이학박사 학위논문

Generation and molecular characterization of the patient derived xenograft(PDX) model of pancreatic cancer and triple negative breast cancer(TNBC) with its application in preclinical study

췌장암과 삼중음성유방암의 환자유래이종이식 동물모델 구축과 분자적특성 분석 및 전임상 연구에의 응용

울 산 대 학 교 대 학 원

의 학 과

정 재 윤

Generation and molecular characterization of the patient derived xenograft(PDX) model of pancreatic cancer and triple negative breast cancer(TNBC) with its application in preclinical study

지 도 교 수　　장　수　환

이 논문을 이학박사 학위 논문으로 제출함

2018 년　8 월

울 산 대 학 교 대 학 원

의　　학　　과

정　재　윤

정재윤의 이학박사학위 논문을 인준함

심사위원장 강상욱 (인)

심사위원 강민지 (인)

심사위원 김상수 (인)

심사위원 김용섭 (인)

심사위원 장수환 (인)

울 산 대 학 교 대 학 원

2018 년 8 월

ABSTRACT

Establishing an appropriate preclinical model is crucial for translational cancer research. The most common way that has been adopted by far is grafting cancer cell lines. Although this xenograft model is easy to generate, but has several limitations because this cancer model could not represent the unique features of each cancer patient sufficiently. In attempt to avoid these discrepancies between xenograft model and patients' tumor, a patient-derived xenograft (PDX) model has been recently generated and applied.

Here we report the generation and characterization of 20 PDX models of PDAC. The success rate of the initial graft was 74%. Histological analysis of the PDXs and primary tumors revealed a conserved expression pattern of p53 and SMAD4. An exome single nucleotide polymorphism (SNP) array and Comprehensive Cancer Panel showed that PDXs retained over 94% of cancer associated variants. In addition, Polyphen2 and the Sorting Intolerant from Tolerant (SIFT) prediction identified 623 variants among the functional SNPs, highlighting the heterologous nature of pancreatic PDXs; an analysis of 409 tumor suppressor genes and oncogenes in Comprehensive Cancer Panel revealed heterologous cancer gene mutation profiles for each PDX-primary tumor pair.

Also, for breast cancer study, to identify novel druggable targets and to understand its unique biology, we tried to implement 39 patient-derived xenografts (PDXs) of TNBC. The overall success rate of PDX implantation was 61.5%. Immunohistochemical analysis revealed conserved ER/PR/HER2 negativity between the original and PDX tumors. Genomic analysis of 10 primary tumor-PDX pairs with Ion AmpliSeq CCP revealed high degree of variant conservation (85.0% to 96.9%) between primary and PDXs. Further analysis showed 44 rare variants with a predicted high impact in 36 genes including TP53, PTEN, NOTCH1, and COL1a1. Furthermore, RNA-seq analysis of 24 PDXs revealed 594 gene fusions, of which 163 were in-frame, including AZGP1-GJC3.

Moreover, this paper presents two preclinical studies to show that PDX models are useful for drug screening, biomarker development, and the preclinical evaluation of personalized medicine strategies. Firstly, we conducted preclinical study investigating the efficacy of GnRH agonist for TNBC. We found that apoptosis was increased in tumor highly

expressing GnRH receptor, and also B-Raf, p-ERK, p-MEK were decreased in the group which are injected GnRH agonist, on the contrary, cleaved PARP was increased.

As another example showing usefulness of PDX model, we studied the relationship of miR-155 and YAP. We conducted reverse phase protein array (RPPA) with 30 human TNBC samples (15 of miR-155 high and 15 of miR-155 low). The results of RPPA analysis showed phosphorylated YAP(S127) level was significantly down-regulated in miR-155 high group. Western blot results confirmed that pYAP(S127) was elevated in miR-155 knock-down cells compared to control. In addition, we selected a couple of protein by calculation of p-value in FPKM data from RNA-seq. As a result, SVIL was considered to be an important mediator in regulating YAP by miR-155. Conversely, we also tested if YAP controls the expression of miR-155. Indeed, the overexpression of YAP increased the level of miR-155, whereas its knockdown decreased miR-155 level.

Altogether, we expect these PDX models are a promising platform for screening novel therapeutic agents and diagnostic markers for the detection and eradication of PDAC and TNBC. Also our report provides a molecular basis for the usefulness of the TNBC PDX model in preclinical study.

Keywords: Patient-derived xenografts, pancreatic cancer, triple-negative breast cancer, Novel variant, Cancer panel, heterogeneity

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

According to the report of World Health Organization and National Cancer Institute, cancer is defined as a group of diseases involving abnormal cell growth with the potential to invade or spread to other parts of the body. In 2012, an estimated 14.1 million new cancer cases and 8.2 million cancer deaths occurred in worldwide[1)], that comprises 14.6% of total human deaths. Cancer is a leading cause of death worldwide in countries of all income levels. The number of cancer cases and deaths is expected to grow rapidly as populations grow, age, and adopt lifestyle behaviors that increase cancer risk[2)]. In this aspect, cancer research needs to be performed in depth based on the model that better represents the exact characteristic of human cancer.

Cancer cell lines derived from patients have been used for not only in vitro but also in vivo experiments[3)]. Especially, a xenograft model generated by the injection of these cell lines subcutaneously into immune-deficient mice is the most commonly used model in preclinical drug development[4)]. This model is easy to generate and shows consistent tumor growth among animals. However, the xenograft model using cancer cell lines has several limitations on the other hand. The first one is this model could not reflect the patient's drug response sufficiently. The fact that successful clinical approval rate for cancer drugs is very low (approximately lower than 15%), especially in solid tumor[5)], supports this idea. The second issue comes from an observation that most tumors contain highly heterogeneous subsets of cancer cells with different characteristics in one tumor. Also, the xenograft tumor often cannot recapitulate the tumor microenvironment although tumor microenvironment affects the growth and metastasis of cancer cells. To overcome these limitations, there has been increasing interest in patient-derived tumor xenograft (PDX) model as a more advanced preclinical cancer model. In the PDX model, a tumor specimen is directly transplanted into immunocompromised mice, providing a faithful representation of individual tumor[6)].

Among various cancers, pancreatic cancer is a fatal disease in humans[7, 8)] and is often referred to as being a silent killer because in general, there are no symptoms until late tumor stages, at which point the tumor cells have metastasized and multiple lesions are

formed [9]. Consequently, only 20% of the tumors are resectable [10], which limits translation research using cancer specimens. Currently, a few chemotherapeutic options are available for pancreatic cancer, such as Gemcitabine or fluorouracil (5-FU). However, these are not greatly effective (extending the survival by only a few months) and produce substantial side effects[7, 11]. Therefore, there is an urgent clinical need for the development of novel diagnostic and therapeutic options. The establishment of a preclinical model for pancreatic cancer is a prerequisite for developing new treatments. A genetically engineered mouse model is currently available for pancreatic cancer, in which activated Kras and/or Trp53 mutant proteins are specifically induced in the pancreatic ductal epithelial cells[12, 13]. However, this model cannot fully reflect human pancreatic cancer, which is genetically heterogeneous. Consequently, the use of patient-derived xenograft (PDX) models is becoming an attractive option because the tumor specimens are directly transplanted into immunocompromised mice, providing a faithful representation of individual tumors as mentioned above[14]. However, establishing PDX can also be a challenge, with the success rate varying according to several factors, including the type of tumor, recipient mouse, transplant technique, and time gap between surgery and transplantation[15]. Recent studies have successfully generated PDXs for pancreatic cancer. For example, Delitto et al. also successfully generated 15 PDXs from 25 specimens and they also found that mouse stromal cells infiltrated the human cancer cells, suggesting active tumor-stromal interactions in pancreatic cancer[16]. Regarding the molecular analysis of Pancreatic Ductal Adenocarcinoma (PDAC) PDXs, it was reported that KRAS and PIK3CA mutation were analyzed up to eight passages and similar mutations frequency in PDXs was found[15]. Therefore, it seems that use of the PDX model is very successful for pancreatic cancer, suggesting that it would be a good preclinical model for understanding this complex disease. Here, we describe 20 pancreatic PDXs originating from PDAC patients who underwent surgery at the Asan Medical Center, Seoul, Korea. Clinical information and analysis of the molecular data revealed that these pancreatic PDXs have novel and heterologous characteristics.

Breast cancer is the second most common cancer among women worldwide. In developed countries, the 5-year survival rate is close to 90%[17]. The high survival rate is due

to the early detection of cancer by ultrasonography or mammography and availability of multiple targeted therapeutic agents for hormone or growth receptors, such as letrozole or trastuzumab. Nevertheless, there is no effective targeted therapy for Triple Negative Breast Cancer (TNBC), therefore, more research is needed to understand and design new therapy for this subtype of breast cancer[18]. Indeed, recent reports focused new therapeutics for TNBC using PDX models[19, 20]. Other reports, however, showed that TNBCs themselves are diverse[21]; accordingly, additional molecular characterization of TNBCsis urgently needed. Furthermore, the sub-classification of TNBCs by molecular profiling[22] or by the response to therapeutic agents will ultimately help researchers to design a strategy to treat this cancer.

Toward this goal, we generated PDXs from TNBC and performed a cancer panel analysis and high-throughput RNA sequencing (RNA-seq) as well as immunohisto-chemical(IHC) analysis. Unlike ER-positive breast tumors, which showed a poor graft success rate even with estrogen medication, triple-negative breast tumors manifested a remarkable success rate (~61.5%) for the initial graft. In this article, we describe molecular characteristics of the successful TNBC PDXs in comparison with their original tumors and demonstrate the potential of this model as a preclinical tool for personalized treatment.

To show the way how to use PDX model in preclinical study, we presented two examples. The one is evaluation of efficacy of GnRH agonist in TNBC PDX mice, the other one is a research on relationship between microRNA-155 and phosphorylated-YAP in breast cancer.

Gonadotropin-releasing hormone (GnRH) is a gonadotropin hormone that is secreted from hypothalamus in the brain[23]. GnRH is secreted to adenohypophysis in the pituitary gland and promotes the secretion of Follicle-stimulating hormone (FSH) and luteinizing hormone (LH)[24]. The blood carries the FSH and LH to ovary or testis which is target organ. GnRH activates its own receptor, gonadotropin-releasing hormone receptor(GnRHR), which is seven-transmembrane G protein-coupled receptor[25]. There are three types of GnRH receptor in humans, first of which is known to be the most important receptor[26]. Recently, GnRH agonist has been used for neoadjuvant therapy which is treatment given before primary therapy. The main reason for the treatment with GnRH agonist is the pregnancy problem in breast cancer patients[27]. Usually, Chemotherapy which is usually

treated to breast cancer has side effect that ovarian follicle is not formed in the ovary[28]. It can cause infertility, which is critical problem for young women. Although the exact mechanism of GnRH agonist therapy has not been clarified in clinical practice, but it has been shown to be able to maintain ovarian function, which is attracting attention as an adjuvant therapy for breast cancer. The second great advantage of GnRH agonist therapy is that it is effective in the TNBC breast cancer type[29, 30]. TNBC is one of several types of breast cancer in which estrogen receptor, progesterone receptor, and HER2 receptor are not expressed well[31]. Unlike other subtype breast cancer which hormone therapy or HER2 targeted therapy can be applied as therapeutics, TNBC has no effective treatment yet and its prognosis is the worst. Therefore, we examined the effect of GnRH agonist (zoladex) and tried to understand their mechanism in breast cancer PDX model. Several signaling pathways associated with GnRH have been identified[32], but it remains to be revealed which pathway mediates the effects of the GnRH agonist on breast cancer. After the GnRH receptor level was measured and analyzed in 10 PDX samples and selected 2 samples whose expression is the highest and the lowest. These samples were transplanted into 10 mice respectively. Doxorubicin and Cyclophosphamide, which are currently used in chemotherapy for breast cancer, with and without GnRH agonist were administered to mice and analyzed the effect of GnRH agonist as a adjuvant therapy.

