Human Pathology (2011) xx, xxx–xxx

ELSEVIER

Human PATHOLOGY

www.elsevier.com/locate/humpath

Original contribution

Illumina whole-genome complementary DNA–mediated annealing, selection, extension and ligation platform: assessing its performance in formalin-fixed, paraffin-embedded samples and identifying invasion pattern–related genes in oral squamous cell carcinoma

Olivier Loudig PhD[a,b,*], Margaret Brandwein-Gensler MD[a,c], Ryung S. Kim PhD[b], Juan Lin PhD[b], Tatyana Isayeva MD, PhD[c], Christina Liu BA[b], Jeffrey E. Segall PhD[d], Paraic A. Kenny PhD[e], Michael B. Prystowsky MD, PhD[a]

[a]*Department of Pathology, Albert Einstein College of Medicine, Bronx, NY, USA*
[b]*Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA*
[d]*Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, NY, USA*
[e]*Developmental and Molecular Biology, Albert Einstein College of Medicine, Bronx, NY, USA*
[c]*Department of Pathology, University of Alabama at Birmingham, Birmingham, AL, USA*

Received 7 December 2010; revised 19 February 2011; accepted 23 February 2011

Keywords:
Messenger RNA expression profiling;
Formalin-fixed paraffin-embedded;
Whole-genome DASL;
Head and neck cancer;
Oral squamous cell carcinoma;
NEDD9;
Invasion

Summary High-throughput gene expression profiling from formalin-fixed, paraffin-embedded tissues has become a reality, and several methods are now commercially available. The Illumina whole-genome complementary DNA–mediated annealing, selection, extension and ligation assay (Illumina, Inc) is a full-transcriptome version of the original 512-gene complementary DNA–mediated annealing, selection, extension and ligation assay, allowing high-throughput profiling of 24 526 annotated genes from degraded and formalin-fixed, paraffin-embedded RNA. This assay has the potential to allow identification of novel gene signatures associated with clinical outcome using banked archival pathology specimen resources. We tested the reproducibility of the whole-genome complementary DNA–mediated annealing, selection, extension and ligation assay and its sensitivity for detecting differentially expressed genes in RNA extracted from matched fresh and formalin-fixed, paraffin-embedded cells, after 1 and 13 months of storage, using the human breast cell lines MCF7 and MCF10A. Then, using tumor worst pattern of invasion as a classifier, 1 component of the "risk model," we selected 12 formalin-fixed, paraffin-embedded oral squamous cell carcinomas for whole-genome complementary DNA–mediated annealing, selection, extension and ligation assay analysis. We profiled 5 tumors with nonaggressive, nondispersed pattern of invasion, and 7 tumors with aggressive dispersed pattern of invasion and satellites scattered at least 1 mm apart. To minimize variability, the formalin-fixed, paraffin-embedded specimens were prepared from snap-frozen tissues, and

* Corresponding author.
E-mail address: olivier.loudig@einstein.yu.edu (O. Loudig).

0046-8177/$ – see front matter
doi:10.1016/j.humpath.2011.02.011

RNA was obtained within 24 hours of fixation. One hundred four down-regulated genes and 72 up-regulated genes in tumors with aggressive dispersed pattern of invasion were identified. We performed quantitative reverse transcriptase polymerase chain reaction validation of 4 genes using Taqman assays and in situ protein detection of 1 gene by immunohistochemistry. Functional cluster analysis of genes up-regulated in tumors with aggressive pattern of invasion suggests presence of genes involved in cellular cytoarchitecture, some of which already associated with tumor invasion. Identification of these genes provides biologic rationale for our histologic classification, with regard to tumor invasion, and demonstrates that the whole-genome complementary DNA–mediated annealing, selection, extension and ligation assay is a powerful assay for profiling degraded RNA from archived specimens when combined with quantitative reverse transcriptase polymerase chain reaction validation.

1. Introduction

Patients with low-stage oral cavity squamous carcinoma are usually treated by primary surgery and generally do not receive adjuvant radiotherapy. Unfortunately, 25% and 37% of stage I and stage II patients, respectively, are expected to fail treatment [1]. Distinguishing which patients with low-stage oral cancer are at high risk for disease progression would represent a major advancement in disease management. If these patients could be identified at diagnosis, more aggressive treatment protocols could be planned from the onset, and they could also be counseled in advance regarding the pathobiologic potential of their cancer.

We have developed and validated a predictive model, referred to as the *risk model*, which we have shown to be highly predictive of outcome for patients with oral and oropharyngeal squamous cell carcinoma (SCC), when adjusted for confounders [2,3]. The risk score is assessed from primary resection specimens by quantifying 3 significant histologic variables: (1) tumor worst pattern of invasion (WPOI) at the advancing tumor edge, (2) perineural invasion (PNI), and (3) lymphocytic host response at the advancing tumor edge. A current limitation of the risk model is that it cannot be reliably assessed on initial biopsies because the small size of these specimens limits full evaluation of the tumor/host interface. We hypothesize that there are tumor-specific genes that may reflect components of the risk model (WPOI, lymphocytic host response, and PNI), which are expressed diffusely throughout the tumors, and that may be assessed in a limited biopsy to serve as surrogates for the risk model. Our long-term goal is to develop a quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) assay that may help predict high-risk oral SCC (OSCC) using formalin-fixed and paraffin-embedded (FFPE) biopsy samples. Development of a predictive panel that can be assessed on recent FFPE biopsies, akin to the breast Oncotype DX test, is a practical approach for pathologists, requiring no additional clinical procedures or technical handling of samples [4].

Archived clinical specimens represent a potentially abundant resource for studying the molecular basis of disease, and adequate high-throughput retrospective studies may help correlate clinical outcome and/or therapeutic response with molecular features. In the past 10 years, research efforts have been directed toward improving recovery of RNA from banked specimens and its analysis using multiplexed qRT-PCR and whole-transcriptome microarray analyses [5,6]. Although conventional RNA amplification methods have remained dependent on the quality of the RNA, more specifically RNA length and presence of a polyadenylated tail, the multiplexed RT-PCR approach from Illumina, Inc, termed *cDNA-mediated annealing, selection, extension and ligation* (DASL) assay, has been designed to target stretches of 50 nucleotides, hence, its adequacy for the analysis of highly degraded RNA and FFPE RNA. After reverse transcription of the RNA using a combination of random and oligo-dT primers, a set of 2 adjacent primers (each 23 nucleotides) is annealed to the reverse-transcribed products, and by elongation and ligation of the 2 primers, a complementary DNA (cDNA) product is synthesized and magnetically purified. The cDNA products are PCR amplified, and the amplicons are quantified on bead arrays [7]. Illumina, Inc, had initially developed a DASL assay for expression profiling of 512 to 1536 genes, which demonstrated high sensitivity and reproducibility while requiring only 50 to 200 ng of total RNA [8]. Although the original 512- to 1536-gene DASL assay was restricted to a small set of chosen genes, the whole-genome DASL (WG-DASL) is designed for the simultaneous analysis of 24 526 annotated transcripts and may become a very powerful new tool for discovery of novel biomarkers using archival FFPE material or even degraded RNA from frozen specimens [9].

In this study, using RNA from matched fresh and FFPE malignant and benign human breast cell lines, MCF7 and MCF10A, we evaluated the reproducibility and sensitivity of the WG-DASL assay. Then, using OSCC specimens and tumor WPOI as a classifier, we tested for the presence of WPOI-associated genes in 2 distinct types of tumors and validated our results using state-of-the-art qRT-PCR experiments and immunohistochemistry (IHC). Our results show the feasibility of using the WG-DASL assay to interrogate degraded RNA recovered from FFPE tissues and demonstrate that we can identify RNA species that are associated with distinct histopathologic patterns of tumor invasion.

