating signal transduction, cell cycle, cell motility, proliferation and migration/invasion as well as the inflammatory response. We further performed an in silico analysis to investigate the effect of the identified genes in the prognosis of breast cancer patients and reported those DEGs that directly correlated with smoking and were upregulated in breast cancer in comparison with normal tissue. Moreover, the Kaplan–Meier curve analysis showed a significant correlation between WPS-dysregulated genes (MX1, CCL8, GNGT1 and MMP9) and relapse-free survival in breast cancer patients. Conclusions: Our data clearly suggest that exposure to WPS can alter the expression of key regulator genes involved in the pathogenesis of breast cancer, thereby affecting the breast cancer prognosis.

Keywords: Breast cancer, gene deregulation, mammary epithelial cells, smoking, water-pipe

Introduction

Tobacco smoking is an avoidable risk factor for various noncommunicable diseases including pulmonary, diabetes, cardiovascular and different types of cancer; it is also responsible for the rise in mortality rates. Various forms of tobacco intake include water-pipe smoking (WPS), cigarettes, cigars and electronic cigarettes (E-cigarettes). Currently, WPS along with E-cigarettes are becoming global trends as they are generally more preferred publicly than cigarette smoking especially among youths and women mainly due to entertainment recreation. On average, around 100 million smokers use WPS on a regular basis, resulting in nearly five million deaths annually. Remarkably, Middle Eastern people as well as those with a Middle Eastern origin residing in the West consume WPS as an integral part of their culture and ethnicity, thus escalating this tendency in Western countries.

In WPS, charcoal-heated air is passed across a perforated aluminum foil and through flavored tobacco to turn into smoke, which consists of toxins similar to those present in cigarettes including carbon monoxide, tar, nicotine, hydrocarbons and other toxicants. Compared with cigarette smokers, the plasma concentration of nicotine in individuals smoking a WP once a day is analogous to consuming 10 cigarettes in a day. Therefore, in spite of the general belief that WPS is less toxic than cigarette smoking, investigations point out that the...
consumption of both WPS and cigarette smoking leads to severe health problems including nicotine/tobacco addiction and an increased risk for a variety of systemic serious human diseases.24,25 WPS also induces significant embryotoxicity on the early stage of embryogenesis, thereby causing serious complications in early pregnancy.26

Previous investigations have revealed that WPS exposure can have an important impact on the development of various human cancers including head and neck, oral and breast cancers.27,28 Prolonged WPS exposure induces gene alterations regulating DNA stability and repair and detoxification as well as xenobiotic metabolism, thereby enhancing cancer susceptibility.29,30 Exposure to WPS can stimulate the transition from an epithelial to a mesenchymal phenotype and increase the cell invasive ability of breast cancer cells through the Erk1/Erk2 pathways.31 Nevertheless, it is important to highlight that altered genes in normal mammary tissues exposed to WPS that can potentially participate in the onset and/or development of human breast cancer have not been explored yet. Therefore, in this investigation, we examined the outcome of chronic exposure to WPS on a set of known carcinogenesis-related gene targets and molecular pathway profiles in human normal mammary epithelial (HNME) cells.

Materials and Methods

Smoking machine protocol and preparation of the water-pipe smoking solution

The Aleppo method, a standard smoking procedure, was used as previously reported.19,22,23 Briefly, the water-pipe was set by packing the head with 10 g of a brand of tobacco mixture (Two Apples, Paterson, NJ, USA) and concealing it with perforated aluminum foil to permit air passage. The quick-light block charcoal (Tree Kings, Paterson, NJ, USA) was inflamed and positioned on top of the head at the start of the smoking session. Post 1 h of smoking, the smoking condensate was collected using a regular laboratory filter paper attached to the mouthpiece. Filters were then parched and weighed before and after collecting the smoke. Later, smoked filters were dissolved in a phosphate-buffer saline (1×) (PBS) or keratinocyte serum-free medium (KSFM) (1×) (Gibco®, Life Technologies, Burlington, ON, Canada) to a final concentration of 20 mg/mL of smoking particles. Several previous investigations outlined the detailed yield of a WPS session,26,27 however, in this study, the overall effect of WPS was under investigation not the individual components. Based on previous studies,19,22,23 the collectable WPS particulates were dissolved in the mentioned solvents. The PBS and KSFM solutions were then filtered using size 0.45 µm filters (Costar, Washington, DC, USA) to obtain the final extractable WPS solution.