As the second example experiment using PDX model, we researched the relationship between miR-155 and YAP. 30 human TNBC samples were divided two group depending on the level of miR-155 and reverse phase protein array (RPPA) was conducted. In this data, miR-155 and phosphoYAP(p127) showed dramatic reverse correlation.

miR-155, representative oncogenic microRNA, is known to promote survival and growth of cancer cell, accelerating angiogenesis and epithelial mesenchymal transition, which causes cancer migration and invasion[33-35]. Increased level of miR-155 was detected in B-cell lymphoma and chronic lymphocytic leukemia[36]. It is reported that miR-155 plays an important role in cancer development in pancreatic[37, 38] and breast cancer[39]. The Hippo pathway is known to regulate cell proliferation and organ size in both Drosophila and mammals. In mammals, the Hippo pathway consists of a highly conserved kinase cascade, whose upstream proteins SAV1, MST1/2, and LATS1/2 together form a complex to

phosphorylate the key transcriptional coactivator Yes-associated protein (YAP), thus facilitating YAP's retention in the cytoplasm by creating 14-3-3 binding site[40]. Consequently, YAP cannot move to the nucleus where it interacts with the TEAD transcription factors and mediates the expression of the downstream growth-promoting and apoptosis-inhibiting genes. YAP was found to be amplified in various human cancers, and a functional analysis demonstrated that its overexpression induced growth factor–independent proliferation, anchorage-independent growth, epithelial-to-mesenchymal transition, and apoptosis suppression in nontransformed human mammary epithelial cells, indicating a potential oncogenic role for this protein[41]. In our study, miR-155 regulates YAP phosphorylation and YAP also regulates the expression of miR-155. Taken altogether, our investigation suggests a novel role of miR-155 on hippo pathway.

MATERIALS AND METHODS

Generation and storage of PDX

The animal care protocol for this study was approved by the International Animal Care and Use Committee (IACUC) of the Laboratory of Animal Research at the Asan Medical Center, Seoul, Korea. Five-week-old male NOD/SCID mice (for PDAC PDX) and female NOD/SCID mice (for TNBC PDX) were used for tumor engraftment and were grown in a specific pathogen-free facility. The surgical specimens were obtained under permission from the institutional review board (IRB) of the Asan Medical Center (No. S2013-0744-0009). IRB number for breast PDX is No. 2014-0800. Fresh tumor tissues were obtained from pancreatic cancer or breast cancer patients who underwent surgery and were immediately placed in RPMI medium (10% FBS, 1% penicillin/streptomycin) at 4℃ in the refrigerator. As soon as possible after this, the samples were spliced into 2-3mm^3 fragments and implanted into the subcutaneous interscapular fat pad for pancreatic cancer PDX or the inguinal mammary fat pad of mice for breast cancer PDX, respectively. All of the animals were anesthetized with 15 mg/kg of Zoletil® (Virbac, USA) and 2.5 mg/kg of Rompun® (Bayer Korea, Korea) by intraperitoneal injection for tumor implantation. Following implantation, the mice were monitored twice per week for at most 12 months. Once the xenograft tumor had attained a size of 300–500 mm^3, the tumor was excised and the mice were euthanized following the protocol of the Laboratory of Animal Research at the Asan Medical Center. Part of the tumor that had been excised from the mouse was engrafted into another NOD/SCID mouse for expansion, while the residual part of the tumor was stored in a freezing medium with dimethyl sulfoxide (DMSO) and kept in a deep freezer. The success or fail was determined whether tumor was generated in mouse and IHC results were the same as primary tumor.

Immunohistochemical analysis

Tumors were fixed in 10% formalin for at least 24 hour then embedded in paraffin. Both human and mouse tumor tissues were sectioned at a 5μm thickness and stained with haematoxylin and eosin stain (H&E). For analysis of pancreatic cancer,

Immunohistochemistry (IHC) was performed to examine the expression of p53 and SMAD4 in the primary human tumors, as previously described[42], following the protocol of the Department of Diagnostic Pathology at the Asan Medical Center. Briefly, after deparaffinization and antigenic retrieval, the slides were labeled with a monoclonal antibody against p53 (cloneDO-7, 1:3,000; DAKO, Glostrup, Denmark) and SMAD4 (clone EP618Y, 1:100; GeneTex, Irvine, CA, USA). Labeling was detected using the avidin-biotin complex staining method. 3, 3-diaminobenzidine (DAB) was used as the chromogen for p53 and 3-amino-9-ethylcarbazole was used for SMAD4. A pathologist who was experienced in pancreatic cancer reviewed the slides to compare the tumor architecture and desmoplastic appearance. For analysis of breast cancer PDX, IHC analysis was performed to examine the expression of estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor (HER2) in the primary human tumors, in accordance with the protocol of the Department of Diagnostic Pathology at the Asan Medical Center. Formalin-fixed paraffin-embedded tissue slices were stained with an automatic immunohistochemical staining device (BenchMark XT; Ventana Medical Systems, Tucson, AZ). Antibodies to ER (diluted 1:50, NCL-6F11, monoclonal; Novocastra, Newcastle, UK) and to PR (diluted 1:100, 1E2, rabbit monoclonal; Roche, Tucson, USA) were used. IHC analysis of the HER2 protein was performed by means of the anti-HER2 antibody (anti-HER2/neu [4B5] rabbit monoclonal antibody; Ventana Medical Systems), and an UltraView™ universal DAB detection kit (Ventana Medical Systems).

Collection of SNP for pancreatic cancer PDX

Genetic variant data for the PDX samples were gathered using the Infinium HumanExome-12 v1.2 BeadChip. This platform targets putative functional exonic variants selected from over 12,000 individual exome and whole-genome sequences. The output data contain both SNP and single base insertion or deletion information. The data also include the GeneCall score for each variant of the samples, which is a quality control measure that was scaled between 0 and 1.

Quality control for genetic data for pancreatic cancer PDX

For each sample, we counted the number of variants that completely failed in genotype calling (GenCall score = 0) (Supplementary Table 1-S7). This resulted in the exclusion of two samples (#12 and 16) that had an exceedingly large number of failed genotypes (>10,000). We then chose 217,793 variants (from a total of 244,770) that had a positive GeneCall score in all remaining 20 samples, and used these variants in the subsequent genetic analysis.

Genetic similarity and MDS analysis for pancreatic cancer PDX

We used plink v1.07 to perform a similarity analysis using the genetic data. We calculated the identify-by-state (IBS) pairwise similarity between samples using the--cluster-distance-matrix options in plink. We then generated a heatmap and dendrogram using R. We also generated an MDS plot using the--cluster--mds-plot options in plink and the R package heatmap v3.

Prediction and selection of deleterious variants of Pancreatic cancer PDX

We used SIFT[43)] and Polyphen2[44)] to predict and select putatively important variants that may cause protein damage. Polyphen2 predicted which variants were possibly damaging, probably damaging, or benign, while SIFT predicted which variants were damaging or tolerated based on the Rapid Stain Identification Series (RSID) of each variant. We defined a variant as deleterious if the Polyphen2 prediction was possibly/probably damaging or if the SIFT prediction was damaging. Among 244,770 variants (i.e., all variants before applying the quality controls), 13,613 were predicted as being deleterious.

Defining the gene disruption variable in pancreatic cancer PDX

To analyze the data at the gene level, we newly defined a genetic variable that indicated whether the gene was disrupted or not. We defined a gene as being disrupted if any variant that was predicted as being deleterious within the gene carried the risk allele. Since Polyphen2 and SIFT did not provide information about the risk allele, we obtained this information from Illumina, and confirmed this by comparing the data to the predictions from

Ensemble.

Confirmation of variants by Sanger sequencing in breast cancer

All the positions of missense variations found in the Notch1 gene were between 139,399,350 and 139,399,405 on chromosome 9. Forward and reverse primers were designed to read the sequences of the area. Primer sequences were as follows: (F) 5′-TCC ACC AGT TTG AAT GGT CA-3′ and (R) 5′-AGC TCA TCA TCT GGG ACA GG-3′. After extraction of DNA from tumor tissue, samples were sequenced by Macrogen Korea (Seoul 08511, Republic of Korea). Sanger sequencing results were analyzed on a chromatogram, and mutations were confirmed.

RNA isolation and Quantitative RT-PCR (qRT-PCR)

RNA extraction was performed using TRizol (Invitrogen) by the instructions from the manufacturer. 1 ug of total RNA was used for cDNA synthesis (Superscript First-Strand Synthesis System; Invitrogen) following the manufacturer's protocol. The expression levels of all genes were measured by SYBR Green PCR Kit, in LightCycler 480 II (Roche). Standard PCR conditions are: 15 minutes at 95℃ followed by 40 cycles of 15-second denaturation at 94℃, 30-second annealing at 60℃, and 30-second extension at 70℃. Human and mouse ribosomal protein gene human RpL13a was used as an internal control. Relative quantification was carried out by the $2-\Delta\Delta Ct$ method[45]. For quantitative assessment of miR-155, RNAs was reverse transcribed by miScriptll RT Kit (Qiagen, Hilden, Germany) and analyzed by miScript SYBR Green PCR Kit (Qiagen) with the instructions from the manufacturer. Samples were monitored using the LightCycler 480 II for an initial denaturation at 95℃ for 15 minutes followed by 40 PCR cycles with 94℃ for 15 seconds, 55℃ for 30 seconds, and 70℃ for 30 seconds. The expression was normalized to the noncoding, small nuclear RNA molecule U6 (RNU6). Data were presented as fold change versus control. The primer sequences were summarized in Supplementary Table 5-1.

RNA-seq analysis for breast cancer PDX

Except for one sample that failed quality control, 24 PDX samples were subjected to TrueSeq mRNA library construction and RNA-seq on an Illumina HiSeq 4000. Paired-end sequencing with a read length of 101 base pair resulted in 60 million reads per sample on average. We employed Bowtie 2 [46] for mapping to the hg19 reference genome. Gene annotation and transcript quantification were performed using RNA-Seq by Expectation-Maximization (RSEM)[47] based on GENCODE v19 [48]. Transcripts per kilobase of a gene per million mapped reads (TPKM) were calculated for normalized quantification of RNA amounts.