2. Materials and methods

2.1. Cell culture and FFPE procedure and RNA extraction

The human nontumorigenic epithelial MCF10A cell line (ATCC, Cat No. CRL-10317) was cultured in DMEM/F12 medium, supplemented with 5% horse serum, 20 ng/mL EGF, 0.5 μg/mL hydrocortisone, 100 ng/mL cholera toxin, 10 μg/mL insulin, 50 U/mL penicillin, and 50 μg/mL streptomycin in a humidified atmosphere composed of 95% air and 5% CO_2. The human breast adenocarcinoma MCF7 cell line (ATCC, Cat No. HTB-22D) was cultured in DMEM with 2 mmol/L L-glutamine and 10% FBS. Cells were provided with fresh medium every 2 days, and after reaching a confluence of 40% to 60% in 100-mm Petri dish, they were scraped and collected in 1× phosphate-buffered saline (PBS) solution. A total of nine 100-mm plates were collected to obtain a cellular pellet in a 15-mL corning flat bottom tube, at 800*g* for 5 minutes at 4°C in a centrifuge. The PBS was removed, and the cells were fixed in 15 mL of 3.7% buffered formaldehyde (2 mL 37% formaldehyde, 2 mL 10× PBS pH7.4, and 16 mL nuclease-free water from Ambion) for 4 hours. The pellet was then placed into a blotting paper and put into a cassette that was stored in 70% ethanol before routine processing and paraffin embedding in the histopathology core laboratory at the Albert Einstein College of Medicine. Total RNA from fresh MCF10A cells was extracted with TRIzol (Invitrogen). For FFPE RNA degradation, total RNA was recovered, from 10 μm sections, at 1, 4, and 13 months of storage at room temperature using the High-pure RNA paraffin kit (Roche), following the manufacturer's instructions. Total RNA from fresh and FFPE RNA was analyzed and quantified on an Agilent Bioanalyzer mRNA nanochip, which provided a quantitative measure of RNA quality (the RNA integrity number [RIN]) and the concentration.

2.2. Tumor procurement from clinical samples and RNA extraction

Twelve OSCC, previously flash frozen and banked in cassettes at −80°C, were selected. Tumors from the 2 extremes were selected for comparison: WPOI type 3 (WPOI-3) represents a nonaggressive pattern of invasion with tumor satellites that are large (>15 cells) and not dispersed, whereas WPOI type 5 (WPOI-5) describes tumors with dispersed satellites scattered at least 1 mm from the next closest tumor satellite [2,3]. Seven tumors were classified as WPOI-5, and 5 were classified as WPOI-3. These banked specimens were thawed at room temperature in 10% buffered formalin, then processed and paraffin embedded in the histopathology laboratory at the Montefiore Hospital, Bronx, NY. Hematoxylin and eosin–stained slides were produced within 24 hours of formalin fixation, and tumor-predominant regions were selected for procurement. A Chemicon microarrayer was used to obtain 3 to 5 morphologically guided 1-mm cores from the FFPE blocks, from the center of the tumors; these samples were placed into DNase-/RNase-free sterile Eppendorfs tubes. Hematoxylin and eosin slides were generated after core procurement to confirm that the cores contained greater than 80% SCC-enriched cells. RNA was extracted from FFPE cores using the Roche High-Pure RNA Paraffin kit, as recommended by Illumina for FFPE RNA analysis on the WG-DASL platform, and the RNA was analyzed and quantified on Agilent Bioanalyzer total RNA chips.

2.3. Illumina WG-DASL assay

The WG-DASL assay was performed as described previously [9]. Briefly, cDNA was generated from total RNA using biotinylated oligo-dT^{18} and random nonamer primers. For each transcript of interest, 2 assay-specific oligonucleotides are designed to target a contiguous 50-nucleotide sequence on each cDNA. A total of 24 526 oligonucleotide pairs (probes) were designed and pooled, which together constituted the WG-DASL assay pool, corresponding to 18 626 unique genes, based on well-annotated content derived from the National Center for Biotechnology Information Reference Sequence Database (Build 36.2, Release 22). The DASL assay pool was then annealed to the targeted cDNAs during a 16-hour temperature-gradient (70°C-30°C) incubation, followed by enzymatic extension and ligation steps, as previously described. Ligated products were PCR amplified and labeled with a universal Cy3-coupled primer after which single-stranded labeled products were precipitated and then hybridized to the beadchips. The bead arrays were washed and coated with a polymer to protect the fluorescence signals. The fluorescent intensities were captured, by scanning the arrays, on a BeadArray Reader (Illumina).

2.4. WG-DASL data analysis

For the first data set on matched fresh and FFPE cell lines, the raw expression intensities were normalized by quantile normalization method using the BeadStudio software. First, we computed Spearman rank correlation between intensities of repeats to measure the reproducibility. Second, we determined the sensitivity of the WG-DASL assay, defined by the overlap of differentially expressed genes between fresh and FFPE RNA obtained from MCF7 and MCF10A cell lines, that is, $m/n \times 100$ (%), if m of n most up-regulated (or down-regulated) genes in FFPE samples overlap with n most up-regulated (or down-regulated) genes in fresh samples. The fold change (FC) in intensities was calculated as the intensity of MCF7 over the intensity of MCF10A.

For the second data set on clinical samples, the raw expression intensities were normalized by robust spline

method [10-12]. We considered a gene to be present in a sample if the detection *P* value computed by BeadStudio was less than .05. Then, the differentially expressed genes were identified, which met all of the following criteria: (1) *P* value by *t* test on log scale to be at most .005, (2) FC between average intensities in 2 groups to be at least 1.2, (3) difference of average normalized intensities between the 2 groups to be at least 500, and (4) intensities of at least 5 of 6 WPOI-3 and 7 of 8 WPOI-5 samples to be detected. The false discovery rate (FDR) in differential expression was estimated by permuting the sample labels 500 times.

2.5. Gene clustering analysis

Functional clustering analysis was performed using online software (http://david.abcc.ncifcrf.gov) with the list of the 75 significant differentially expressed genes up-regulated in aggressive WPOI-5 tumors. The Database for Annotation, Visualization, and Integrated Discovery gene set analysis generates an enrichment score by testing the relatedness of different combinations of genes according to common biologic function, chromosomal location, or regulation [13,14]. A high enrichment score for a group of genes indicates that annotation term members are playing important roles in a given study, and an enrichment score of 1.3 is equivalent to a non–log scale $P = .05$. The *P* value is also termed *the EASE score* and represents a modified Fisher exact test; the smaller the *P* value, the more significant the gene association. The Benjamini test globally corrects enrichment *P* values to control for family-wide FDRs less than .05.

2.6. qRT-PCR validation

Four differentially expressed genes were selected for qRT-PCR validation using Taqman reagents (CEACAM1 cat# hs00236077_m1, NEDD9 cat# hs00610590_m1, RABAC1 cat# hs00197506_m1, and AKT3 Cat# hs00178533_m1), and their expression levels were validated in the same five WPOI-3 and six WPOI-5 FFPE tumor-enriched tissue samples. For qRT-PCR experiments, the Taqman RNA-to-CT 1-Step Kit (Applied Biosystems) was used with RNA extracted from FFPE tissue. The RNA used in these experiments was quantified using the Agilent Bioanalyzer total RNA chip and aliquoted at 50 ng/μL. Each reaction was performed in triplicate (total reaction volume, 20 μL) from a master mix of 66 μL prepared for each tumor (33 μL Taqman RT-PCR mix, 3.3 μL RNA (50 ng/μL), 3.3 μL primers, 1.65 μL enzyme, 24.75 μL water). qRT-PCR measurements were performed using the StepOnePlus real-time PCR instrument (Applied Biosystems), following 1 step at 48°C for 15 minutes, 1 step at 95°C for 10 minutes, 40 cycles at 95°C for 15 seconds, and 60°C for 1 minute. The RT-PCR instrument determined the comparative threshold (Ct) for each gene as the average of triplicate measurements. The ΔCt was calculated by subtracting the Ct value of the endogenous control GAPDH from the Ct of each gene (v-akt murine thymoma viral oncogene homolog 3 [AKT3], carcinoembryonic antigen-related cell adhesion molecule 1 [CEACAM1], neural precursor cell expressed developmentally down-regulated 9 [NEDD9], and Rab acceptor 1 [RABAC1]) for each RNA sample, and the FC was calculated using the $^{\Delta\Delta Ct}$ formula. Gene expression differences between the WPOI-3 and WPOI-5 groups were calculated using the nonparametric Mann-Whitney *U* test [15]. All statistical analyses were made using the R statistical environment [16].

