Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

Articles

AKR1C1 as a Biomarker for Differentiating the Biological Effects of Combustible from Non-Combustible Tobacco Products

Sangsoon Woo [1,†], Hong Gao [2,†], David Henderson [1], Wolfgang Zacharias [2,3], Gang Liu [4], Quynh T. Tran [4,*] and G. L. Prasad [4]

[1] Statistical Genetics, Axio Research LLC, 4th Ave. Suite 200, Seattle, WA 98121, USA; sangsoonw@axioresearch.com (S.W.); davidh@axioresearch.com (D.H.)
[2] Department of Medicine, James Graham Brown Cancer Center, University of Louisville School of Medicine, Louisville, KY 40202, USA; kygaohong@gmail.com (H.G.); wolfgang.zacharias@louisville.edu (W.Z.)
[3] Department of Pharmacology and Toxicology, University of Louisville School of Medicine, Louisville, KY 40202, USA
[4] RAI Services Company, 401 N. Main Street, Winston-Salem, NC 27101, USA; liug@rjrt.com (G.L.); prasadg@rjrt.com (G.L.P.)
* Correspondence: trann1@rjrt.com; Tel.: +01-336-741-2772
† Co-first authors

Abstract Smoking has been established as a major risk factor for developing oral squamous cell carcinoma (OSCC), but less attention has been paid to the effects of smokeless tobacco products. Our objective is to identify potential biomarkers to distinguish the biological effects of combustible tobacco products from those of non-combustible using oral cell lines. Normal human gingival epithelial cells (HGEC), non-metastatic (101A) and metastatic (101B) OSCC cell lines were exposed to different tobacco product preparations (TPPs) including cigarette smoke total particulate matter (TPM), whole-smoke conditioned media (WS-CM), smokeless tobacco extract in complete artificial saliva (STE), or nicotine (NIC) alone. We performed microarray-based gene expression profiling and found 3456 probe sets from 101A, 1432 probe sets from 101B, and 2717 probe sets from HGEC to be differentially expressed. Gene Set Enrichment Analysis (GSEA) revealed xenobiotic metabolism and steroid biosynthesis were the top two pathways that were upregulated by combustible but not by non-combustible TPPs. Notably, aldo-keto reductase genes, AKR1C1 and AKR1C2, were the core genes in the top enriched pathways and were statistically upregulated more than 8 fold by combustible TPPs. Our qRT-PCR results statistically support AKR1C1 as a potential biomarker for differentiating the biological effects of combustible from non-combustible tobacco products.

Keywords: aldo-keto reductases; cigarette smoke; smokeless tobacco products; nicotine; oral cavity cells; xenobiotic metabolism

1. Introduction

Oral squamous cell carcinoma (OSCC) is the most common malignancy of the head and neck, with a worldwide incidence of 300,000 new cases annually. One of the major risk factors for OSCC is exposure to cigarette smoke, considered to be responsible for 50-90% of cases worldwide, and the incidence of OSCC in cigarette smokers is 7–10 times higher than never smokers [1-3]. Cigarette smoke is primarily considered to be a tumor-promoting and co-carcinogenic agent and only a weak complete carcinogen [4-6]. Cigarette smoke or its components cause DNA mutations and chromosomal damage, protein modifications, and expression changes of genes involved in cell death, inflammation, DNA repair and cell cycle regulation [4,7-11]. In addition, the molecular response to cigarette smoke exposure is cell- and tissue-specific and varies with the type of exposed target organ and tobacco product [7,12,13].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

Cigarette smoke contains thousands of chemicals, including those designated as Harmful and Potentially Harmful Compounds [14]. Metabolic activation by Phase I detoxification enzymes such as cytochromes P450 (CYP) and aldo-keto reductases (AKRs) that produce highly reactive carcinogenic electrophiles is an important step in the metabolism of the smoke toxicants [15-19]. For example, Polcyclic Aromatic Hydrocarbons (PAHs) are activated to genotoxic intermediates through different primary pathways, One of the pathways involves dihydrodiol dehydrogenases, members of the aldo-keto reductase (AKR) superfamily that includes AKR1A1, AKR1C1, AKR1C2, AKR1C3 and AKR1C4 [20]. Activation of PAHs by the AKR family of enzymes leads to PAH metabolites that can form DNA adducts or reactive oxygen species (ROS) resulting in oxidative DNA damage [19,21-25].

Relative to cigarette smoking, the use of smokeless tobacco is recognized as less harmful [26,27]. While a significant body of knowledge exists from in vitro and clinical studies on the effects of exposure to cigarette smoke constituents on oral cells, there is limited information on how smokeless tobacco alters oral cell biology. We have utilized a set of normal and malignant oral carcinoma cell lines to investigate how exposure to moist snuff (used as a representative of smokeless tobacco) impacts oral biology. Our previous work indicated that exposure to moist snuff is significantly less cytotoxic [13], and that moist snuff minimally causes DNA damage [28] compared to the cytotoxicity and DNA damage caused by the constituents of cigarette smoke. In addition, nicotine exerted detectable adverse effects only at very high doses [28]. In this study, we characterized the effects of different tobacco products on the transcriptome of human oral cavity cells in order to identify gene-based biomarker(s) [13,28].

While evidence for the effects of exposure to smoke (or its components) on transcriptome of normal or oral cavity cancer cell lines is abundant, not much attention has been given to the effects of exposure to smokeless tobacco products [29-32]. In this study, our objectives were to characterize the effects of different tobacco products on the transcriptome of human oral cavity cells and to identify a transcriptome-based biomarker(s) and ultimately provide the basis for risk reduction strategies in smokeless tobacco products. Our group has shown that combusted tobacco products produced much higher cytotoxicity, cell death, and DNA damage in normal human oral epithelial cells and malignant oral carcinoma cell lines compared to non-combusted smokeless tobacco products including nicotine itself [13,28]. In this study, the total particulate matter (TPM) fraction prepared from cigarettes, whole-smoke conditioned media (WS-CM), low and high dose smokeless tobacco extract in complete artificial saliva (low-STE and high-STE) and nicotine at low and high dose (low-NIC and high-NIC) were used for exposure of human gingival epithelial cells (HGEC), non-metastatic (101A), and metastatic oral carcinoma (101B) cells. Our transcriptomic and gene set enrichment analysis revealed that the top two upregulated KEGG pathways by combustible TPPs (TPM and WS-CM) were metabolism of xenobiotics by cytochrome P450 pathway and steroid hormone biosynthesis. The analysis also suggested that AKR1C1 and AKR1C2, members of the aldo-keto reductase family, were the two core genes induced by combustible, but not by smokeless TPP products. Although our qRT-PCR results confirmed the induction of these two genes, only the increase of AKR1C1 gene expression level was statistically significant and thus, supporting this gene as a potential biomarker candidate for harm reduction of smokeless tobacco products. Further, identification of genes whose expression is specifically modified by exposure to certain tobacco product preparations (TPPs) will provide a better understanding of their mechanisms of action, and allow the development of sensitive and specific biomarkers for both exposure and any potential reduced harm effects of tobacco products.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

2. Materials and Methods

2.1. Reagents:

RNeasy mini kit was purchased from Qiagen (Valencia, CA). qPCR reagents SuperScript® VILO™ cDNA synthesis kit, TaqMan® Fast Advanced Master Mix, AKR1C1 and AKR1C2 TaqMan® Gene Expression Assay were obtained from Life Technologies (Carlsbad, CA). Complete Lysis-M was obtained from Roche (Indianapolis, IN). Pierce ECL Western Blotting Substrate and Restore Western Blot Stripping Buffer were from Thermo Fisher Scientific (Rockford, IL. NIC (nicotine) was obtained from Sigma Chemical Company (St. Louis, MO).

