processes

MDPI

Article

Comprehensive Analysis of Mutation-Based and Expressed Genes-Based Pathways in Head and Neck Squamous Cell Carcinoma

Bhumsuk Keam [1], Jin-Young Park [2], Jin-Pyo Kim [2], Gun-Do Kim [2], Yun-Suk Yu [2], Sang-Hee Cho [3], Sangwoo Kim [4], Hee-Kyung Ahn [5], Sang Hoon Chun [6], Jung-Hye Kwon [7], Tak Yun [8], Ji-Won Kim [9], Ji-Eun Kim [10], Myung-Ju Ahn [11], Joo-Hang Kim [12] and Hwan-Jung Yun [13,*]

1 Department of Internal Medicine, Seoul National University Hospital, Seoul 03080, Korea; bhumsuk@snu.ac.kr
2 Cbs Bioscience Inc., Daejeon-si 34036, Korea; jonnypark@cbsbio.com (J.-Y.P.); kjp@cbsbio.com (J.-P.K.); gundokim@cbsbio.com (G.-D.K.); yuyunsuk@cbsbio.com (Y.-S.Y.)
3 Department of Hemato-Oncology, Chonnam National University Hwasun Hospital, Chonnam 58128, Korea; shcho@chonnam.ac.kr
4 Department of Biomedical Systems Informatics and Brain Korea 21 PLUS Project for Medical Science, Yonsei University College of Medicine, Seoul 03722, Korea; SWKIM@yuhs.ac
5 Department of Medical Oncology, Gachon University Gil Medical Center, Incheon-si 21565, Korea; hkahn@gilhospital.com
6 Department of Internal Medicine, Bucheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Bucheon 14647, Korea; rowett1004@gmail.com
7 Division of Hemato-Oncology, Department of Internal Medicine, Hallym University College of Medicine, Kangdong Sacred Heart Hospital, Seoul 05355, Korea; kwonjhye@naver.com
8 Rare Cancers Clinic, Center for Specific Organs Cancer, National Cancer Center, Goyang-si 10408, Korea; hmotakyun@ncc.re.kr
9 Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam 13620, Korea; werbinig@gmail.com
10 Department of Pathology, Seoul National University College of Medicine, SMG-SNU Boramae Hospital, Seoul 07061, Korea; npol181@snu.ac.kr
11 Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, Korea; silkahn@skku.edu
12 Division of Medical Oncology, Department of Internal Medicine, CHA Bundang Medical Center, Seongnam-si 13496, Korea; kim123@cha.ac.kr
13 Department of Internal Medicine, Chungnam National University Hospital, Daejeon-si 35015, Korea
* Correspondence: hjyun@cnuh.co.kr; Tel.: +82-42-280-7157; Fax: +82-42-257-5753

Abstract: Over- or under-expression of mRNA results from genetic alterations. Comprehensive pathway analyses based on mRNA expression are as important as single gene level mutations. This study aimed to compare the mutation- and mRNA expression-based signaling pathways in head and neck squamous cell carcinoma (HNSCC) and to match these with potential drug or druggable pathways. Altogether, 93 recurrent/metastatic HNSCC patients were enrolled. We performed targeted gene sequencing using Illumina HiSeq-2500 for NGS, and nanostring nCounter® for mRNA expression; mRNA expression was classified into over- or under-expression groups based on the expression. We investigated mutational and nanostring data using the CBSJukebox® system, which is a big-data driven platform to analyze druggable pathways, genes, and protein-protein interaction. We calculated a Treatment Benefit Prediction Score (TBPS) to identify suitable drugs. By mapping the high score interaction genes to identify druggable pathways, we found highly related signaling pathways with mutations. Based on the mRNA expression and interaction gene scoring model, several pathways were found to be associated with over- and under-expression. Mutation-based pathways were associated with mRNA under-expressed genes-based pathways. These results suggest that HNSCCs are mainly caused by the loss-of-function mutations. TBPS found several matching drugs such as immune checkpoint inhibitors, EGFR inhibitors, and FGFR inhibitors.

Keywords: head and neck squamous cell carcinoma (HNSCC); precision medicine; pathway analysis; CBSJukebox

Citation: Keam, B.; Park, J.-Y.; Kim, J.-P.; Kim, G.-D.; Yu, Y.-S.; Cho, S.-H.; Kim, S.; Ahn, H.-K.; Chun, S.-H.; Kwon, J.-H.; et al. Comprehensive Analysis of Mutation-Based and Expressed Genes-Based Pathways in Head and Neck Squamous Cell Carcinoma. *Processes* **2021**, *9*, 792. https://doi.org/10.3390/pr9050792

Academic Editor: Bonglee Kim

Received: 27 March 2021
Accepted: 16 April 2021
Published: 30 April 2021

1. Introduction

Head and neck squamous cell carcinoma (HNSCC) is not a single disease entity, but a highly heterogeneous group of diseases categorized by diverse tumor types arising from various anatomic structures including oral cavity, oropharynx, hypopharynx, larynx, and paranasal sinus. In the era of precision oncology, traditional classification based on pathology is not sufficient to achieve accurate clinical diagnostics [1]. Next generation sequencing (NGS) revealed that HNSCC is more heterogeneous based on mutational and molecular subtypes [2–4].

Recently, we also found several targetable genetic alterations in HNSCC, suggesting that implementation of precision medicine in HNSCC was feasible [5]. Based on this feasibility, we designed an umbrella trial for recurred/metastatic HNSCC, consisting of five targeted therapies including PI3K inhibitor, pan-HER inhibitor, FGFR inhibitor, CDK4/6 inhibitor, and immune checkpoint inhibitor (ClinicalTrials.gov: NCT03292250) [6]. Although potentially targetable genetic alterations in genes such as PIK3CA, EGFR, and FGFR have been identified in HNSCC, in-depth functional studies to validate their roles as predictive biomarkers have not been performed.

Integration of cancer genes into networks offers opportunities to reveal protein–protein interactions (PPIs) with functional and therapeutic significance. PPI networks based on cancer gene landscapes can give us insight into how these genes contribute to deregulated oncogenic pathways [7–10]. Pathological over- or under-expression of mRNA results from cancer specific genetic alterations. Genetic mutation without change in mRNA expression might not result in the functional change at the protein level; mRNA expression based pathways are as important as single gene level mutational analysis.