Cell Lines

HNME cells28 were grown and maintained in KSFM (1×) (Gibco®, Life Technologies, Burlington, ON, Canada) with heregulin (5 ng/mL), bovine pituitary extract (BPE) (5 mg/100 mL) (Life Technologies, Burlington, ON, Canada) and penicillin-streptomycin (100 µg/mL) (Invitrogen, Life Technologies, Burlington, ON, Canada). Cells were exposed to 150 µg/mL of WPS dissolved in either the PBS or KSFM solution for 48 h and maintained at a temperature of 37°C in a 5% CO2 humidified atmosphere.

NanoString

The analysis of gene expression was performed using the NanoString PanCancer Pathways Panel (NanoString 125 Technologies, Seattle, WA, USA) comprising of 770 gene probes associated with tumorigenic pathways derived from The Cancer Genome Atlas (TCGA) data. Raw data (RCC files) from NanoString runs were processed and normalized using the standard protocols (nSolver User Manual) of the nSolver Analysis Software (NanoString Technologies, Seattle, WA, USA) as previously described by our group.23 The obtained data were normalized once again to the geometric mean of the housekeeping genes. Following normalization, data were log2-transformed and then transported to Microsoft Excel for an analysis.

Based on previous studies, a fold-change analysis of 1.5- or 2-fold in addition to a P < 0.05 is frequently used as the cutoff value for identifying differentially expressed genes (DEGs).28,30 Hence, genes were chosen based on a 1.5-fold change or higher with P < 0.05.

RNA extraction and quantitative reverse transcriptase real-time polymerase chain reaction

The extraction of total RNA was done from WPS exposed and unexposed HNME cells using RNeasy Mini Kit spin columns (Qiagen, Valencia, CA, USA) as previously described by our group.23 In brief, a synthesis of the first strand of cDNA was performed using the 5X All-In-One MasterMix (MasterMix-LR, Diaimed, Mississauga, Ontario, Canada according to the manufacturer’s instructions. Quantitative reverse transcriptase real-time PCR (qRT-PCR) was carried out using iTaq Universal SYBR Green Supermix (BioRad, Hercules, CA, USA). The primer sequences used in this study were designed using Primer ExpressTM Software v3.0.1 (ThermoFisher Scientific, Franklin, MA, USA) [Table 1].

Gene profile and in silico analysis

The DEGs identified by the NanoString study were then subjected to an in silico analysis used to further support and validate our findings. The Oncomine TM database (http://www.oncomine.org, November 14, 2020) is a large, public and widely available database that consists of around 65 gene expression datasets,31 we investigated the differential
gene expression in breast cancer in comparison with normal tissues and clinicopathological parameters. From the Oncomine TM database, we analyzed the mRNA expression of the identified DEGs in normal versus malignant patient samples. Furthermore, the Bittner breast dataset was used to evaluate the difference in the log2 median-concentrated intensity between smoker breast cancer patients compared with nonsmokers with breast cancer. Parameters were set and the program produced levels of gene expression per dataset. Based on the analysis, statistically significant deregulated genes were selected. Moreover, we used a cohort of breast carcinoma samples from the PanCancer RNA-seq dataset (Kaplan–Meier plotter database) to evaluate the clinical outcome of patients in relation to individual genes.\[^{[32]}\]

The GOBO database\[^{[33]}\] was then used to evaluate the association between WPS-deregulated genes and breast cancer molecular subtypes in 1881 breast cancer patient samples according to PAM50 or Hu subclassifications.

