Title
A Pilot Study Comparing HPV-Positive and HPV-Negative Head and Neck Squamous Cell Carcinomas by Whole Exome Sequencing.

Permalink
https://escholarship.org/uc/item/73t1x4ks

Authors
Nichols, Anthony C
Chan-Seng-Yue, Michelle
Yoo, John
et al.

Publication Date
2012

DOI
10.5402/2012/809370

Peer reviewed
Research Article

A Pilot Study Comparing HPV-Positive and HPV-Negative Head and Neck Squamous Cell Carcinomas by Whole Exome Sequencing

Anthony C. Nichols, 1, 2, 3, 4, 5 Michelle Chan-Seng-Yue, 6 John Yoo, 1, 2, 4 Wei Xu, 7 Sandeep Dhaliwal, 1 John Basmaji, 1 Christopher C. T. Szeto, 1 Samuel Dowthwaite, 1 Biljana Todorovic, 3, 8 Maud H. W. Starmans, 6, 9 Philippe Lambin, 9 David A. Palma, 2, 4 Kevin Fung, 1, 2, 4 Jason H. Franklin, 1, 2, 4 Bret Wehrli, 5 Keith Kwan, 5 James Koropatnick, 2, 3, 4, 8 Joe S. Mymryk, 2, 3, 4, 8 Paul Boutros, 6, 10 and John W. Barrett 1, 2, 3

1 Department of Otolaryngology-Head and Neck Surgery, Western University, Victoria Hospital, London Health Science Centre, Room B3-431A, 800 Commissioners Road East, London, ON, Canada N6A 5W9
2 London Regional Cancer Program, London, ON, Canada N6A 4L6
3 Lawson Health Research Institute, London, ON, Canada N6C 2R5
4 Department of Oncology, Western University, London, ON, Canada N6A 4L6
5 Department of Pathology, Western University, London, ON, Canada N6A 5C1
6 Informatics and Biocomputing Platform, Ontario Institute for Cancer Research, Toronto, ON, Canada M5G 0A3
7 Department of Biostatistics, University of Toronto, Toronto, ON, Canada M5T 3M7
8 Department of Biostatistics, University of Toronto, Toronto, ON, Canada M5T 3M7
9 Department of Radiation Oncology (MAASTRO), GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center, P.O. Box 616, 6200 MD Maastricht, The Netherlands
10 Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada M5G 2M9

Correspondence should be addressed to Anthony C. Nichols, anthony.nichols@lhsc.on.ca

Received 31 October 2012; Accepted 22 November 2012

Academic Editors: R.-J. Bensadoun and H.-W. Lo

Copyright © 2012 Anthony C. Nichols et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Next-generation sequencing of cancers has identified important therapeutic targets and biomarkers. The goal of this pilot study was to compare the genetic changes in a human papillomavirus- (HPV-) positive and an HPV-negative head and neck tumor. Methods. DNA was extracted from the blood and primary tumor of a patient with an HPV-positive tonsillar cancer and those of a patient with an HPV-negative oral tongue tumor. Exome enrichment was performed using the Agilent SureSelect All Exon Kit, followed by sequencing on the ABI SOLiD platform. Results. Exome sequencing revealed slightly more mutations in the HPV-negative tumor (73) in contrast to the HPV-positive tumor (58). Multiple mutations were noted in zinc finger genes (ZNF3, 10, 229, 470, 543, 616, 664, 638, 716, and 799) and mucin genes (MUC4, 6, 12, and 16). Mutations were noted in MUC12 in both tumors. Conclusions. HPV-positive HNSCC is distinct from HPV-negative disease in terms of evidence of viral infection, p16 status, and frequency of mutations. Next-generation sequencing has the potential to identify novel therapeutic targets and biomarkers in HNSCC.

1. Introduction

Tobacco use has steadily declined over the last four decades [1]. In parallel, there has been a decline in cancers of most sites in the upper aerodigestive tract [2]. The exception to this trend is cancers of the oropharynx, particularly those of the palatine and lingual tonsils, which are caused by oncogenic subtypes of the human papillomavirus (HPV) [3]. The rise
in incidence of HPV-positive head and neck squamous cell carcinoma (HNSCC) has been dramatic, causing the rates of tonsillar cancer to increase by as much as threefold in some countries [3, 4]. HPV-positive patients experience markedly better survival, and their tumors are molecularly distinct from traditional head and neck cancers [5]. Overexpression of p16 and proteolysis of p53 are nearly universal in HPV-positive tumors, in contrast to frequent loss of p16 and point mutations in p53 that are found in HPV-negative cancers [5]. However, the specific mechanisms responsible for improved survival in HPV-positive patients have not been fully elucidated.

Next-generation sequencing has yielded important insights into the pathogenesis of other cancers by identifying biomarkers and therapeutic targets. High-throughput sequencing of HNSCC tumors has recently been reported, and NOTCH inactivation was the most significant finding [6, 7]. This pilot study aims to contrast the mutations seen in an HPV-positive and an HPV-negative tumor using whole exome sequencing and further our understanding about the mutations that define HNSCC.