2.2 Tobacco product preparations and chemical analysis:

TPM was prepared by Labstat International (Kitchener, Ontario, Canada) from 3R4F reference cigarettes with a 20-port Borgwaldt-KC smoke machine using the International Organization for Standardization smoking regime (35 ml puff volume, 60 sec puff intervals, 2 sec puff duration). TPM was collected on glass fiber filters (Cambridge), and the amount of TPM obtained determined by weight increase of the filter. The collected particulate phase was dissolved in dimethylsulfoxide (DMSO) to a tar concentration of 20 mg/ml stock solution; DMSO was used as solvent control [12,13].

ST/CAS, also prepared by Labstat International, was obtained by extracting 2.5 g of 2S3 smokeless tobacco (reference moist snuff) with 25 ml complete artificial saliva (CAS) for 2 hours, yielding a 10% (w/v) ST/CAS stock solution; CAS was used as solvent control [13,33,34].

WS-CM was freshly prepared from two 3R4F cigarettes with a Borgwaldt-KC single-port smoke machine (Borgwaldt, Richmond, VA) through 10 ml cell culture medium using the same smoke regime as above for TPM. Following previously established protocols [12], Phenol Red-free media were used for these preparations (DMEM without Phenol Red for 101A, 101B cells; Invitrogen Epi-Life media for HGECs). Mainstream smoke was passed through the respective media yielding a 20% stock solution (2 cigarettes per 10 ml media); parent media was used as solvent control [12,13]. For each condition, the dilution factors and NIC equivalents delivered are listed in **Table S1**.

Chemical analyses were performed on representative batches of the respective reagent stocks (TPM, 20 mg/ml; ST/CAS, 10%; WS-CM, 20%) for nicotine, pH, and key TSNAs. The amounts of nicotine contained in TPM and ST/CAS were in the same range, and were substantially higher than those contained in the WS-CM preparation. Similarly, the levels for selected TSNAs were in the low ng - range for TPM and ST/CAS, but were below detection limit for both WS-CM preparations. The TPPs were applied as dilutions of the respective stocks defined in Materials and Methods per the data acquired from the chemical analyses. This allowed for exposures to be conducted on relative nicotine levels in addition to the percent concentration or dilution of the test preparations. Chemical analyses for nicotine and tobacco-specific nitrosamine (TSNA) content in all preparations were performed by Labstat International, Kitchener, Ontario, Canada [34] (**Table S2**).

2.3. Cell lines and cultures:

Human oral squamous cell carcinoma cell lines 101A (UM-SCC-101A, primary tonsil tumor) and 101B (UM-SCC-101B, lymph node metastasis) were obtained from Dr. T. Carey (Univ. of Michigan) [35,36]. Both cells were grown in DMEM with 10% fetal bovine serum (Atlanta Biologicals, Lawrenceville, GA), 100 U/ml of penicillin-streptomycin, 2 mM glutamine, and 100 nM nonessential amino acids. Primary human gingival epithelial cells (HGEC) were obtained from the University of Louisville School of Dentistry (Dr. D. Kinane) [37,38]. HGECs were grown in keratinocyte-serum-free medium (Invitrogen, Carlsbad, CA) containing 10 µg/ml of insulin, 5 µg/ml of transferrin, 10 µM of 2-mercaptoethanol, 10 µM of 2-aminoethanol, 10 mM of sodium selenite, 50 µg/ml of bovine pituitary extract, 100 U/ml of penicillin-streptomycin, and 50 ng/ml of fungizone. All cells were grown at 37ºC in a humidified atmosphere with 5% CO_2.

2.4 Treatments:

Target cells were exposed for 24 hours to different TPPs or solvent only as vehicle controls. Doses were chosen based on previous systematic dose- and time-dependent cytotoxicity studies [13].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

For TPM and WS-CM, EC-30 doses were applied; for ST/CAS, the dose with the same amount of NIC as that in TPM at EC-30 was applied since no EC-30 could be determined. In addition, low and high doses of NIC alone were used as controls. The low dose (14 ug/ml) represented the level of nicotine in the combustible TPP exposures, while the high dose (474 ug/ml) was used to reach exposure level nearing the cytotoxic response in the TPP exposures. The high NIC dose was chosen in this range because it was previously shown that NIC at close to millimolar doses had some cytotoxic effects on other cell types [12] and since we had determined such range yielded 20% cytotoxicity in oral cells [13]. NIC vs DMSO group comparisons were used mainly as controls for exposure and were not included for microarray expression profiling in the 101B cell line.

2.5 Microarray gene expression profiling:

Cells were grown to approximately 75% confluence in 6-well plates, treated for 24 hours, and total RNA was subsequently isolated with TriZol™ reagent (Invitrogen, Carlsbad, CA). The purity and integrity of each RNA preparation was evaluated by using RNA Nano Chips on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Micoarray expression analysis was performed on a Microarray Analysis Suite instrument system (Affymetrix, Santa Clara, CA) using Affymetrix HG-U133A v2 arrays. Sample preparations, labeling, hybridizations, and scanning were performed according to the established Affymetrix protocols.

All sample data were scaled to a uniform target intensity of 500 to allow quantitative comparisons across treatments. RMA normalization was applied to raw data across all treatment and control samples to normalize data at the probe set level. Probe sets with too small or large variation across all samples were filtered out and control adjustment within treatment was applied using treatment specific control samples. The differences in logarithmic intensities between each treatment and its own control were calculated. Principal Component Analysis (PCA) was performed and identified no outliers (data not shown). The data was submitted to Gene Expression Omnibus repository (GSE89923).

To identify differentially expressed probe sets in any TPPs, a linear regression model was applied on the calculated differences to each probe set using an Empirical Bayes method (see Appendix A) for obtaining moderated estimates of the model F-statistic which uses information from variation across all probe sets as developed in the limma Bioconductor package in R [39]. This allowed us to test whether any of these 6 differences was different from 0; in other words, if any of the 6 treatments had an effect on the mRNA abundance levels in each cell line. The Benjamini & Yekutieli (BY) false discovery rate (FDR) procedure was performed to correct for multiple hypothesis testing [40]. An adjusted p-value < 0.01 was considered to be statistically significant when applicable. Fold change was calculated by raising the log fold change value to the power of 2. A two tailed Student's t-test was used to determine the significance of the treatment effects within probe set: a value of $p<0.05$ was considered to be statistically significant when applicable.

2.6 Gene Set Enrichment Analysis (GSEA):

To assess the association of a collection of gene sets with sensitivity to each of the treatment, we used version 2.2.2 of the GSEA tool developed by the Broad Institute [41]. Gene lists were ranked by fold-change prior to running GSEA at the default setting except with the permutation type set to "gene_set" since our sample size in each treatment was less than 7; the seed for permutation was set to 149 for reproducibility purposes. A suggestive false discovery rate (FDR) of 0.25 was used in this exploratory discovery phase of identifying candidate gene sets that were influenced by each TPP.