Variant calling in breast cancer PDX

For variant calling from RNA-seq, not exome sequencing, Splice Transcripts Alignment to a Reference (STAR)[49] has been reported to have greater sensitivity than other aligners including Burrows–Wheeler Alignment (BWA). To carry out sample-wise splicing junction analysis, we performed the STAR 2-pass alignment steps. Variant calling and filtering were done using Genome Analysis Toolkit (GATK)[50]. We employed SplitNCigarReads to split reads into exon segments and to hard-clip any sequences overhanging with the intronic regions. We filtered clusters of at least three single-nucleotide polymorphisms (SNPs) that are within a window of 35 bases between them. As in exome variant filtering, variants showing Fisher Strand (FS) > 30.0 and Qual By Depth (QD) < 2.0 were selected. For functional variant annotation, ANNOVAR (Annotate Variation)[51] was applied. We performed ANNOVAR filtering based on allele frequencies from public databases. For frequency of variants in whole-genome data, we used 1000g2015aug (10000 Genome Project dataset). For frequency of variants in whole-exome data, we chose exac03 (Exome Aggregation Consortium dataset) and esp6500siv2 (NHLBI-ESP project dataset). For variant identification, dbSNP (Single Nucleotide Polymorphism database) version 147 was searched.

Gene fusion analysis in breast cancer PDX

To detect gene fusion events, we availed ourselves of defuse[52] with the hg19 human reference genome. For alignment, Bowtie 2[46] and GMAP (genomic mapping and alignment

program)[53] were chosen. The filtering parameters included clustering_precision = 0.95, span_count_threshold = 5, percent_identity_threshold = 0.90, split_min_anchor = 4, splice_bias = 10, and probability_threshold = 0.50. To confirm the gene fusion events, primers were designed to help read the junction sequences by Sanger sequencing. DNA extracted from tumor tissue was sent to Macrogen Korea (Seoul, Republic of Korea) and sequenced. FASTA sequences were analyzed by Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/), and fusion was confirmed by the results of BLAST. The primer sequences were AZGP1-GJC3: 5′-CTC TGG AGA CAG GTG GAA GG-3′ and 5′-GTC CCA GTT GTC GGT TAT GC-3′; for VDAC2-ADK: 5′-CGC GTC CAA TGT GTA TTC CT-3′ and 5′-GTG CAG CCA GTC CGT CTA AT-3′.

Western blot analysis

Western blot analysis was performed as previously described[6]. Briefly, cells were lysed in lysis buffer (150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 5mM Tris-HCl [pH 7.5], 2 mM ethylenediaminetetraacetic acid [EDTA;pH 8.0], and 0.1% SDS). Following this, 10~50 μg of protein were separated on SDS PAGE, transferred to a nitrocellulose membrane, and probed with anti-EGFR, p-ERK, p-B-Raf, p-MEK, p-AKT, p53, SMAD4, MTAP, and p16 (1:1,000; Cell Signaling Technology). The membranes were then stripped and reprobed with anti β-actin antibody (1:1,000; Santa Cruz Biotechnology, CA, USA) to ensure equal loading. For analysis of breast cancer PDX, antibodies against AR, p-AR, p-MEK, PI3K, EGFR, p-AKT, ERK, p-ERK (1:1000, Cell Signaling Technology), or GnRHR (1:1000, Abcam) are used. Membranes were stripped and reprobed with an anti-GAPDH antibody (1:1000; Santa Cruz Biotechnology, Dallas, TX, USA) to ensure equal loading.

Statistics

For the analysis of clinical factors affecting successful xenograft, we applied a univariate and multivariate statistical models. For univariate statistical analysis, the statistical significance was measured by a t-test or a chi-square test. For multivariate analysis, a logistic regression method was used to determine the effect of multiple clinical factors. The survival curve was plotted using Kaplan-Meier method and the significance of the differences between the two

curves was calculated by a log-rank test. Cox proportional hazards regression model was also used both for individual variable and for the multivariate analysis. All the statistical analysis was also carried out by Microsoft Excel or the R package (ver.3.3).

GnRH agonist preclinical study

We measured the expression level of GnRH receptor from 10 tumor tissues. We chose 2 samples whose GnRH receptor expression level was relatively high, and low. We generated 10 TNBC breasat cancer PDX model for each sample. Aniamls were divided into three groups (PBS, doxorubicin and cyclophosphamide(AC), AC plus zoladex). The concentration of drug was 5mg/kg, 150mg/kg and 8mg/kg for doxorubicin, cyclophosphamide and zoladex respectively. Drug was injected subcutaneously every other day, however doxorubicin was injected by tail vein injection every four day. After that, when tumor reached $100mm^3$~$150mm^3$ volume, we started to measure tumor growth and animal weight. If tumor overgrows $500mm^3$, we sacrificed that mouse.

Cell culture and transfection

Human breast cancer cell lines such as MDA-MB-436, MDA-MB-231, HS578T were cultured in DMEM media supplemented with 10% FBS and Penicillin-streptomycin (P/S). Patient-derived cell (harvest from PDX tumor) were maintained in RPMI media containing 5% FBS, 1% penicillin/streptomycin, hEGF, hydrocortisone and transferrin. For overexpression, cells were transfected using lipofectamine 3000 (Invitrogen) with the each plasmid containing gene or promoter of interest. For knockdown of YAP, we ordered siRNA to Genolution and treated into each cells. siRNA Sequences are summarized in Supplementary Table 5-3.

Luciferase assay

Cells were seeded in 24-well plate and transfected with 150ng of BIC-Luc (PGL3 enhancer vector containing BIC promoter) with or without YAP plasmid. As an internal control, 10 ng of SV40-Luc (Renilla luciferase) per well was used. Forty-eight hours after transfection, cells were lysed and luciferase activity was measured using Dual-Luciferase Reporter Assay

System (Promega) and GloMax Luminometer (Promega), according to the manufacturer's protocol.

Annexin V / PI staining

A part of tumor tissue from PDX was chopped with sharp scissor and minced with homogenizer. Fragmented tissue was incubated in 37C shaking incubator with collagenase. Then, tissue was put in strainer and ground with rubber part of 2ml syringe. Cells with media passed through out of strainer were collected and stained according to the protocol of FITC-Annexin V Apoptosis Detection kit (BD Pharmingen). Briefly, trypsinized cells were washed twice with cold PBS and then resuspended in 1X binding buffer at a concentration of $1x10^6$ cells/ml. 100ul of the solution was transferred to 1.5ml tube. 5ul of FITC Annexin V and 5ul PI were added to each tube and incubated for 15mins at RT in the dark. Then 400ul of 1X binding buffer was added to each tube. Stained cells were analyzed by using an Accuri Flow Cytometry (BD Biosciences). The percentage of FITC-positive cells was calculated using the CFlow software.

Reverse Phase Protein Array (RPPA)

For RPPA analysis, 30 triple negative breast cancer (TNBC) samples were lysed and protein was extracted using RPPA lysis buffer. [1% Triton X-100, 50mM HEPES, pH 7.4, 150mM NaCl, 1.5mM MgCl2, 1mM EGTA, 100mM NaF, 10mM Na pyrophosphate,1mM Na3VO4, 10% glycerol, containing freshly added protease and phosphatase inhibitors]. Lysates were manually serial-diluted in 5 two-fold dilutions with lysis buffer and analyzed at the RPPA Core, MD Anderson Cancer Center, Texas, USA. Briefly, samples printed on nitrocellulose-coated slides using an Aushon Biosystems 2470 arrayer. Slides were probed with approximately 300 validated primary antibodies followed by detection with appropriate Biotinylated secondary antibodies (Goat anti-Rabbit IgG, Goat anti-Mouse IgG, or Rabbit anti-Goat IgG). The signal obtained was amplified using a Cytomation–catalyzed system of Avidin-Biotinylated Peroxidase (Vectastain Elite ABC kit from Vector Lab) binding to the secondary antibody and catalyzing Tyramide-Biotin conjugation to form insoluble biotinylated phenols. Signals were visualized by a secondary streptavidin-conjugated HRP

and DAB colorimetric reaction. The slides were scanned, analyzed, and quantified using Array-Pro Analyzer software (MediaCybernetics) to generate spot intensity.

Chromatin immunoprecipitation (ChIP) assay

ChIP was performed by an online protocol (http://cshprotocols.cshlp.org/content/ 2009/9/pdb prot5279. full) with minor modifications. Briefly, MDA-MB-436, MDA-MB-231 and HS578T cells were cross-linked with 1% formaldehyde for 10 minutes at RT. Then the cells were collected, lysed, and the chromatin was sonicated, and incubated with Flag or YAP antibody. PCR was used to amplify DNA bound to the immunoprecipitated miR-155 promoter, after reversing the protein-DNA crosslinks. Primer sequences for CHIP were summarized in Supplementary Table 5-2.

RESULTS

Chapter 1. Generation and molecular characterization of pancreatic cancer patient-derived xenografts reveals their heterologous nature

1) Generation of pancreatic patient-derived xenografts (PDXs) and primary cells

In total, we obtained 29 freshly dissected specimens from surgery, in which we carefully selected the region that exhibited enriched tumor cells. Approximately 1cm^3 of tumor tissue was obtained and cut into small pieces (1–2mm^3 on average). Three or four of these pieces were then subcutaneously transferred into NOD/SCID mice under anaesthetized conditions. From here, it usually takes 1~2 months for the tumor to grow. Using this method, we successfully produced 20 PDXs, representing a 72.4% success rate. Representative pictures of these PDXs are shown in Figure 1-1A and 1-1B, and Supplementary Figure S1-1. We also obtained six primary cancer cell lines from the PDXs (Supplementary Figure S1-2; also see Supplementary Table S1-1 for clinical information) and utilized some of these cells in our genomic analysis, along with human pancreatic ductal epithelial (HPDE) cells and PDAC (Panc1) cells.

2) Analysis of clinical data reveals several clinical parameters affecting the success of PDX

Next, we checked the clinical information to determine which factor(s) affected the success of PDX (see Table 1-1 for a summary). Due to the highly metastatic properties of PDAC, most of our PDX samples fell into Stage IIA or IIB, exhibiting lymph node metastasis but not distance metastasis. Here, we specifically focused on tumor size at surgery, recurrence, gender, and survival/death of the patient. Other factors such as lymphatic/vascular invasion, histological type, and distant metastasis were not considered due to the limited number of cases for each. Among the clinical characteristics analyzed, we found survival/death of the patient was significantly associated with the success rate of PDX (P=0.023 by Cox proportional hazard regression analyses, Figure 1-1C). In addition, tumor size (P=0.059) and recurrence showed a positive correlation but was not significant (Figure 1-1D and 1-1E). Multivariate analysis of the recurrence and tumor size, however, revealed the tumor size is a significant factor for the success of PDX (p=0.048, Table 1-2).