2. Methods

2.1. Patient Selection and Tumor and Blood Sample Collection. Ethics approval was obtained from Western University Health Sciences Research Ethics Board. Informed consent was obtained from patients undergoing ablative surgery for head and neck cancer to have a portion of their tumor stored, a 10 mL blood sample taken, and their clinical parameters prospectively collected. Two patients were identified for this pilot study: a 49-year-old nonsmoking male with a T2N0 tonsillar cancer treated with transoral robotic surgery and neck dissection, and an 81-year-old female with a history of heavy smoking with a T2N0 oral tongue cancer treated with transoral robotic surgery and neck dissection and an 81-year-old female with a history of heavy smoking with a T2N0 oral tongue cancer treated with partial glossectomy, neck dissection, and free flap reconstruction. Primary site tumor specimens were taken from the center of the resection specimen. Ten mL of venous blood were drawn intraoperatively into heparinized collection tubes.

2.2. p16 Immunohistochemistry. For each patient, a portion of the primary tumor was fixed in formalin and embedded in paraffin. The blocks were then sectioned (5 μm thick). p16 immunohistochemistry was performed as previously described using a mouse monoclonal antibody against p16 (MTM Laboratories, Heidelberg, Germany) at 1:500 dilution [8]. Immunohistochemistry scoring was conducted by two study pathologists (BW and KK) blinded to HPV status and patient information. Scoring was as described by Begum et al. with strong and diffuse staining (>80 percent of tumor cells) regarded as a positive result, and negative if absent or focal [9].

2.3. DNA Extraction from Blood and Tumor Tissue. DNA was extracted from 10 mL of whole blood using the QIAmp Blood Maxi kit following instructions provided by the manufacturer (Qiagen, Valencia, CA, USA). DNA was extracted from approximately 25 mg of primary tumor using the AllPrep DNA/RNA/Protein kit (Qiagen).

2.4. In Situ Hybridization for Human Papillomavirus Testing. Slides were deparaffinized by immersion in xylene, rehydrated in alcohol, and rinsed in water. Slides were then treated with 20 μg/mL proteinase K (Sigma, St. Louis, MO) for 30 minutes, followed by immersion in 0.3% H2O2 in methanol at room temperature for 20 minutes. Slides were then treated for 10 minutes with avidin solution followed by biotin solution. The DakoGenpoint (Dako, Carpinteria, CA, USA) biotinylated probe that identifies high-risk subtypes 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68 was added to the slides. The slides were then covered and heated to 92°C for 5 minutes and then incubated at 37°C for 18 hours. DNA-DNA hybrids were detected by successive incubation with 1:100 diluted primary horseradish peroxidase-conjugated streptavidin (streptavidin-HRP) for 15 min, with biotinyl-tyramide for 15 min, and with secondary streptavidin-HRP for 15 min. A cervical cancer was used as a positive control, and a tonsil specimen from a healthy child undergoing tonsillectomy for sleep apnea was used as a negative control. Punctuate hybridization signals localized to the tumor cell nuclei defined an HPV-positive tumor. Scoring was conducted by the two study pathologists (BW and KK).

2.5. PCR Confirmation of Patient HPV Status. Patient DNA was extracted from thin section tissue slices. Briefly, a single pathology slide from each patient was deparaffinized, and then the tumor tissue was scraped into a 1.5 mL eppendorf tube containing 50 μL of Arcturus PicoPure extraction buffer, containing proteinase K (Applied Biosystems). The sample was digested at 65°C for 16 hours. The proteinase K was inactivated at 95°C for 10 minutes, and DNA was used directly in PCR. Primers were designed against unique regions of the E6-E7 loci of HPV type 16 and type 18 and synthesized by Sigma Genosys (Oakville, Canada, Table 1). Primers were also synthesized against GAPDH, a cellular gene used as a positive control for the PCR reactions. 0.2 μL of DNA extracted from the tumor tissue was added to the appropriate reaction tubes. PCR products were amplified with DNA Phusion polymerase (Thermo Scientific, Nepean, Canada) in 20 μL reactions following the manufacturer’s instructions.

2.6. Exome Sequencing. Exome libraries were created at The Centre for Advanced Genomics (Toronto, Canada) according to the manufacturer’s standard protocol for SOLiD library

---

Table 1: Primers for HPV testing.

| Name     | Sequence 5’ to 3’       |
|----------|------------------------|
| GAPDH F  | GCTATTGCCAGGGGGAGCC     |
| GAPDH R  | CGTATGATCTTTGACTGTTG    |
| HPV 16 F | TTGCCAGGATCAAGAACAGCTAG |
| HPV 16 R | GAGAGATCACTGTCTCTGTTGC  |
| HPV 18 F | CAGCGGAGGCAAGAGACG     |
| HPV 18 R | TAAAGGTCACCCGAATTTTCAT  |

F: forward, R: reverse.
preparation (Applied Biosystems, Carlsbad, CA, USA). Three µg of genomic DNA extracted from matched patient blood and tumor samples was sheared via sonication using the Covaris (S-Series) instrument. The ends of fragmented DNA were repaired and ligated to SOLiD P1 and A1 adapters provided in the Agilent Human All Exon 50 Mb Kit following the manufacturers protocol (Agilent, Santa Clara, CA, USA). The exomes were then captured using the Agilent Human All Exon 50 Mb kit, and the amplified library was purified with AMPure XP beads (Beckman Coulter Genomics, Danvers, MA). Sequencing was performed with the SOLiDTOP Paired End Sequencing Kit (Applied Biosystems). The image data collected was analyzed using the ABI corona pipeline to generate DNA reads that were mapped to the reference human genome (UCSC’s hg19) using BFAST [10].