2.5 Quantitative RT-PCR:

RNA was isolated using RNeasy mini kit (Qiagen) following the manufacturer's instructions, with the exception that TriZol™ reagent (Invitrogen) was used to disrupt and homogenize cells instead of Buffer RLT provided by the kit. The integrity of RNA was confirmed by Agilent Bioanalyzer analysis. RNA was reverse transcribed into first strand cDNA using SuperScript® VILO™ cDNA synthesis kit (Thermo Fisher). Specifically, 100 ng RNA was mixed with 4 µl 5X

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

VILO™ Reaction Mix, 2 µl 10X SuperScript® Enzyme Mix, 2 µl 50 ng/µl RNA, and 12 µl RNase free water, for a 20 µl reaction. The reaction mixture was incubated at 42°C for 60 min and then heated at 85°C for 5 min to terminate the reaction.

qRT-PCR was performed using the TaqMan® Gene Expression system according to the manufacturer's instruction (Applied Biosystems). Briefly, 10 µl reaction mixture contained 5 µl 2X TaqMan® Fast Advanced Master Mix, 0.5 µl AKR1C1 or AKR1C2 TaqMan® Gene Expression Assay, 1 µl of cDNA template, 3.5 µl RNase free water. The thermal cycling conditions included an initial denaturation step at 95°C for 20 sec, 40 cycles at 95°C denature for 1 s, 60°C annealing and extension for 20 s. Each reaction was performed in triplicate and no-template controls were included in each experiment. Reactions were run in ViiA™ 7 Real-Time PCR System (Life Technologies). The cycle threshold (C_T) values were normalized to 18S ribosomal RNA and the fold change was calculated using $2^{-\Delta\Delta C_T}$ method [42]. A linear regression model was fitted for AKR1C1 and AKR1C2 expression levels to identify statistical significance of treatment effects.

3. Results

3.1. Microarray expression profiling:

We sought to evaluate the gene expression changes in HGEC, 101A, and 101B cell lines after treating with combustible and non-combustible TPPs. The regression model identified 2,717 probe sets in HGEC, 3,456 probe sets differentially expressed in 101A, and 1,432 probe sets in 101B (**Fig. 1**). In general, the gene expression patterns affected by TPPs were more similar between HGEC (**Fig. 1A**) and 101A (**Fig. 1B**) cell lines, but they were different from those expressed in 101B cell lines (**Fig. 1C**). This suggests that once the cells were metastatic as in 101B, they responded very differently from normal gingival epithelial cells and non-metastatic OSCC cells when treated by different TPPs. In HGEC, the gene expression profiles affected by combustible TPM and WS-CM clustered together and separated from the cluster of non-combustible TPPs, in which low-NIC and high-NIC grouped with low-STE (**Fig. 1A**). In 101A and 101B, the grouping of gene expression changes by WS-CM was more similar to those changes by low-STE (**Fig. 1A and 1B**). In all three cell lines, high-STE exerted an opposite effect on gene expressions compared to the effects produced by other TPPs; indicating that at this high dose of STE, distinct biological effects may occur (**Fig. 1**).

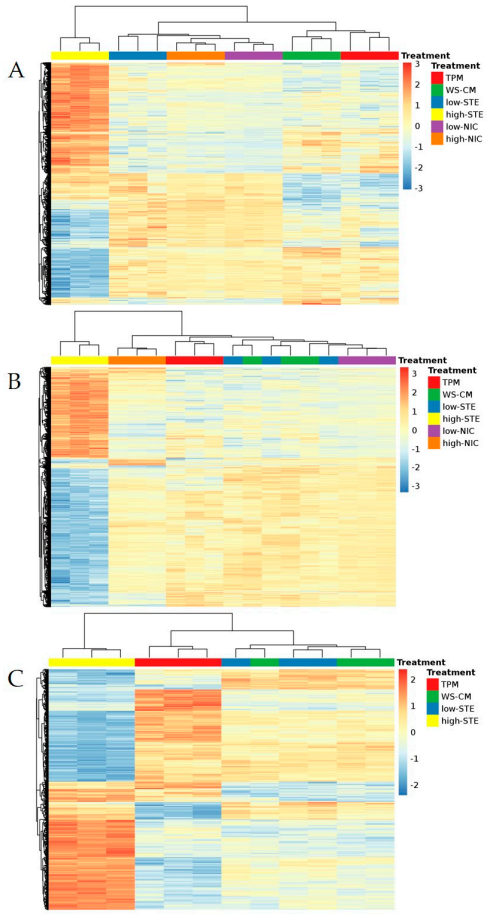

Figure 1. Heat maps of the differentially expressed probe sets in each cell line: A) HGEC cell line, B) non-metastatic 101A cell line, C) metastatic 101B cell line. Hierarchical clustering was performed to cluster samples and probe sets into similar clusters. The color scale represents the log2 fold change ranging from -3 (dark blue) to 3 (dark red). Different treatments were color coded as follow: red=TPM, blue=low-STE, green=WS-CM, purple=low-NIC, orange = high-NIC, and yellow=high-STE.

Combusted TPM affected 1264 (of which 438 were upregulated), 846 (of which 288 were upregulated), and 861 (of which 455 were upregulated) probe sets accordingly in HGEC, 101A, and 101B (**Fig. 2A**). WS-CM statistically altered 1434 (of which 594 were upregulated), 175 (of which 123 were upregulated), and 72 (of which 31 were upregulated) probe sets in HGEC, 101A, and 101B cell lines, respectively (**Fig. 2B**). We observed that high-STE affected mRNA levels of thousands of genes, accounting for almost all the changes in gene expression while the low-STE and low-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

NIC itself affected the least number of genes in each cell line. Specifically, high-STE led to at least a 1.3-fold change in the expression of 2415 (of which 1323 were upregulated), 3369 (of which 1288

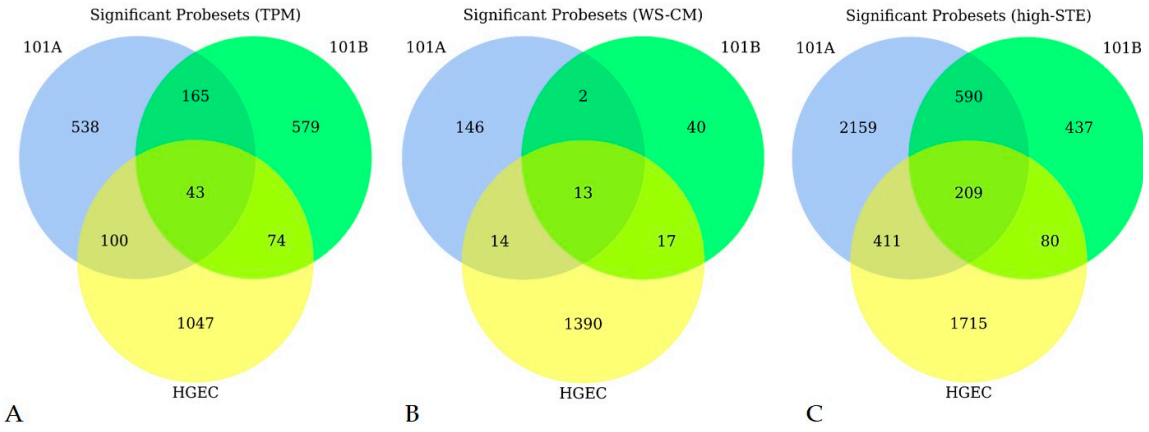

Figure 2. Venn diagrams showing the number of significant probe sets for each treatment in each cell line. A) Under TPM condition, B) Under WS-CM condition, C) Under high-STE. Blue circle=101A, green circle=101B, and yellow circle=HGEC.