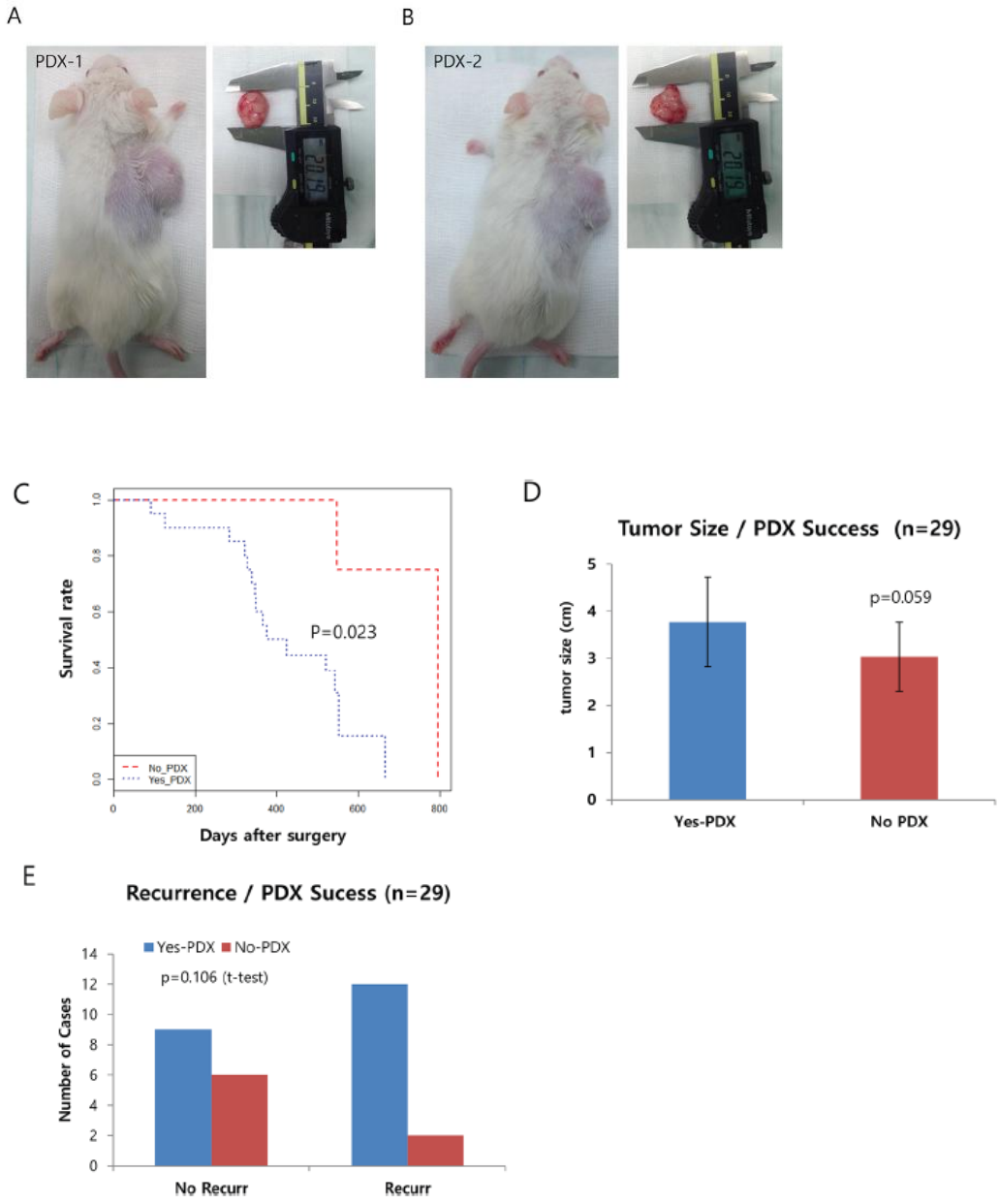


Figure 1-1: Generation of pancreatic patient-derived xenografts (PDXs) and the clinical features affecting their success. A-B. Representative pictures of PDXs in NOD/SCID mice. The right panels show the dissected tumors being measured with calipers. C. Kaplan-Meyer curve of the two groups of successful xenograft (Yes PDX in Blue) or failed xenograft (No PDX in Red) (by Song Cheol Kim.). D and E. Graphs showing a positive correlation between the success of the xenograft and other clinical factors, including tumor size (D) and recurrence (E).(by Suhwan Chang)

Table 1-1: Clinical characteristics of the parental tumors of the patient-derived xenografts (PDXs) (by Song Cheol Kim)

ID	Tumor Size after Surgery	pT	pN	M	M1 Site	Stage	Lymphatic Invasion	Vascular Invasion	Histologic_type	Recur	Recur_type	Death_data	Age	Gender
AMC001	3.5	3	0	0		ⅡA	1	1	DUCTAL ADENOCARCINOMA	1	remnant pancreas	2014-04-08	66	M
AMC002	2.1	3	0	0		ⅡA	0	0	DUCTAL ADENOCARCINOMA	1	liver meta	2014-09-26	30	M
AMC003	3.5	3	1	0		IIB	0	0	DUCTAL ADENOCARCINOMA				66	M
AMC004	3.1	3	1	0		IIB	0	0	DUCTAL ADENOCARCINOMA				53	M
AMC005	5.7`	3	1	0		IIB	1	1	DUCTAL ADENOCARCINOMA	1	liver meta	2014-12-01	70	M
AMC006	2.2	3	0	0		ⅡA	0	0	DUCTAL ADENOCARCINOMA	1	encasing the celiac trunk and SMA	2015-05-30	58	F
AMC007	4.3	3	1	0		IIB	1	1	ADENOCARCINOMA	1	remnant pancreas(tail)		53	M
AMC008	3.5	3	0	0		ⅡA	1	1	ADENOCARCINOMA, AOV				50	M
AMC009	3.7	3	1	0		IIB	1	1	DUCTAL ADENOCARCINOMA	1	liver meta	2015-02-20	57	M
AMC010	4.3	3	1	0		IIB	1	1	DUCTAL ADENOCARCINOMA	1	liver meta	2015-04-26	48	F
AMC011	3.7	3	1	0		IIB	1	1	DUCTAL ADENOCARCINOMA	1	liver meta		61	M
AMC012	5.5	3	0	0		ⅡA	0	0	DUCTAL ADENOCARCINOMA	1	remnant pancreas(head)	2015-03-21	64	F

Table 1-1. (continued)

AMC013	5.1	3	1	0		IIB	1	1	DUCTAL ADENOCARCINOMA			2015-03-15	50	M
AMC014	5	3	1	0		IIB	1	1	SARCOMATOID CARCINOMA	1	liver meta and peritoneal seeding	2014-07-22	59	F
AMC015	3.6	3	1	0		IIB	1	1	DUCTAL ADENOCARCINOMA			2015-04-06	73	F
AMC016	3.5	3	1	1	liver	Ⅳ	1	1	DUCTAL ADENOCARCINOMA				61	M
AMC017	4.5	3	1	0		IIB	0	0	DUCTAL ADENOCARCINOMA			2015-04-29	56	M
AMC018	3.7	3	0	0		ⅡA	0	0	DUCTAL ADENOCARCINOMA				50	M
AMC019	2.1	3	1	0		ⅡA	1	1	DUCTAL ADENOCARCINOMA	1	hepatoduodenal ligament and around pancreaticojejunostomy site	2015-03-26	71	F
AMC020	3.9	3	0	0		ⅡA	0	0	ADENOSQUAMOUS CARCINOMA	1	liver meta	Lost	77	F

Tumor size, TMN stage, histological type, invasion, recurrence, and death data are summarized. Note that most of the PDX tumors were at stage IIA to IIB as they were operable upon diagnosis. AOV: ampullar of vater

Table 1-2: Summary of the multivariate analysis affecting the success rate of PDX

	Coefficient	Std. error	p-value
Tumor_size	1.8052	0.9135	0.0481
Recur	2.6846	1.4763	0.069
(intercept)	−6.0514	3.2551	0.063

For the tumor size and recurrence, a logistic regression method was used to determine the effect of multiple clinical factors on the success of PDX.

3) The comparison of histological features of PDX with its original primary tumor

To confirm that the gross histology of the primary tumor was conserved in the PDXs, we performed hematoxylin and eosin (H&E) staining and immunostaining with anti-p53 and SMAD4 antibodies. Overall, we observed similarities between the gross histology of the primary tumors and PDX tumors (Figure 1-2A, 1-2D, and 1-2G, and Supplementary Figure S1-3). In addition, p53 and SMAD4 staining showed a comparable reactivity in most cases (13 out of 16) of PDX-primary tumor pairs (Figure 1-2B, 1-2C, 1-2E, 1-2F, 1-2H, and 1-2I, and Supplementary Figure S1-3). These results show that the PDAC PDXs generated in this study recapitulated the primary tumors histologically.

4) An exome single nucleotide polymorphism (SNP) array enables grouping of the PDXs and identifies putative, functional SNPs

To characterize the PDXs at the molecular level, we performed an exome SNP array. Among the 20 PDXs, we excluded #12 and 16 due to the poor data quality. Instead, we included primary cancer cell line (59390), HPDE cells and Panc1 cells as cancer and normal cell controls. We aimed to compare the SNP profile of each sample so that we could subcategorize PDXs, and discover putatively functional SNPs. These functional SNPs could help us to better understand the molecular mechanisms of tumorigenesis as well as tumor heterogeneity.

We first selected 24,000 non-rare variants from 244,770 variants using plink (option – maf 0.1). We then found 1,385 deleterious variants, as predicted by Polyphen2 and Sorting Intolerant from Tolerant (SIFT) (see methods). Following this, we removed variants whose risk alleles were present in HPDE to obtain only cancer-specific variants, which left us with 623 variants (Supplementary Table S1-2). Table 1-3 summarizes the top 10 genes for each PDX that showed a high number of deleterious variants. We found that there was little overlap between these variants among the PDXs, with the sum of the top 10 variants for each PDX tumor comprising only a minor portion, ranging from 62 (10.9%) to 102 (16.4%), which implied that pancreatic PDXs are heterogeneous However, a phylogenetic tree

