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

Bhumsuk Keam (bhumsuk@snu.ac.kr)  
Seoul National University Hospital  
https://orcid.org/0000-0001-8196-4247

Jin Young Park  
CBS Bioscience

Jinpyo Kim  
CBS Bioscience

Gun Do Kim  
CBS Bioscience

Yunsuk Yu  
CBS Bioscience

Sang Hee Cho  
Chonnam National University Hospital

Sangwoo Kim  
Severance Hospital

Hee Kyung Ahn  
Gachon University Gil Medical Center

Sang Hoon Chun  
Catholic University of Korea Bucheon Saint Mary’s Hospital

Jung Hye Kwon  
Hallym University College of Medicine

Tak Yun  
National Cancer Center

Ji-Won Kim  
Seoul National University Bundang Hospital

Ji Eun Kim  
Seoul National University College of Medicine

Myung-Ju Ahn  
Samsung Medical Center

Joo-Hang Kim  
CHA Bundang Medical Center

Hwan Jung Yun
Research article

Keywords: Head and neck squamous cell carcinoma (HNSCC); Precision Medicine; Pathway Analysis; CBSJukebox

DOI: https://doi.org/10.21203/rs.3.rs-41520/v1

License: ☑️ ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: 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.

Methods: Altogether, 93 recurrent/metastatic HNSCC patients were enrolled. We performed targeted gene sequencing using Agilent 244 customized gene panel for NGS, and nanostring nCounter® for mRNA expression. mRNA expression was classified into over- or under expression groups based on the expression levels that showed 2-fold change in tumor compared with non-tumor tissues. 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. Analyses were based on 14 databases related to protein interaction, signaling pathways and drugs such as NCBI, Uniprot, KEGG, Biogrid, DIP, HPRD, and Drugbank. High score interaction genes were mapped to investigate druggable pathways. We calculated Treatment Benefit Prediction Score (TBPS) to identify suitable drugs.

Results: By mapping the high score interaction genes to identify druggable pathways, we found that highly related signaling pathways with mutation mapping were Influenza A, pathways in cancer, EBV, HPV, HTLV-1 pathways. Based on mRNA expression and interaction gene scoring model, several pathways were found to be associated with over- and under-expression. Comparison between mutation-based pathway and over-/under-expressed genes-based pathways, we observed that 19.8% of NGS based pathways matched with mRNA over-expressed genes-based pathways, and 43.2% of NGS based pathways consist with mRNA under-expressed genes-based pathways. TBPS found several matching drugs such as immune checkpoint inhibitors, EGFR inhibitors, and FGFR inhibitors.

Conclusions: Mutation-based pathway was associated with mRNA under-expressed genes-based pathway. These results suggest that HNSCCs are mainly caused by the loss-of-function mutations. However, big data based platform for druggable pathway can find potential matching drugs.

Background

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 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 performed.

Integration of cancer genes into networks offers opportunities to reveal protein–protein interactions (PPIs) with functional and therapeutic significance. PPIs network based on cancer gene landscape 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 and 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 is to integrate the DNA mutational landscape and mRNA expression patterns into the PPI network pathways, which could be used to match potential drug or druggable pathway in HNSCC.

Methods

Patients and data collection

Altogether, 93 recurrent/metastatic HNSCC patients from 19 institutions were enrolled. The details of the study population were 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. Institutional Review Boards of each institute approved this study protocol.

Targeted gene sequencing and mRNA expression assay

Genomic DNA was isolated from formalin-fixed paraffin-embedded 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 > 1000x. 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.

Basic scheme of protein-protein interaction network analysis

To analyze with a deep insight of combinatorial signaling events evolved in cell communication, we applied novel PPI method named CBSJukebox®. Figure 1 shows the analytic flow in CBSJukebox®. In
brief, CBSJukebox® enabled to compare DNA mutation-based pathways and over/under expression based pathways leveraged PPI analysis; and further, CBSJukebox® enabled to perform a simple signal pathway analysis as well as high interaction frequency ratio genes analysis.

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 |fold-change| > 2) and compared with those expressed in normal tissue for further over/under expression based pathway analysis.