In this study, we aimed to compare the mutational and mRNA expression based signal transduction pathways in HNSCC and establish a cancer-associated PPI network in an efficient high throughput format. The objective of this study was to integrate the DNA mutational landscape and mRNA expression patterns into the PPI network pathways, which could then be used to match potential drug or druggable pathway in HNSCC.

2. Patients and Methods

2.1. Patients and Data Collection

Altogether, 93 recurrent/metastatic HNSCC patients from 19 institutions were enrolled. The details of the study population have been described in our previous report [5]. In brief, pretreatment tumor tissues (somatic) and matched normal DNA (germline) from prospectively recruited patients with HNSCC were used for the analysis. Clinicopathological data were collected from patient medical records. Informed consents were obtained. The Institutional Review Boards of each institute approved this study protocol.

2.2. Targeted Gene Sequencing and mRNA Expression Assay

Genomic DNA was isolated from formalin-fixed paraffin-embedded (FFPE) tissue samples for the targeted sequencing of 244 head and neck cancer-related genes. The genomic regions of the 244 genes were captured by the customized SureSelectXT Target Enrichment library generation kit (Agilent, Santa Clara, CA, USA) and sequenced using the Illumina HiSeq 2500 platform with a depth of coverage >1000×. The nCounter Analysis System (Nanostring Technologies, Seattle, WA, USA) was used to screen for the expression of 93 immune-related genes. Counts were normalized to the internal controls and reference genes using the nSolver software, version 4.0 (NanoString, Seattle, WA, USA).

2.3. Basic Scheme of Protein-Protein Interaction Network Analysis

To analyze with a deep insight of the combinatorial signaling events evolved in cell communication, we applied a novel PPI method called CBSJukebox®. Figure 1 shows the analytic flow in CBSJukebox®. In brief, CBSJukebox® enabled us to compare DNA

mutation-based pathways and over- or under-expression based pathways by using PPI analysis; further, CBSJukebox® enabled us to perform a simple signal pathway analysis as well as high interaction frequency ratio genes analysis.

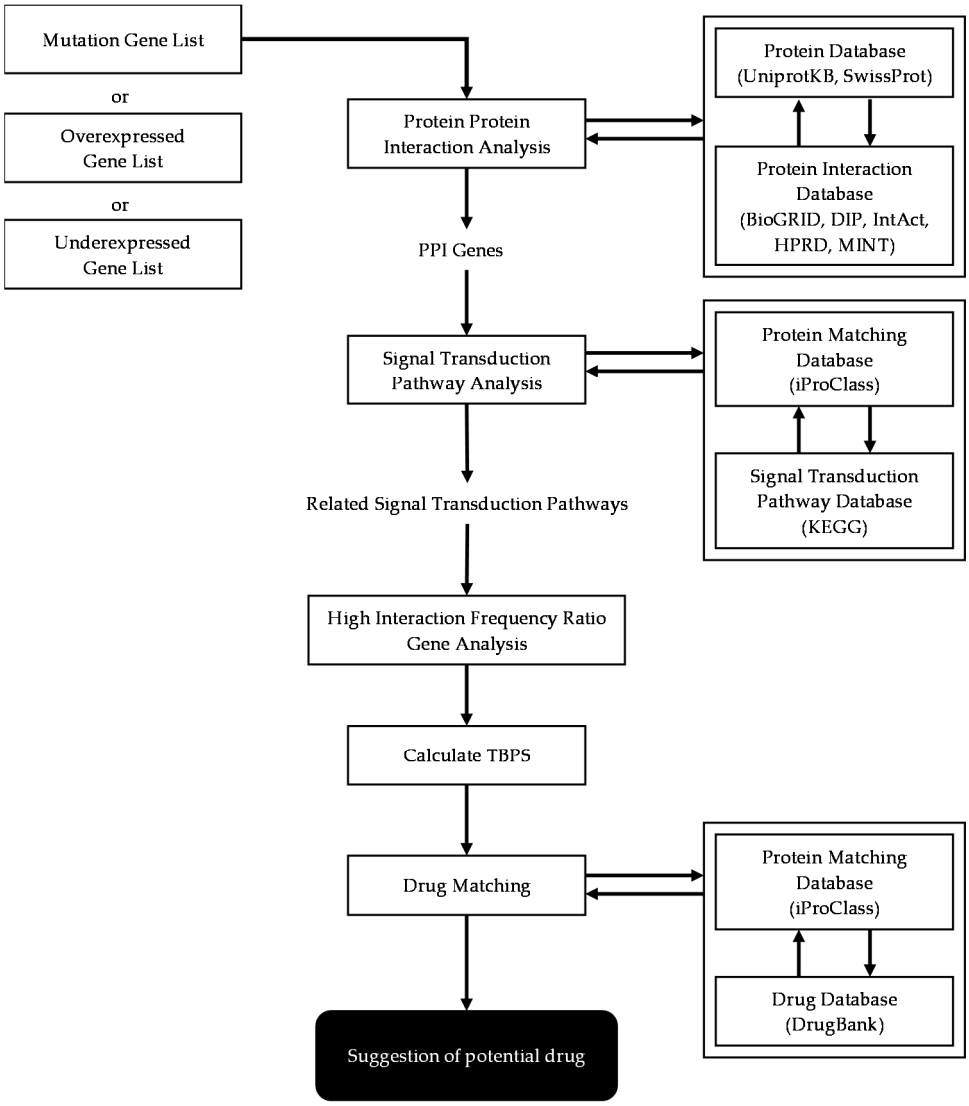

Figure 1. Overflow of protein-protein interaction analysis (CBSJukebox® analysis).

2.4. Gene List Enrichment

The variants selected for DNA mutation-based analysis included nonsynonymous single-nucleotide variant (SNV), frameshift inserts, frameshift deletions, stop-gain, stop-loss, and copy number variation (CNV). The significantly different mRNAs expression subtypes were identified as over- or under-expressed genes based on Student's *t* test (p-value < 0.05 and $|\text{fold-change}| > 2$) and compared with those expressed in normal tissue for further over- or under-expression based pathway analysis.