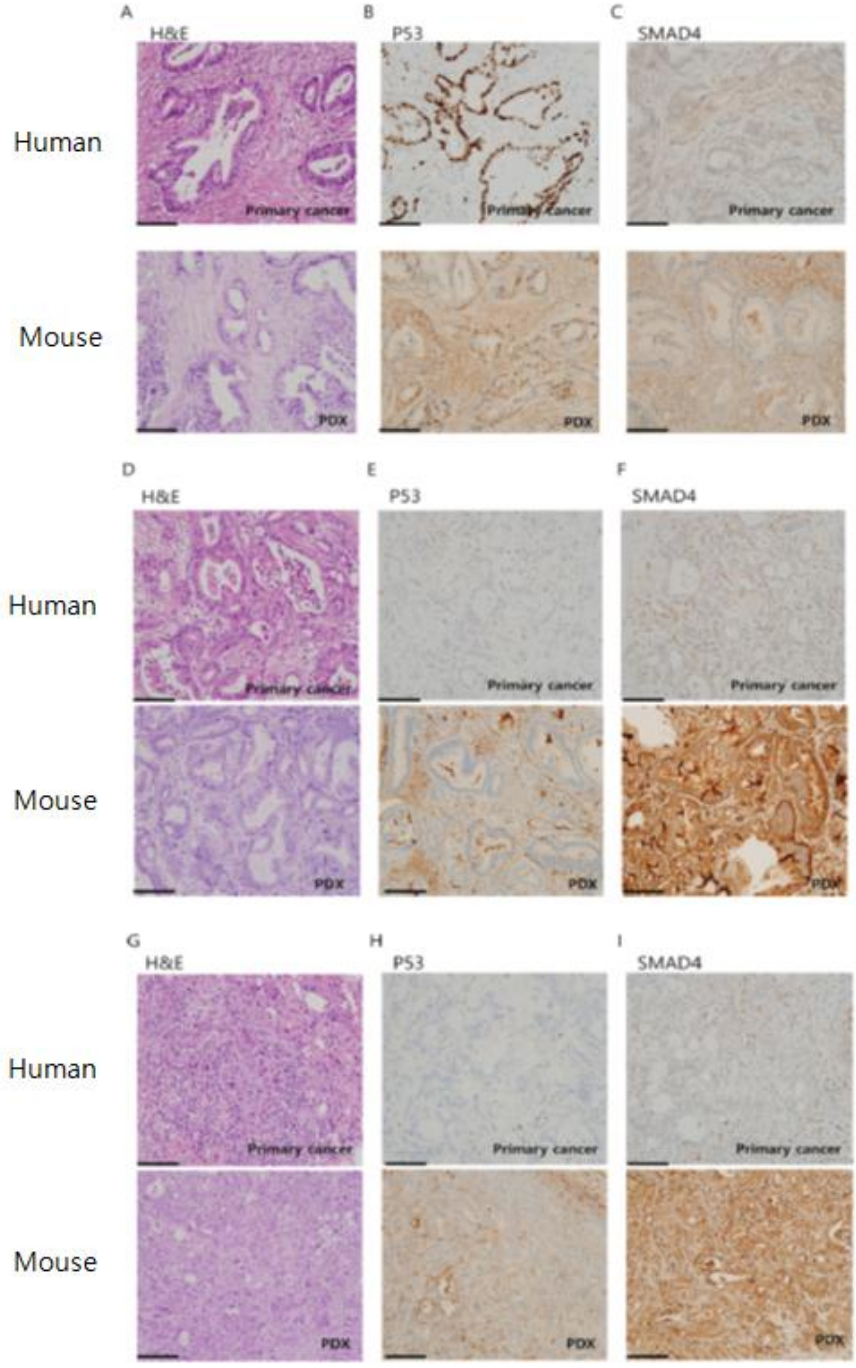


Figure 1-2: Immunohistochemistry (IHC) analysis of patient-derived xenograft (PDX)-primary tumor pairs reveals a conserved histology. IHC images of three PDX-primary tumor pairs: AMC001 A-C. AMC002 B-D. and AMC003 E-G. stained with hematoxylin and eosin (H&E) (A, D, G), p53 (B, E, H), or SMAD4 (C, F, I). Scale bar:100um.

Table 1-3. List of genes containing deleterious variants that are frequently found in pancreatic patient-derived xenografts (PDXs) (by Je-Keun Rhee and Suhwan Chang)

All									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
LAMB3	*TPO*	*MYBPC1*	*CD101*	*SPINK5*	*MOV10L1*	*C3orf20*	*RAET1E*	*TTN*	*PKD1L2*
222	132	132	126	114	110	104	102	98	92
PDX 1									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TPO	*ADAM15*	*MOV10L1*	*TMEM176B*	*PASK*	*DLG1*	*SLC7A9*	*RAET1E*	*PRR5*	*PAK6*
12	12	10	8	8	8	6	6	6	6
PDX 2									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TTN	*CPT1B*	*MYBPC1*	*LAMB3*	*ZNF484*	*TMEM176B*	*FAM129C*	*DLG1*	*ANKRD6*	*RAET1E*
14	14	12	12	8	8	8	8	8	6
PDX 3									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TTN	*LAMB3*	*MOV10L1*	*ZNF484*	*TMEM176B*	*SPINK5*	*SLC7A9*	*RAET1E*	*PCDHGA1*	*NFATC3*
14	12	10	8	8	6	6	6	6	6
PDX 4									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TTN	*CPT1B*	*LAMB3*	*TMEM176B*	*PRICKLE1*	*PASK*	*CX3CR1*	*SPINK5*	*RAET1E*	*PPP2R4*
14	14	12	8	8	8	8	6	6	6
PDX 5									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TTN	*TPO*	*MYBPC1*	*NUP62*	*MOV10L1*	*ZNF484*	*TMEM176*B	*TLR10*	*FAM129C*	*DLG1*
14	12	12	10	10	8	8	8	8	8
PDX 6									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
CAMKK2	*NUP62*	*MOV10L1*	*ZNF484*	*PASK*	*FAM129C*	*CX3CR1*	*SPINK5*	*SP110*	*SLC7A9*
18	10	10	8	8	8	8	6	6	6
PDX 7									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TMEM176B	*TLR10*	*SWT1*	*FAM129C*	*SPINK5*	*SLC7A9*	*PAK6*	*LAMB3*	*GFAP*	*CD101*
8	8	8	8	6	6	6	6	6	6
PDX 8									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TPO	*MYBPC1*	*LAMB3*	*DLG1*	*ANKRD6*	*SPINK5*	*SP110*	*RAET1E*	*PMEL*	*NRAP*
12	12	12	8	8	6	6	6	6	6
PDX 9									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TPO	*MYBPC1*	*LAMB3*	*NUP62*	*TLR10*	*PASK*	*FAM129C*	*SPINK5*	*SLC7A9*	*RAET1E*
12	12	12	10	8	8	8	6	6	6
PDX 10									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
LAMB3	*TMEM176B*	*TLR10*	*SLC7A9*	*PPP2R4*	*NBR1*	*MUSK*	*EGFLAM*	*CD101*	*CAPG*
12	8	8	6	6	6	6	6	6	6

Table 1-3 (Continued)

PDX 11									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
CAMKK2	*CPT1B*	*NUP62*	*MOV10L1*	*SWT1*	*PASK*	*FAM129C*	*SLC7A9*	*RAET1E*	*PRR5*
18	14	10	10	8	8	8	6	6	6
PDX 13									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TPO	*LAMB3*	*ADAM15*	*ZNF484*	*TLR10*	*CX3CR1*	*SP110*	*SLC7A9*	*RAET1E*	*NRAP*
12	12	12	8	8	8	6	6	6	6
PDX 14									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
ZNF484	*SPINK5*	*SLC7A9*	*PPP2R4*	*NBR1*	*MUSK*	*LAMB3*	*DFNA5*	*CD101*	*C3orf20*
8	6	6	6	6	6	6	6	6	6
PDX 15									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TTN	*CPT1B*	*LAMB3*	*ADAM15*	*NUP62*	*MOV10L1*	*PRICKLE1*	*SPINK5*	*NBR1*	*GBP3*
14	14	12	12	10	10	8	6	6	6
PDX 17									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
CAMKK2	*ZNF484*	*TMEM176B*	*SPINK5*	*PPP2R4*	*NBR1*	*MUSK*	*GFAP*	*EGFLAM*	*CHIA*
18	8	8	6	6	6	6	6	6	6
PDX 18									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
CAMKK2	*MYBPC1*	*MOV10L1*	*ZNF484*	*TLR10*	*SWT1*	*PRICKLE1*	*FAM129C*	*CX3CR1*	*SPINK5*
18	12	10	8	8	8	8	8	8	6
PDX 19									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
CAMKK2	*CPT1B*	*LAMB3*	*MOV10L1*	*ZNF484*	*TLR10*	*SPINK5*	*RAET1E*	*NRAP*	*MUSK*
18	14	12	10	8	8	6	6	6	6
PDX 20									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
TTN	*MYBPC1*	*LAMB3*	*ZNF484*	*TMEM176B*	*DLG1*	*SP110*	*PRR5*	*PAK6*	*NBR1*
14	12	12	8	8	8	6	6	6	6
HPDE									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
ZZZ3	*ZZEF1*	*ZYX*	*ZYG11A*	*ZXDC*	*ZSWIM6*	*ZSWIM4*	*ZSWIM2*	*ZSCAN5B*	*ZSCAN5A*
0	0	0	0	0	0	0	0	0	0
Pancl									
[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
MOV10L1	*TLR10*	*SWT1*	*CX3CR1*	*SPINK5*	*RAET1E*	*NBR1*	*GBP3*	*CHIA*	*ZNF229*
10	8	8	8	6	6	6	6	6	4

analysis of Single nucleotide polymorphism (SNP) array data were obtained from 18 PDXs and two primary tumors, as well as pancreatic ductal adenocarcinoma (Panc1) and human pancreatic ductal epithelial (HPDE) cells. The top 10 ranked genes with the highest SNP frequencies are listed for each of the PDX-primary tumor pairs. “All” denotes the combined data from all samples. The functional SNPs yielded three groups of clusters (Figure 1-3A). An Information-Based Similarity (IBS) matrix analysis of the deleterious SNPs (Figure 1-3B) and a multidimensional scaling (MDS) plot analysis (Figure 1-3C) showed 70~80% similarity (with the exception of #8), confirming the diversity of genetic variants among the pancreatic PDXs.

5) Comprehensive Cancer Panel reveals unknown genetic alterations specific to pancreatic cancer

Although the data shown in Figure 1-3 and Table 1-3 generated by the exome SNP array provided useful information to classify the18 PDXs along with the primary cancer cell lines, they were insufficient for determining the molecular characteristics of the PDX-primary tumor pairs in terms of cancer-related genes. Therefore, to examine how the cancer-related mutations were conserved between the PDXs and primary tumors, we conducted an analysis of eight PDX-primary tumor pairs using Ion Ampliseq Comprehensive Cancer Panel, which covers 409 cancer related genes (Supplementary Figure S1-4 for general data; for the gene list, see Thermofisher.com). The total number of variants was 40,827, of which 10,031 were novel (Supplementary Table S1-3). There were up to 1,804 variants in the coding region and untranslated region (UTR), and 13 of the genes with these variants were predicted to be highly affected by them. Table 1-4 shows examples of the variants that had a large impact. Notably, we found that PTEN, SMAD4, and TP53 were in this list, confirming previous findings(54, 55) (for raw data, see Supplementary Table S1-4).

Clustering analysis (Figure 1-4A) showed that there was a high similarity between each PDX and primary tumor, with the exception of PDX #20. Furthermore, in the similarity matrix (Figure 1-4B), we could clearly see the conservation of most cancer gene variants between each pair of PDX-primary tumors (ranging from 90.2% to 97.4%). Interestingly, however, all other combinations among the 18 primary tumors showed much less similarity

(from 59% to 67.7%), suggesting heterogeneity of the PDX tumors. The numbers of variants found in the tumors were very close to each other (around 700; Figure 1-4C and Supplementary Table S1-5, nTotal column), implying that there was comparable genetic alteration among the tumors. This was further confirmed by counting the number of novel variants (Supplementary Table S1-6).

Lastly, we measured the degree of mouse cell infiltration by measuring the relative mouse RPL13a expression to human RPL13a in the PDXs. We also included a control comprised of 95% HPDE mixed with 5% mouse fibroblast cells. This showed that there was 1–12% mouse RPL13a expression (Figure 1-4D), suggesting variable mouse cell infiltration in the PDXs.