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 record (NCBI ID, https://www.ncbi.nlm.nih.gov/gene) from iProClass (https://www.ncbi.nlm.nih.gov/pubmed/15022647) database, and the over/under-expressed genes were matched with gene name and synonym in Uniprot /Swiss-Prot (Uniprot Knowledgebase, https://www.ncbi.nlm.nih.gov/pubmed/27899622) further to interchange with identification factor “Uniprot Ac” in CBSJukebox®. We then conducted the interactive proteins network analysis leveraging IntAct (IntAct, http://europepmc.org/abstract/MED/24234451), BioGRID (Biological General Repository for Interaction Datasets, https://www.ncbi.nlm.nih.gov/pubmed/30476227), DIP (Database of interacting proteins, https://www.ncbi.nlm.nih.gov/pubmed/10592249), HPRD (Human Protein Reference Database, https://www.ncbi.nlm.nih.gov/pubmed/18988627) and MINT (The Molecular INTeraction, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1751541/) database accordingly. The selectable identification included interaction distance, interaction type, interaction detection method, number of interactive information related database, number of related literature and number of interaction detection method [11]. In this study, we investigated the directly interacting genes with the start genes (mutation genes, over/under-expressed genes), and the organism chosen was Homo sapiens.

Signal Transduction Pathway analysis

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

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 deems the gene that has highest interaction frequency with start genes within each signal transduction pathway. We calculated the interaction frequency ratio of each gene lied within each signal transduction pathway and set the high interaction frequency ratio cut-off of 75%.

**Treatment Benefit Prediction Score (TBPS) calculation**

We applied a novel algorithm that calculated 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 a sum of gene interaction scores included in the top 10 signal transduction pathways [12].

**Potential treatment recommendation for patient**

CBSJukebox® current version enables to suggest potential treatment options in the order of the genes’ TBPS. The genes with high TBPS considered as potential targets for patient treatment could match with the drug target genes from DrugBank (Drugbank, https://academic.oup.com/nar/article/46/D1/D1074/4602867) database. In this study, we only considered drug targets, not limited to drug conditions of approval, indication, and non-prescription.

**Comparison of mutation-based pathway and over-/under expressed genes-based pathways**

The top 10 mutation-based pathways (MBP), top 10 mRNA over-expressed genes-based pathway (OEBP) and top 10 mRNA under-expressed genes-based pathways (UEBP) for each patient were analyzed. The matching rate of MBP and OEBP, and the matching rate of MBP and OEBP were compared.

**Validation of TBPS in two HNSCC patients treated with targeted agents**

We validated 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 TRIUMPH trial (NCT03292250) [6], 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 TBPS in these two patients and analyzed the correlation between TBPS and drug matching results.

**Results**
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.

We excluded tumor samples without any mutations, because such tumors cannot perform in pathway mapping analysis. We also excluded tumor sample with FoxoG error [13] and QC flag. Altogether, 77 samples which were available with regard to both mutational data and over-/under-expression mRNA data were finally analyzed.
### Table 1
Baseline characteristics in all patients

|                         | N  | %   |
|-------------------------|----|-----|
| **N = 93**              |    |     |
| Age, median (range)     | 59 | 28–80 |
| Gender                  |    |     |
| Female                  | 18 | 19  |
| Male                    | 75 | 81  |
| Anatomic site           |    |     |
| Oropharynx              | 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  |
**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 overlapping both in mutation-based analysis and in over/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, UEBP were more frequently overlapped overlapping with MBP. Among the 60 MBPs, 18 MBPs were not overlapping with OEBP or UEBP. Among the 82 OEBPs, 39 OEBPs were not overlapping with MBP or UEBP. All of the UEBP were overlapping with either MBP or OEBPs (Fig. 2).
Table 2
Top 10 pathway discovered by mutation-based and mRNA over/under expressed genes-based analysis

| Rank | Number of patients (n = 77) | Mutated based analysis | mRNA over-expressed genes-based analysis | mRNA under-expressed genes-based analysis |
|------|-----------------------------|------------------------|------------------------------------------|------------------------------------------|
|      |                             | 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                       |

Overlapping of mutation-based pathway and over/under-expressed genes based- pathways
Comparing MBP and mRNA of over/under-expression genes-based pathway, we observed that 19.1% (147/770) of MBP were overlapping with OEBP, and 42.7% (329/770) of MBP were overlapping with UEBP (Table 3).