2.7. Immunohistochemistry

One protein, differentially up-regulated in WPOI-5 tumors, was selected for validation: NEDD 9 (Hef-1, monoclonal mouse anti-human antibody clone 2G9; Novus Biologicals, Littleton, CO). This monoclonal antibody (clone 2G9) has previously been validated for specific recognition of NEDD9 by Western blot and IHC [17,18]. Briefly, 5 µm sections of paraffin-embedded tumor samples from all 12 carcinomas were deparaffinized and hydrated through graded alcohol. Antigen retrieval was performed using citrate buffer (pH 6.0) and steam for 20 minutes. Sections were cooled, and endogenous peroxidases were removed using 0.3% H_2O_2 in methanol for 30 minutes and blocked with 3% mouse serum or rabbit serum for 30 minutes. Tissue sections were incubated with primary antibodies, NEDD9 at 1:500, overnight at 4°C. Sections were washed in PBS with 0.05% Tween-20 and incubated at room temperature with biotin-conjugated goat antimouse or goat antirabbit secondary antibody for 20 minutes. After washing, sections were incubated with streptavidin-conjugated horseradish peroxidase for 20 minutes at room temperature. After another wash with PBS with 0.05% Tween-20, immunodetection was done using 3,3′-diaminobenzidine H_2O_2 (Vector Labs) and counterstained with hematoxylin. The positive control consisted of FFPE sections of 2 pancreatic adenocarcinomas, which expressed NEDD9; the negative controls consisted of the same test carcinoma samples, with omission of the primary antibody incubation step.

3. Results

3.1. Assessing reproducibility and sensitivity of the WG-DASL using matched fresh and FFPE RNA from monoclonal cell lines

To determine the sensitivity of the WG-DASL assay and its applicability to FFPE RNA, we compared the genes that were differentially expressed between malignant (MCF7) and benign (MCF10A) breast cell lines using matched fresh and FFPE total RNA. Total RNA was recovered from fresh

cells (Fig. 1A) and showed no degradation (RIN, 10.0). Total RNA from the FFPE blocks recovered after 1 month of storage (Fig. 1B) displayed a profile of degradation (RIN, 3.2). The extent of degradation was increased after 13 months of storage (RIN, 2.3) (Fig. 1C). The Bioanalyzer data show that both formalin fixation and storage of FFPE blocks at room temperature contribute to the process of RNA degradation.

We assessed the reproducibility of the WG-DASL assay by using fresh RNA and FFPE RNA obtained from replicate blocks, derived from the same cells, and processed at the same time under the same conditions (Fig. 1D and E). Using fresh RNA, we obtained a Spearman rank correlation coefficient of 0.987 between duplicate measures (Fig. 1D). Using FFPE RNA samples, from 2 individual blocks (4-month-old MCF10A blocks), we obtained a Spearman rank correlation coefficient of 0.986 between these repeats (Fig. 1E). These results demonstrate the high reproducibility of the WG-DASL assay using either fresh or FFPE RNA.

We then assessed the sensitivity of the WG-DASL assay for identifying genes differentially expressed between 2 breast cell lines (MCF7 and MCF10A) in FFPE RNA recovered after 1 and 13 months of storage, by comparison with RNA from