6) Western blot analysis for the major growth signaling/cell cycle regulatory proteins reveals the heterologous nature of PDX tumors

In addition to the genetic analyses described above, which used an SNP array and Comprehensive Cancer Panel, we performed a series of western blot analyses to check the levels of the major growth signaling and cell cycle regulatory proteins that have previously been implicated in pancreatic cancer[56-59] Accordingly, we found heterogeneous expression levels of these proteins (Figure 1-5). In particular, we observed the frequent loss of p53 expression (by approximately 50%), as well as the minimal expression of p16. In contrast, we detected various levels of p-B-Raf and p-MEK, which are major downstream effectors of K-Ras[60]. Interestingly, some of the PDXs (#4, 9, and 15) showed discordant p-B-Raf and p-MEK levels, suggesting that some alternative pathway activates p-MEK in these tumors. We detected a relatively consistent level of p-AKT and SMAD4, whereas the levels of p-ERK and MTAP varied greatly. Therefore, our protein analysis revealed that the PDXs have a heterologous molecular nature that resembles the known heterologous character of primary tumors, supporting the strategy of using PDX as a preclinical model in pancreatic cancer.

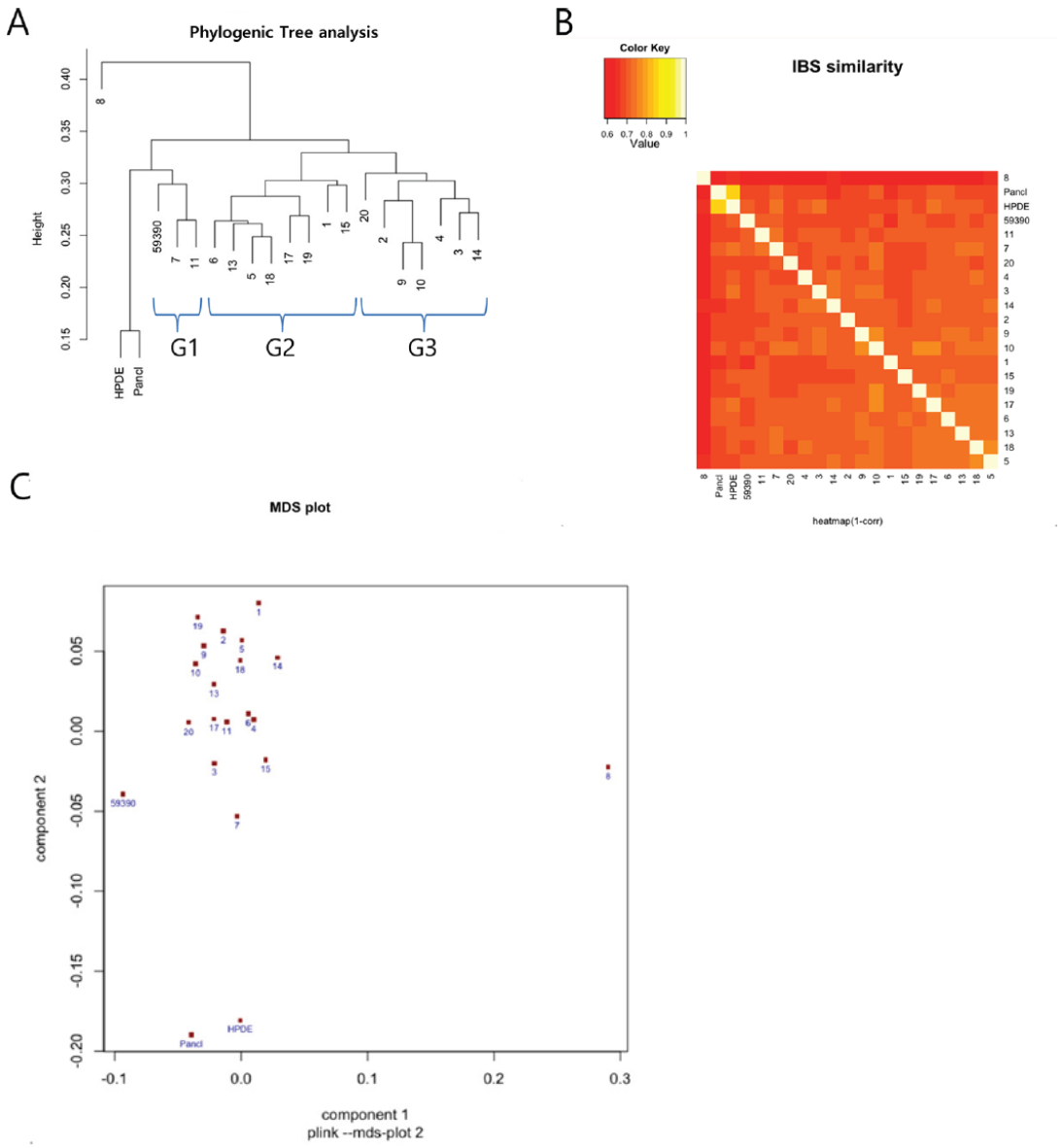


Figure 1-3: Summary of a single nucleotide polymorphism (SNP) array analysis from 18 patient-derived xenografts (PDXs), a primary tumor cell line (59390), pancreatic ductal adenocarcinoma (Panc1) cells, and human pancreatic ductal epithelial (HPDE) cells. The results were obtained from 623 deleterious cancer-specific SNPs. A. Phylogenetic tree showing three main clusters of the variants (marked G1, G2, and G3) occurring among the PDXs. B. Information-Based Similarity (IBS) matrix based on the SNP variants among the PDXs. C. Multidimensional scaling (MDS) plot showing a clustering pattern.(By Eunji Kim, Hyun Cue Lee and Buhm Han)

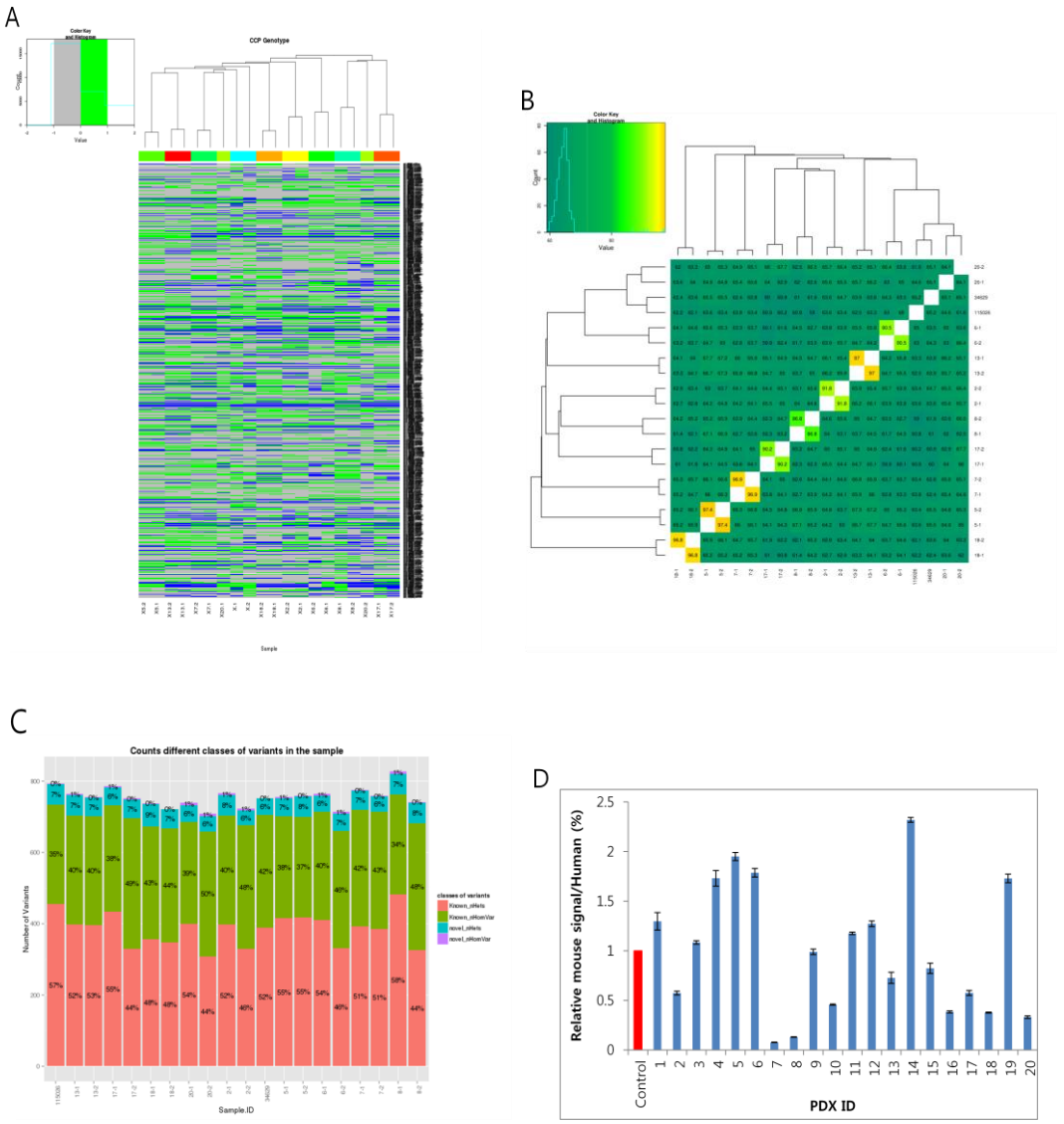


Figure 1-4: Summary of the Ion-Ampliseq Comprehensive Cancer Panel analysis for eight patient-derived xenograft (PDX)-primary tumor pairs. A. Cluster analysis of 18 samples (8 pairs and two cell lines) based on the variants found in 402 cancer genes. B. Similarity matrix showing conservation of the variants between PDX and primary tumors ranging from 90.5% to 97.4%. C. Number of variants found in each PDX and primary tumor sample. The numbers in the bar denote the proportion of known/novel and homologous/heterologous variants, respectively. D. Estimation of the proportion of infiltrated mouse cells in the PDXs calculated by dividing mouse RP13a expression by humanRP13a expression. The control was 5% mouse cells mixed with human pancreatic ductal epithelial (HPDE) cells. (By Macrogen)

Table 1-4: Examples of variants with a high impact, as identified from the Comprehensive Cancer Panel analysis (By Macrogen)

CHROM	POS	REF	ALT	Variant type	QUAL	DP	Allele	Effect	Impact	Gene_ Name	HGVS.c	HGVS.p
chr5	55243415	G	A	SNP	535.96	1513	A	stop_ gained&splice_ region_variant	HIGH	`IL6ST	c.1843C>T	p.Gln615*
chr3	37818889	C	T	SNP	43280.76	29332	T	stop_gained	HIGH	`ITGA9	c.2548C>T	p.Gln850*
chr9	134073362	A	AC	INS	4361.47	934	AC	frameshift_ variant	HIGH	`NUP214	c.4484_ 4485insC	p.Glu1495fs
chr10	89720633	CT	C, CTT	DEL, INS	25093.71	7744	T	splice_acceptor_ variant&intron_ variant	HIGH	`PTEN	c.802-3dupT	
chr8	145738599	A	C	SNP	241.71	1903	C	splice_donor_ variant&intron_ variant	HIGH	`RECQL4	c.2463 + 2T>G	
chr18	48573537	G	T	SNP	22955.8	30701	T	stop_gained	HIGH	`SMAD4	c.121G>T	p.Glu41*
chr17	7579470	C	CG	INS	6935.51	6514	CG	frameshift_ variant	HIGH	`TP53	c.216dupC	p.Val73fs