The association between gene expression and molecular subtypes was presented as a boxplot where the band inside the box exemplified the median and the top (high expression) and bottom (low expression) of the box implied the distance between the different quantiles. Outliers were presented as circles. The level of significance provided by the database was calculated using an ANOVA test.\[^{[33]}\]

Network and pathway interaction

We used the Search Tool for the Retrieval of Interacting Genes (STRING v9.1) (https://string-db.org/, November 10, 2020) to analyze the network and interaction between the altered WPS deregulated genes and their biological function as previously performed by our group.\[^{[23]}\] This tool was used to underline the vitality of plausible networks linking obtained genes to understand the underlying mechanisms of breast cancer progression under the effect of smoking.

Statistical analysis

All in vitro experiments were carried out in triplicates of at least three independent experiments and the results were expressed as means ± standard error mean. The Student’s t-test was performed to calculate the statistical significance. GraphPad Prism (Version 8.4.3) and nSolver analysis software were used to carry out the statistical analysis. A Kaplan–Meier survival analysis was done to determine the association between WPS-dysregulated genes and survival (relapse-free survival [RFS] and overall survival (OS)); significance was achieved at a $P < 0.05$ (log-rank test).

Results

Identification of a set of breast cancer-associated differentially expressed genes deregulated by water-pipe smoking in human normal mammary epithelial cells

To analyze the detrimental outcome of WPS exposure on human breast carcinogenesis, we investigated the effect of WPS on HNME cells.\[^{[28]}\] Our data showed that exposure to WPS slightly induces Epithelial Mesenchymal Transition (EMT) where HNME cells display a mesenchymal phenotype compared with the matched unexposed controls. As shown in Figure 1, compared with the unexposed cells, WPS exposed ones were found to become elongated in appearance with a diminished cell-to-cell contact [Figure 1]. We found that exposure of HNME cells to 100 µg/mL of a WPS solution for 48 h disrupted the regulation of cell proliferation and the progression of the cell cycle in HNME cells in comparison with untreated ones (Data not shown).

We further identified gene deregulated by WPS exposure in the development of human breast cancer. We performed a differential gene expression analysis on HNME cells (exposed and unexposed to WPS) using the NanoString nCounter PanCancer Pathways Panel comprising of probes for 770 genes involved in tumorigenic pathways. A NanoString analysis revealed 13 DEGs in

### Table 1: List of primer sequences used for reverse transcriptase real-time polymerase chain reaction

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward (5'-3')</th>
<th>Reverse (3'-5')</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCL5</td>
<td>GGTGCCAGCAAGATAACCCT</td>
<td>GCTTGCCGTACCTCCCTTT</td>
</tr>
<tr>
<td>MXI</td>
<td>AGGTTCCAGTAGGGCATGTG</td>
<td>TTTGAAAGAAGGTCCTGTTGCT</td>
</tr>
<tr>
<td>CCL21</td>
<td>CTGGACAAGACACACATCCCC</td>
<td>TGATCTGGGAGCCCGATCA</td>
</tr>
<tr>
<td>IFNγ</td>
<td>CTCATGTAAGCCGCCAGAAA</td>
<td>GCCAGTT CCTGCAGAGTAG</td>
</tr>
<tr>
<td>ALOX5</td>
<td>ACTTGGCCGCTTTGAGAA</td>
<td>CAGGAGGTACCAACAGTGA</td>
</tr>
<tr>
<td>CCL8</td>
<td>GCGGCAAGTGTAATAGAGG</td>
<td>CAGGTAAAGACAGAGGTAAG</td>
</tr>
<tr>
<td>GNGT1</td>
<td>CAGCACCCTTCAAACAAAT</td>
<td>CAGGAAC ATTTGTCAGGT</td>
</tr>
<tr>
<td>MMP9</td>
<td>GTCTGTGGAGGCTTTGAGG</td>
<td>CAGGAGTCCCTTCCTT</td>
</tr>
<tr>
<td>TNFSF14</td>
<td>CTGCAAGACAGGGTAAAAAGC</td>
<td>TGGAGGTGGGCTTCACA</td>
</tr>
<tr>
<td>PTGR1</td>
<td>GAAAGTCAGTAGGGCACA</td>
<td>TCCCTCTCTTGGCCTCTCA</td>
</tr>
<tr>
<td>CCL4</td>
<td>GCTAATCCAGTGGTGGA</td>
<td>GCTGCTTCTTTGTTT</td>
</tr>
<tr>
<td>IL3</td>
<td>GTAGAGACGGGGTTTACCCA</td>
<td>GGGACAGGGCCTAGAAGTGA</td>
</tr>
<tr>
<td>TLR9</td>
<td>CACGAGCTCAGTGCAGTACGC</td>
<td>AAGGCCAGTAAATTGTCACG</td>
</tr>
</tbody>
</table>
WPS exposed versus unexposed HNME cells: CCL5, MX1, CCL21, IFNγ, ALOX5, CCL8, GNNT1, MMP9, TNFSF14, PTGR1, CCL4, IL3 and TLR9 (1.5-fold change or higher, \( P < 0.05 \)).