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

| Matching rate comparison                                                                 | Number of patients |
|-----------------------------------------------------------------------------------------|--------------------|
| Comparison of mutation-based pathways                                                    |                    |
| Over-expressed based genes pathways > Under-expressed based genes pathways                | 6                  |
| Over-expressed based genes pathways < Under-expressed based genes pathways                | 64                 |
| Over-expressed based genes pathways = Under-expressed based genes pathways                | 7                  |

| Average matching rate                                                                   |                    |
| Average matching rate with mutation-based pathways                                       |                    |
| Over-expressed based genes pathways                                                     | 19.09%             |
| Under-expressed based genes pathways                                                     | 42.73%             |

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, the patient #4 had alteration in T cell receptor signaling pathway, and CD3E gene was identified as druggable gene. Muromonab was suggested as targeted agent with TBPS 72.7 for the patient #4. The patient #5 had alteration in cell adhesion molecules (CAMs) pathway, and CD274 (PD-L1) over-expression. PD-L1 inhibitors, such as atezolizumab, avelumab, durvalumab were suggested for the patient #5. Interestingly, JAK1, which is not a well-known target for HNSCC, was identified in the patients #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 Score (TBPS) and suggestion of specific drugs

1) 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           |

2) 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     |

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
1) Over-expression genes related analysis

| Pathway                                                                 | Gene   | Score | Matched Drug         |
|------------------------------------------------------------------------|--------|-------|----------------------|
| 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            |

3) 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 between MBP, OEBP, and UEBP, pathways in cancer were the most commonly overlapping ones (23 times [29.87%]). Second most commonly overlapping pathway was HTLV-I infection pathway (15 times [19.48%] followed by Human papillomavirus infection pathway (12 times [15.58%]).

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 times [14.29%]). PI3K-Akt signaling pathway was most commonly overlapping (52 times [68.83%]) pathway between MBP and UEBP (Fig. 3, Supplementary Figure S1).

To validate TBPS and suggest the matching drug, we analyzed the data of two HNSCC patients who showed good response to the treatment with PIK3CA inhibitor and FGFR inhibitor. When analyzing data of the nintedanib responding patient with FGFR1 mutation using cutoff of frequency ratio as 75%, nintedanib was suggested in mRNA expression-based analysis with 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).
| Pathway name                      | High frequency gene | Frequency ratio | TBPS | Matching_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                   |
| Breast cancer                    | EGFR                | 50.0           | 3.3  | Lapatinib, Neratinib, Trastuzumab |
| Regaulation 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                   |

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

**Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)**
| Mutated based analysis |
|------------------------|
| **PI3K-Akt signaling pathway** |
| PDGFRB | 50.0 | 2.6 | Erdfatinib, Midostaurin |
| HSP90AB1 | 50.0 | 2.6 | Tanespimycin |
| **MAPK signaling pathway** |
| EGFR | 50.0 | 2.5 | Afatinib, Canertinib, Cetuximab, Erlotinib, Gefitinib, Lapatinib, Necitumumab, Olimtinib, Osimertinib, Panitumumab, Pelitinib, Rindopepimut, Vandetanib, Zalutumumab |
| PDGFRB | 50.0 | 2.5 | Becaplermin, Dasatinib, Imatinib, Midostaurin, Pazopanib, Regorafenib, Sorafenib, Sunitinib |
| RAF1 | 50.0 | 2.5 | Dabrafenib, Regorafenib, Sorafenib |
| EPHA2 | 50.0 | 2.5 | Dasatinib, Regorafenib |
| FGFR2 | 50.0 | 2.5 | Lenvatinib, Nintedanib, Regorafenib |
| FGFR1 | 50.0 | 2.5 | Lenvatinib, Nintedanib, Regorafenib, Sorafenib |

mRNA based analysis

**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

**Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)**
### Mutated based analysis

| Pathway Name             | High Frequency Gene | Frequency Ratio | TBPS | Matching_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                                         |

1) **Overexpression 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, | TNF                                                | 135.3| Adalimumab                |
|             |                                                                                  |                                                    |      |                           |
|             |                                                                                  |                                                    |      |                           |
|             |                                                                                  |                                                    |      |                           |