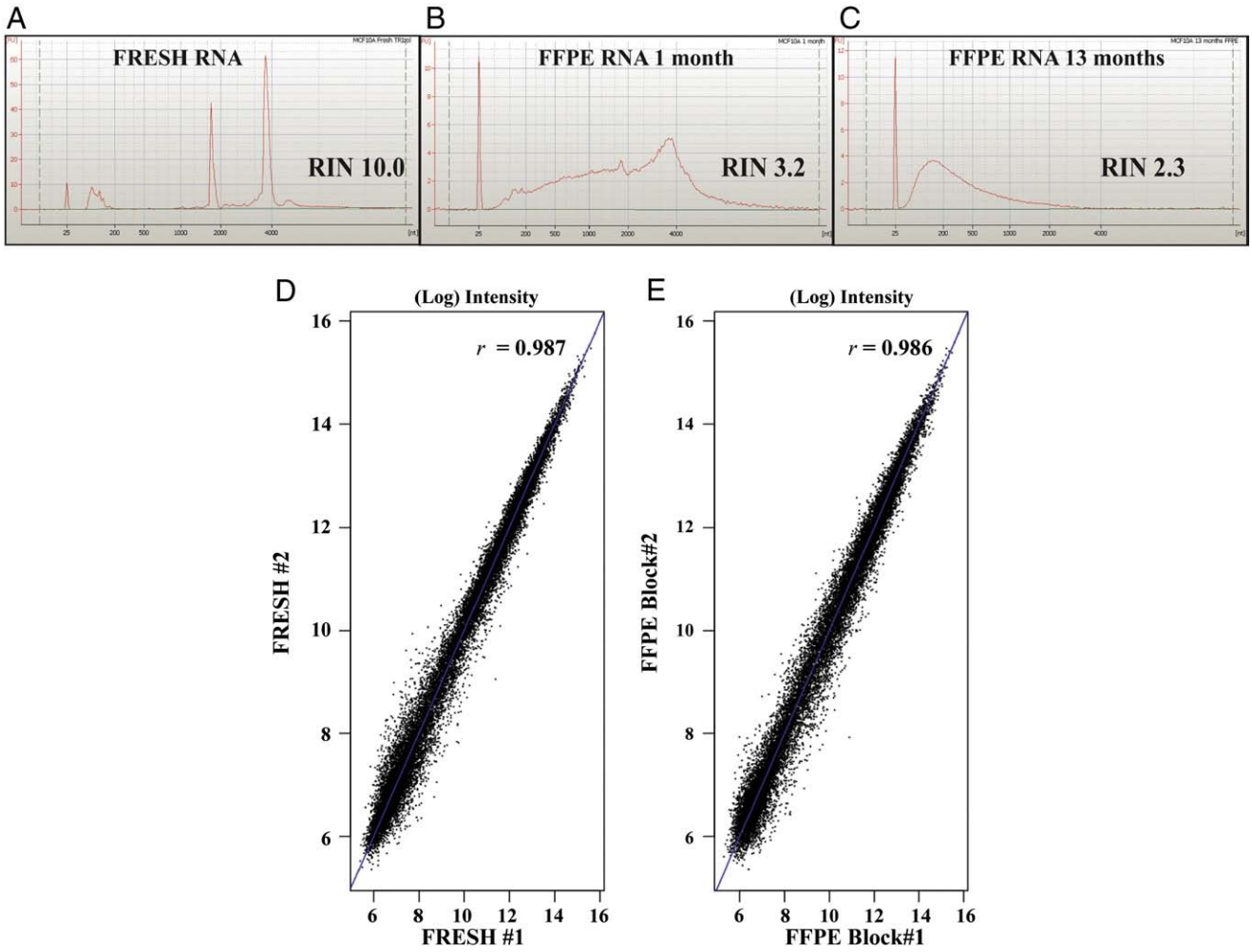

Fig. 1 A to E, FFPE RNA degradation and its impact on the WG-DASL platform performance. Human breast cancer cells (MCF10A and MCF7) were cultured, and total RNA was extracted from either fresh cells or fresh cells pelleted, FFPE, and stored at room temperature for up to 13 months. A, Total RNA extracted from fresh cancer cells (MCF10A) observed on the Agilent 2100 Bioanalyzer and displaying an RIN of 10.0 (maximum). B, Total RNA recovered after 1 month of storage of FFPE cancer cells (MCF10A) and displaying an RIN of 3.2. C, Total RNA recovered after 13 months of storage of FFPE cancer cells and displaying an RIN of 2.3. Panels D and E show the repeatability of the WG-DASL assay using fresh RNA (panel D, $r = 0.99$) and FFPE RNA extracted from 2 distinct blocks that were fixed at the same time, with the same cells, and stored in the same conditions (panel E, $r = 0.98$). F to I, Panels F and G show the detection in FC differences (0.01- to 100-fold) for transcripts detected in MCF7 and MCF10A cells, between fresh and FFPE samples after 1 month of fixation (panel F) and after 13 months of fixation (panel G). Panel H shows the overlap for the top 1000 up-regulated genes (MCF7/MCF10A) between fresh (x-axis) and FFPE cells (y-axis), after 1 month of fixation (black line) and 13 months of fixation (red-dotted line). Panel I shows the overlap for the top 1000 down-regulated genes (MCF7/MCF10A) between fresh (x-axis) and FFPE cells (y-axis), after 1 month of fixation (green-dotted line) and 13 months of fixation (blue-dotted line). The WG-DASL measurements for fresh, 1-month-old, and 13-month-old FFPE cells were performed in duplicates, and the average data were used for analysis.

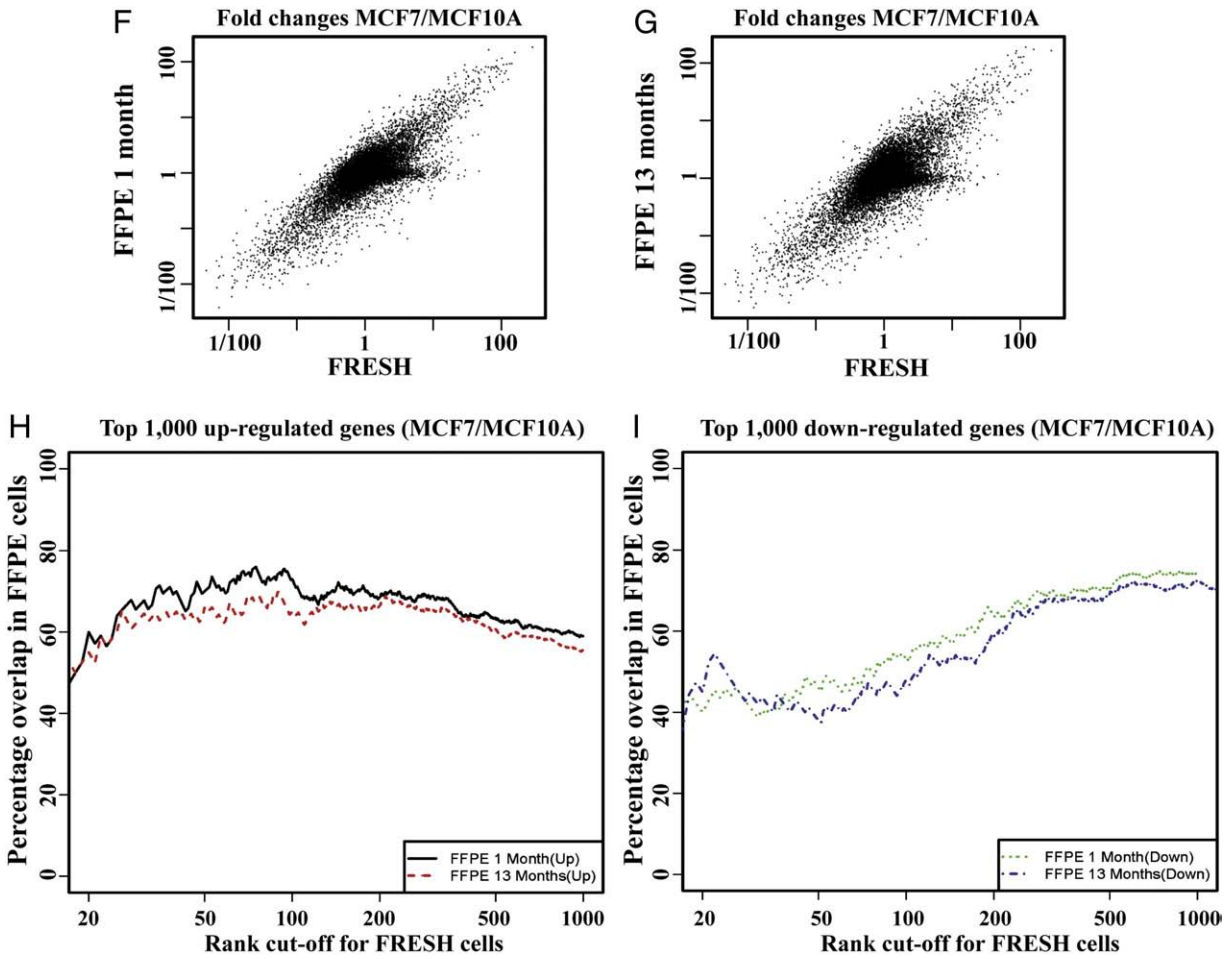

Fig. 1 *(continued)*

fresh cells (Fig. 1F and G). Genes differentially expressed in fresh and FFPE RNA, between MCF7 and MCF10A, are displayed in FCs (MCF7/MCF10A; Fig. 1F and G). Our results show that the WG-DASL displays a similar sensitivity after 1 (Fig. 1F) and 13 months of storage (Fig. 1G).

We further evaluated the sensitivity of the WG-DASL assay by measuring the overlap between the top ranked 1000 up-regulated (Fig. 1H) and top ranked 1000 down-regulated genes (Fig. 1I) between fresh and FFPE RNA at 1 and 13 months of block storage. Fig. 1H and I displays the percentage of overlap between the genes with high FCs in fresh samples and those in FFPE samples, up-regulated and down-regulated, respectively. For example, among the 20 most up-regulated genes in fresh samples, 12 also showed highest FCs in 1-month-old FFPE samples. Thus, the percentage of overlap at 20th rank was computed as 60% (or [12/20] × 100 [%]). Although our results show a slight decrease in the overlap between fresh and FFPE cells between 1 and 13 months of storage, both for up-regulated (Fig. 1H, black and red lines) and down-regulated (Fig. 1I, green and blue lines) genes, they demonstrate that the WG-DASL is a very stable assay with degraded RNA, which allows successful quantification of 60% of the genes measured as differentially expressed. These data demonstrate that the WG-DASL assay is a highly reproducible and sensitive assay for detection of differentially expressed genes using FFPE-derived RNA.