2.7. Bioinformatics. Samples were processed as matched sets through the Genome Analysis ToolKit (GATK) v1.3-16 pipeline [11]. Samples were initially locally realigned using the IndelRealigner walker from the GATK package with known insertions and deletions found in dbSNP 135. This was followed by base quality recalibration from GATK. Default parameters were used for both steps except for SOLiD specific parameters in the recalibration step. Reads without any color space calls were marked as failing vendor quality and thus were removed from further downstream analysis. In addition, reads that had a reference base inserted into the reads due to inconsistent color space calls had those bases set to Ns with base qualities of zero. Finally variants were called and filtered using the GATK UnifiedGenotyper and VariantFiltration walkers again with default settings.

To be considered for further downstream analysis, a tumor variant had to have at least 8x coverage within the target regions 37,038,261 sites (71.86%) for the HPV-positive tumor and 38,058,450 sites (73.84%) for the HPV-negative tumors, resp.). This presented 36,020,799 and 37,049,778 comparable sites in the HPV-positive and HPV-negative tumors, respectively. This presented 36,020,799 and 37,049,778 comparable sites in the HPV-positive and HPV-negative tumors. Using in-house custom Perl code, somatic variants within the targeted regions were repaired and ligated to SOLiD P1 and A1 adapters provided in the Agilent Human All Exon 50 Mb kit following the manufacturers protocol (Agilent, Santa Clara, CA, USA). The exomes were then captured using the Agilent Human All Exon 50 Mb kit, and the amplified library was purified with AMPure XP beads (Beckman Coulter Genomics, Danvers, MA). Sequencing was performed with the SOLiDTOP Paired End Sequencing Kit (Applied Biosystems). The image data collected was analyzed using the ABI corona pipeline to generate DNA reads that were mapped to the reference human genome (UCSC’s hg19) using BFAST [10].

A reference variant required a minimum read depth of 8x within the target region for further consideration (38,673,520 sites (75.03%) and 38,058,450 sites (73.84%) for HPV-positive and HPV-negative tumors, resp.). This presented 36,020,799 and 37,049,778 comparable sites in the HPV-positive and HPV-negative tumors. Using in-house custom Perl code, somatic variants within the targeted regions were identified. To be classified as a somatic variant the following conditions had to be met: (1) a tumor variant was identified by GATK that met the above filtration requirements and (2) the corresponding position in the normal sample had 8x coverage and did not have a GATK variant call.

| Table 2: Patient demographics. | Patient 1 | Patient 2 |
|--------------------------------|-----------|-----------|
| Age                            | 49        | 81        |
| Gender                         | Male      | Female    |
| Primary site                   | Tonsil    | Oral tongue |
| Stage                          | T2N0      | T2N0      |
| Smoking                        | Nonsmoker | 50 pack years |
| Alcohol                        | Nondrinker | Rare  |
| Differentiation                | Moderate  | Moderate to poorly |
| Adverse features               | Perineural invasion | Perineural invasion |
| p16                            | Positive  | Negative  |
| HPV                            | Positive  | Negative  |
| Treatment                      | TORS + ND | Transoral resection, ND, RFFF |

HPV: human papillomavirus, TORS: transoral robotic surgery, ND: neck dissection, and RFFF: radial forearm free flap.

Somatic variants were annotated with refGene annotations (http://varianttools.sourceforge.net/Annotation/RefGene), and consequences were identified using ANNOVAR v2012-03-08 [12].

3. Results

3.1. p16 Immunohistochemistry and HPV Testing. Genomic DNA was extracted from matching tumor and blood samples from two head and neck cancer patients: patient 1 was a 49-year-old nonsmoking and nondrinking male, and patient 2 was an 81-year-old female smoker. Patient demographics, treatment details, and histopathologic parameters are outlined in Table 2. Tumor sections from each patient were stained with hematoxylin and eosin (Figures 1(a) and 1(b)). Patient 1 stained diffusely positive for p16 (Figure 1(c)), while the tumor tissue from patient 2 was negative for p16 (Figure 1(d)). In situ hybridization testing with the broad-spectrum HPV probe demonstrated strong punctate staining within nuclei of the tumor of patient 1, consistent with high-risk HPV infection (Figure 1(e)). HPV-specific, punctate nuclear staining was absent in the tumor of patient 2 (Figure 1(f)).

We employed primers designed specifically against unique portions of the E6-E7 region of HPV type 16 and type 18 to confirm the HPV status of the patients in this study. The GAPDH control was amplified from both patients; as expected, only patient 1 was HPV type 16 positive (Figure 2). Patient 2 was negative for HPV type 16, and both patients were HPV type 18 negative (data not shown).

3.2. Exome Capture and Raw Sequencing Results. The exomes from tumor tissue and matched blood samples from each patient were sequenced. For each tumor or blood sample, approximately 1.2 billion bases were sequenced, 86% of which were specific for exome sequences. The mean coverage of the exome targets was 28.1-fold, with 91.6% of the targets being sequenced at least once and 67.4% sequenced at
Figure 1: Tumors from two patients were sectioned. Slices were stained with ((a) and (b)) H&E, ((c) and (d)) p16, or ((e) and (f)) HPV in situ hybridization. Panels represent magnified images of the complete section (inset).

Figure 2: PCR confirmation of the presence of HPV type 16 DNA in patient 1 and the absence of HPV type 16 sequences in patient 2.