Post the identification of plausible candidate DEGs, we performed qRT-PCR to validate our obtained gene panel from NanoString data. The panel of DEGs correlated with the NanoString analysis with 13 genes (CCL5, MX1, CCL21, IFNγ, ALOX5, CCL8, GNNT1, MMP9, TNFSF14, PTGR1, CCL4, IL3 and TLR9) upregulated by a fold-change varying from 1.6-to 24-fold [Figure 2].

Moreover, based on functional annotations and molecular pathways underpinning carcinogenesis, we found the 13 DEGs to directly regulate cell cycle, cell proliferation, cell survival, cell migration/invasion, cell death (apoptosis), signal transduction and the inflammatory response [Table 2].

**Differentially expressed genes by water-pipe smoking are upregulated in invasive breast cancer samples in comparison with normal tissue**

For further evaluation of the role of our top DEGs deregulated by WPS in our *in vitro* study, we try to validate those DEGs in patients’ samples using *in silico* approach. To achieve this, we primarily investigated the expression patterns of those DEGs in samples obtained from normal tissues and compare its expression from samples obtained from invasive breast tumor patients using many databases included in the publicly available Oncomine database.

The TCGA dataset (137 patient samples) revealed that the expression of CCL5 \( (P < 0.001) \), MX1 \( (P < 0.001) \), MMP9 \( (P < 0.001) \), IFNγ \( (P < 0.001) \), ALOX5 \( (P < 0.001) \), GNNT1 \( (P < 0.001) \), TNFSF14 \( (P = 0.031) \), IL3 \( (P < 0.001) \) and TLR9 \( (P = 0.004) \) were significantly higher in invasive breast carcinoma compared with the normal tissue [Supplementary Figure 1a]. Moreover, the Finak dataset (59 patient samples) showed CCL4 \( (P < 0.001) \), CCL8 \( (P < 0.001) \) and CCL21 \( (P < 0.001) \) to be upregulated in invasive breast carcinoma [Supplementary Figure 1b]. The Zhao dataset (39 patient samples) revealed PTGR1 \( (P = 0.018) \) to be overexpressed in invasive breast carcinomas [Supplementary Figure 1c].

**Differentially expressed genes are highly expressed in smoking breast cancer patients in comparison with nonsmoker patients**

To further analyze the correlation amongst our identified WPS-deregulated genes and smoking as a risk factor in breast cancer development, we explored the fold-change of the 13 deregulated genes in breast cancer samples in smoker against nonsmoker breast cancer patients using the Bittner breast dataset of the Oncomine database. Remarkably, our data confirmed that of the 13 identified WPS-deregulated genes, the expression of 9 genes were upregulated in smoking patients with breast cancer in comparison with those who had never smoked. These genes included CCL5, MXI, CCL21, ALOX5, PTGR1, TNFSF14, CCL4, IL3 and TLR9 \( (P ≤ 0.05) \). Unfortunately, the smoking status was unavailable for MMP9, IFNγ, GNNT1 and CCL8 in the database [Supplementary Figure 2].