3.2. Application of the WG-DASL to FFPE RNA obtained from nonaggressive (WPOI-3) and aggressive (WPOI-5) OSCC tumor samples

We selected primary OSCC, and the strategy was to use the 2 extremes of WPOI, the nonaggressive WPOI-3 and aggressive WPOI-5 (Fig. 2A and B, respectively) to determine if invasion pattern–related genes differentially expressed between these tumors could be identified. Total RNA from nonaggressive WPOI-3 and aggressive WPOI-5 FFPE samples with RIN greater than 2.0, meeting quality control for reproducible and sensitive performance of the WG-DASL (Fig. 1C, G, H, and I), was selected for the analysis [9]. Five WPOI-3 and seven WPOI-5 tumor-enriched samples were analyzed, and reproducibility of the WG-DASL assay, in these experiments, was assessed by analyzing HN#5 (WPOI-3) and HN#9 (WPOI-5) in duplicate experiments (Fig. 3A, green and red rectangles above

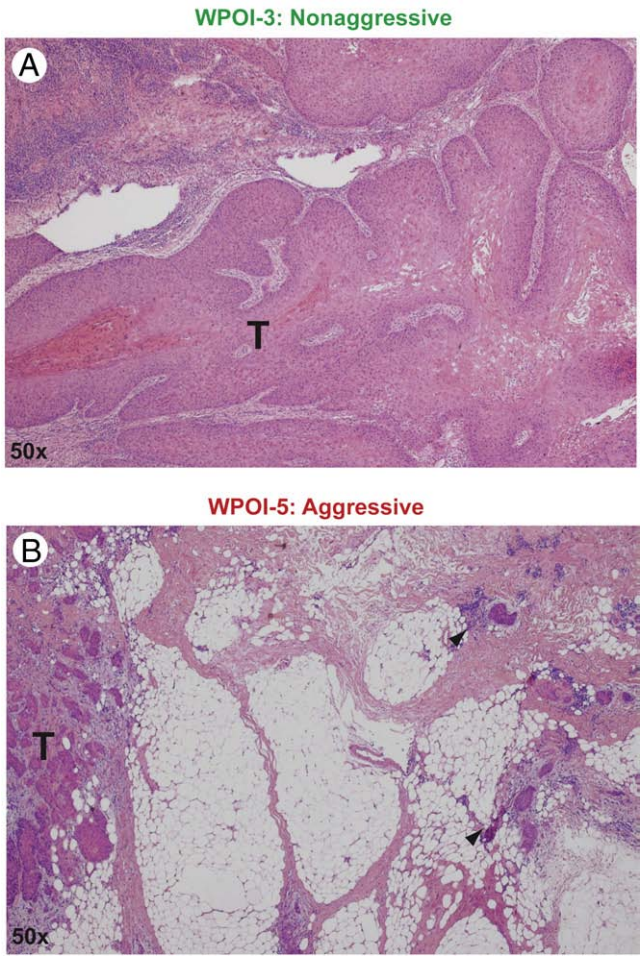

Fig. 2 Images of 2 different histologic types of OSCC tumors. In A, the panel displays a nonaggressive WPOI-3 tumor with a finger-like morphology (T represents the tumor). In B, the panel displays an aggressive WPOI-5 tumor, with disperse tumor satellites (see arrowheads) scattered at least 1 mm from the main tumor (T).

heatmap, respectively). These duplicates generated high correlation coefficient (data not shown), similar to those observed with breast cell lines (Fig. 1E). A total of 104 down-regulated and 72 up-regulated genes were identified that distinguished nonaggressive (WPOI-3) from aggressive (WPOI-5) tumor samples (Fig. 3A). The median FDR was 0.0043 (or 4.3%).

3.3. Functional cluster analysis and qRT-PCR validation of putative nonaggressive WPOI-3– and aggressive WPOI-5–associated genes

Functional cluster analysis for the subgroup of 72 genes significantly up-regulated in aggressive WPOI-5 tumors revealed 3 significant gene clusters. The highest enrichment score (2.87) grouped genes associated with cell adhesion, followed by genes encoding calcium-binding proteins with EF-hand motif (enrichment score of 1.47), followed by signal anchor genes (enrichment score 1.33) (see Table 1).

We then used qRT-PCR experiments to validate data obtained from the WG-DASL platform on 4 of the 176 genes identified as differentially expressed between WPOI-3 and WPOI-5 (Fig. 3A, see yellow highlighted gene symbols). Our results show that CEACAM1 displays higher expression in WPOI-3 FFPE samples than in aggressive WPOI-5 FFPE samples; however, its differential expression did not appear significant likely because of the small sample size ($P < .1775$; Fig. 3B, panel CEACAM1). The qRT-PCR results indicated that in 4 of the 6 aggressive WPOI-5 FFPE tumors analyzed (mean FC, 2.47), the trend of expression of CEACAM1 was lower than in nonaggressive WPOI-3 samples (mean FC, 15.35). We then selected 3 genes identified for their up-regulation in WPOI-5 FFPE tumors, NEDD9, AKT3, and RABAC1. NEDD9 showed significantly higher expression in WPOI-5 samples (mean FC, 21.17) compared with WPOI-3 samples (mean FC, 0.67; $P < .0043$; Fig. 3B, panel NEDD9). For AKT3 (Fig. 3B, panel AKT3), the qRT-PCR experiments also confirmed the increased expression in WPOI-5 tumor samples ($P < .02$; mean FC, 5.05) compared with WPOI-3 tumor samples (mean FC, 1.62). For RABAC1 (Fig. 3B, panel RABAC1), the qRT-PCR confirmed the trend of increased expression in WPOI-5 tumor samples ($P < .1775$; mean FC, 2.07) compared with WPOI-3 tumor samples (mean FC, 0.85). The qRT-PCR analyses not only validated differential expression identified with the WG-DASL but also demonstrated the importance of qRT-PCR validation experiments after WG-DASL profiling of FFPE samples. These experiments clearly indicate the potential usefulness of FFPE OSCC samples for identifying differentially expressed genes between different histologic phenotypes.

3.4. IHC validation of putative invasion-related gene NEDD9 expressed in WPOI-5 tumors

We further explored the potential of the WG-DASL assay for potentially identifying different protein expression levels between different tumor phenotypes. Considering the correlation between the WG-DASL assay and the qRT-PCR as well as the high FC differences between WPOI-3 and WPOI-5 tumors, we selected NEDD9 for IHC experiments (Fig. 4). We used a monoclonal antibody (clone 2G9) already validated for NEDD9 recognition in Western blot and IHC experiments [17,18]. The 2 top panels of Fig. 4 demonstrate absence of NEDD9 expression in sample HN#1 (WPOI-3), and the 2 lower panels of Fig. 4 show strong nuclear and cytoplasmic protein expression of NEDD9 in sample HN#6 (WPOI-5). The IHC data confirm the specificity of the WG-DASL assay for discovering biologically relevant differentially expressed tumor specific genes between different FFPE samples.

4. Discussion

High-throughput tumor messenger RNA (mRNA) expression profiling offers the potential of classifying biologic