3.3. Bioinformatic Interpretation of Sequencing Results. We compared the sequencing results of each patient’s tumor to their matched blood samples in order to eliminate background germline variations and to focus on somatic alterations unique to the tumor genome. Although the exome capture is designed to target coding regions, some intergenic and intron regions adjacent exons are captured in the process. A complete listing of the identified variants in coding and noncoding regions for the HPV-positive and HPV-negative tumors is reported in Tables S1 and S2, respectively (see Supplementary Material available online at doi: 10.5402/2012/809370). Only the variants that occurred within exons are listed in Tables 3 and 4. Fifty-eight somatic mutations were noted in the HPV-positive tumor, 32 of which were nonsynonymous mutations within exons. Seventy-three mutations were observed in the HPV-negative tumor, including 36 coding mutations. Forty-nine of the mutated genes identified in this study were also...
shown to harbor mutations in large-scale sequencing studies (Tables S1 and S2) [6, 7].

No mutations were noted in TP53, CDKN2A (p16), or the NOTCH receptors in either tumor. However, multiple mutations were noted in zinc finger genes (ZNF3, 10, 229, 470, 543, 616, 664, 638, 716, and 799) and mucin genes (MUC4, 6, 12, and 16). Mutations were noted in MUC12 in both tumors.

Patient characteristics, PCR analysis, in situ hybridization testing, and immunohistochemistry all indicated that patient 1 was HPV type 16 positive (HPV type 18 negative by PCR) and that patient 2 was HPV negative (both type 16 and type 18 negative). When we used our four compiled exome sequences (blood and tumor from both patients) as queries against the HPV type 16 genomic sequence (RefSeq NC.001526.2), the tumor sequence from patient 1 matched numerous regions (39 hits) of the HPV 16 genome (Figure 3). Matches were identified to all the HPV type 16 genes (except E4) suggesting that the HPV type 16 genome had integrated into the tumor genome of patient 1. The tumor sample from patient 2 and the blood samples from both patients did not align to any specific HPV sequences.

4. Discussion

HPV-positive head and neck squamous cell carcinoma (HNSCC) has been described as molecularly distinct from traditional head and neck cancer [5]. The human papillomavirus (HPV) oncoproteins E6 and E7 promote carcinogenesis by degrading the tumor suppressors p53 and retinoblastoma protein (Rb), respectively. In contrast, p53 is not degraded in HPV-negative HNSCC but is frequently mutated, and p16 is often lost through homozygous deletion, methylation, or, less frequently, point mutation [5, 14]. This might lead one to believe that carcinogens like tobacco and alcohol would promote HNSCC comprised of a large number of mutations in many different pathways. In contrast HPV-positive cancers, modulated by the activities of viral oncoproteins, might not accumulate a large number of cellular mutations. In our study, we provided quadruple confirmation of tumor HPV status with p16 immunohistochemistry, HPV in situ hybridization, HPV detection by PCR, and detection of the HPV 16 genome sequences within patient 1’s sequenced exome. We observed more mutations in the HPV-negative tumor when compared to the HPV-positive tumor, although the absolute difference was not dramatic (73 versus 58, resp.). Two large-scale exome sequencing efforts characterizing HNSCC have been reported recently [6, 7]. The study led by Stransky et al. reported twice as many mutations in the HPV-negative samples (4.83 mutations/Mb versus 2.28 mutations/Mb) [7]. The second group examined a set of 32 patients, four of which were HPV positive and reported on a subset of mutations that were identified by exome sequencing and confirmed by PCR. In this subset of genes, there were four times as many mutations in the HPV-negative tumors (20.6 ± 16.7 versus 4.8 ± 3 mutations in the HPV-positive tumors) [6]. Given the broad range of mutations seen in the HPV-negative cancers, our finding of slightly more mutations in the HPV-negative tumor is consistent with their results. As expected we did not identify TP53 or p16 mutations in the HPV-positive tumor; however these two genes appeared as wild type in the HPV-negative tumor as well. The lack of a p16 mutation in the setting of low expression levels as evidenced by immunohistochemistry may reflect that it has been inactivated by promoter methylation, the second most common cause of p16 loss [14].

Only a single genetic mutation (Muc12) was shared by both HPV-positive and HPV-negative tumor samples. The cell surface associated Muc12 was the only mucin identified in the HPV-negative tumor. In contrast, the HPV-positive tumor had five mutations in four different mucin genes, including the secreted Muc6, and the transmembrane bound Muc4, Muc12 and Muc16. Stransky et al. reported mutations in all the above mucins except for Muc12 [7]. Mucins are known to be involved in the development of epithelial cancer where they are often overexpressed, disrupting the epithelial cell polarity and promoting the epithelial to mesenchymal transition (EMT) phenotype [15]. Multiple damaging mutations within the mucins of HPV-positive tumor may suggest another cellular difference between these two distinct tumor types.