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

**Supplement** Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)
| Mutated based analysis                      | Herpes simplex infection, Influenza A | 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 |
| 11                                        | MAPK signaling pathway, Hepatitis C, Hepatitis B, T cell receptor signaling pathway, Toll-like receptor signaling pathway, NF-kappa B signaling pathway, Herpes simplex infection, Cytokine-cytokine receptor interaction | TNF        | 120.2 | Adalimumab |
|                                           |                                       |            |      | Golimumab   |
|                                           |                                       |            |      | Infliximab  |
|                                           | Ras signaling pathway, Rap1 signaling pathway, MAPK signaling pathway | FGF1       | 19.3 | Pazopanib   |
|                                           | Ras signaling pathway, Rap1 signaling pathway, MAPK signaling pathway | FGFR4      | 19.3 | Lenvatinib  |
| 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 |

2) Underexpression genes related analysis

| Patient No. | Druggable Pathway | Druggable Gene | TBPS | Matched Drug |
|-------------|-------------------|----------------|------|--------------|

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

**Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)**
## Mutated based analysis

| 14 | Endocytosis, Human papillomavirus infection, MAPK signaling pathway, MicroRNAs in cancer, PI3K-Akt signaling pathway, Pathways in cancer | EGFR | 18.4 | Afatinib |
|    |                                                                                                                   |      |      | Cetuximab |
|    |                                                                                                                   |      |      | Erlotinib |
|    |                                                                                                                   |      |      | Gefitinib |
|    |                                                                                                                   |      |      | Icotinib |
|    |                                                                                                                   |      |      | Lapatinib |
|    |                                                                                                                   |      |      | Lidocaine |
|    |                                                                                                                   |      |      | Necitumumab |
|    |                                                                                                                   |      |      | Olmutinib |
|    |                                                                                                                   |      |      | Osimertinib |
|    |                                                                                                                   |      |      | Panitumumab |
|    |                                                                                                                   |      |      | Trastuzumab |
|    |                                                                                                                   |      |      | Vandetanib |

| 57 | Human papillomavirus infection, MAPK signaling pathway, MicroRNAs in cancer, PI3K-Akt signaling pathway, Pathways in cancer | GRB2 | 17.8 | Pegademas bovine |
|    | Kaposi's sarcoma-associated herpesvirus infection, Human papillomavirus infection, MicroRNAs in cancer, PI3K-Akt signaling pathway, Pathways in cancer | PIK3R2 | 16.5 | Isoprenaline |
|    | Axon guidance, T cell receptor signaling pathway, Measles, Natural killer cell mediated cytotoxicity | FYN | 30.2 | Dasatinib |

**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

**Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)**
## Mutated based analysis

| Druggable Pathway                                                                 | Druggable Gene | TBPS | Matched Drug                  |
|----------------------------------------------------------------------------------|----------------|------|-------------------------------|
| 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                    |
| HTLV-I infection, Influenza A, Human papillomavirus infection, Viral carcinogenesis, Epstein-Barr virus infection, Ras signaling pathway, IL-17 signaling pathway, cAMP signaling pathway, Herpes simplex infection, Pathways in cancer | NFKB1          | 124.9 | Acetylsalicylic acid          |
|                                                                                 |                |      | Pranlukast                   |
|                                                                                 |                |      | Thalidomide                  |
|                                                                                 |                |      | Triflusal                    |
| IL-17 signaling pathway, Pathways in cancer                                      | HSP90AA1       | 26.2 | Nedocromil                   |
|                                                                                 |                |      | Rifabutin                    |