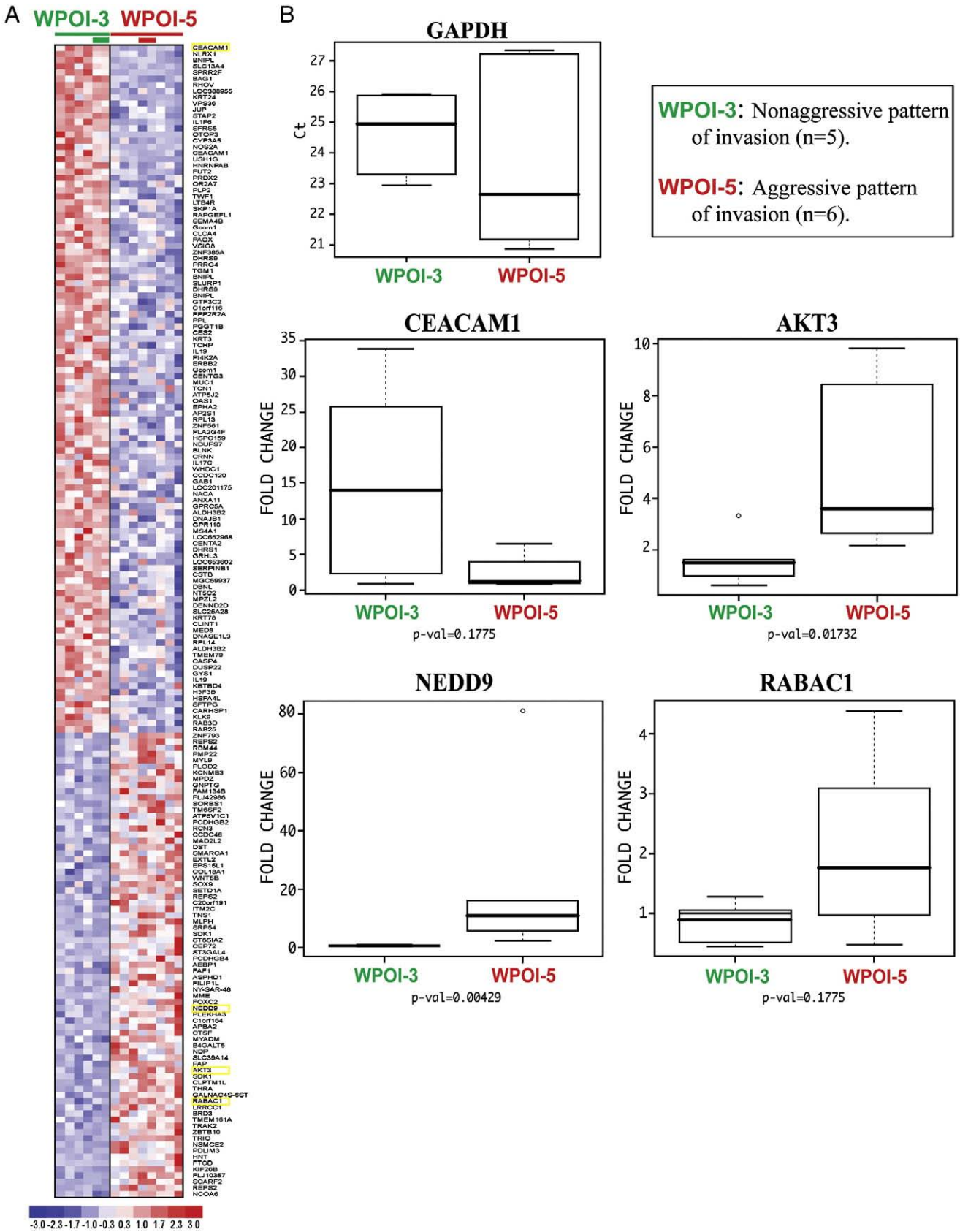
A
WPOI-3
WPOI-5
-3.0 -2.3 -1.7 -1.0 -0.3 0.3 1.0 1.7 2.3 3.0
B
GAPDH
Ct
WPOI-3
WPOI-5
WPOI-3: Nonaggressive pattern of invasion (n=5).
WPOI-5: Aggressive pattern of invasion (n=6).
CEACAM1
FOLD CHANGE
p-val=0.1775
AKT3
p-val=0.01732
NEDD9
p-val=0.00429
RABAC1
p-val=0.1775

Table 1 Significant functional clusters for 72 genes up-regulated in tumors with aggressive pattern of invasion (WPOI-5)

Gene symbol	Gene name	Protein function
Annotation cluster 1, cell adhesion: enrichment score, 2.87; P = .00024; and Benjamini, 0.063		
AEBP1	AE binding protein 1	Cell adhesion protein
SOX9	SRY (sex-determining region Y)-box 9	Transcription factor
COL18A1	Collagen, type XVIII, alpha 1	Extracellular matrix constituent
DST	Dystonin	Adhesion junction protein
NEDD9	Neural precursor cell expressed, developmentally down-regulated 9	Docking protein
PCDHGB2	Protocadherin gamma subfamily B, 2	Cellular adhesion protein
PCDHGB4	Protocadherin gamma subfamily B, 4	Cellular adhesion protein
SCARF2	Scavenger receptor class F, member 2	Probable adhesion protein
SDK1	Sidekick homolog 1, cell adhesion molecule (chicken); hypothetical LOC730351	Cell adhesion protein
SORBS1	Sorbin and SH3 domain containing 1	Cytoskeletal protein binding
THRA	Thyroid hormone receptor, alpha (erythroblastic leukemia viral (v-erb-a) oncogene homolog, avian)	Nuclear hormone receptor
Annotation cluster 2, calcium-binding proteins with EF-hand motif: enrichment score, 1.47; P = .0079; and Benjamini, 0.47		
REPS2	RALBP1-associated Eps domain containing 2	Growth factor signaling
DST	Dystonin	Adhesion junction protein
EPS15L1	Epidermal growth factor receptor pathway substrate 15–like 1	Receptor-mediated endocytosis
MYL9	Myosin, light chain 9, regulatory	Cellular locomotion
RCN3	Reticulocalbin 3, EF-hand calcium-binding domain	Signal transduction
Annotation cluster 3, signal anchor: enrichment score, 1.33; P = .00049; and Benjamini, 0.06		
ST3GAL4	ST3 beta-galactoside alpha-2,3-sialyltransferase 4	Catalytic enzyme
ST8SIA2	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 2	Sialyl transferase
B4GALT5	UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 5	Galactosyl transferase
ASPHD1	Aspartate beta-hydroxylase domain containing 1	Hydroxylase
EXTL2	Exostoses (multiple)-like 2	Glycosyl transferase
FAP	Fibroblast activation protein, alpha	Extracellular matrix proteolysis
ITM2C	Integral membrane protein 2C	Transmembrane protein
MME	Membrane metalloendopeptidase	Membrane catalytic enzyme

malignancies into clinically meaningful subgroups. For instance, expression profiling of breast carcinomas has brought to light the diversity of molecular subtypes and defined the concept of *molecular portrait* or *gene signatures*, which have improved diagnosis and brought a more personalized approach to treatment [19,20]. An important goal of cancer molecular profiling is to develop clinically useful assays that may be performed on FFPE samples as an adjunct to current pathologic diagnostic tools. Initial gene discovery and gene expression profiling, however, have most successfully been performed with high-quality RNA, usually extracted from fresh or snap-frozen samples. The integrity of mRNA is profoundly affected by ubiquitous RNases, delaying extraction or allowing specimens to remain at room temperature; together, these factors can cause progressive RNA degradation [21]. Nonetheless, RNA extraction from fresh or snap-frozen samples is a standard procedure in most laboratories and yields material that can be reliably analyzed [22,23]. Although the use of these types of samples has helped decipher cancer transcriptomes, they have rarely been correlated to long-term outcome. Large retrospective cohorts with longer outcome data are more commonly assembled with archived FFPE specimens. However, formalin fixation and paraffin embedding of clinical samples have been shown to lead to RNA degradation, secondary chemical modifications, and cross-linking of nucleic acids

Fig. 3 Supervised clustering of genes specific to nonaggressive and aggressive tumors, WPOI-3 and WPOI-5 tumors, and qRT-PCR validations. A, The heat map shows 104 down-regulated genes (blue; 111 probe sets) and 72 up-regulated genes (red; 75 probe sets) sorted by the FC between nonaggressive (WPOI-3) and aggressive (WPOI-5) tumor samples. The median FDR was 0.0043 (or 4.3%). The list of the abbreviated genes located on the right of the cluster also displays the 4 genes selected for qRT-PCR validations (yellow highlights). B, qRT-PCR validations of CEACAM1, AKT3, NEDD9, and RABAC1 mRNA expression in WPOI-3 and WPOI-5 tumors. The housekeeping gene GAPDH was used as an endogenous control for normalization of the expression of the 4 different genes, and its expression is represented by the mean Ct, in WPOI-3 and WPOI-5 tumor groups. The mean normalized FCs ($2^{\Delta\Delta Ct}$) of WPOI-3 and WPOI-5 samples were determined for CEACAM1, NEDD9, AKT3, and RABAC1. The different box plots display the difference between nonaggressive (WPOI-3) and aggressive (WPOI-5) patterns of invasion in FCs, and the P value between the 2 groups for each gene is indicated below each panel.