**Table 2: Classification of deregulated genes based on their functional annotations**

<table>
<thead>
<tr>
<th>Molecular and cellular functions</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular processes (cell cycle, cell proliferation, cell migration, cell invasion, cell apoptosis and angiogenesis)</td>
<td>IL3, MMP9, TNFSF14</td>
</tr>
<tr>
<td>Signal transduction (NF-κB signaling, TLR signaling, cytosolic DNA-sensing, GPCR signaling, Erk1/2 signaling, Ras signaling and PI3K-Akt signaling pathways)</td>
<td>CCL4, CCL5, CCL8, CCL21, GNNT1, IFNγ, MX1, PTGR1, TLR9, TNFSF14</td>
</tr>
<tr>
<td>Inflammatory response</td>
<td>ALOX5, CCL4, CCL5, CCL8, IFNγ</td>
</tr>
</tbody>
</table>
Water-pipe smoking-deregulated genes and their relation to breast cancer molecular subtypes

Breast cancer, a heterogenous disease, is categorized into five intrinsic molecular subtypes (Luminal A, Luminal B, HER2-positive, normal-like and basal-type).\[34,35\] Hence, we further analyzed WPS-deregulated gene expressions in relation to different breast cancer molecular subtypes using clinical cases available from the GOBO database (1881 patients). Remarkably, using the PAM50 subclassification, we found most of those genes to be more expressed in the highly aggressive basal subtype including CCL5, MX1, MMP9, CCL8 and CCL4 \[Figure 3a\]. Similarly, CCL21, TNFSF14 and IL3 expressions showed higher expression in the basal subtypes according to the Hu subclassification \[Figure 3b\].

Water-pipe smoking-deregulated genes have a direct impact on the prognosis of breast cancer patients

We then assessed the plausible impact of WPS-deregulated DEGs on the prognosis of breast cancer patients. We evaluated the correlation between the expression of the DEGs’ mRNA levels and the outcome of patients, described as RFS, using a large breast cancer cohort \(n = 1764\) patients from the Kaplan–Meier plotter database.

Our results showed conflicting data regarding the association between individual genes and the survival of patients. While MXI \(P = 0.0049\), CCL8 \(P < 0.001\), GNGT1 \(P = 0.012\) and MMP9 \(P = 0.0039\) showed a significant association with a poor outcome of patients presented as a shortened RFS, other genes showed a significant association with a favorable outcome presented as prolonged patient survival \[Figure 4\]. These findings clearly indicated the central role of WPS in modulating breast cancer cells that might affect their behavior leading to a more aggressive phenotype and a worse outcome.

On the other hand, WPS-induced genes were not associated with OS \[Supplementary Figure 3\].

Water-pipe smoking-deregulated genes are commonly involved in immune response pathways

Subsequently, we further investigated major gene interactions between WPS-deregulated DEGs and their plausible pathway enrichment \[Figure 5\].

We found that these genes cooperated in major pathways including signal transduction, ligand bindings and the synthesis of lipoxins, leukotrienes, interleukins and interferon \[Table 3\]. Moreover, these DEGs were also part of molecular functions that included chemokine, cytokine and protein binding receptors, phospholipases, phosphotransferases and kinases with catalytic activity \[Table 3\].

Discussion

To our knowledge, this study was the first cancer gene expression profiling study on the effect of WPS treatment in HNME cells. Similar to the present data, WPS enhanced the progression of EMT and invasion of breast cancer through the Erk1/2 pathway accompanied by E-cadherin and FAK gene deregulation in human breast cancer cells.\[22\] Moreover, cigarette smoking enhanced EMT in several human carcinoma cells\[36-40\] and, hence, smoking was a significant etiological factor in the onset and progression of various human cancers including breast.\[22,41-44\] Our present study implied that WPS exposure could play a vital role in the onset and possible progression of human breast cancer.

Indeed, in this investigation, the NanoString nCounter PanCancer Pathways Panel of 770 gene transcripts scattered in 13 biological pathways was used to identify the gene targets of WPS exposed HNME cells. Our data revealed significant alterations in the expression of 13 genes as targets for WPS exposure in HNME cells, which were further confirmed by qRT-PCR in addition to the Oncomine TM database. Subsequently, we determined the prognostic outcome of WPS-deregulated genes on breast cancer prognosis using the PanCancer RNA-seq dataset of the Kaplan–Meier plotter database. More significantly, this study indicated that these DEGs were found for the first time as targets of WPS exposure in human normal mammary cells. The discovered genes were involved in cell cycle, cell proliferation, cell migration/invasion, cell apoptosis, signal transduction and the inflammatory response and were thus likely involved in the neoplastic transformation of normal mammary epithelial cells leading to the onset of breast cancer.