### 3) 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          |

**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

**Supplement** Table 1. **Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)**
| Pathways in cancer, PI3K-Akt signaling pathway, Human papillomavirus infection, MicroRNAs in cancer, Focal adhesion, Breast cancer, Prostate cancer | GRB2 | 27.7 | Pegademase bovine |
|---|---|---|---|
| Tight junction, Rap1 signaling pathway, Endocytosis, Adherens junction | SRC | 26.6 | Nintedanib |
| Human papillomavirus infection, Transcriptional misregulation in cancer, Alcoholism | HDAC1 | 14 | Romidepsin |
| Human papillomavirus infection, Transcriptional misregulation in cancer, Alcoholism | HDAC2 | 11.3 | Aminophylline |
| Human papillomavirus infection, Transcriptional misregulation in cancer, Herpes simplex infection | TP53 | 10.8 | Acetylsalicylic acid |
| Human papillomavirus infection, Alcoholism | GRB2 | 7.8 | Pegademase bovine |
| Human papillomavirus infection, Alcoholism | HDAC3 | 7.8 | Vorinostat |
| Human papillomavirus infection, Rap1 signaling pathway | PIK3R1 | 7 | Isoprenaline |
| Human papillomavirus infection | AKT1 | 2 | Arsenic trioxide |

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

**Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)**
### Mutated based analysis

| Human papillomavirus infection | CTNNB1 | 2 | Urea |
|--------------------------------|--------|---|------|
| Human papillomavirus infection | EGFR   | 2 | Afatinib |
|                                 |        |   | Cetuximab |
|                                 |        |   | Erlotinib |
|                                 |        |   | Gefitinib |
|                                 |        |   | Icotinib |
|                                 |        |   | Lapatinib |
|                                 |        |   | Lidocaine |
|                                 |        |   | Necitumumab |
|                                 |        |   | Olmutinib |
|                                 |        |   | Osimertinib |
|                                 |        |   | Panitumumab |
|                                 |        |   | Trastuzumab |
|                                 |        |   | Vandetanib |

| Human papillomavirus infection | GSK3B  | 2 | Lithium |
|--------------------------------|--------|---|---------|
| 18 Pathways in cancer, PI3K-Akt signaling pathway, Prostate cancer, Th17 cell differentiation, Fluid shear stress and atherosclerosis | HSP90AA1 | 29.1 | Nedocromil |
|                                 |        |   | Rifabutin |
|                                 | MAPK signaling pathway, Osteoclast differentiation | JUN | 6.9 | Arsenic trioxide |
|                                 |        |   | Irbesartan |
|                                 |        |   | Pseudoephedrine |
|                                 |        |   | Vinblastine |
|                                 | MAPK signaling pathway, FoxO signaling pathway | EGFR | 6.1 | Afatinib |

**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

**Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)**
| Mutated based analysis | Cetuximab | Erlotinib | Gefitinib | Icotinib | Lapatinib | Lidocaine | Necitumumab | Olmutinib | Osimertinib | Panitumumab | Trastuzumab | Vandetanib |
|------------------------|-----------|-----------|-----------|----------|-----------|-----------|-------------|-----------|-------------|-------------|-------------|-----------|
| MAPK signaling pathway, Cellular senescence | MYC | 5.9 | Nadroparin |
| Osteoclast differentiation | FYN | 4.3 | Dasatinib |
| Osteoclast differentiation | NFKBIA | 4.3 | Acetylsalicylic acid |

AKT1: AKT Serine/Threonine Kinase 1, CD274: Programmed cell death 1 ligand 1, CD3E: T-cell surface glycoprotein CD3 epsilon chain, CD80: T-lymphocyte activation antigen CD80, CTNNB1: Catenin Beta 1, EGFR: Epidermal Growth Factor Receptor, EGFR: Epidermal Growth Factor Receptor, FGF1: Fibroblast growth factor 1, FGF4: Fibroblast growth factor receptor 4, FYN: FYN Proto-Oncogene, FGFR1: Fibroblast growth factor receptor 1, FGFR2: Fibroblast Growth Factor Receptor 2, GRB2: Growth Factor Receptor Bound Protein 2, GSK3B: Glycogen Synthase Kinase 3 Beta, HDAC1: Histone Deacetylase 1, HDAC2: Histone Deacetylase 2, HDAC3: Histone Deacetylase 3, HSP90AA1: Heat Shock Protein 90 Alpha Family Class A Member 1, HSP90AA1: Heat Shock Protein 90 Alpha Family Class A Member 1, JAK1: Tyrosine-protein kinase JAK1, JUN: Jun Proto-Oncogene, AP-1 Transcription Factor Subunit, LCK: LCK Proto-Oncogene, Src Family Tyrosine Kinase, MYC: MYC Proto-Oncogene, BHLH Transcription Factor, NFKB1: Nuclear Factor kappa B Subunit 1, NFKBIA: NFKB Inhibitor Alpha, Non-Receptor Tyrosine Kinase, PDCD1: Programmed cell death protein 1, PIK3R1: Phosphoinositide-3-Kinase Regulatory Subunit 1, PIK3R2: Phosphoinositide-3-Kinase Regulatory Subunit 2, SRC: SRC Proto-Oncogene, SRC: SRC Proto-Oncogene, TNF: Tumor necrosis factor, TP53: Tumor Protein P53

Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)
Mutated based analysis

Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)

Supplement Table 2. The results of pathway analysis in PIK3CA Inhibitor responding patient

| Pathway Name                  | High Frequency Gene | Frequency Ratio | TBPS | Matching_drug_names                  |
|-------------------------------|---------------------|-----------------|------|--------------------------------------|
| PI3K-Akt signaling pathway    | HSP90AA1            | 100.0           | 3.7  | Alvespimycin, Tanespimycin           |
| PI3K-Akt signaling pathway    | PIK3CA              | 66.7            | 2.4  | Alpelisib, Copanlisib               |
| PI3K-Akt signaling pathway    | PDGFRA              | 66.7            | 2.4  | Erdafitinib, Midostaurin            |
| PI3K-Akt signaling pathway    | KIT                 | 33.3            | 1.2  | Ancestim, Erdafitinib, Midostaurin  |
| PI3K-Akt signaling pathway    | CSF1R               | 33.3            | 1.2  | Erdafitinib                         |
| PI3K-Akt signaling pathway    | FGFR1               | 33.3            | 1.2  | Erdafitinib                         |
| PI3K-Akt signaling pathway    | PDGFRB              | 33.3            | 1.2  | Erdafitinib, Midostaurin            |
| PI3K-Akt signaling pathway    | CSF3R               | 33.3            | 1.2  | Filgrastim, Lenograstim, Lipeglgrastim, Pegfilgrastim |
| PI3K-Akt signaling pathway    | HSP90AB1            | 33.3            | 1.2  | Tanespimycin                        |
| PI3K-Akt signaling pathway    | FLT1                | 33.3            | 1.2  | Vatalanib                           |

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
### Mutated based analysis

#### nCounter based analysis

| Pathway Name                          | High Frequency Gene | Frequency Ratio | TBPS | Matching_drug_names                                                                 |
|---------------------------------------|---------------------|----------------|------|-------------------------------------------------------------------------------------|
| PI3K-Akt signaling pathway            | HSP90AA1            | 80.0           | 2.6  | Alvespimycin, Tanespimycin                                                            |
| Proteoglycans in cancer               | EGFR                | 58.3           | 2.5  | Dacomitinib                                                                         |
| PI3K-Akt signaling pathway            | HSP90AB1            | 70.0           | 2.3  | Tanespimycin                                                                         |
| Pathways in cancer                    | HSP90AA1            | 80.0           | 1.9  | Alvespimycin                                                                         |
| Pathways in cancer                    | PIK3R1              | 80.0           | 1.9  | Enzastaurin                                                                          |
| Pathways in cancer                    | EGFR                | 70.0           | 1.6  | Afatinib, Brigatinib, Canertinib, Cetuximab, Erlotinib, Gefitinib, Icotinib, Lapatinib, Neratinib, Olmutinib, Osimertinib, Panitumumab, Pelitinib, Rindopepimut, Trastuzumab, Vandetanib, Varlitinib, Zalutumumab |
| PI3K-Akt signaling pathway            | MET                 | 50.0           | 1.6  | Cabozantinib                                                                        |
| Proteoglycans in cancer               | ESR1                | 33.3           | 1.4  | Toremifene                                                                           |
| PI3K-Akt signaling pathway            | FGFR1               | 40.0           | 1.3  | Erdaftinib                                                                           |