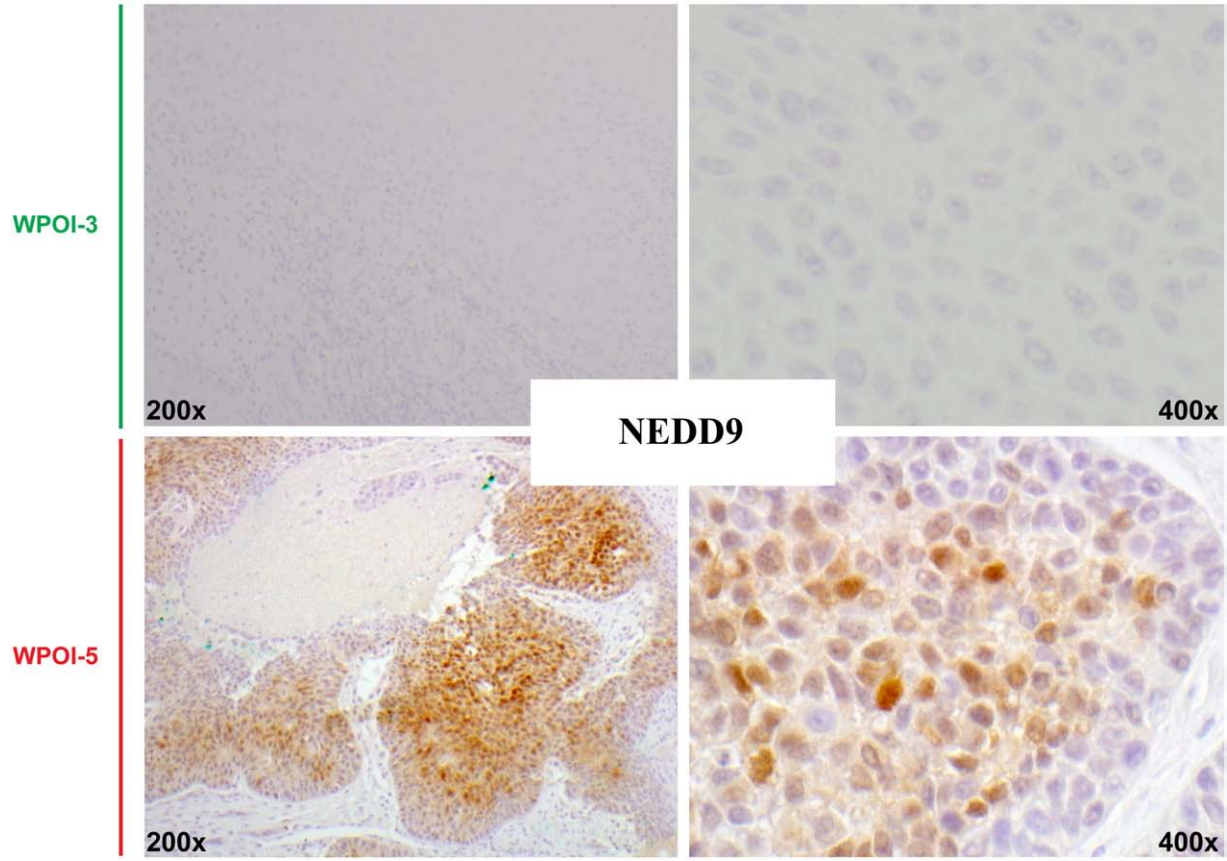

Fig. 4 IHC validation of the expression of NEDD9 in nonaggressive (WPOI-3) and aggressive (WPOI-5) tumors. The 2 upper panels represent the IHC detection of NEDD9 in WPOI-3 sample HN#1 (left and right panels represent low [×200] and high magnification [×400], respectively). The 2 lower panels represent the IHC detection of NEDD9 in WPOI-5 sample HN#6 (left and right panels represent low [×200] and high magnification [×400], respectively).

and proteins [5,24,25]. Although many factors affect the quality and availability of FFPE RNA, a growing body of evidence suggests that degraded RNA recovered from archival specimens is a valuable and informative resource [6-9,24-26]. Technical advances over the last decade in mRNA purification, amplification, and analysis have lead to improved gene expression analyses from archival specimens [27,28]. State-of-the-art in vitro transcription amplification and microarray technologies have also been applied to the analysis of polyadenylated mRNAs from archival FFPE specimens, but the success is dependent on the quality of the mRNA recovered [29-31]. In our laboratory, we established a robust method for restoring mRNA transcripts from highly degraded FFPE samples and demonstrated that specific technologies were required for high-throughput analysis of degraded FFPE RNA [32]. Illumina, Inc, specifically designed the DASL assay, which has been shown to be extremely robust, for the analysis of up to 512 genes in degraded RNA and FFPE RNA [7,8]. Development of the WG-DASL assay for simultaneous analysis of 24 526 annotated transcripts provides an opportunity to discover signatures, which may correlate with tumor behavior and treatment outcome.

SCC of the oral cavity is a disease treated by primary surgical resection. Pathologic staging and margin status will determine whether adjuvant therapy would be indicated for each individual. At the time of initial presentation, the preoperative biopsy cannot predict which oral carcinomas are particularly aggressive. Tumor WPOI refers to a classification schema determined by histologic examination of the resection specimen and is part of a multifactor predictive model [2,3]. The most aggressive WPOI, named WPOI-5, is characterized by dispersed, scattered tumor satellites. WPOI-5 is histomorphologically heterogeneous as tumor scatter owing to dispersed islands within the tongue muscle or multifocal PNI, dispersed lymphovascular tumor emboli, or a combination of these features. Patients with WPOI-5 squamous carcinoma are automatically classified as having high-risk disease and are significantly more likely to develop disease progression. Thus, the ability to predict WPOI-5 at initial biopsy, which is not feasible based purely on histology, could impact initial treatment planning and patient counseling.

Using histologically guided selection of FFPE tumor, we identified 176 genes differentially expressed between 5

nonaggressive WPOI-3 and 7 aggressive WPOI-5 OSCC tumors (104 down-regulated and 72 up-regulated genes). Interestingly, a number of the genes up-regulated in aggressive WPOI-5 FFPE–enriched tumors have already been linked with tumor invasion in other contexts (Table 1) [33-35]. For instance, NEDD9, also known as human enhancer of filamentation 1 and also known as Crk-associated substrate-related protein, lymphocyte type, is a scaffolding protein, which expression is up-regulated in epithelial cells and that regulates cell attachment, migration, and invasion, in addition to other processes (apoptosis and cell cycle) [33]. NEDD9 has been shown to be regulated by TGF-β; and in vitro up-regulation of NEDD9 in human breast adenocarcinoma MCF7 cells demonstrated that it promotes migration and invasion through actin polymerization but can also simultaneously induce apoptosis and cell cycle arrest [34-36]. Recently, Lucas et al [37] demonstrated that VEGF-stimulated tumor invasion for the head and neck squamous carcinoma cell line SCC 9 was dependent on the expression of NEDD9 (human enhancer of filamentation 1), which was associated with downstream expression of metalloproteases. Thus, using antibodies directed against NEDD9 in WPOI-3 and WPOI-5 tissues, we validated high transcript expression of NEDD9 in the 7 aggressive tumor tissues, which were analyzed using the WG-DASL platform (WPOI-5). These experiments demonstrated that tumor tissue selection, using core procurement, is a valid and useful approach for identification of tumor-associated genes and demonstrated that FFPE specimens are a valuable source of RNA for gene expression profiling, followed by qRT-PCR validation experiments.

To our knowledge, only 1 other publication has profiled archived FFPE oral squamous carcinomas with the DASL platform [38]. These investigators used the previous DASL platform to query a panel of 502 genes. They studied 31 patient specimens, which had been stored for up to 5 years, and compared cancer with noncancer expression profiles. Thus, no comparison can be made between their data and that of the current study.