We also found multiple mutations in the zinc finger (ZNF) family genes in both tumor types. The ZNF family represents a large group of molecules which are involved in various aspects of transcriptional regulation [16]. There were almost twice as many ZNF mutated genes in the HPV-positive sample. Although there were a total of 11 ZNF mutations between the two tumor types, there were no shared ZNF members mutated in both cancers. Stransky et al.
| Chr | Position | Reference allele | Tumor allele | Zygosity | dhSNP ID  | Region | Type          | Gene     | Transcript name | Exon | CDS position | Protein change | Stansky* | Agrawal** |
|-----|----------|-----------------|-------------|----------|-----------|--------|--------------|----------|----------------|------|-------------|----------------|----------|-----------|
| 19  | 15789140 | A               | G           | Homo     | rs609290  | Exonic | Nonsynonymous | CYP4F12  | NM_023944      | 3    | c.A268G     | p.190V         | Yes      |           |
| 10  | 88946876 | G               | A           | Hetero   | rs1950902 | Exonic | Nonsynonymous | FAM35A   | NM_019054      | 8    | c.G2227A    | p.D743N        | Yes      |           |
| 3   | 51930850 | T               | C           | Homo     | rs57859638| Exonic | Nonsynonymous | IQCF1    | NM_152397      | 3    | c.A169G     | p.K57E         | Yes      |           |
| 7   | 20180717 | C               | G           | Homo     | rs3735615 | Exonic | Nonsynonymous | MACC1    | NM_182762      | 7    | c.G2411C    | p.R804T        | Yes      |           |
| 14  | 64882380 | A               | G           | Hetero   | rs1950902 | Exonic | Nonsynonymous | MTHFD1   | NM_005956      | 6    | c.A401G     | p.K134R        | Yes      |           |
| 1    | 100634194| C               | G           | Hetero   | MUC12     | NM_001164462| Exonic | Nonsynonymous | MUC12    | 2            | c.C350G        | p.A117G  |           |
| 19  | 9006667  | A               | G           | Homo     | rs75101943| Exonic | Nonsynonymous | MUC16    | NM_024690      | 44   | c.T39581C   | p.I13194T      | Yes      |           |
| 9   | 9074122  | T               | G           | Homo     | MUC16     | NM_024690 | Exonic | Nonsynonymous | MUC4     | 2            | c.G698A        | p.G233E  |           |
| 3   | 195517753| C               | T           | Hetero   | MUC4      | NM_018406 | Exonic | Nonsynonymous | MUC1     | 1            | c.C5240G       | p.T180S  | Yes       |
| 11  | 1017381  | G               | C           | Hetero   | rs34912894| Exonic | Nonsynonymous | NAPSA    | NM_004851      | 2    | c.T119C     | p.H1278D       | Yes      |           |
| 4   | 50885355 | A               | G           | Homo     | rs676314  | Exonic | Nonsynonymous | FAM35A   | NM_019054      | 8    | c.G2227A    | p.D743N        | Yes      |           |
| 17   | 170345835| G               | C           | Hetero   | NEK1      | NM_001199397| Exonic | Nonsynonymous | NEK1     | 31           | c.C3901G       | p.R131Q  | Yes       |
| 3   | 3195485  | C               | T           | Hetero   | OR3A1     | NM_002550 | Exonic | Nonsynonymous | OR3A1    | 1            | c.G392A        | p.R314Q  | Yes       |
| 9   | 98073075 | A               | G           | Hetero   | OR5K4     | NM_001005517| Exonic | Nonsynonymous | OR5K4    | 1            | c.A378G        | p.N126M  | Yes       |
| 14  | 140590766| A               | G           | Homo     | rs2910006 | Exonic | Nonsynonymous | PCDH12   | NM_018932      | 1    | c.A2287G    | p.K763E        | Yes      |           |
| 12  | 27787878 | G               | A           | Hetero   | PPFIB1    | NM_001198916| Exonic | Nonsynonymous | PPFIB1   | 4            | c.G100A        | p.D34N   |           |
| 15  | 43827261 | G               | C           | Hetero   | PPi5K1    | NM_001190214| Exonic | Nonsynonymous | PPi5K1   | 30           | p.C3832G       | p.H1278D| Yes       |
| 1    | 12853378 | T               | G           | Hetero   | PRAMEF1   | NM_023013 | Exonic | Nonsynonymous | PRAMEF1  | 2            | c.T2G          | p.MIR    | Yes       |
| X    | 13595675 | G               | A           | Hetero   | RBMX      | NM_002139 | Exonic | Nonsynonymous | RBMX     | 9            | c.G92T         | p.P310L  | Yes       |
| 6   | 33272855 | G               | C           | Homo     | TAPBP     | NM_172208 | Exonic | Nonsynonymous | TAPBP    | 4            | c.G779G        | p.T260R  | Yes       |
| 14  | 11183046 | C               | T           | Hetero   | TAS2R31   | NM_176885 | Exonic | Nonsynonymous | TAS2R31  | 1            | c.G889A        | p.V297M  |           |
| 12  | 11213969 | A               | G           | Hetero   | TAS2R46   | NM_176887 | Exonic | Nonsynonymous | TOMM70A  | 1            | c.T925C        | p.S309P  | Yes       |
| 3   | 100084425| T               | G           | Homo     | TOMM70A   | NM_014820 | Exonic | Nonsynonymous | TOMM70A  | 12           | c.A1810C       | p.K640Q  | Yes       |
| 15  | 135277007| C               | T           | Homo     | rs3739916 | Exonic | Nonsynonymous | TTF1     | NM_007344      | 2    | c.G1202A    | p.R401Q        | Yes      |           |
| 19  | 38578569 | A               | T           | Hetero   | WDR87     | NM_031951 | Exonic | Nonsynonymous | WDR87    | 6            | c.T5353A       | p.N1845K| Yes       |
| 7   | 74975645 | C               | G           | Hetero   | ZFAND5    | NM_001102421| Exonic | Nonsynonymous | ZFAND5   | 2            | c.G50C         | p.G17A   | Yes       |
| 12  | 133732818| C               | T           | Hetero   | ZNF10     | NM_015394 | Exonic | Nonsynonymous | ZNF10    | 5            | c.C986T        | p.S329F  | Yes       |
| 4    | 44937006 | C               | T           | Hetero   | ZNF229    | NM_014518 | Exonic | Nonsynonymous | ZNF229   | 6            | c.G1250A       | p.S417N  | Yes       |
| 5    | 57089050 | C               | T           | Hetero   | ZNF470    | NM_001001668| Exonic | Nonsynonymous | ZNF470   | 6            | c.C1253T       | p.T4181  | Yes       |
| 5    | 52618497 | A               | T           | Hetero   | ZNF616    | NM_178523 | Exonic | Nonsynonymous | ZNF616   | 4            | c.T1920A       | p.N640K  | Yes       |
| 7    | 71651475 | G               | A           | Homo     | ZNF638    | NM_001252612| Exonic | Nonsynonymous | ZNF638   | 24           | c.G5176A       | p.V1726M| Yes       |
| 7    | 57528923 | G               | C           | Hetero   | ZNF716    | NM_001159279| Exonic | Nonsynonymous | ZNF716   | 4            | c.G756C        | p.W252C  |           |