**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

**Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)**
| Mutated based analysis                                      | Pathways in cancer | Treatment Benet Prediction Score (TBPS) and suggestion of drug (Full data) |
|------------------------------------------------------------|--------------------|----------------------------------------------------------------------------|
| PI3K-Akt signaling pathway                                 | FGFR2              | 40.0 1.3 Erdafitinib                                                      |
| PI3K-Akt signaling pathway                                 | FGFR4              | 40.0 1.3 Erdafitinib                                                      |
| Pathways in cancer                                         | ERBB2              | 50.0 1.2 Afatinib, Brigatinib, Lapatinib, Pertuzumab, Trastuzumab, Trastuzumab emtansine, Tucatinib, Varlitinib |
| Pathways in cancer                                         | MET                | 50.0 1.2 Amuvatinib, Brigatinib, Cabozantinib, Crizotinib, Tivantinib     |
| Pathways in cancer                                         | AKT1               | 50.0 1.2 Arsenic trioxide, Enzastaurin                                   |
| Pathways in cancer                                         | JAK2               | 50.0 1.2 Ruxolitinib                                                     |
| PI3K-Akt signaling pathway                                 | CDK4               | 30.0 1.0 Abemaciclib, Palbociclib, Ribociclib                            |
| PI3K-Akt signaling pathway                                 | PIK3CA             | 30.0 1.0 **Alpelisib**, Copanlisib                                       |

AKT1: AKT Serine/Threonine Kinase 1, CDK4: Cyclin-dependent kinase 4, CSF1R: Macrophage colony-stimulating factor 1 receptor, CSF3R: Granulocyte colony-stimulating factor receptor, EGFR: Epidermal growth factor receptor, ERBB2: Receptor tyrosine-protein kinase erbB-2, ESR1: Estrogen receptor, FGFR1: Fibroblast growth factor receptor 1, FGFR2: Fibroblast growth factor receptor 2, FGFR4: Fibroblast growth factor receptor 4, FLT1: Vascular endothelial growth factor receptor 1, HSP90AA1: Heat shock protein HSP 90-alpha, HSP90AB1: Heat shock protein HSP 90-beta, JAK2: Tyrosine-protein kinase JAK2, KIT: Mast/stem cell growth factor receptor Kit, MET: Hepatocyte growth factor receptor, PDGFRA: Platelet-derived growth factor receptor alpha, PDGFRB: Platelet-derived growth factor receptor beta, PIK3CA: Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform, PIK3R1: Phosphatidylinositol 3-kinase regulatory subunit alpha isoform.

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.

**Supplement Table 1. Treatment Benefit Prediction Score (TBPS) and suggestion of drug (Full data)**
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 platform 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 integration of each mutation and the respective mRNA expression into signaling pathway can identify their functional significance and therapeutic target. Pathway network based on cancer gene landscape 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 measures the overlap between mutation and mRNA expression data, and calculates 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 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, so 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. TBPS suggests the matching drug, and visualize the likelihood of responding probability.

During the experiment, we also observed that not only oncogenic pathways but also non-oncogenic pathways were deregulated and activated in HNSCC. These multiple pathways involvement implies that targeting multiple pathway is useful for further refining the anti-cancer chemotherapy. We also found that overly activated pathway measured by mRNA over-expression and suppressed pathway 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 FGFR inhibitor and PIK3 inhibitor. Our developed CBSJukebox® system suggested both the FGFR inhibitor and PIK3
inhibitor. However, when we applied interaction frequency ratio cut-off of 75%, the PIK3 inhibitor was excluded. This might be have been caused by the insufficient availability of gene data that interacted with PIK3CA pathway in 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 large sample size. Regarding future work, our Bayesian network model offers an easy way of incorporating additional data types such as CNV, proteomics data, methylation data, and so on, and such model extension should be attempted.

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 HNSCC are mainly caused by the loss-of-function mutations. However, big data based platform for druggable pathways can find potential matching drugs.

**Abbreviations**

- cell adhesion molecules (CAMs)
- copy number variation (CNV).
- head and neck squamous cell carcinoma (HNSCC)
- Korean Cancer Study Group (KCSG)
- next generation sequencing (NGS)
- mutation-based pathways (MBP),
- over-expressed genes-based pathway (OEBP)
- single-nucleotide variant (SNV)
- protein–protein interactions (PPIs)
- Treatment Benefit Prediction Score (TBPS)
- under-expressed genes-based pathways (UEBP)

**Declarations**

*Ethics approval and consent to participate*
Institutional Review Boards of each institute approved this study protocol. (List of IRB: IRB of Seoul National University Hospital, IRB of Chonnam National University Hwasun Hospital, IRB of Severance Hospital, IRB of Gachon University Gil Medical Center, IRB of Bucheon St. Mary's Hospital, IRB of Kangdong Sacred Heart Hospital, IRB of National Cancer Center, IRB of Seoul National University Bundang Hospital, IRB of SMG-SNU Boramae Hospital, IRB of Samsung Medical Center, IRB of CHA Bundang Medical Center, IRB of Chungnam National University Hospital) Written informed consents were obtained prior study enrollment from all participants.