5. Conclusion

This preliminary study demonstrates the feasibility of using FFPE specimens for gene expression profiling using the WG-DASL platform from Illumina, followed by qRT-PCR experiments for validation of gene candidates, to identify genes associated with pathologic features using high-throughput expression profiling. We have found that some genes differentially up-regulated in WPOI-5 oral squamous carcinomas have already been linked to tumor invasion, potentially providing a biologic rationale for our WPOI classification schema. A larger study to examine the association of these genes with clinical outcome using archival samples from an adequately powered cohort of patients with OSCC is necessary to identify and validate novel gene biomarkers.

Acknowledgment

The work that is described in this article was supported by the Department of Pathology at the Albert Einstein College of Medicine and the Montefiore Medical Center, Bronx, NY. The authors thank Michael Ronan, Field Application Specialist with Illumina, Inc, for providing technical support and training with the WG-DASL platform. The authors also thank Dr Qiulu Pan, Director of the Molecular Pathology Laboratory, Montefiore Medical Center, Bronx, NY, for providing access to the Illumina equipment.

References

[1] Sessions DG, Spector GJ, Lenox J, et al. Analysis of treatment results for oral tongue cancer. Laryngoscope 2002;112:616-25.
[2] Brandwein-Gensler M, Lewis C, Lee B, et al. Oral squamous cell carcinoma: histological risk assessment, but not margin status, is strongly predictive of local disease-free and overall survival. Am J Surg Pathol 2005;20:167-78.
[3] Brandwein-Gensler M, Smith RV, Wang B, et al. Validation of the histological risk model in a new patient cohort with primary head and neck squamous cell carcinoma. Am J Surg Pathol 2010;34:676-88.
[4] Cobleigh MA, Tabesh B, Bitterman P, et al. Tumor gene expression and prognosis in breast cancer patients with 10 or more positive lymph nodes. Clin Cancer Re 2005;11:8623-31.
[5] Benchekroun M, DeGraw J, Gao J, et al. Impact of fixative on recovery of mRNA from paraffin-embedded tissue. Diagn Mol Pathol 2004;13:116-25.
[6] Roberts L, Bowers J, Sensinger K, et al. Identification of methods for use of formalin-fixed, paraffin-embedded tissue samples in RNA expression profiling. Genomics 2009;94:341-8.
[7] Bibikova M, Talantov D, Chudin E, et al. Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays. Am J Pathol 2004;165:1799-807.
[8] Ravo M, Mutarelli M, Ferraro L, et al. Quantitative expression profiling of highly degraded RNA from formalin-fixed, paraffin-embedded breast tumor biopsies by oligonucleotide microarrays. Lab Invest 2008;88:430-40.
[9] April C, Klotzle B, Royce T, et al. Whole-genome gene expression profiling of formalin-fixed, paraffin-embedded tissue samples. PLoS One 2009;4:e8162.
[10] Xie Y, Wang X, Story M. Statistical methods of background correction for Illumina BeadArray data. Bioinformatics 2009;25:751-7.
[11] Li C, Wong WH. DNA-Chip Analyzer (dChip). In: Parmigiani G, Garett ES, Irizarry RA, Zeger SL, editors. Chapter 1: "The analysis of gene expression data: methods and software" statistics for biology and health. New York: Springer; 2003.
[12] Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics 2008;24:1547-8.
[13] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc 2009;4:44-57.
[14] Dennis Jr G, Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4:P3.

[15] Hollander MW, Douglas A. Nonparametric statistical methods. New York: John Wiley & Sons; 1973.
[16] Team, R.D.C.. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing3-900051-07-0; 2009.
[17] Natarajan M, Stewart JE, Golemis EA, et al. HEF1 is a necessary and specific downstream effector of FAK that promotes the migration of glioblastoma cells. Oncogene 2006;25:1721-32.
[18] Izumchenko E, Singh MK, Plotnikova OV, et al. NEDD9 promotes oncogenic signaling in mammary tumor development. Cancer Res 2009;69:7198-206.
[19] Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 2001;98:10869-74.
[20] Kim C, Taniyama Y, Paik S. Gene expression–based prognostic and predictive markers for breast cancer: a primer for practicing pathologists. Arch Pathol Lab Med 2009;133:855-9.
[21] Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. Mol Aspects of Med 2006;27:126-39.
[22] Medeiros F, Rigl CT, Anderson GG, et al. Tissue handling for genome-wide expression analysis: a review of the issues, evidence, and opportunities. Arch Pathol Lab Med 2007;131:1805-16.
[23] Wang SS, Sherman ME, Rader JS, et al. Cervical tissue collection methods for RNA preservation: comparison of snap-frozen, ethanol-fixed, and RNAlater-fixation. Diagn Mol Pathol 2006;15:144-8.
[24] Hewitt SM, Lewis FA, Cao Y, et al. Tissue handling and specimen preparation in surgical pathology: issues concerning the recovery of nucleic acids from formalin-fixed, paraffin-embedded tissue. Arch Pathol Lab Med 2008;132:1929-35.
[25] Masuda N, Ohnishi T, Kawamoto S, et al. Analysis of chemical modification of RNA from formalin- fixed samples and optimization of molecular biology applications for such samples. Nucleic Acids Res 1999;27:4436-43.
[26] Farragher SM, Tanney A, Kennedy RD, et al. RNA expression analysis from formalin fixed paraffin embedded tissues. Histochem Cell Biol 2008;130:435-45.
[27] Chung JY, Braunschweig T, Hewitt SM. Optimization of recovery of RNA from formalin-fixed, paraffin-embedded tissue. Diagn Mol Pathol 2006;15:229-36.
[28] Ribeiro-Silva A, Zhang H, Jeffrey SS. RNA extraction from ten year old formalin-fixed paraffin- embedded breast cancer samples: a comparison of column purification and magnetic bead-based technologies. BMC Mol Biol 2007;8:118, doi:10.1186/1471-2199-8-118.
[29] Penland SK, Keku TO, Torrice C, et al. RNA expression analysis of formalin-fixed paraffin-embedded tumors. Lab Invest 2007;87:383-91.
[30] Coudry AR, Meireles RS, Cooper HS, et al. Successful application of microarray technology to microdissected formalin-fixed, paraffin-embedded tissue. J Mol Diagn 2007;9:70-9.
[31] Duenwald S, Zhou M, Wang Y, et al. Development of a microarray platform for FFPET profiling: application to the classification of human tumors. J Transl Med 2009;7:65, doi:10.1186/1479-5876-7-65.
[32] Loudig O, Milova E, Brandwein-Gensler M, et al. Molecular restoration of archived transcriptional profiles by complementary-template reverse-transcription (CT-RT). Nucleic Acids Res 2007;35:e94.
[33] Singh M, Cowell L, Seo S, et al. Molecular basis for HEF1/NEDD9/Cas-L action as a multifunctional co-ordinator of invasion, apoptosis and cell cycle. Cell Biochem Biophys 2007;48:54-72.
[34] O'Neill GM, Seo S, Serebriiskii IG, et al. A new central scaffold for metastasis: parsing HEF1/Cas-L/NEDD9. Cancer Res 2007;67:8975-9.
[35] Fashena SJ, Einarson MB, O'Neill GM, et al. Dissection of HEF1-dependent functions in motility and transcriptional regulation. J Cell Sci 2002;115:99-111.
[36] Giampieri S, Manning C, Hooper S, et al. Localized and reversible TGF beta signaling switches breast cancer cells from cohesive to single cell motility. Nat Cell Biol 2009;11:1287-96.
[37] Lucas Jr JT, Salimath BP, Slomiany MG, Rosenzweig SA. Regulation of invasive behavior by vascular endothelial growth factor is HEF1-dependent. Oncogene 2010;29:4449-59.
[38] Saleh A, Zain RB, Hussaini H, et al. Transcriptional profiling of oral squamous carcinoma using formalin-fixed paraffin-embedded samples. Oral Oncol 2010;46:379-86.