Chr: chromosome, *mutations present in Stransky et al., Science 2011 [7], **mutations present in Agrawal et al., Science 2011 [6], Homo: homozygous, Hetero: heterozygous.
| Chr | Position | Reference allele | Tumor allele | Zygosity | dbSNP ID      | Region | Type              | Gene     | Transcript name | Exon | CDS position | Protein change | Stansky* | Agrawal** |
|-----|----------|------------------|--------------|-----------|---------------|--------|-------------------|----------|----------------|------|--------------|----------------|----------|-----------|
| 2   | 73679256 | C                | G            | Homo      |               |        | Exonic Nonsynonymous | ALMS1    | NM_015120      | 8    | C.S5599G     | p.L1867V       | Yes      |           |
| 10  | 37451586 | G                | A            | Hetero    | rs1276684     | Exonic | Nonsynonymous      | ANKRD30A | NM_052997      | 16   | C.G1742A     | p.G581E        | Yes      |           |
| 9   | 33385733 | C                | T            | Hetero    |               | Exonic | Nonsynonymous      | AQP7     | NM_001170      | 7    | C.G657A      | p.M219I         |         |           |
| 17  | 42271663 | G                | C            | Homo      |               | Exonic | Nonsynonymous      | ATXN7L3  | NM_01098833    | 12   | C.I102G      | p.P338A         |         |           |
| 6   | 136599913| A                | T            | Hetero    |               | Exonic | Nonsynonymous      | BCLA1F1 | NM_00177441    | 4    | C.T106A      | p.S6T           |         |           |
| 1   | 26647626 | A                | G            | Hetero    | rs1071849     | Exonic | Nonsynonymous      | CD52     | NM_001803      | 2    | C.A119G      | p.N40S          |         |           |
| 1   | 26647630 | A                | G            | Homo      | rs17645      | Exonic | Nonsynonymous      | CD52     | NM_001803      | 2    | C.A123G      | p.I41M          |         |           |
| 19  | 42213948 | C                | G            | Hetero    |               | Exonic | Nonsynonymous      | CEACAM5 | NM_004363      | 2    | C.C141G      | p.F138L         |         |           |
| 12  | 31242362 | A                | G            | Hetero    | rs395088     | Exonic | Nonsynonymous      | DD1X11   | NM_003653      | 8    | C.A818G      | p.K273R         | Yes      |           |
| 8   | 83657820 | G                | C            | Hetero    |               | Exonic | Nonsynonymous      | FAM103A1 | NM_031452      | 3    | C.G50C       | p.X17S          |         |           |
| 6   | 76023130 | T                | G            | Homo      |               | Exonic | Nonsynonymous      | FILL1P   | NM_015687      | 5    | C.A2418C     | p.Q806H         | Yes      |           |
| 1   | 152191709| C                | G            | Hetero    | rs6662450     | Exonic | Nonsynonymous      | HRNR     | NM_01009931    | 3    | C.G2396C     | p.S799T         | Yes      |           |
| 9   | 21202136 | A                | G            | Hetero    | rs145794215   | Exonic | Nonsynonymous      | IFNA7    | NM_021057      | 1    | C.T29C       | p.V10A          |         |           |
| 10  | 100635103| A                | C            | Hetero    |               | Exonic | Nonsynonymous      | MUC12    | NM_001164462   | 2    | C.A1259C     | p.E420A         |         |           |
| 17  | 10404435 | T                | G            | Hetero    |               | Exonic | Nonsynonymous      | MYH1     | NM_005963      | 27   | C.A3730C     | p.K124Q         | Yes      |           |
| 18  | 47360806 | C                | T            | Hetero    |               | Exonic | Nonsynonymous      | MYO5B    | NM_00108467    | 37   | C.G4939A     | p.D1647N        | Yes      |           |
| 12  | 57106595 | C                | G            | Hetero    |               | Exonic | Nonsynonymous      | NACA     | NM_00113203    | 10   | C.G2674C     | p.E892Q         |         |           |
| 1   | 21795333 | C                | G            | Hetero    | rs6662450     | Exonic | Nonsynonymous      | NBP3F3   | NM_0012264     | 3    | C.C286G      | p.Q69E          | Yes      | Yes       |
| 11  | 113570274| T                | C            | Homo      | rs10749138    | Exonic | Nonsynonymous      | NRA1     | NM_0006175     | 30   | C.A3442G     | p.I1148V        | Yes      | Yes       |
| 11  | 48346916 | G                | C            | Homo      | rs77069283    | Exonic | Nonsynonymous      | OR4C3    | NM_001044072   | 1    | C.G424C      | p.V142L         |         |           |
| 11  | 4608046  | T                | A            | Hetero    |               | Exonic | Nonsynonymous      | OR5212   | NM_00105170    | 1    | C.T4A        | p.C2S           |         |           |
| 11  | 56128081 | A                | G            | Hetero    | rs10896200    | Exonic | Nonsynonymous      | OR8J1    | NM_00105205    | 1    | C.A359G      | p.Y120C         | Yes      |           |
| 11  | 56143823 | G                | A            | Homo      | rs77614949    | Exonic | Nonsynonymous      | OR8U1,   | NM_00105204    | 1    | C.G72A       | p.G242S         |         |           |
| 11  | 124497119| T                | A            | Hetero    |               | Exonic | Nonsynonymous      | OR8U2    | NM_00105204    | 1    | C.G72A       | p.G242S         |         |           |
| 12  | 1174327  | C                | T            | Hetero    | rs72475481    | Exonic | Nonsynonymous      | OR8U3    | NM_00105204    | 1    | C.G72A       | p.G242S         |         |           |
| 12  | 1183676  | C                | T            | Hetero    | rs73049072    | Exonic | Nonsynonymous      | OR8U3    | NM_00105204    | 1    | C.G72A       | p.G242S         |         |           |
| 12  | 1174327  | C                | T            | Hetero    | rs72475481    | Exonic | Nonsynonymous      | OR8U3    | NM_00105204    | 1    | C.G72A       | p.G242S         |         |           |
| 7   | 3524086  | A                | T            | Hetero    |               | Exonic | Nonsynonymous      | TBX20    | NM_00107653    | 7    | C.T999A      | p.N333K         |         |           |
| 10  | 10190347 | A                | G            | Hetero    | rs150482     | Exonic | Nonsynonymous      | TPTE     | NM_199261     | 22   | C.T1409C     | p.L470P         |         |           |
| 9   | 99669149 | C                | T            | Hetero    |               | Exonic | Nonsynonymous      | ZNF3     | NM_0032942     | 6    | C.G958A      | p.A320T         |         |           |
| 19  | 57839150 | A                | G            | Homo      | rs8100491    | Exonic | Nonsynonymous      | ZNF543   | NM_213598      | 4    | C.A320G      | p.Q107R         | Yes      |           |
| 19  | 124497119| T                | A            | Hetero    |               | Exonic | Nonsynonymous      | ZNF543   | NM_213598      | 4    | C.A320G      | p.Q107R         | Yes      |           |
| 19  | 12501852 | C                | T            | Hetero    |               | Exonic | Nonsynonymous      | ZNF599   | NM_001080821   | 4    | C.G1360A     | p.G454R         | Yes      |           |