2.5. PPI Mapping of Mutated Genes and Over- or Under-Expressed Genes

A multi-functional analytical tool, CBSJukebox®, was used to match DNA mutated genes with Entrez Gene records (NCBI ID, https://www.ncbi.nlm.nih.gov/gene, accessed on 25 March 2021) from the iProClass (https://www.ncbi.nlm.nih.gov/pubmed/15022647, accessed on 25 March 2021) database, and the over- or under-expressed genes were matched with gene name and synonym in Uniprot/Swiss-Prot (Uniprot Knowledgebase, https://www.ncbi.nlm.nih.gov/pubmed/27899622, accessed on 25 March 2021) further to interchange with identification factor "Uniprot Ac" in CBSJukebox®. We then conducted

the interactive protein network analysis using the IntAct (IntAct, http://europepmc.org/abstract/MED/24234451, accessed on 25 March 2021) database, Biological General Repository for Interaction Datasets (BioGRID, https://www.ncbi.nlm.nih.gov/pubmed/30476227, accessed on 25 March 2021) the Database of Interacting Proteins (DIP, https://www.ncbi.nlm.nih.gov/pubmed/10592249, accessed on 25 March 2021), the Human Protein Reference Database (HPRD, https://www.ncbi.nlm.nih.gov/pubmed/18988627, accessed on 25 March 2021), and the Molecular INTeraction (MINT, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1751541/, accessed on 25 March 2021) database, accordingly. The selectable identification included interaction distance, interaction type, interaction detection method, number of the interactive information-related database, number of related literature, and number of interaction detection methods [11]. In this study, we investigated the directly interacting genes with the start genes (mutation genes, over- or under-expressed genes), and the organism chosen was Homo sapiens.

2.6. Signal Transduction Pathway Analysis

For each patient, CBSJukebox® identified genes that interacted with start genes and mapped genes in signal transduction pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.ncbi.nlm.nih.gov/pubmed/11752249, accessed on 25 March 2021) database and provided the type of interaction information (interaction distance and ratio etc.). We selected the top 10 signal transduction pathways among all of the recorded pathways in the KEGG database based on the weight of the number of interactions as well as the interacting genes.

2.7. High Interaction Frequency Ratio Genes Analysis

For each signal transduction pathway, CBSJukebox® calculated the interaction frequency ratio of interacting genes that interacted with start genes. A 100% interaction frequency gene is deemed by the gene that has the highest interaction frequency with start genes within each signal transduction pathway. We calculated that the interaction frequency ratio of each gene lay within each signal transduction pathway and set the high interaction frequency ratio cut-off as 75%.

2.8. Treatment Benefit Prediction Score (TBPS) Calculation

We applied a novel algorithm that calculated the gene interaction score for the top 10 signal transduction pathways that divided the number of interactions for each interacting gene (between start genes and interacting genes) in a specific signal transduction pathway by the total number of interactions. Then we calculated each gene's treatment benefit prediction score (TBPS) by the sum of gene interaction scores included in the top 10 signal transduction pathways [12].

2.9. Potential Treatment Recommendation for Patient

The CBSJuekbox® current version enables us to suggest potential treatment options in the order of the genes' TBPS. The genes with a high TBPS that were considered as potential targets for patient treatment could be matched with the drug target genes from the DrugBank (Drugbank, https://academic.oup.com/nar/article/46/D1/D1074/4602867, accessed on 25 March 2021) database. In this study, we only considered drug targets not limited to drug conditions of approval, indication, and non-prescription.

2.10. Comparison of Mutation-Based pathway and Over- or Under-Expressed Genes-Based Pathways

The top 10 mutation-based pathways (MBPs), the top 10 mRNA over-expressed genes-based pathways (OEBPs) and the top 10 mRNA under-expressed genes-based pathways (UEBPs) for each patient were analyzed. The matching rate of MBPs and OEBPs and the matching rate of MBPs and OEBPs were compared.

2.11. Validation of TBPS in Two HNSCC Patients Treated with Targeted Agents

We validated the TBPS in two HNSCC patients who were treated with molecular targeted gene therapies. One patient was a 55-year-old male patient. He had recurrent cancer and metastatic oral cavity cancer with Q75E mutation in PIK3CA. The other patient was a 38-year-old female patient, and she had recurrent and metastatic paranasal sinus squamous cell carcinoma with frame shift mutations in FGFR1. These two patients were enrolled in the TRIUMPH trial (NCT03292250) [6], an umbrella trial for recurrent/metastatic HNSCC consisting of five targeted therapies including PI3K inhibitor, pan-HER inhibitor, FGFR inhibitor, CDK4/6 inhibitor, and immune checkpoint inhibitor. These two patients received alpelisib (BYL719) monotherapy and nintedanib monotherapy, respectively, and showed partial responses. We calculated the TBPS in these two patients and analyzed the correlation between TBPS and drug matching results.

3. Results

3.1. Clinical Characteristics

Altogether, 93 patients were enrolled. Clinical characteristics are summarized in Table 1; the median age was 59 years (range, 28–80), and 39 patients (42%) had stage 4 disease at the initial diagnosis. Median overall survival (OS) was 70.0 months (95% confidence interval (CI), 57.4–84.4). Oral cavity (38%) was the most frequent location of HNSCC.

Table 1. Baseline characteristics in all patients.

***n* = 93**	***n***	**%**
Age, median (range)	59 (28–80)	
Gender		
Female	18	19
Male	75	81
Anatomic site		
Oropharnx	26	28
Oral cavity	35	38
Hypopharynx	15	16
Glottic larynx	9	10
Supraglottic larynx	3	3
Maxillary sinus	5	5
Tobacco use		
Never	26	28
Former	49	53
Current	18	19
Alcohol use		
Never	34	37
Former	33	35
Current	26	28
Initial clinical stage		
I–III	54	58
IV	39	42
HPV status		
Positive	20	22
Negative	56	60
Unknown	17	18

We excluded tumor samples without any mutations because such tumors cannot perform in pathway mapping analysis. We also excluded tumor samples with the FoxoG error [13] and QC flag. Altogether, 77 samples that were available with regard to both mutational data and over- or under-expression mRNA data were finally analyzed.