were upregulated), and 1316 (of which 696 were upregulated) probe sets in HGEC, 101A, and 101B, respectively (**Fig. 2C** and **Tables S3-S5**). Similarly, high-NIC affected 1040 (of which 404 were upregulated) and 284 (of which 190 were upregulated) probe sets in HGEC and 101A cell lines, respectively (**Tables S3-S4**). The top ten probe sets that were significantly changed (t-test p-values <0.01) by TPM in HGEC and 101A were mapped to the aldo-keto reductase genes (AKR1C1 and AKR1C2), cytochrome P450 genes (CYP1B1, CYP24A1), NAD(P)H quinone dehydrogenase 1 (NQO1), and E74-like ETS transcription factor 3 gene (ELF3) (**Tables S3 and S4**). On the other hand, in 101B, TPM significantly changed probe sets that were mapped to matrix metalloprotease genes (MMP1, MMP3), proteasome subunit beta genes (PSMB9, PSMB8), butyrophilin subfamily 3 genes (BTN3A3, BTN3A2), and interleukin genes (IL6, IL6R) (**Table S5**). We also found a set of genes that were robustly affected by combustible TPPs across all three cell lines. **Fig. 2A** shows that TPM commonly affected 43 probe sets which were mapped to 36 genes. Similarly, WS-CM robustly changed 13 probe sets or 7 genes in all 3 cell lines (**Fig. 2B**). Six of the 7 genes upregulated by WS-CM were also included in the 36 genes differentially changed by TPM after FDR adjusted. Notably, none of these 6 genes were affected by low-STE nor low-NIC (**Table 1**). They include detoxification and anti-oxidant genes such as AKR1C1, AKR1C2, SLC7A11 (solute carrier family 7 member 11), HMOX1 (heme oxygenase 1), GPX2 (glutathione peroxidase 2), and PTGR1 (prostaglandin reductase 1). When the 3 cell lines were treated with high-STE, 209 probe sets were commonly affected (**Fig. 2C**). However, there were no common genes affected by low-STE across all three cells lines, nor by low-NIC in 101A and HGEC (**Fig. S1A and S1B**). High-NIC commonly affected only 43 probe sets (**Fig. S1C**).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

Table 1. Six common genes that are robustly upregulated by TPM and WS-CM in all 3 cell lines

Affy ID	Gene Name	Log_2 FC		Adjusted p-values	
		TPM	WS-CM	TPM	WS-CM
1562102_at		3.45	4.96	4.4591E-07	3.0208E-15
1555854_at	AKR1C1, AKR1C2	3.12	3.52	1.5291E-07	2.9126E-08
216594_x_at	AKR1C1, AKR1C2	2.37	3.01	1.4115E-10	1.9914E-07
204151_x_at	AKR1C1, AKR1C2	2.38	2.65	5.6067E-12	9.951E-08
211653_x_at	AKR1C1, AKR1C2	2.28	2.75	4.4159E-10	7.8189E-07
209699_x_at	AKR1C1, AKR1C2	2.01	2.90	1.8013E-09	3.075E-07
203665_at	HMOX1	3.87	4.72	1.7717E-10	8.392E-12
207528_s_at	SLC7A11	3.21	3.31	1.0043E-09	6.3878E-10
217678_at	SLC7A11	2.96	2.59	8.7581E-06	2.4987E-09
209921_at	SLC7A11	2.94	2.55	8.7581E-06	2.7297E-09
202831_at	GPX2	1.45	2.38	0.00188523	8.0032E-10
231897_at	PTGR1	0.80	1.73	0.00504145	5.9494E-11

3.2. Gene Set Enrichment Analysis (GSEA):

We observed that only in HGEC, the changes on gene expression by combustible TPPs can be distinctively separated from those changes by non-combustile TPPs (**Fig. 1**). As such, we focused our next pathway analysis in HGEC only. To identify candidate KEGG gene sets that were associated with each treatment, we performed GSEA. We ranked each gene list based on fold change, and then loaded each of them to GSEA tool. **Table 2** shows the top 3 positive and 3 negative gene sets that were overrepresented by each TPP in HGEC. Among these top gene sets, the metabolism of xenobiotics by cytochrome P450 and steroid hormone biosynthesis KEGG pathways were upregulated by combustible TPM and WS-CM and by high dose of 2S3 smokeless tobacco (high-STE). The core genes for this enriched pathway includes cytochrome P450 family 1 (CYP1A1, CYP1B1), aldo-keto reductase family 1 (AKR1C1, AKR1C2), and UDP-glucosyltransferase (UGTA1) (**Tables S5**). TPM and WS-CM commonly downregulated gene sets involved in cell cycle and DNA replciation.

The low-STE did not enrich any KEGG pathway. The 3 pathways that were most negatively affected by low-STE are gene sets involved in phenylalanine metabolism, glycerine, serine, and theronine metabolism, and olfactory transduction (**Table 2**). Besides xenobiotic metabolism and steroid hormone biosynthesis, high-STE also upregulated gene sets involved in arachidonic acid metabolism while it down-regulated gene sets in peroxisome, cytosolic DNA sensing and RIG1 like receptor signaling pathways. High-NIC did not induced any KEGG gene set, but it suppressed DNA replication, systemic lupus erythematosus, and fatty acid metabolism (**Table 2**). None of the KEGG gene sets was significantly affected by low-NIC (**Table 2**). The complete GSEA results is provided in **Table S6**.

Table 2. The top 3 up- and 3 down-regulated KEGG Pathways from Gene Set Enrichment Analysis for each TPP in HGEC

KEGG PATHWAY	TPM	WS-CM	Low-STE	High-STE	Low-NIC	High-NIC
	q-value (NES)*	q-value (NES)*	q-value (NES)*	q-value (NES)*	q-value (NES)*	q-value (NES)*
METABOLISM_OF_XENOBIOTICS_BY_CYTOCHROME_P450	0 (2.41)	0 (2.37)	--	0.0095 (1.90)	--	--
STEROID_HORMONE_BIOSYNTHESIS	0 (2.36)	0.0014 (2.05)	--	0.0050 (2.00)	--	--
RETINOL_METABOLISM	0.0018 (2.02)	--	--	--	--	--
PHENYLALANINE_METABOLISM	--	--	0.0475 (-1.73)	--	--	--
PORPHYRIN_AND_CHLOROPHYLL_METABOLISM	--	0.0616 (1.78)	--	--	--	--
SYSTEMIC_LUPUS_ERYTHEMATOSUS	--	--	--	--	--	0.1501 (-1.72)
CELL_CYCLE	0 (-2.48)	0 (-2.30)	--	--	--	--
DNA_REPLICATION	0 (-2.30)	0 (-2.18)	--	--	--	0.1628 (-1.67)
OOCYTE_MEIOSIS	--	0.0007 (-2.05)	--	--	--	--
MISMATCH_REPAIR	0.0016 (-1.98)	--	--	--	--	--
GLYCINE_SERINE_AND_THREONINE_METABOLISM	--	--	0.0273 (-1.85)	--	--	--
OLFACTORY_TRANSDUCTION	--	--	0.0286 (-1.80)	--	--	--
ARACHIDONIC_ACID_METABOLISM	--	--	--	0.0050 (1.95)	--	--
CYTOSOLIC_DNA_SENSING_PATHWAY	--	--	--	0.0186 (-1.93)	--	--
PEROXISOME	--	--	--	0.0942 (-1.75)	--	--
RIG_I_LIKE_RECEPTOR_SIGNALING_PATHWAY	--	--	--	0.0676 (-1.74)	--	--
FATTY_ACID_METABOLISM	--	--	--	--	--	0.2090 (-1.75)

*NES: normalized enrichment score. It reflects the degree to which a gene set is overrepresented at the top or bottom of a ranked genes list and accounts for differences in gene set size and in correlations between gene sets and the expression dataset. Positive NES indicates gene set enrichment at the top of the ranked list while a negative NES indicates gene set enrichment at the bottom.