Chr: chromosome, *mutations present in Stransky et al., Science 2011 [7], **mutations present in Agrawal et al., Science 2011 [6], Homo: homozygous, Hetero: heterozygous.
reported 50% of the same ZNF mutations that we found in both tumor samples suggesting that genetic changes to this family of transcription regulators may be important in the development of HNSCC [7]. Not enough is known about the role of mucins and ZNF proteins in HNSCC. These molecules may warrant further study.

We confirmed that the sequence from the human pathogen HPV type 16 was identified within exome sequence of a HNSCC tumor. In order for HPV to be oncogenic, the viral E2 protein, which represses the expression of E6 and E7, must be lost [17]. This only occurs during integration when the episomal HPV DNA breaks within the E2 gene. PCR detection of E6 and E7 can detect both episomal and integrated forms and thus cannot distinguish between a superficial HPV infection and integrated viral DNA causing the cancer [17]. An additional benefit of whole exome sequencing is the detection of integrated HPV DNA. In Stranksy’s study, exome sequencing appeared even more sensitive than PCR for detection of HPV, as it identified the presence of HPV 16 in 14 of 73 cases versus 11 for PCR [7]. Perhaps more interesting is the concept of screening human disease genomes against pathogen datasets. In fact, it was this exact strategy that led to the discovery of Merkel cell polyomavirus in 2008 [18]. It may be that a subset of other cancers have a yet undiscovered viral etiology.

This study represents a pilot effort to gain experience with this exciting new technology, which was instructive as our group moves forward with large-scale projects. In addition to the small sample size, the quality of data generated limited by the ABI SOLiD platform with an average 30-fold coverage with 50 base pair paired-end reads yielded only 10-fold coverage over approximately two-thirds of the coding sequence. Thus, approximately a third of the exome was not adequately evaluated and important mutations could have been missed. We have recently completed characterizing a panel of head and neck cancer cell lines with 100-fold coverage with 100 base pair paired-end reads, and the results were vastly superior [19]. An average of 99% of the targeted exome had at least 10 reads and 90% had fiftyfold coverage. Perhaps more importantly, an expert bioinformatics team is critical to properly analyze the data. Although there are standard steps involved with aligning the sequencing data to the reference genome, false positive results can be frequent without adequate quality control measures. A carefully validated pipeline is necessary to filter spurious results in order to generate valid data.

It should be noted that tremendous insights can be gained by exome sequencing; however, whole genome sequencing offers the advantage to identify other genetic changes that can lead to tumorigenesis including copy number variation and translocations, in addition to point mutations, insertions, and deletions. Alterations in noncoding regions that may be important, such as promoters and miRNAs, would also be identified. The study by Stranksy et al. reported whole genome sequencing in two patients and revealed markedly more translocations in the HPV-negative versus the HPV-positive tumors [7]. Ideally, future large-scale initiatives will be carried out using this more extensive but also more expensive technique to identify additional important genetic changes underlying HNSCC.