3.2. Top 10 Signaling Pathway Discovered by Mutation-Based Analysis and mRNA Expressed Genes-Based Analysis

We compared the top 10 pathways frequently discovered by mutation-based analysis and mRNA gene expressed-based analyses. Two pathways, Kaposi's sarcoma associated herpes virus infection and HTLV-I infection pathways, were found to be overlapping, both in mutation-based analysis and in over- or under-expression genes-based analysis (Table 2). It was found that the following five pathways were overlapping, both in mutation based analysis and under-expression based analysis: (1) Pathways in cancer, (2) Human papillomavirus infection, (3) PI3K-Akt signaling pathway, (4) HTLV-I infection, (5) Kaposi's sarcoma associated herpesvirus infection. Overall, UEBPs were more frequently overlapping with MBPs. Among the 60 MBPs, 18 MBPs were not overlapping with OEBPs or UEBPs. Among the 82 OEBPs, 39 OEBPs were not overlapping with MBPs or UEBPs. All of the UEBPs were overlapping with either MBPs or OEBPs (Figure 2).

Table 2. Top 10 pathways discovered by mutation-based and mRNA over- or under-expressed genes-based analysis.

	Mutated Based Analysis		**mRNA Over-Expressed Genes-Based Analysis**		**mRNA Under-Expressed Genes-Based Analysis**	
Rank	**Number of Patients (*n* = 77)**	**Pathway Name**	**Number of Patients**	**Pathway Name**	**Number of Patients**	**Pathway Name**
1	74	Pathways in cancer	47	Herpes simplex infection	77	Pathways in cancer
2	63	PI3K-Akt signaling pathway	47	Kaposi's sarcoma-associated herpesvirus infection	74	Human papillomavirus infection
3	62	HTLV-I infection	43	T cell receptor signaling pathway	71	PI3K-Akt signaling pathway
4	61	Human papillomavirus infection	42	NF-kappa B signaling pathway	68	Herpes simplex infection
5	41	MicroRNAs in cancer	42	Cytokine-cytokine receptor interaction	48	Ras signaling pathway
6	41	Viral carcinogenesis	38	HTLV-I infection	46	HTLV-I infection
7	35	Kaposi's sarcoma-associated herpesvirus infection	36	Cell adhesion molecules (CAMs)	46	Kaposi's sarcoma-associated herpesvirus infection
8	34	Epstein–Barr virus infection	31	Influenza A	40	Natural killer cell mediated cytotoxicity
9	33	MAPK signaling pathway	31	Toll-like receptor signaling pathway	40	Endocytosis
10	29	Proteoglycans in cancer	25	Measles	34	MicroRNAs in cancer

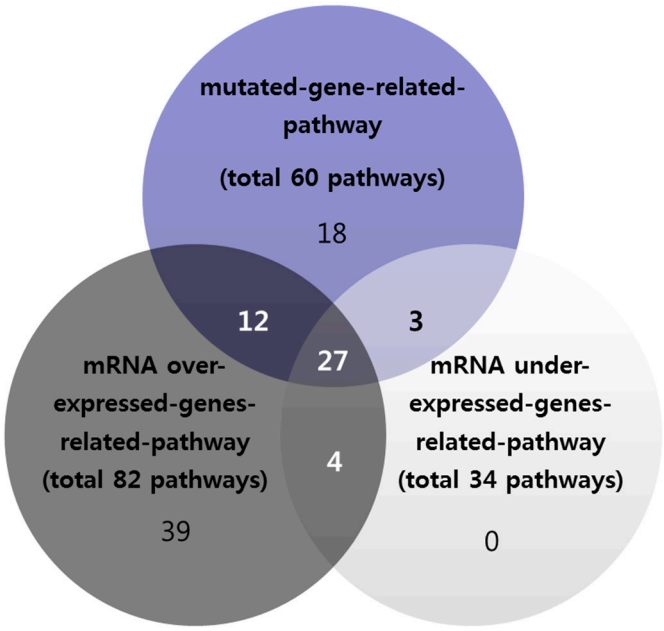

Figure 2. Comparison of mutation-based pathways, over-expressed genes-based pathways, and under-expressed genes-based pathways.

3.3. Overlapping of Mutation-Based Pathway and Over- or Under-Expressed Genes-Based Pathways

Comparing MBPs and mRNA of over- or under-expression genes-based pathway, we observed that 19.1% (147/770) of MBPs were overlapping with OEBPs, and 42.7% (329/770) of MBPs were overlapping with UEBPs (Table 3).

Table 3. Overlapping of mutation-based pathway and mRNA over- or under-expressed genes-based pathways, and their matching rate comparison.

Matching Rate Comparison	
Comparison of mutation-based pathways	Number of patients
Over-expressed genes-based pathways > Under-expressed genes-based pathways	6
Over-expressed genes-based pathways < Under-expressed genes-based pathways	64
Over-expressed genes-based pathways = Under-expressed genes-based pathways	7
Average Matching Rate	
Average matching rate with mutation-based pathways	Average matching rate
Over-expressed genes-based pathways	19.09%
Under-expressed genes-based pathways	42.73%

3.4. Calculation of Treatment Benefit Prediction Score (TBPS)

Table 4 (and Supplementary Table S1) shows the results of TBPS and suggested drug. In the OEBP results, Patient 4 had an alteration in the T cell receptor signaling pathway, and the CD3E gene was identified as a druggable gene. Muromonab was suggested as a targeted agent with TBPS 72.7 for Patient 4. Patient 5 had an alteration in the cell adhesion molecules (CAMs) pathway and CD274 (PD-L1) over-expression. PD-L1 inhibitors, such as atezolizumab, avelumab, and durvalumab were suggested for Patient 5. Interestingly, JAK1, which is not a well-known target for HNSCC, was identified in patient 16, and roxilitinib was suggested. In the UEBP results, FYN was identified as a candidate gene in Patient 57, and dasatinib was suggested as the matching drug.

Table 4. Treatment Benefit Prediction Scores (TBPSs) and suggestion of specific drugs: (A) Over-expression genes-related analysis, (B) Under-expression genes-related analysis, (C) Mutation genes-related analysis.