*FDR q-value is the estimated probability that a gene set with a given NES represents a false positive finding. The suggestive FDR cutoff of 25% was utilized.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

3.4. AKR1C1 and AKR1C2 as candidate biomarkers for the effects of combustible tobacco products:

From the GSEA results, we observed that the gene set in the metabolism of xenobiotics by cytochrome P450 KEGG pathway was upregulated in all cell lines only by combusted TPM or WS-CM, but not in non-combusted STE or NIC itself except at the high-STE. Among the core genes that contributed the most to the enrichment of the xenobiotic metabolism pathway and steroid hormone biosynthesis, AKR1C1 and AKR1C2 were the only two common genes that were statistically upregulated more than 2 fold by combusted TPM and WS-CM (t-test p-values < 0.01), but not by non-combusted low-STE or NIC in all 3 cell lines (**Table S3-S5**). These two genes were also among the 6 most robust genes affected by combustible TPPs (**Table 1**). At the high dose of non-combusted STE, the levels of AKR1C1 and AKR1C2 expression were also upregulated; however, this effect and the effect exerted on other thousands of genes under the treatment of this high dosage may be related to the increase in concentrations of other ST extract components that were not present at the low dosage of ST/CAS. Thus, the upregulation of AKR1C1 and AKR1C2 mRNA levels and their association with the metabolism of xenobiotics by cytochrome P450 KEGG pathway and steroid hormone biosynthesis appeared to be strongly regulated by combustible TPPs. As such, they can serve as the candidate biomarkers for the effects of combustible tobacco products on normal gingival epithelial cells (HGEC).

Because the analysis suggested that AKR1C1 and AKR1C2 were the two genes most statistically significant and induced by combustible TPPs across three cell lines, we performed quantitative real-time PCR (qRT-PCR) to validate the changes in the mRNA levels of AKR1C1 and AKR1C2 in HGEC when treated with TPM, WS-CM, low-STE, and high-NIC. We observed that TPM and WS-CM significantly induced the mRNA levels of AKR1C1 to 16 and 33 fold, respectively (**Table 3**). The combustible TPPs also induced AKR1C2 mRNA levels about 2 fold. Although this induction was not statistically significant, it agrees with the upregulated trend detected by microarray. In agreement with microarray, low-STE and high-NIC did not change the mRNA levels of AKR1C1 and AKR1C2 (**Table 3**).

Table 3. Microarray results and real time qRT-PCR results with Student's t-test p-value for AKR1C1 and AKR1C2 mRNA level in HGEC

TPPs	microarray		qRT-PCR			
	AKR1C1/AKR1C2		AKR1C1		AKR1C2	
	Fold change	Adj. p-value	Fold change	p-value	Fold change	p-value
TPM	8.69	<0.001	16.43	<0.001	1.95	0.3121
WS-CM	11.47	<0.001	33.28	<0.001	2.07	0.2719
low-STE	1.08	1	1.23	0.5689	1.61	0.8227
high-NIC	1.59	0.5413	1.61	0.2052	0.59	0.4112

4. Discussion

In this study, we further characterized the effects of combustible (TPM and WS-CM) and non-combustible TPPs (ST/CAS low and high dose, NIC alone at low and high dose) on gene expression and cellular pathways in normal human oral cavity cells as well as non-metastatic and metastatic oral squamous cell carcinoma cells. As these were shown to relate to xenobiotic metabolism, steroid biosynthesis, arachidonic acid metabolism, and DNA replication/repair, the results can potentially bring a new level of understanding to the complex biological effects of smoke and smokeless tobacco products in healthy and cancerous conditions. Of special note in this context is that tobacco smoke induced AKR1C1 and AKR1C2 both in normal and non-metastatic cancerous conditions. These genes are in the aldo-keto reductase family that are involved in the metabolism of polycyclic aromatic hydrocarbons (PAH), a family of ubiquitous environmental carcinogens contained in cigarette smoke [4,18,34,43,44]. Cigarette smoke has been shown to induce cyptochrome P450 (CYP1A1, CYP1B1) and aldo-keto reductase (AKR1C1, AKR1C3, AKR1B10)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

genes in oral dysplasia and primary oral carcinomal cell lines [45]. In our study, the data suggested that AKR1C1 and possibly AKR1C2 can be candidate biomarkers for the effects of combustible tobacco products.

Changes in AKR1 gene expression have been reported in diseases linked to smoking, and in vitro treatment with combustible TPPs. For example, overexpression of AKR1C1/2 was observed in non-small cell lung carcinoma [46,47]. *AKR1C1* and *AKR1C3* were also upregulated 15-30 fold in OSCC, and induced by treatment with cigarette smoke condensate in oral dysplasic cells [45]. In addition, *AKR1C1* was part of a gene battery upregulated in buccal oral samples of smokers [29,30]. In bronchial epithelial cell brushes of smokers, *AKR1C1* and *AKR1C2* were two of the most upregulated genes, but their expressions were downregulated in smokers who quit [48,49]. Thus, AKR1C1 and AKR1C2 are consistently overexpressed, in smoking-related cancer and upon smoke exposure. AKR1C1 and AKR1C2 are members of the aldo/keto reductase superfamily that consists of more than 40 known enzymes. They share 98% homology, with only seven different amino acids [50]. They also share high sequence identity with other gene family members, and display overlapping but distinct substrate specificity [21,22,51]. They are monomeric intracellular enzymes that catalyze the conversion of aldehydes and ketones to their corresponding alcohols by utilizing NADH and/or NADPH as cofactors, and bind bile acid with high affinity. Their enzymatic activities play crucial roles in balancing malignant transformation, cancer progression, and response to cytotoxic therapies [25,52-55].

Our data are in agreement with previous observations on the action of xenobiotics on AKR expression. It is known that, in order to minimize the insults caused by cigarette smoke toxicants and other xenobiotics, various tissues are equipped with diverse Phase I (cytochrome p450, CYP) [45,56,57] and Phase II (AKR) [16,21,25,45,50,58] enzymes which are present in abundance either at the basal level or after induction following xenobiotic exposure. Xenobiotics can be conjugated and eliminated by reduction of their carbonyl groups; however, the same reactions can lead to the activation of a pro-drug as well as the inactivation of a drug [17]. On one hand, AKR1C isozymes protect against the harmful effects of reactive oxygen species (ROS), since they catalyze the reduction of 4-hydroxy-2-nonenal, a product of lipid peroxidation [21,52,58]. On the other hand, induced AKR1C isozymes can convert PAH trans-dihydrodiols produced by CYPs to deleterious o-quinones via their dihydrodiol dehydrogenase activity [19,22,23,50]. It was shown that AKR-derived PAH o-quinones will cause change-in-function mutations in the p53 tumor suppressor gene, and that these mutations result from ROS [19]. Also, AKR1C genes can be upregulated by the transcription factors NRF2 and NF-Y which are activated during oxidative stress [59-61]. If o-quinones and the ROS they generate are not eliminated, they can potentially cause covalent and oxidative DNA lesions, increasing the mutational load of PAH-exposed cells. Together, this sequence of events may contribute to PAH-induced oral carcinogenesis and other malignancies [15,23,25,53-55,58,59,62].