Of the thirteen differentially expressed genes, five \(\text{CCL5, CCL4, CCL8, CCL21 and TNFSF14}\) of these genes were part of the chemokine family. Upregulated levels of CCL5 significantly correlated with breast cancer...
progression, metastasis and/or relapse in addition to drug resistance, signifying a fundamental role in cancer progression. Recent investigations have shown an upregulation of CCL5 in breast cancer tissues compared with normal ones. Increased CCL5 levels can recruit monocytes in the tumor microenvironment, thus promoting breast cancer progression. CCL5 also enhances breast cancer progression in a p53-dependent manner through CCR5. Similar to our data obtained from the PAM50 classification analysis, other studies have also found elevated CCL5 levels in triple-negative breast cancer (TNBC). Concordantly, our results showed an enhanced CCL5 expression, thus indicating a plausible association of CCL5 with breast cancer progression upon WPS use. Indeed, this association between CCL5 expression and aggressive and non-remissive breast cancer might be due to its ability to trigger the release of matrix-metalloproteinase; our data analysis identified MMP9 as a target gene. An earlier study showed that the overexpression of MMP9 linked with the progression of dysplasia to breast cancer; its elevated expression is found in breast cancer and correlates with poor disease prognosis. On the other hand, the overexpression of CCL5 is involved in enhancing tumor tolerance leading to poor prognosis in...
CCL5 upregulation is also associated with non-remissive and later stage breast cancer. This could be due to its ability to augment MMP9 and monocyte migration, thus promoting angiogenesis and tumor

Figure 4: Association between deregulated genes under the effect of water-pipe smoking and prognosis in breast cancer patients using the Kaplan–Meier plotter database expressed by relapse-free survival
Previous reports have shown a positive MMP9 to be associated with a shorter RFS and breast cancer-related survival.\(^{[58,59]}\) Our results also demonstrated that CCL5 as well as MMP9 were targets of WPS in human normal mammary cells. Intriguingly, ALOX5 facilitated an invasion via MMP9 stimulation; enhanced ALOX5 expression plays a role in tumor pathogenesis.\(^{[60]}\) Furthermore, tumor-initiating genes associated with ALOX5 expression enhance mitogenesis, mutagenesis, angiogenesis, cell survival, immunosuppression and metastasis in breast cancer.\(^{[61]}\) An earlier study by Wculek et al. reported that neutrophils enhanced ALOX5-dependent breast cancer lung metastasis.\(^{[62]}\) Moreover, the ALOX5 inhibitor, Zileuton, significantly decreased breast cancer metastasis, further supporting our finding of a suggestive role of ALOX5 in breast cancer initiation and progression.\(^{[62]}\) This was similar to our data where we found upregulated

![Image](image-url)

**Figure 5:** Schematic protein interaction analysis of water-pipe smoking -deregulated genes using the Search Tool for the Retrieval of Interacting Genes (STRING v9.1). The enriched biological process and molecular functions of those proteins are included.
ALOX5 significantly correlated with the HER2-positive breast cancer subtype.