(A) Over-Expression Genes-Related Analysis				
Patient No.	**Druggable Pathway**	**Druggable Gene**	**TBPS**	**Matched Drug**
4	Measles, Hematopoietic cell lineage, Chagas disease (American trypanosomiasis), T cell receptor signaling pathway, HTLV-I infection	CD3E	72.7	Muromonab
5	RIG-I-like receptor signaling pathway, Hepatitis C, IL-17 signaling pathway, MAPK signaling pathway, Toll-like receptor signaling pathway, Herpes simplex infection, Influenza A	TNF	135.3	Adalimumab
				Golimumab
				Infliximab
	Toll-like receptor signaling pathway, Cell adhesion molecules (CAMs)	CD80	23.4	Durvalumab
	Cell adhesion molecules (CAMs)	CD274	9.1	Atezolizumab
				Avelumab
				Durvalumab
	Cell adhesion molecules (CAMs)	PDCD1	9.1	Nivolumab
				Pembrolizumab
16	Influenza A, Epstein–Barr virus infection, Kaposi's sarcoma-associated herpesvirus infection, Human papillomavirus infection, Pathways in cancer, Tuberculosis, Herpes simplex infection, HTLV-I infection	JAK1	41.4	Ruxolitinib
(B) Under-Expression Genes-Related Analysis				
Patient No.	**Druggable Pathway**	**Druggable Gene**	**TBPS**	**Matched Drug**
57	Axon guidance, T cell receptor signaling pathway, Measles, Natural killer cell mediated cytotoxicity	FYN	30.2	Dasatinib
	T cell receptor signaling pathway, Pathways in cancer, Human papillomavirus infection, Natural killer cell mediated cytotoxicity	GRB2	18.9	Pegademase bovine
	Kaposi's sarcoma-associated herpesvirus infection, Pathways in cancer, Human papillomavirus infection, HTLV-I infection	PIK3R1	16.9	Isoprenaline
	T cell receptor signaling pathway, HTLV-I infection, Natural killer cell mediated cytotoxicity	LCK	16	Dasatinib
				Nintedanib
				Ponatinib
(C) Mutation Genes-Related Analysis				
Patient No.	**Druggable Pathway**	**Druggable Gene**	**TBPS**	**Matched Drug**
3	Pathways in cancer, PI3K-Akt signaling pathway, HTLV-I infection, Human papillomavirus infection, MicroRNAs in cancer, Kaposi's sarcoma-associated herpesvirus infection, Epstein–Barr virus infection, Breast cancer, Prostate cancer	TP53	42.5	Acetylsalicylic acid
	Pathways in cancer, PI3K-Akt signaling pathway, Human papillomavirus infection, MicroRNAs in cancer, Focal adhesion, Breast cancer, Prostate cancer	GRB2	27.7	Pegademase bovine

CD274: Programmed cell death 1 ligand 1, CD3E: T-cell surface glycoprotein CD3 epsilon chain, CD80: T-lymphocyte activation antigen CD80, FYN: FYN Proto-Oncogene, FYN: FYN Proto-Oncogene, GRB2: Growth Factor Receptor Bound Protein 2, GRB2: Growth Factor Receptor Bound Protein 2, JAK1: Tyrosine-protein kinase JAK1, LCK: LCK Proto-Oncogene, Src Family Tyrosine Kinase, PDCD1: Programmed cell death protein 1, PIK3R1: Phosphoinositide-3-Kinase Regulatory Subunit 1, TNF: Tumor necrosis factor, TP53: Tumor Protein P53.

When counting overlapping pathways, MBPs, OEBPs, and UEBPs in cancer were the most commonly overlapping ones (23 times (29.87%)). The second most commonly overlapping pathway was the HTLV-I infection pathway (15 time (19.48%)), followed by the Human papillomavirus infection pathway (12 times (15.58%)).

The HTLV-I infection pathway was most commonly overlapping (16 times (20.78%)) pathway between MBP and OEBP, followed by Kaposi's sarcoma-associated herpesvirus infection pathway (11 time, (14.29%)). The PI3K-Akt signaling pathway was the most commonly overlapping (52 times (68.83%)) pathway between MBP and UEBP (Figure 3, Supplementary Figure S1).