The above described interplay between ROS, PAHs, oxidative stress, DNA damage, and AKR enzymes could explain the differential responses of the oral cavity cells to combusted versus non-combusted tobacco products. Clearly, the chemical compositions and carcinogen profiles of burnt tobacco smoke, either collected on filters as TPM or absorbed in media as WS-CM, will be different from those extracted at ambient temperature from non-combusted tobacco material [4,18,34,43]. In general, TPM contains an abundance of low molecular weight organic combustion products like aldehydes, aromatic amines and nitrosamines, phenols, and polycyclic aromatic hydrocarbons [63]. On the other hand, ST preparations contain mostly uncombusted organic compounds like monosaccharides, glycerol, fatty acids, as well as inorganic salts, chlorides, carbonates, and transition metal ions [43].

Chemical analysis showed that in the TPM, ST/CAS, and WS-CM preparations, the TSNAs analyzed (NNN, NAT, NAB, NNK) were consistently several-fold lower in ST/CAS compared to TPM, and were below quantitation limit in WS-CM (**Table S2**). On the other hand, the NIC

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

equivalents contained in each preparation treatment condition were similar in TPM and ST/CAS but several-fold lower in WS-CM (**Table S1**). This indicates that NIC itself has no or only a minimal role in the cellular responses described here, in agreement with our previous observations [12,13].

5. Conclusions

In summary, our global analysis of gene expression profiles in TPP-treated versus untreated human oral cavity cells revealed that two members of the AKR1 gene family are highly induced by short-term exposure to combusted tobacco product components. Induction of some of these genes was previously observed in lung, colon, or rectal tumor tissues of smokers [13,45,64]. However, their selective induction by combusted but not by non-combusted TPPs has not been characterized before. Our results suggest that AKR1C1, and possibly AKR1C2, may be potential biomarkers for reduced harm effects in oral cavity cells when treated with non-combustible (ST/CAS; NIC) compared to combustible (TPM; WS-CM) tobacco products.

Supplementary Materials: The following are available online at www.mdpi.com/link, **Figure S1**: Venn diagram of the significant probe sets in 101A and HGEC cell lines treated under high dose nicotine, **Table S1**: Treatment conditions for microarray gene expression profiling samples. **Table S2**: Summary of chemical analyses for nicotine and TSNAs in representative batches of the TPPs used, **Table S3**: The linear regression results with logFC and p-value for 2717 significant probe sets in HGEC, **Table S4**: The linear regression results with logFC and p-value for 3456 significant probe sets in 101A, **Table S5**: The linear regression results with logFC and p-value for 1432 significant probe sets in 101B, **Table S6**: GSEA results for all treatments in HGEC,

Acknowledgments: This work was supported by a collaborative contract between Reynolds American Inc. (RAI) Services Company and the University of Louisville. Part of this work was performed with the assistance of the University of Louisville Genomics Facility, which is supported by NIH/NIGMS KY-INBRE P20 GM103434 and NIH/NIGMS COBRE-P30 GM106396 awards, the James Graham Brown Foundation, and user fees.

Author Contributions S.W. analyzed the data and drafted the manuscript; H.G. performed the experiments and drafted the manuscript; D.H. contributed to the statistical analysis and the interpretation; W.Z and G.L.P conceived, designed the experiments, interpreted the data, and edited the manuscript; L.G. critically reviewed and edited the manuscript; Q.T.T. analyzed the data, interpreted the results, organized, and wrote the manuscript.

Conflicts of Interest: Three of the authors (Q.T.T., G.L., and G.L.P.) are full time employees of RAI Services Company. This research was funded by RAIS. None of the authors has any other conflict of interest to declare.

Appendix A

Linear modeling with empirical Bayes method at the probe set level

Statistical analysis is performed at the probe set level instead of gene level because some of probe sets mapped onto multiple genes. To identify differentially expressed probe sets, a linear regression model was applied to each probe set using an Empirical Bayes method for obtaining moderated estimates of the model F-statistic which uses information from variation across all probe sets as developed in the Bioconductor package limma in R [39].

The linear regression model for each probe set is

$$log_2\left(Gene\ Expression_{ij}\right) \sim TPM + STE14 + STE474 + WS\_CM + NIC14 + NIC474 + \varepsilon_{ij}$$

$$where\ i = 1, \dots, n\ probe\ sets, j = 1, 2\ for\ cell\ lines\ of\ 101A\ and\ HGEC.$$

$$log_2\left(Gene\ Expression_{ij}\right) \sim TPM + STE14 + STE474 + WS\_CM + \varepsilon_{ij}$$

$$where\ i = 1, \dots, n\ probe\ sets, j = 1\ \ for\ cell\ lines\ of\ 101B.$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

References

1. Jha, P.; Ramasundarahettige, C.; Landsman, V.; Rostron, B.; Thun, M.; Anderson, R.N.; McAfee, T.; Peto, R. 21st-century hazards of smoking and benefits of cessation in the united states. *The New England journal of medicine* **2013**, *368*, 341-350.
2. Lee, P.N.; Hamling, J. Systematic review of the relation between smokeless tobacco and cancer in europe and north america. *BMC medicine* **2009**, *7*, 36.
3. Report, S.G. *How tobacco smoke causes disease: The biology and behavioral basis for smoking-attributable disease*; Centers for Disease Control and Prevention (US): Atlanta (GA), 2010.
4. DeMarini, D.M. Genotoxicity of tobacco smoke and tobacco smoke condensate: A review. *Mutation research* **2004**, *567*, 447-474.
5. Le Houezec, J.; McNeill, A.; Britton, J. Tobacco, nicotine and harm reduction. *Drug and alcohol review* **2011**, *30*, 119-123.
6. Nagler, R.M. Oral cancer induced by cigarette smoke--the mediatory role of the salivary antioxidant system. *Anticancer research* **2003**, *23*, 5125-5128.
7. Rubin, H. Selective clonal expansion and microenvironmental permissiveness in tobacco carcinogenesis. *Oncogene* **2002**, *21*, 7392-7411.
8. Fields, W.R.; Leonard, R.M.; Odom, P.S.; Nordskog, B.K.; Ogden, M.W.; Doolittle, D.J. Gene expression in normal human bronchial epithelial (nhbe) cells following in vitro exposure to cigarette smoke condensate. *Toxicological sciences : an official journal of the Society of Toxicology* **2005**, *86*, 84-91.
9. Nordskog, B.K.; Blixt, A.D.; Morgan, W.T.; Fields, W.R.; Hellmann, G.M. Matrix-degrading and pro-inflammatory changes in human vascular endothelial cells exposed to cigarette smoke condensate. *Cardiovascular toxicology* **2003**, *3*, 101-117.
10. Semlali, A.; Witoled, C.; Alanazi, M.; Rouabhia, M. Whole cigarette smoke increased the expression of tlrs, hbds, and proinflammory cytokines by human gingival epithelial cells through different signaling pathways. *PloS one* **2012**, *7*, e52614.
11. Shishodia, S.; Aggarwal, B.B. Cyclooxygenase (cox)-2 inhibitor celecoxib abrogates activation of cigarette smoke-induced nuclear factor (nf)-kappab by suppressing activation of ikappabalpha kinase in human non-small cell lung carcinoma: Correlation with suppression of cyclin d1, cox-2, and matrix metalloproteinase-9. *Cancer research* **2004**, *64*, 5004-5012.
12. Arimilli, S.; Damratoski, B.E.; Bombick, B.; Borgerding, M.F.; Prasad, G.L. Evaluation of cytotoxicity of different tobacco product preparations. *Regulatory toxicology and pharmacology : RTP* **2012**, *64*, 350-360.
13. Gao, H.; Prasad, G.L.; Zacharias, W. Differential cell-specific cytotoxic responses of oral cavity cells to tobacco preparations. *Toxicology in vitro : an international journal published in association with BIBRA* **2013**, *27*, 282-291.
14. FDA. Harmful and potentially harmful constituents in tobacco products and tobacco smoke; established list. Services, H.a.H., Ed. FDA: 2012; Vol. 77.
15. Beranic, N.; Brozic, P.; Brus, B.; Sosic, I.; Gobec, S.; Lanisnik Rizner, T. Expression of human aldo-keto reductase 1c2 in cell lines of peritoneal endometriosis: Potential implications in metabolism of progesterone and dydrogesterone and inhibition by progestins. *The Journal of steroid biochemistry and molecular biology* **2012**, *130*, 16-25.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