Analogous to CCL5, CCL4 has a comparable role in cancer progression; CCL4 induces breast cancer metastasis.\[64\] Recently, a study revealed that smoking in addition to CCL4 polymorphism could pose an elevated risk of breast cancer.\[65\] Similarly, we reported that WPS enhanced the CCL4 expression resulting in an augmented inflammatory response, thus promoting tumor development and progression. Remarkably, CCL8, a monocyte chemo-attractant protein-2, deregulates several cellular processes including proliferation, apoptosis and differentiation as well as enhances the progression of EMT.\[66,67\] CCL8 can also trigger fibroblasts, thus creating a pro-tumor environment specifically in the TNBC stroma and promotes breast cancer metastasis.\[69,68\] Our data were confirmed by previous investigations that found elevated CCL8 expression in breast cancer tissues to be significantly associated with negative hormone receptors, TNBC subtypes, basal-like subtypes, high grade breast cancers and a worse prognosis.\[69,69\] The other identified chemokine, CCL21, also plays a vital role in regulating cellular proliferation, invasion, apoptosis and metastasis.\[69,71\] Smoking enhances blood and bronchoalveolar lavage fluid levels of the CCR7 ligands CCL19 and CCL21\[55\] as well as contributing to the migration of lung cancer cells\[72\] via the EMT and ERK1/2 signaling pathways.\[51\] Numerous reports have shown that CCL21 plays a role in the migratory properties of breast cancer cells.\[73\] Concordant to the data reported by Chen and colleagues, in our study high levels of CCL21 significantly correlated with the basal-like subtype.\[69\] Interestingly, previous studies reported cross-talk of various CC chemokines in breast cancer including CCL8/21; cross-talk between CCL8 and CCL21 is involved in the development and progression of breast cancer and correlates with patient prognosis.\[69\] In concordance with our data, we showed the presence of CCL8/21 in normal mammary epithelial cells when exposed to WPS indicating its role in the transition to cancerous ones. Tumor-necrosis factor superfamily member 14 (TNFSF14), also known as LIGHT, is an inflammatory cytokine and plays a role in the anti-tumor immune response.\[74\] Ganstev and colleagues reported an upregulation of TNFSF14 in newly formed lymph nodes in breast cancer\[75\] this was in concordance with our data and thus suggested a role of TNFSF14 in the onset and progression of breast cancer. Moreover, studies have shown that smoking enhances the expression of TNFSF14,\[76,77\] which is upregulated in female smokers while the expression of TNFSF14 is absent in male smokers.\[76\] This finding supported our data as TNGSF14 expression was enhanced in WPS-induced breast cancer.\[78,79\]

Subsequently, in our study we identified TLR9, a gene involved in the innate immune system. Studies have shown an elevated expression of TLR9 in breast cancer, which was associated with tumor grade.\[80-82\] An in vitro study by Merrell et al. reported upregulated TLR9 expression in the TNBC cell line (MDA-MB-231) and indicated a plausible role of TLR9 in tumor growth, progression and metastasis.\[81\] Intriguingly, several investigations have reported cigarette/e-cigarette smoke to elevate TLR9 expression;\[83-85\] these studies correlated with our data where we demonstrated exposure to WPS smoke-induced TLR9 expression in breast cancer. Furthermore, we identified prostaglandin reductase 1 (PTGR1), a metabolic enzyme involved in blocking a chemotactic factor, leukotriene B4. A previous study showed elevated PTGR1 expression in several breast cancer cell lines including HER2-positive and TNBC cell lines with the highest expression present in the TNBC cell line, HCC1937\[86\] further supporting our data. Another investigation also reported the expression of PTGR1 to correlate with TNBC pathogenicity; the silencing of PTGR1 with licochalcone A decreased the TNBC pathogenicity.\[87\]

On the contrary, cytokine IL3, a selective growth factor, is released by a subset of tumor-infiltrating T-cells in breast cancer tissues stimulating tumor angiogenesis.\[88\] In our study we found an increase of IL3 expression in human mammary epithelial cells upon WPS exposure. Additionally, the upregulated expression of IL3 has been shown to play a role in breast cancer bone metastasis;\[89\] this further confirmed the strong association between WPS upregulated genes and breast cancer tumor progression. In this study, we also identified another type of cytokine, interferon gamma (IFNG). Our data showed elevated IFNG expression in mammary epithelial cells exposed to WPS. An earlier investigation showed that breast cancer cells exhibited enhanced IFNG expression\[90\] thus promoting cancer invasion and angiogenesis.\[91\] In breast cancer, the key pathway associated with a prolonged RFS focuses on the immune response with IFNG signaling being one crucial pathway.\[92\] On the other hand, we also identified an interferon-related gene, MX1, which is upregulated in breast cancer,\[93\] concordant with our finding, MX1 levels were elevated in mammary epithelial cells exposed to smoke. Furthermore, similar to our data, mRNA and protein levels of MX1 were found to be increased in both in vivo and in vitro tamoxifen and fulvestrant resistance experimental models\[94-96\] indicating MX1 involvement in RFS and a poor prognosis. Recently, a study showed the involvement of the PI3K/AKT pathway in enhancing MX1 expression, which can be linked with the stimulation of growth signaling pathways in relapsing patients.\[97\]