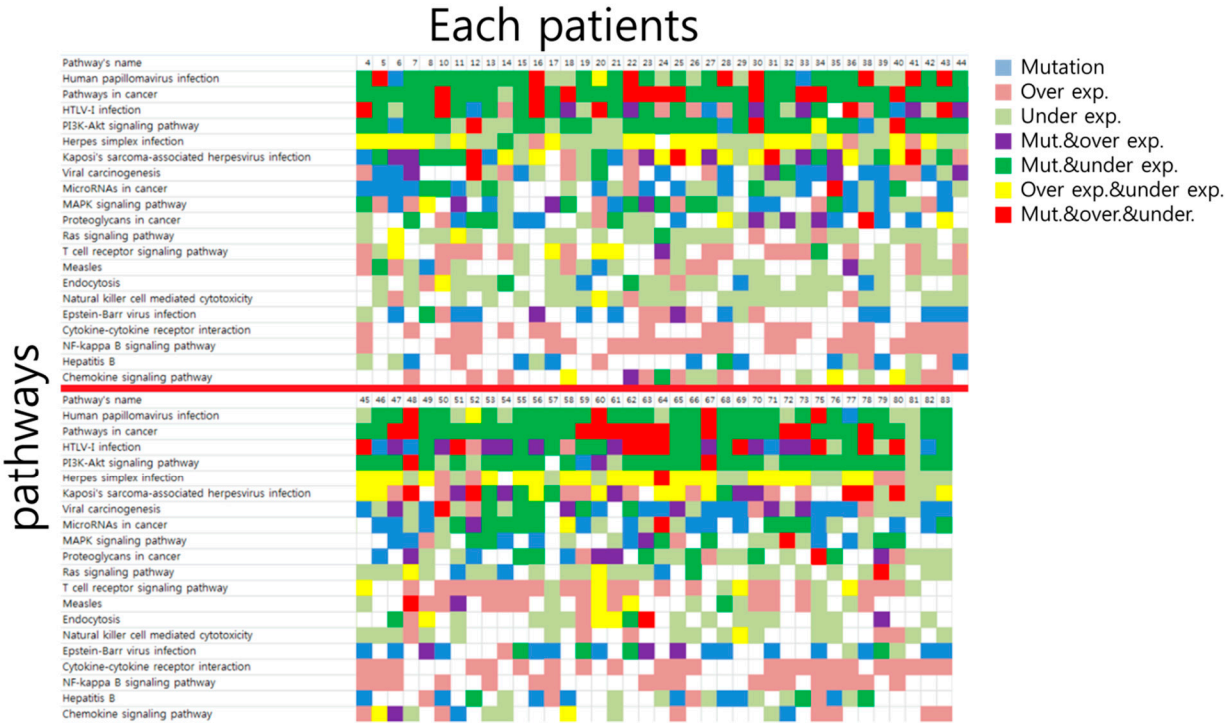

Figure 3. Oncoplot for top 20 pathway analyses in all patients.

To validate TBPSs and suggest the matching drug, we analyzed the data of two HNSCC patients who showed a good response to the treatment with the PIK3CA inhibitor and the FGFR inhibitor. When analyzing data of the nintedanib responding patient with the FGFR1 mutation using a cutoff of frequency ratio of 75%, nintedanib was suggested in mRNA expression-based analysis with a TBPS of 2.5 (Table 5). Alpelisib (BYL719) was suggested for the alpelisib responding patient at the level of 66% frequency ratio in mutation-based analysis (Supplementary Table S2).

Table 5. The results of pathway analysis in the FGFR Inhibitor, nintedanib responding patient.

Mutated Based Analysis				
Pathway Name	**High Frequency Gene**	**Frequency Ratio**	**TBPS**	**Maching_Drug_Names**
Ras signaling pathway	AKT1	100.0	4.8	Enzastaurin
Pathways in cancer	AKT1	100.0	3.9	Arsenic trioxide, Enzastaurin
Melanoma	RAF1	50.0	3.7	Dabrafenib
Proteoglycans in cancer	EGFR	50.0	3.6	Dacomitinib
Regulation of actin cytoskeleton	PDGFRB	100.0	3.3	Becaplermin

Table 5. *Cont.*

Mutated Based Analysis				
Pathway Name	**High Frequency Gene**	**Frequency Ratio**	**TBPS**	**Maching_Drug_Names**
Breast cancer	EGFR	50.0	3.3	Lapatinib, Neratinib, Trastuzumab
Regulation of actin cytoskeleton	FGFR1	100.0	3.3	Palifermin
Regulation of actin cytoskeleton	FGFR2	100.0	3.3	Palifermin
Breast cancer	ESR2	50.0	3.3	Tamoxifen
Gastric cancer	EGFR	50.0	3.1	Trastuzumab
PI3K-Akt signaling pathway	HSP90AA1	50.0	2.6	Alvespimycin, Tanespimycin
PI3K-Akt signaling pathway	FGFR1	50.0	2.6	Erdafitinib
PI3K-Akt signaling pathway	FGFR2	50.0	2.6	Erdafitinib
PI3K-Akt signaling pathway	PDGFRB	50.0	2.6	Erdafitinib, Midostaurin
PI3K-Akt signaling pathway	HSP90AB1	50.0	2.6	Tanespimycin
MAPK signaling pathway	EGFR	50.0	2.5	Afatinib, Canertinib, Cetuximab, Erlotinib, Gefitinib, Lapatinib, Necitumumab, Olmutinib, Osimertinib, Panitumumab, Pelitinib, Rindopepimut, Vandetanib, Zalutumumab
MAPK signaling pathway	PDGFRB	50.0	2.5	Becaplermin, Dasatinib, Imatinib, Midostaurin, Pazopanib, Regorafenib, Sorafenib, Sunitinib
MAPK signaling pathway	RAF1	50.0	2.5	Dabrafenib, Regorafenib, Sorafenib
MAPK signaling pathway	EPHA2	50.0	2.5	Dasatinib, Regorafenib
MAPK signaling pathway	FGFR2	**50.0**	2.5	Lenvatinib, **Nintedanib**, Regorafenib
MAPK signaling pathway	FGFR1	**50.0**	2.5	Lenvatinib, **Nintedanib**, Regorafenib, Sorafenib
mRNA Based Analysis				
Pathway Name	**High Frequency Gene**	**Frequency Ratio**	**TBPS**	**Maching_Drug_Names**
Proteoglycans in cancer	EGFR	40.0	2.9	Dacomitinib
MAPK signaling pathway	EGFR	80.0	2.6	Afatinib, Canertinib, Cetuximab, Erlotinib, Gefitinib, Lapatinib, Necitumumab, Olmutinib, Osimertinib, Panitumumab, Pelitinib, Rindopepimut, Vandetanib, Zalutumumab
MAPK signaling pathway	FGFR3	**80.0**	2.6	Lenvatinib, **Nintedanib**, Pazopanib
MAPK signaling pathway	FGFR2	**80.0**	2.6	Lenvatinib, **Nintedanib**, Regorafenib
MAPK signaling pathway	FGFR1	**80.0**	2.6	Lenvatinib, **Nintedanib**, Regorafenib, Sorafenib

AKT1: RAC-alpha serine/threonine-protein kinase, RAF1: RAF proto-oncogene serine/threonine-protein kinase, EGFR: Epidermal Growth Factor Receptor, EPHA2: Ephrin type-A receptor 2, ESR2: Estrogen receptor beta, FGFR1: Fibroblast growth factor receptor 1, FGFR2: Fibroblast Growth Factor Receptor 2, FGFR3: Fibroblast Growth Factor Receptor 3, HSP90AA1: Heat Shock Protein 90 Alpha Family Class A Member 1, HSP90AB1: Heat shock protein HSP 90-beta, PDGFRB: Platelet-derived growth factor receptor beta.

4. Discussion

In this study, we described a novel approach for pathway analysis using mutation data and mRNA expression data. Mutated gene-related pathways were associated mainly with mRNA under-expression genes-related pathways. These results suggest that HNSCC are mainly related to loss-of-function mutations. However, big data based platforms for druggable pathways can find potential matching drugs.