16. Brozic, P.; Turk, S.; Rizner, T.L.; Gobec, S. Inhibitors of aldo-keto reductases akr1c1-akr1c4. *Current medicinal chemistry* **2011**, *18*, 2554-2565.
17. Guengerich, F.P. Metabolism of chemical carcinogens. *Carcinogenesis* **2000**, *21*, 345-351.
18. Rubin, H. Synergistic mechanisms in carcinogenesis by polycyclic aromatic hydrocarbons and by tobacco smoke: A bio-historical perspective with updates. *Carcinogenesis* **2001**, *22*, 1903-1930.
19. Yu, D.; Berlin, J.A.; Penning, T.M.; Field, J. Reactive oxygen species generated by pah o-quinones cause change-in-function mutations in p53. *Chemical research in toxicology* **2002**, *15*, 832-842.
20. Penning, T.M. Aldo-keto reductase regulation by the nrf2 system: Implications for stress response, chemotherapy drug resistance, and carcinogenesis. *Chemical research in toxicology* **2016**.
21. Jin, Y.; Duan, L.; Lee, S.H.; Kloosterboer, H.J.; Blair, I.A.; Penning, T.M. Human cytosolic hydroxysteroid dehydrogenases of the aldo-ketoreductase superfamily catalyze reduction of conjugated steroids: Implications for phase i and phase ii steroid hormone metabolism. *The Journal of biological chemistry* **2009**, *284*, 10013-10022.
22. Jin, Y.; Mesaros, A.C.; Blair, I.A.; Penning, T.M. Stereospecific reduction of 5beta-reduced steroids by human ketosteroid reductases of the akr (aldo-keto reductase) superfamily: Role of akr1c1-akr1c4 in the metabolism of testosterone and progesterone via the 5beta-reductase pathway. *The Biochemical journal* **2011**, *437*, 53-61.
23. Palackal, N.T.; Lee, S.H.; Harvey, R.G.; Blair, I.A.; Penning, T.M. Activation of polycyclic aromatic hydrocarbon trans-dihydrodiol proximate carcinogens by human aldo-keto reductase (akr1c) enzymes and their functional overexpression in human lung carcinoma (a549) cells. *The Journal of biological chemistry* **2002**, *277*, 24799-24808.
24. Penning, T.M.; Burczynski, M.E.; Jez, J.M.; Hung, C.F.; Lin, H.K.; Ma, H.; Moore, M.; Palackal, N.; Ratnam, K. Human 3alpha-hydroxysteroid dehydrogenase isoforms (akr1c1-akr1c4) of the aldo-keto reductase superfamily: Functional plasticity and tissue distribution reveals roles in the inactivation and formation of male and female sex hormones. *The Biochemical journal* **2000**, *351*, 67-77.
25. Penning, T.M.; Byrns, M.C. Steroid hormone transforming aldo-keto reductases and cancer. *Annals of the New York Academy of Sciences* **2009**, *1155*, 33-42.
26. Rodu, B. An alternative approach to smoking control. *Am J Med Sci* **1994**, *308*, 32-34.
27. Savitz, D.; Meyer, R.; Tanzer, J.; Mirvish, S.; Lewin, F. Public health implications of smokeless tobacco use as a harm reduction strategy. *Am J Public Health* **2006**, *96*, 1934-1939.
28. Gao, H.; Prasad, G.L.; Zacharias, W. Combusted but not smokeless tobacco products cause DNA damage in oral cavity cells. *Environmental toxicology and pharmacology* **2014**, *37*, 1079-1089.
29. Boyle, J.O.; Gumus, Z.H.; Kacker, A.; Choksi, V.L.; Bocker, J.M.; Zhou, X.K.; Yantiss, R.K.; Hughes, D.B.; Du, B.; Judson, B.L., *et al.* Effects of cigarette smoke on the human oral mucosal transcriptome. *Cancer prevention research* **2010**, *3*, 266-278.
30. Gumus, Z.H.; Du, B.; Kacker, A.; Boyle, J.O.; Bocker, J.M.; Mukherjee, P.; Subbaramaiah, K.; Dannenberg, A.J.; Weinstein, H. Effects of tobacco smoke on gene expression and cellular pathways in a cellular model of oral leukoplakia. *Cancer prevention research* **2008**, *1*, 100-111.
31. Severino, P.; Alvares, A.M.; Michaluart, P., Jr.; Okamoto, O.K.; Nunes, F.D.; Moreira-Filho, C.A.; Tajara, E.H.; Head; Neck Genome Project, G. Global gene expression profiling of oral cavity

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

cancers suggests molecular heterogeneity within anatomic subsites. *BMC research notes* **2008**, *1*, 113.

32. Sridhar, S.; Schembri, F.; Zeskind, J.; Shah, V.; Gustafson, A.M.; Steiling, K.; Liu, G.; Dumas, Y.M.; Zhang, X.; Brody, J.S., *et al.* Smoking-induced gene expression changes in the bronchial airway are reflected in nasal and buccal epithelium. *BMC genomics* **2008**, *9*, 259.
33. Pappas, R.S.; Stanfill, S.B.; Watson, C.H.; Ashley, D.L. Analysis of toxic metals in commercial moist snuff and alaskan iqmik. *Journal of analytical toxicology* **2008**, *32*, 281-291.
34. Rickert, W.S.; Joza, P.J.; Trivedi, A.H.; Momin, R.A.; Wagstaff, W.G.; Lauterbach, J.H. Chemical and toxicological characterization of commercial smokeless tobacco products available on the canadian market. *Regulatory toxicology and pharmacology : RTP* **2009**, *53*, 121-133.
35. Lansford, C.D.; Grenman, R.; Bier, H.; Somers, K.D.; Kim, S.-Y.; Whiteside, T.L.; Clayman, G.L.; Welkoborsky, H.J.; Carey, T.E. Head and neck cancers. In *Human cell culture - cancer cell lines part 2*, Masters, J.; Palsson, B., Eds. Kluwer Academic Press, : Dordrecht (Holland), 1999; Vol. 2, pp 185-255.
36. Takebayashi, S.; Ogawa, T.; Jung, K.Y.; Muallem, A.; Mineta, H.; Fisher, S.G.; Grenman, R.; Carey, T.E. Identification of new minimally lost regions on 18q in head and neck squamous cell carcinoma. *Cancer research* **2000**, *60*, 3397-3403.
37. Eskan, M.A.; Hajishengallis, G.; Kinane, D.F. Differential activation of human gingival epithelial cells and monocytes by porphyromonas gingivalis fimbriae. *Infection and immunity* **2007**, *75*, 892-898.
38. Kinane, D.F.; Shiba, H.; Stathopoulou, P.G.; Zhao, H.; Lappin, D.F.; Singh, A.; Eskan, M.A.; Beckers, S.; Waigel, S.; Alpert, B., *et al.* Gingival epithelial cells heterozygous for toll-like receptor 4 polymorphisms asp299gly and thr399ile are hypo-responsive to porphyromonas gingivalis. *Genes and immunity* **2006**, *7*, 190-200.
39. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic acids research* **2015**, *43*, e47.
40. Benjamini, Y.; Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* **2001**, *29*, 1165-1188.
41. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S., *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 15545-15550.
42. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative pcr and the 2(-delta delta c(t)) method. *Methods* **2001**, *25*, 402-408.
43. Borgerding, M.F.; Bodnar, J.A.; Curtin, G.M.; Swauger, J.E. The chemical composition of smokeless tobacco: A survey of products sold in the united states in 2006 and 2007. *Regulatory toxicology and pharmacology : RTP* **2012**, *64*, 367-387.
44. Hyndman, D.; Flynn, T.G. The aldo-keto reductases and their role in cancer. *Advances in experimental medicine and biology* **1999**, *463*, 427-434.
45. Nagaraj, N.S.; Beckers, S.; Mensah, J.K.; Waigel, S.; Vigneswaran, N.; Zacharias, W. Cigarette smoke condensate induces cytochromes p450 and aldo-keto reductases in oral cancer cells. *Toxicology letters* **2006**, *165*, 182-194.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