We herein identified the G Protein Subunit Gamma Transducin 1 (GNGT1) gene, which encodes for the guanine nucleotide binding protein (G protein) as a target of WPS. Although previous studies have demonstrated an enhanced expression of GNGT1 in head and neck squamous cell carcinomas,\[98\] lung cancer\[99,100\] as well as liver cancer,\[101\] this was the first study that reported the
overexpression of GNGT1 in breast cancer. As smoking is considered to be a key risk factor for lung cancer,[102] we suggested a plausible role for GNGT1 in WPS-induced breast cancer. Furthermore, GNGT1 correlated with poor overall and progression-free survival in serous ovarian cancer,[103,104] thus supporting our data.

While smoking is a vital etiological factor in the onset and progression of various human cancers including lung and oral as well as breast,[22,19–41,44,105,106] a previous study demonstrated that WPS exposure can stimulate the cell invasion of breast cancer cells.[22] However, an increase in WPS consumption leads to rising levels of toxicant intake; it is postulated that WPS can be a carcinogenic and therefore it can play a plausible role in the development and progression of various types of human cancers as well as cancer-related mortality in comparison with cigarette smoking. Moreover, in this study we identified DEGs that could be plausible therapeutic targets; nevertheless, future studies are essential to validate and determine the mechanisms underpinning WPS-induced breast carcinogenesis.

**Conclusions**

In our study, we revealed for the first time that WPS could plausibly play a role in inducing EMT in HNME cells along with the deregulation of a set of genes responsible of the development and progression of human breast carcinogenesis and RFS. Thus, WPS could enhance breast cancer development and/or its progression predominantly due to its effect on key regulatory carcinogenic genes that have a direct impact on the outcome of breast cancer patients. However, further research is warranted to further elucidate the underlying mechanism underpinning WPS-induced human breast carcinogenesis.

**Data availability statement**

Data supporting the reported results are contained within the article or Supplementary Materials.

**Acknowledgments**

We would like to thank A. Kassab for her critical reading of the manuscript.

**Financial support and sponsorship**

This work is supported by Qatar University, grant numbers: QUCG-CMED-20/21-2 and QUHI-CMED-19/20-1. The funders had no role in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript or in the decision to publish the results.

**Conflicts of interest**

There are no conflicts of interest.

**References**


79. Maker AV. Precise identification of immunotherapeutic targets for solid malignancies using clues within the tumor microenvironment-evidence to turn on the LIGHT. Oncoimmunology 2016;5:e1069937.


Supplementary Figure 1: MRNA expression levels of water-pipe smoking -deregulated differentially expressed genes in normal tissue in comparison with invasive breast cancer using (a) TCGA dataset, (b) Finak dataset and (c) Zhao dataset comprised in the Oncomine database.
Supplementary Figure 3: Association between deregulated genes under the effect of water-pipe smoking and prognosis in breast cancer patients using the Kaplan–Meier plotter database expressed by overall survival.

Supplementary Figure 2: DNA copy number of the top water-pipe smoking-deregulated differentially expressed genes in smoker versus never-smoked breast cancer patients using the Bittner Breast dataset and the Oncomine database. The band in the middle of the box represents the median DNA copy number, while the top and bottom of each box represents the distance between quartile 1, quartile 3 as well as 1.5 times the interquartile range.

Supplementary Figure 3: Association between deregulated genes under the effect of water-pipe smoking and prognosis in breast cancer patients using the Kaplan–Meier plotter database expressed by overall survival.