Consent for publication

Not applicable.

Availability of data and materials

The data that support the findings of this study are available from KCSG (Korean Cancer Study Group) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of KCSG.

Competing interests

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.

Funding

TRIUMPH trial was supported by a grant from the National R&D Program for Cancer Control, 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.

Author Contributions

Conceptualization, B.K., J.Y.P, H.J.Y.; Resources and data curation, B.K., S.H.C.3, S.K., H.K.A., S.H.C.6, J.H.K., T.Y., J.W.K., J.E.K., M.J.A., J.H.K.; Methodology, B.K., J.Y.P, G.D.K.; Formal analysis, J.K., Y.Y.; Writing-original draft, B.K.; Writing-review and editing, G.D.K., H.J.Y.; All authors have read and agreed to the published version of the manuscript.

Acknowledgement

TRIUMPH trial was supported by a grant from the National R&D Program for Cancer Control, 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

References

1. Sheikine Y, Kuo FC, Lindeman NI. Clinical and Technical Aspects of Genomic Diagnostics for Precision Oncology. J Clin Oncol. 2017;35(9):929–33.

2. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015;517(7536):576–82.

3. Agrawal N, Frederick MJ, Pickering CR, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science. 2011;333(6046):1154–7.

4. Stransky N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. Science. 2011;333(6046):1157–60.

5. Lim SM, Cho SH, Hwang IG, 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(1):300–12.

6. Keam B, Kim HR, Yun HJ. TRIUMPH Trial: One Small Step Could Become One Giant Leap for Precision Oncology in Head and Neck Cancer. Cancer Res Treat. 2019;51(1):413–4.

7. Li Z, Ivanov AA, Su R, et al. The OncoPPI network of cancer-focused protein-protein interactions to inform biological insights and therapeutic strategies. Nat Commun. 2017;8:14356.

8. Ivanov AA, Revennaugh B, Rusnak L, et al. The OncoPPI Portal: an integrative resource to explore and prioritize protein-protein interactions for cancer target discovery. Bioinformatics. 2018;34(7):1183–91.

9. Ivanov AA, Khuri FR, Fu H. Targeting protein-protein interactions as an anticancer strategy. Trends Pharmacol Sci. 2013;34(7):393–400.

10. Murakami Y, Tripathi LP, Prathipati P, Mizuguchi K. Network analysis and in silico prediction of protein-protein interactions with applications in drug discovery. Curr Opin Struct Biol. 2017;44:134–42.

11. Park IJ, Yu YS, Mustafa B, et al. A Nine-Gene Signature for Predicting the Response to Preoperative Chemoradiotherapy in Patients with Locally Advanced Rectal Cancer. Cancers (Basel). 2020. doi:10.3390/cancers12040800.

12. Kim CM, Hwang S, Keam B, et al. Gene Signature for Sorafenib Susceptibility in Hepatocellular Carcinoma: Different Approach with a Predictive Biomarker. LIC. 2020;9(2):182–92.

13. Costello M, Pugh TJ, Fennell TJ, 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(6):e67.

14. Zhao Y, Hoang TH, Joshi P, et al. A route-based pathway analysis framework integrating mutation information and gene expression data. Methods. 2017;124:3–12.
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.

16. Creixell P, Reimand J, Haider S, et al. Pathway and network analysis of cancer genomes. Nat Methods. 2015;12(7):615–21.

Figures
Figure 1

Overflow of protein-protein interaction analysis (CBSJukebox® Analysis)
Comparison of mutation-based pathways, over-expressed genes- based pathways, and under-expressed genes-based pathways
Figure 3

Oncoplot for top 20 pathways analyses in all patients

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- figureS1full.tif