46. Fukumoto, S.; Yamauchi, N.; Moriguchi, H.; Hippo, Y.; Watanbe, A.; Shibahara, J.; Taniguichi, H.; Ishikawa, S.; Ito, H.; Yamamato, S., *et al.* Overexpression of the aldo-keto reductase family protein akr1b10 is highly correlated with smokers non-small cell lung carcinoma. *Clin. Cancer Res.* **2005**, *11*, 1776-1785.
47. Hsu, N.; Ho, H.; Chow, K.; Lin, T.; Shih, C.; Wang, L.; Tsai, C. Overexpression of dihydrodiol dehydrogenase as a prognostic marker of non-small cell lung cnacer. *Cancer Res.* **2001**, *61*, 2727-2731.
48. Zhang, L.; Lee, J.J.; Tang, H.; Fan, Y.H.; Xiao, L.; Ren, H.; Kurie, J.; Morice, R.C.; Hong, W.K.; Mao, L. Impact of smoking cessation on global gene expression in the bronchial epithelium of chronic smokers. *Cancer Prev. Res.* **2008**, *1*, 112.
49. Zhang, X.; Sebastiani, P.; Liu, G.; Schembri, F.; Dumas, Y.M.; Langer, E.M.; Alekseyev, Y.; O'Connor, G.T.; Brooks, D.R.; Lenburg, M.E., *et al.* Similarities and differences between smoking-related gene expression in nasal and bronchial epithelium. *Physiol. Genomics* **2010**, *41*, 1-8.
50. Penning, T.M.; Drury, J.E. Human aldo-keto reductases: Function, gene regulation, and single nucleotide polymorphisms. *Archives of biochemistry and biophysics* **2007**, *464*, 241-250.
51. Arthur, J.W.; Reichardt, J.K. Modeling single nucleotide polymorphisms in the human akr1c1 and akr1c2 genes: Implications for functional and genotyping analyses. *PloS one* **2010**, *5*, e15604.
52. Abedin, Z.; Sen, S.; Field, J. Aldo-keto reductases protect lung adenocarcinoma cells from the acute toxicity of b[a]p-7,8-trans-dihydrodiol. *Chemical research in toxicology* **2012**, *25*, 113-121.
53. Chien, C.W.; Ho, I.C.; Lee, T.C. Induction of neoplastic transformation by ectopic expression of human aldo-keto reductase 1c isoforms in nih3t3 cells. *Carcinogenesis* **2009**, *30*, 1813-1820.
54. Huang, K.H.; Chiou, S.H.; Chow, K.C.; Lin, T.Y.; Chang, H.W.; Chiang, I.P.; Lee, M.C. Overexpression of aldo-keto reductase 1c2 is associated with disease progression in patients with prostatic cancer. *Histopathology* **2010**, *57*, 384-394.
55. Le Calve, B.; Rynkowski, M.; Le Mercier, M.; Bruyere, C.; Lonez, C.; Gras, T.; Haibe-Kains, B.; Bontempi, G.; Decaestecker, C.; Ruysschaert, J.M., *et al.* Long-term in vitro treatment of human glioblastoma cells with temozolomide increases resistance in vivo through up-regulation of glut transporter and aldo-keto reductase enzyme akr1c expression. *Neoplasia* **2010**, *12*, 727-739.
56. Nebert, D.W.; Dalton, T.P.; Okey, A.B.; Gonzalez, F.J. Role of aryl hydrocarbon receptor-mediated induction of the cyp1 enzymes in environmental toxicity and cancer. *The Journal of biological chemistry* **2004**, *279*, 23847-23850.
57. Shimada, T.; Fujii-Kuriyama, Y. Metabolic activation of polycyclic aromatic hydrocarbons to carcinogens by cytochromes p450 1a1 and 1b1. *Cancer science* **2004**, *95*, 1-6.
58. Wang, R.; Wang, G.; Ricard, M.J.; Ferris, B.; Strulovici-Barel, Y.; Salit, J.; Hackett, N.R.; Gudas, L.J.; Crystal, R.G. Smoking-induced upregulation of akr1b10 expression in the airway epithelium of healthy individuals. *Chest* **2010**, *138*, 1402-1410.
59. Ebert, B.; Kisiela, M.; Wsol, V.; Maser, E. Proteasome inhibitors mg-132 and bortezomib induce akr1c1, akr1c3, akr1b1, and akr1b10 in human colon cancer cell lines sw-480 and ht-29. *Chemico-biological interactions* **2011**, *191*, 239-249.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 January 2017 doi:10.20944/preprints201701.0111.v1

Peer-reviewed version available at Genes 2017, 8, , 132; doi:10.3390/genes8050132

60. Jung, K.A.; Choi, B.H.; Nam, C.W.; Song, M.; Kim, S.T.; Lee, J.Y.; Kwak, M.K. Identification of aldo-keto reductases as nrf2-target marker genes in human cells. *Toxicology letters* **2013**, *218*, 39-49.
61. Pallai, R.; Simpkins, H.; Chen, J.; Parekh, H.K. The ccaat box binding transcription factor, nuclear factor-y (nf-y) regulates transcription of human aldo-keto reductase 1c1 (akr1c1) gene. *Gene* **2010**, *459*, 11-23.
62. Chen, J.; Emara, N.; Solomides, C.; Parekh, H.; Simpkins, H. Resistance to platinum-based chemotherapy in lung cancer cell lines. *Cancer chemotherapy and pharmacology* **2010**, *66*, 1103-1111.
63. Chen, J.; Higby, R.; Tian, D.; Tan, D.; Johnson, M.D.; Xiao, Y.; Kellar, K.J.; Feng, S.; Shields, P.G. Toxicological analysis of low-nicotine and nicotine-free cigarettes. *Toxicology* **2008**, *249*, 194-203.
64. Khariwala, S.S.; Hatsukami, D.; Hecht, S.S. Tobacco carcinogen metabolites and DNA adducts as biomarkers in head and neck cancer: Potential screening tools and prognostic indicators. *Head & neck* **2012**, *34*, 441-447.

© 2017 by the authors; licensee *Preprints*, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).