5. Conclusions

Whole exome sequencing of head and neck cancers can provide important insights into the molecular biology of the disease. HPV-positive and negative head and neck cancers are molecularly distinct, and HPV-negative cancers tend to harbor more mutations. Multiple, integrated HPV 16 sequences were identified in the exome targets from the HPV positive patient. These matches were restricted to the HPV-positive patient’s tumor profile providing evidence of the utility of screening exome sequences against pathogen databases.

Disclosure

None of the authors have financial relationships with any of the commercial identities mentioned in this paper.

Authors’ Contribution

Anthony C. Nichols, Michelle Chan-Seng-Yue, John Yoo, Paul Boutros, and John W. Barrett all contributed equally to this work.

Acknowledgment

Exome sequencing was performed by The Centre for Advanced Genomics (TCAG) at the University of Toronto. Funding was provided through the translational Head and Neck Cancer Program fund, Western University (ACN), and the support of the Ontario Institute for Cancer Research (PCB) by the Government of Ontario are gratefully acknowledged.

References

[1] A. Jemal, M. J. Thun, L. A. G. Ries et al., “Annual report to the nation on the status of cancer, 1975–2005, featuring trends in lung cancer, tobacco use, and tobacco control,” Journal of the National Cancer Institute, vol. 100, no. 23, pp. 1672–1694, 2008.
[2] L. A. Ries, D. Melbert, M. Krapcho et al., “SEER Cancer Statistics Review, 1975–2004. Periodical,” National Cancer Institute, Bethesda, Md, USA, 2007.
[3] A. K. Chaturvedi, E. A. Engels, R. M. Pfeiffer et al., “Human papillomavirus and rising oropharyngeal cancer incidence in the United States,” Journal of Clinical Oncology, vol. 29, no. 32, pp. 4294–4301, 2011.
[4] A. Näsmann, P. Attner, L. Hammarstedt et al., “Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma?” International Journal of Cancer, vol. 125, no. 2, pp. 362–366, 2009.
[5] M. L. Gillison, “Human papillomavirus-associated head and neck cancer is a distinct epidemiologic, clinical, and molecular entity,” Seminars in Oncology, vol. 31, no. 6, pp. 744–754, 2004.
[6] N. Agrawal, M. J. Frederick, C. R. Pickering et al., “Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1,” *Science*, vol. 333, no. 6046, pp. 1154–1157, 2011.

[7] N. Stransky, A. M. Egloff, A. D. Tward et al., “The mutational landscape of head and neck squamous cell carcinoma,” *Science*, vol. 333, no. 6046, pp. 1157–1160, 2011.

[8] A. C. Nichols, W. C. Faquin, W. H. Westra et al., “HPV-16 infection predicts treatment outcome in oropharyngeal squamous cell carcinoma,” *Otolaryngology—Head and Neck Surgery*, vol. 140, no. 2, pp. 228–234, 2009.

[9] S. Begum, M. L. Gillison, M. A. Ansari-Lari, K. Shah, and W. H. Westra, “Detection of human papillomavirus in cervical lymph nodes: a highly effective strategy for localizing site of tumor origin,” *Clinical Cancer Research*, vol. 9, no. 17, pp. 6469–6475, 2003.

[10] N. Homer, B. Merriman, and S. F. Nelson, “BFAST: an alignment tool for large scale genome resequencing,” *PLoS ONE*, vol. 4, no. 11, Article ID e7767, 2009.

[11] A. McKenna, M. Hanna, E. Banks et al., “The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data,” *Genome Research*, vol. 20, no. 9, pp. 1297–1303, 2010.

[12] K. Wang, M. Li, and H. Hakonarson, “ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data,” *Nucleic Acids Research*, vol. 38, no. 16, p. e164, 2010.

[13] D. J. Hedges, T. Guettouche, S. Yang et al., “Comparison of three targeted enrichment strategies on the SOLiD sequencing platform,” *PLoS ONE*, vol. 6, no. 4, Article ID e18595, 2011.

[14] A. L. Reed, J. Califano, P. Cairns et al., “High frequency of p16 (CDKN2/MTS-1/INK4A) inactivation in head and neck squamous cell carcinoma,” *Cancer Research*, vol. 56, no. 16, pp. 3630–3633, 1996.

[15] D. W. Kufe, “Mucins in cancer: function, prognosis and therapy,” *Nature Reviews Cancer*, vol. 9, no. 12, pp. 874–885, 2009.

[16] S. S. Krishna, I. Majumdar, and N. V. Grishin, “Structural classification of zinc fingers: survey and summary,” *Nucleic Acids Research*, vol. 31, no. 2, pp. 532–550, 2003.

[17] M. J. Lace, J. R. Anson, J. P. Klussmann et al., “Human papillomavirus type 16 (HPV-16) genomes integrated in head and neck cancers and in HPV-16-immortalized human keratinocyte clones express chimeric virus-cell mRNAs similar to those found in cervical cancers,” *Journal of Virology*, vol. 85, no. 4, pp. 1645–1654, 2011.

[18] H. Feng, M. Shuda, Y. Chang, and P. S. Moore, “Clonal integration of a polyomavirus in human Merkel cell carcinoma,” *Science*, vol. 319, no. 5866, pp. 1096–1100, 2008.

[19] A. C. Nichols, J. Yoo, D. A. Palma et al., “Frequent mutations in TP53 and CDKN2A found by next-generation sequencing of head and neck cancer cell lines,” *Archives of Otolaryngology—Head & Neck Surgery*, vol. 138, no. 8, pp. 732–739, 2012.