Our model is based on 14 open databases for protein, interaction, and signaling pathways such as NCBI, Uniprot, KEGG, Biogrid, DIP, HPRD, and Drugbank. High interaction genes were mapped to investigate druggable pathways. We hypothesized that integra-

tion of each mutation and the respective mRNA expression into signaling pathway can identify their functional significance and therapeutic targets. Pathway networks based on cancer gene landscapes can give us insight into how these genes contribute to deregulated oncogenic pathways. Several studies [5,14–16] had similar approaches based on pathway analysis. However, we developed a novel scoring model that measured the overlap between mutation and mRNA expression data, and calculated the interaction relationship score for discovering a potential target drug.

Each mutation and mRNA expression data from signaling nodes and hubs transmit pathological cues along molecular networks to achieve integrated tumorigenic pathways. From the interaction of receptors with deregulated growth factors to dimerization of receptor tyrosine kinases triggered by gene mutations, PPIs initiate a cascade of interactions to promote uncontrolled cell proliferation [9]. In response to oncogenic stimulation, PPIs play essential roles in linking networks that relay oncogenic signals, and therefore allow for the suggestion of the target drug.

Unlike existing methods, our model is capable of ranking and scoring the significant KEGG pathways reported in the cancer research literature. We used the prior knowledge specified in the pathway in order to identify the particular pathway in gene/protein interaction that could explain the molecular basis of carcinogenesis. Our novel algorithm, called CBSJukebox®, calculates the interaction frequency ratio of interacting genes. Based on the interaction frequency ratio, we can calculate each gene's TBPS using a sum of gene interaction scores. The TBPS suggests the matching drug and visualizes the responding probability.

During the experiment, we also observed that not only oncogenic pathways but also non-oncogenic pathways were deregulated and activated in HNSCC. This multiple pathway involvement implies that targeting multiple pathways is useful for further refining the anti-cancer chemotherapy. We also found that overly activated pathways measured by mRNA over-expression and suppressed pathways measured by mRNA under-expression were quite different. However, biologically important pathways were overlapping in both mutation-based and expression-based pathways.

This study has limitations. This model was developed in silico and has not yet been fully validated in the patients. We tried to validate TBPS in two HNSCC patients who showed good response to the FGFR inhibitor and the PIK3 inhibitor. Our developed CBSJukebox® system suggested both the FGFR inhibitor and the PIK3 inhibitor. However, when we applied an interaction frequency ratio cut-off of 75%, the PIK3 inhibitor was excluded. This might have been caused by the insufficient availability of gene data that interacted with the PIK3CA pathway in the public database. We will expand and use the updated public database in the future to refine our CBSJukebox® system.

Future work will focus on validation of the suggested drugs that were identified in this study with a larger sample size. Regarding future work, our Bayesian network model offers an easy way of incorporating additional data types such as CNV, proteomics data, and methylation data, and so on, and such model extensions should be attempted.

5. Conclusions

In conclusion, our pathway based systematic analysis of mutational and mRNA expression pathways provides novel mechanistic and clinical insights into the precision therapeutics for HNSCC. NGS-based mutated gene-related pathways were associated with mRNA under-expression genes-related pathways. These results suggest that HNSCCs are mainly caused by the loss-of-function mutations. However, big data based platforms for druggable pathways can find potential matching drugs.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/pr9050792/s1, Figure S1. Oncoplot for pathway analysis in all patients. Table S1: Treatment Benefit Prediction Score (TBPS) and suggestion of drug: (1) Over-expression genes related analysis, (2) Under-expression genes related analysis, (3) Mutation genes related analysis (Full data). Table S2: The results of pathway analysis in PIK3CA Inhibitor responding patient.

Author Contributions: Conceptualization, B.K., J.-Y.P. and H.-J.Y.; resources and data curation, B.K., S.-H.C. (Sang-Hee Cho), S.K., H.-K.A., S.-H.C. (Sang-Hoon Chun), J.-H.K., T.Y., J.-W.K., J.-E.K., M.-J.A. and J.-H.K.; methodology, B.K., J.-Y.P. and G.-D.K.; formal analysis, J.-P.K. and Y.-S.Y.; writing—original draft, B.K.; writing—review and editing, G.-D.K. and H.-J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: The TRIUMPH trial was supported by a grant from the National R& D Program for Cancer Control, the Ministry of Health and Welfare, Republic of Korea (HA16C0015). The funder had no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of all the participating institutes. Institutional Review Board of Catholic Kwandong University International St. Mary's Hospital (protocol code 16IRB037-1, date of approval 2016.11.14), Institutional Review Board of The Catholic University of Korea, Bucheon ST. Mary's Hospital (protocol code HC16PIMI0085, date of approval 2016.10.12), Institutional Review Board of The Catholic University of Korea, Incheon St. Mary's Hospital (protocol code OIRB-00262_1-002, date of approval 2016.11.22), Institutional Review Board of Kangdong Sacred Heart Hospital (protocol code KANGDONG 2016-09-005-001, date of approval 2016.10.24), Institutional Review Board of Konyang University Hospital (protocol code KYUH 2016-08-008-002, date of approval 2016.9.27), Institutional Review Board of Korea University Guro Hospital KUGH16137-002 date of approval 2016.9.20), Institutional Review Board of National Cancer Center (protocol code NCC2016-0275, date of approval 2016.11.22), Institutional Review Board of Seoul National University Bundang Hospital (B-1611-370-401, date of approval 2016.10.24), Institutional Review Board of Bundang CHA Hospital (protocol code CHAMC 2016-12-016, date of approval 2016.12.29), Institutional Review Board of Seoul National University Hospital (protocol code H-1609-057-791 date of approval 2016.10.03), Institutional Review Board of SNU Boramae Medical Center (protocol code 26-2016-150, date of approval 2016.11.7), Institutional Review Board of Yonsei University Hospital (protocol code 4-2014-0775, date of approval 2014.10.30), Institutional Review Board of Ajou University Medical Center (protocol code AJIRB-MED-SMP-16-432, date of approval 2017.1.2), Institutional Review Board of Yeungnam University Medical Center (protocol code YUMC 2016-10-007, date of approval 2016.10.24), Institutional Review Board of Chung-Ang University Hospital (protocol code 1603-001-258, date of approval 2016.12.6), Institutional Review Board of Chungnam National University Hospital (protocol code CNUH 2016-08-026, date of approval 2016.9.26), Institutional Review Board of Chonnam National University Hwasun Hospital (protocol code CNUHH-2016-129, date of approval 2017.10.11).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: B.K. received research funding from MSD (Merck Sharp & Dohme Corp.) MSD and Ono Pharmaceutical Co., Ltd. and served as an advisor for AstraZeneca, MSD, and Genexine outside of the current work. B.K served as an advisor for Cbs Bioscience Inc. inside of the current work.