SNP: Single nucleotide polymorphism; INS: insertion; DEL: deletion; HGVS: Human Genome Variation Society

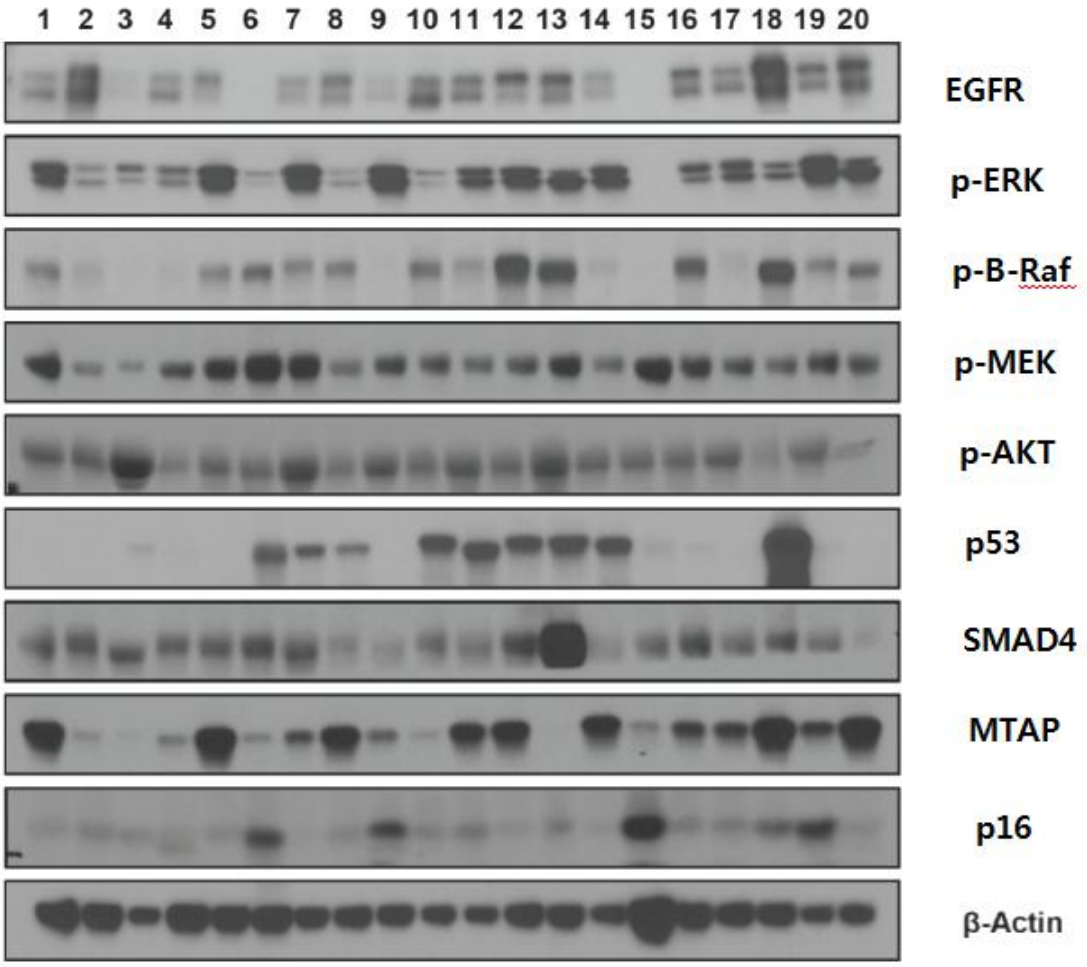


Figure 1-5: Western blots showing the expression levels of various growth signaling and cell cycle regulatory proteins. In total, 20 pancreatic patient-derived xenografts (PDXs) were analyzed. The name of each protein is marked on the right. The beta-actin antibody was used to ensure equal loading

Chapter 2. Novel cancer gene variants and gene fusions of triple-negative breast cancers (TNBCs) reveal their molecular diversity conserved in the patient-derived xenograft (PDX) model

1) Generation of TNBC PDXs

To establish PDXs, we obtained fresh tumor specimens dissected by a pathologist, thereby ensuring that the tumor samples had high cellularity. The tissue samples were stored in the RPMI 1640 medium in 4℃ refrigerator until grafting. Tumor tissue samples 1–2mm^3 in size were introduced into a mammary fat pad and skin of female NOD/SCID mice (Figure 2-1A). On average, we detected tumor growth 138 days after the grafting (Figure 2-1B, Supplementary Table 2-1). The following, second xenograft was performed on NOD/SCID mice. Supplementary Table 2-2 shows the list of PDXs successfully grown until the retransplant. We found that a retransplant takes less time (86 days on average) for the tumor to appear (Figure 2-1C). Among 39 TNBC tumor specimens tested, we obtained 24 successful PDX (Figure 2-1D~2-1G), resulting in a 56.4% success rate (see Discussion).

2) Clinical factors affecting the success rate of PDX

Based on the graft results described above, we analyzed patient's clinical information (Supplementary Table 2-3) to determine which clinical factor is associated with the success rate of the PDX. Our results in the Table 2-1 show that N(positivity for lymph node metastasis) is the only statistically significant parameter affecting the success of PDX (see Discussion). We noticed that primary tumors with a higher N score tended to succeed as a PDX with a higher probability than did primary tumors with a lower N score (Figure 2-2A).

3) IHC analysis of primary tumor–PDX pairs reveal conserved histological features in PDX tissue

One of the important advantages of using the PDX model is that each PDX retains the characteristics of its original tumor. To validate this advantage, we first examined histological features of primary tumor–PDX pairs. Because TNBC lacks ER and PR

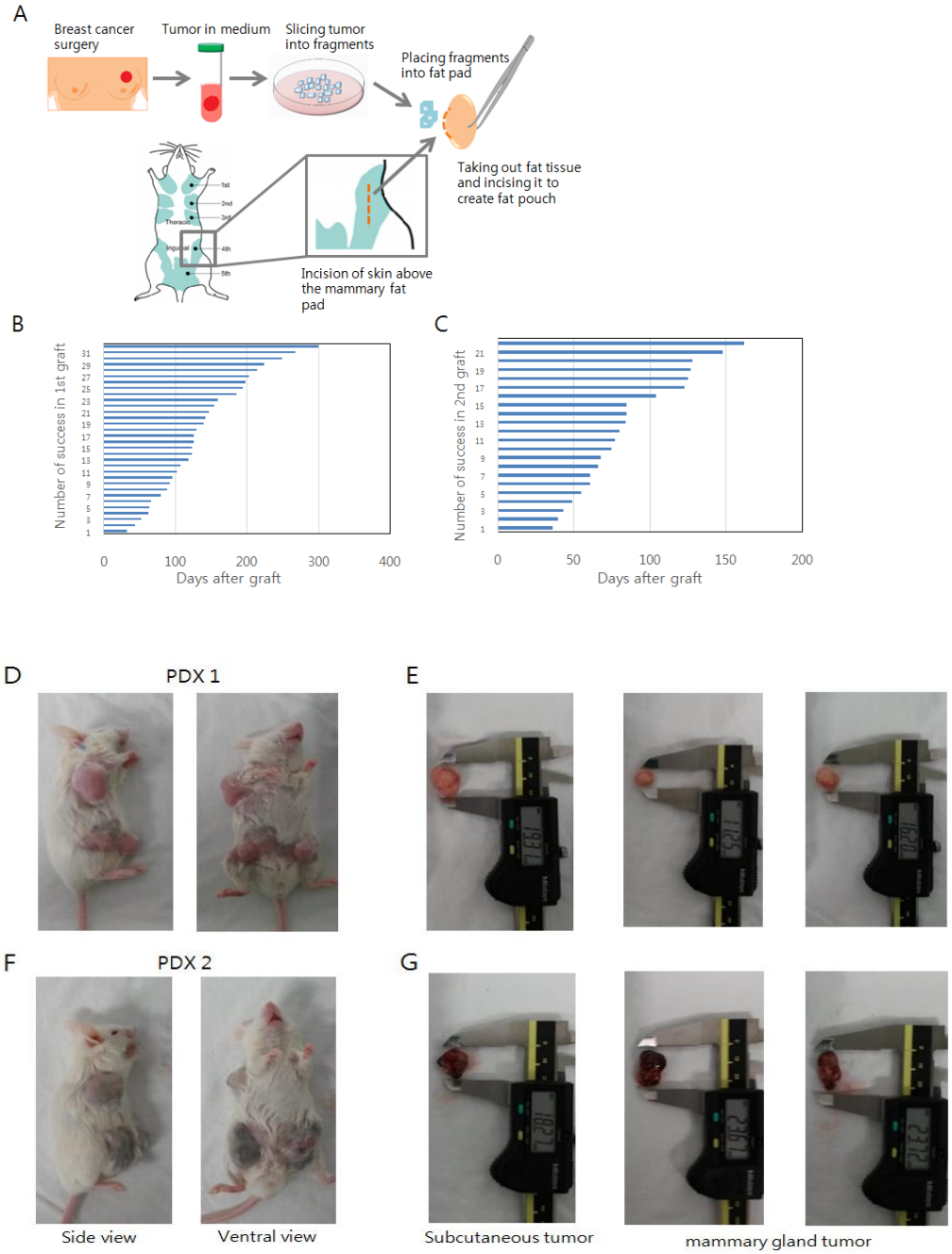


Fig.2-1. Establishment of a patient-derived xenograft (PDX) model for triple-negative breast cancer (TNBC). (A) Schematic diagram for the process of PDX creation. A fresh tissue specimen is minced and directly transferred to the inguinal mammary gland of an immunocompromised mouse.(B) and (C) Days until tumor was generated after graft and (D) (F) Representative pictures of a PDX grown orthotopically (in mammary fat pads) or subcutaneously (below of the right armpit). (E) and (G) Measurement of the dissected tumor using calipers.(B, C by Suhwan Chang)