References

1. Cancer Genome Atlas Network; Sheikine, Y.; Kuo, F.C.; Lindeman, N.I. Clinical and Technical Aspects of Genomic Diagnostics for Precision Oncology. *J. Clin. Oncol.* **2017**, *35*, 929–933. [CrossRef] [PubMed]
2. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **2015**, *517*, 576–582. [CrossRef] [PubMed]
3. Agrawal, N.; Frederick, M.J.; Pickering, C.R.; Bettegowda, C.; Chang, K.; Li, R.J.; Fakhry, C.; Xie, T.X.; Zhang, J.; Wang, J.; et al. Exome sequencing of head and neck squamous cell carcinoma reveals inacti-vating mutations in NOTCH1. *Science* **2011**, *333*, 1154–1157. [CrossRef] [PubMed]
4. Stransky, N.; Egloff, A.M.; Tward, A.D.; Kostic, A.D.; Cibulskis, K.; Sivachenko, A.; Kryukov, G.V.; Lawrence, M.S.; Sougnez, C.; McKenna, A.; et al. The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. *Science* **2011**, *333*, 1157–1160. [CrossRef] [PubMed]
5. Lim, S.M.; Cho, S.H.; Hwang, I.G.; Choi, J.W.; Chang, H.; Ahn, M.-J.; Park, K.U.; Kim, J.-W.; Ko, Y.H.; Ahn, H.K.; et al. Investigating the Feasibility of Targeted Next-Generation Sequencing to Guide the Treatment of Head and Neck Squamous Cell Carcinoma. *Cancer Res. Treat.* **2019**, *51*, 300–312. [CrossRef] [PubMed]
6. Keam, B.; Kim, H.R.; Yun, H.J. Triumph Trial: One Small Step Could Become One Giant Leap for Precision Oncology in Head and Neck Cancer. *Cancer Res. Treat.* **2019**, *51*, 413–414. [CrossRef] [PubMed]

7. Li, Z.; Ivanov, A.A.; Su, R.; Gonzalez-Pecchi, V.; Qi, Q.; Liu, S.; Webber, P.; McMillan, E.; Rusnak, L.; Pham, C.; et al. The OncoPPi network of cancer-focused protein–protein interactions to inform biological insights and therapeutic strategies. *Nat. Commun.* **2017**, *8*, 14356. [CrossRef] [PubMed]
8. Ivanov, A.A.; Revennaugh, B.; Rusnak, L.; Gonzalez-Pecchi, V.; Mo, X.; Johns, M.A.; Du, Y.; Cooper, L.A.; Moreno, C.S.; Khuri, F.R.; et al. The OncoPPi Portal: An integrative resource to explore and prioritize pro-tein-protein interactions for cancer target discovery. *Bioinformatics* **2018**, *34*, 1183–1191. [CrossRef] [PubMed]
9. Ivanov, A.A.; Khuri, F.R.; Fu, H. Targeting protein–protein interactions as an anticancer strategy. *Trends Pharmacol. Sci.* **2013**, *34*, 393–400. [CrossRef] [PubMed]
10. Murakami, Y.; Tripathi, L.P.; Prathipati, P.; Mizuguchi, K. Network analysis and in silico prediction of protein-protein interac-tions with applications in drug discovery. *Curr. Opin. Struct. Biol.* **2017**, *44*, 134–142. [CrossRef] [PubMed]
11. Park, I.J.; Yu, Y.S.; Mustafa, B.; Park, J.Y.; Seo, Y.B.; Kim, G.-D.; Kim, J.; Kim, C.M.; Noh, H.D.; Hong, S.-M.; et al. A Nine-Gene Signature for Predicting the Response to Preoperative Chemoradiotherapy in Patients with Locally Advanced Rectal Cancer. *Cancers* **2020**, *12*, 800. [CrossRef] [PubMed]
12. Kim, C.M.; Hwang, S.; Keam, B.; Yu, Y.S.; Kim, J.H.; Kim, D.S.; Bae, S.H.; Kim, G.D.; Lee, J.K.; Seo, Y.B.; et al. Gene Signature for Sorafenib Susceptibility in Hepatocellular Carcinoma: Different Ap-proach with a Predictive Biomarker. *Liver Cancer* **2020**, *9*, 182–192. [CrossRef] [PubMed]
13. Costello, M.; Pugh, T.J.; Fennell, T.J.; Stewart, C.; Lichtenstein, L.; Meldrim, J.C.; Fostel, J.L.; Friedrich, D.C.; Perrin, D.; Dionne, D.; et al. Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Res.* **2013**, *41*, e67. [CrossRef] [PubMed]
14. Zhao, Y.; Hoang, T.H.; Joshi, P.; Hong, S.-H.; Giardina, C.; Shin, N.-G. A route-based pathway analysis framework integrating mutation information and gene expression data. *Methods* **2017**, *124*, 3–12. [CrossRef] [PubMed]
15. Lee, S.; Park, Y.; Kim, S. MIDAS: Mining differentially activated subpaths of KEGG pathways from multi-class RNA-seq data. *Methods* **2017**, *124*, 13–24. [CrossRef] [PubMed]
16. The Mutation Consequences and Pathway Analysis working group of the International Cancer Genome Consortium; Creixell, P.; Reimand, J.; Haider, S.; Wu, G.; Shibata, T.; Vazquez, M.; Mustonen, V.; Gonzalez-Perez, A.; Pearson, J.; et al. Pathway and network analysis of cancer genomes. *Nat. Methods* **2015**, *12*, 615–621. [CrossRef] [PubMed]