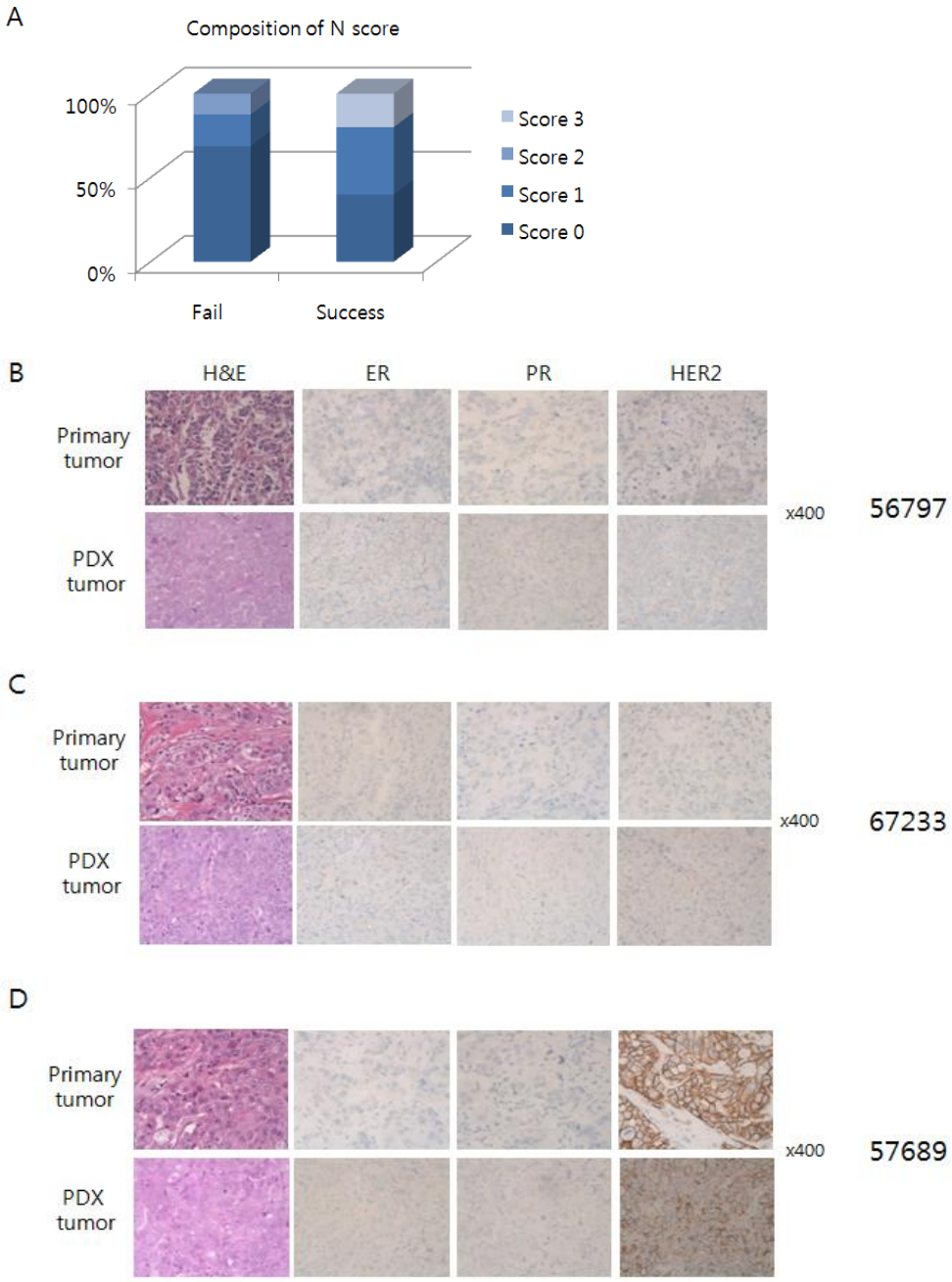


Fig.2-2. Immunohistological analysis of the PDXs compared with their original primary tumor. (A) N score percentage of primary tumor sample (positivity for lymph node metastasis). (B-D) Representative images of H&E staining (left), ER staining (middle left), PR staining (middle right), and HER2 staining (right) in the primary tumor 56797 (upper panels) or PDX tumor (lower panels); 400× magnification. Panels B, C and D show similar images from primary tumor 57689 (for B) or 67233 (for C) or 57689(for D). Note that HER2 expression is conserved in 67689 PDX (C, right panel).

Table 2-1. Clinical factor affecting the success of PDX

Variable	P-value	Test type
Nuclear Grade	0.7417	Chi square test
Histological Grade	0.7417	Chi square test
T (size)	0.5054	Chi square test
N (lymph node involvement)	0.03833	Chi square test
M (distant metastasis)	0.7669	Chi square test
Age at diagnosis	0.376	Logistic regression
Tumor size (in cm)	0.921	Logistic regression

expression and HER2 amplification, we tested whether such triple negativity persists after grafting. As shown in Figure 2-2B ~ 2-2D and Supplementary Figure 2-1, we found that gross histological features (shown by H&E staining) and ER/PR/HER2 status (detected by specific antibodies) did not change from primary to PDX tumors in most cases. The results of comparative IHC analysis of 13 primary tumor–PDX pairs are summarized in Supplementary Table 2-4. In some cases, however, we found that the receptor status of a PDX does not match that of its primary tumor (Supplementary Figure 2-1F and 2-1R, see Discussion).

4) Comprehensive Cancer Panel analysis shows molecular conservation of primary tumor–PDX pairs and identifies rare, unique variants for each PDX

In addition to the histology, we tested whether the molecular signature is conserved between a primary tumor and PDX. For this purpose, we used an Ion AmpliSeq comprehensive cancer panel (CCP) that covers exons of 402 cancer-related genes. We obtained more than 4000 million bases of sequence reads (on average; Fig 2-3A) of which 62% to 76% yielded proper alignments. The distribution of these variants (classified by genomic context) is summarized in Supplementary Figure 2-2. In these CCP data, we observed that variants within 10 primary tumor–PDX pairs had a high rate of matching, ranging from 85% to 96.9% of identity, whereas approximately 60% of identity was observed between other PDXs and their primary tumors (Fig. 2-3B). Clustering analysis of the variants also confirmed an outstanding degree of conservation between primary tumors and corresponding PDX tumors (Fig. 2-3C). Further analysis revealed 45 rare variants with a predicted high impact (Supplementary Table 2-5). Among them, the c.802-3dupT variant in the PTEN gene and the c.2719A>T variant in the COL1A1 gene showed the most frequent alterations (Supplementary Table 2-6). In addition, variants in NOTCH1, NUP214, PMS1, and EPHA7 were frequently found. On the other hand, we found moderate-impact variants in genes APC, MSH2, ATR, and FGFR3(Supplementary Table 2-7), confirming the previous findings[61-65]. The most different genes between primary and PDX tissue were BRAF, FGFR3a and MYH11 (Supplementary Table 2-8). Especially, variants of FGFR3 and MYH11 were not found in primary tissue, on the contrary, variants were found in all PDX tumor tissue.

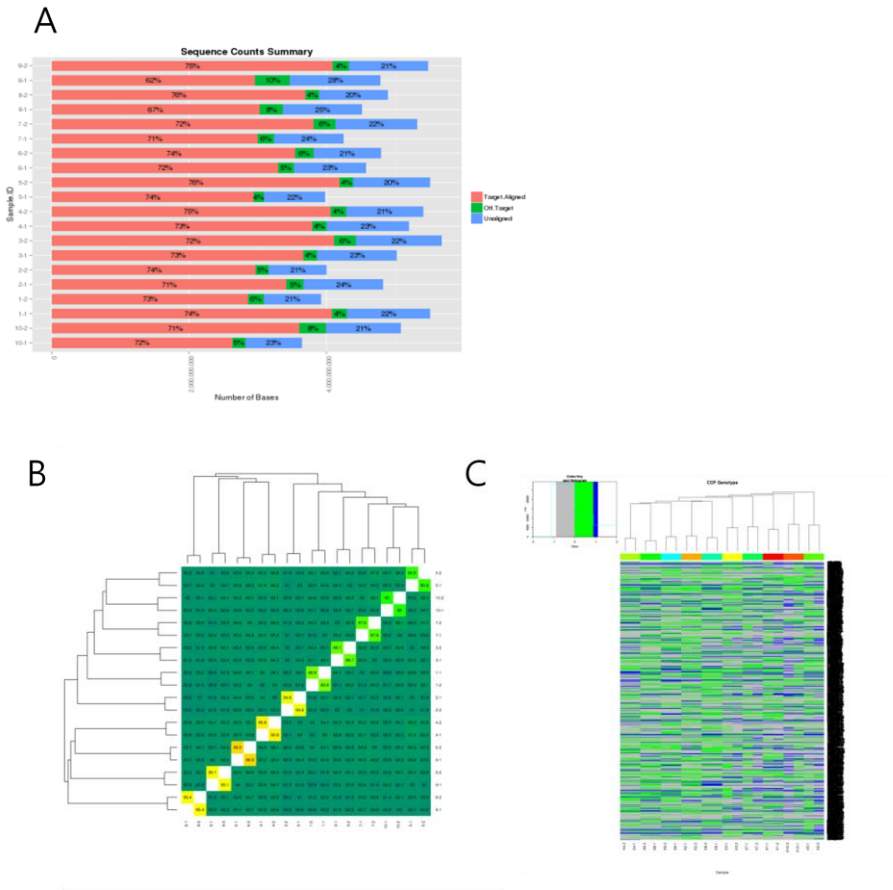


Fig. 2-3. Molecular analysis of the TNBC PDXs by CCP. (A)The graph showing a summary of sequence counts for the Ion AmpliSeq comprehensive cancer panel (CCP). (B) IBS (Identity By Status) cluster analysis showing genotypic similarity of the CCP. The number in each cell indicates similarity between two paired samples. Number1 means a primary tumor, and number2 refers to a PDX tumor. (C) Heat map analysis of CCP genotypes. Note the highly-conserved genotypes between primary (number1) and PDX (number2) tumors (A,B,C by Macrogen)

5) Novel, missense variations found in the NOTCH1 gene inactivate its signaling in primary breast cancer cells

Among the variants found, we further studied missense mutations in the NOTCH1 gene, because these variants in breast cancer have not been previously reported. Sanger sequencing of the NOTCH1 regions coding M1580, E1583, S1588, and R1598 confirmed the missense variation (Figure 2-4A~2-4D, Supplementary Figure 2-3). Interestingly, one of the PDXs (#002079) had homozygous mutations at these sites whereas other PDXs carried heterozygous mutations (Supplementary Table 2-9). To understand the functional impact of these variations, we treated three primary cell lines harboring NOTCH1 variants with resveratrol, which is known to activate NOTCH signaling[66] As a control, we tested the MDA-MB-436 cell line and #67233 PDX lacking variation in the NOTCH1 gene. In these cells, we measured the expression of two known NOTCH1 response genes, HER2 and HEY1[67], to check the NOTCH1 signaling activity in the mutant cells. The graph in Figure 2-4E and 2-4F showed that the expression of the two genes markedly decreased in the three primary cell lines with the NOTCH1 variations. Besides, we observed that resveratrol treatment markedly activated HER2 and HEY1 gene expression in MDA-MB-436 cells but not in the NOTCH1-mutant cells, thus pointing to compromised NOTCH1 signaling in the mutant cells. Also, these differences were checked by performing western blotting analysis. Similar to real-time PCR results, NOTCH1 signaling was amplified in the control cells, but not in mutant cells.(Figure 2-4G)

6) RNA-seq analysis of 24 PDXs revealed their diverse variants and identified novel in-frame gene fusions

To identify the molecular characteristics of the PDX model in more detail, we next performed RNA-seq analysis of 24 PDXs. Supplementary Table 2-10 shows the number of mutated genes and variants located in either a coding region or untranslated region (UTR), detected in each PDX. On average, 5082 variants in coding regions and 21042 variants in UTRs were found (Supplementary Table 2-10). Clustering analysis of 100 (Figure 2-5A) or 500 genes (Figure 2-5B)—that showed high variability—yielded three major groups (indicated as 1, 2, and 3). Nonetheless, there was no clear distinct expression pattern

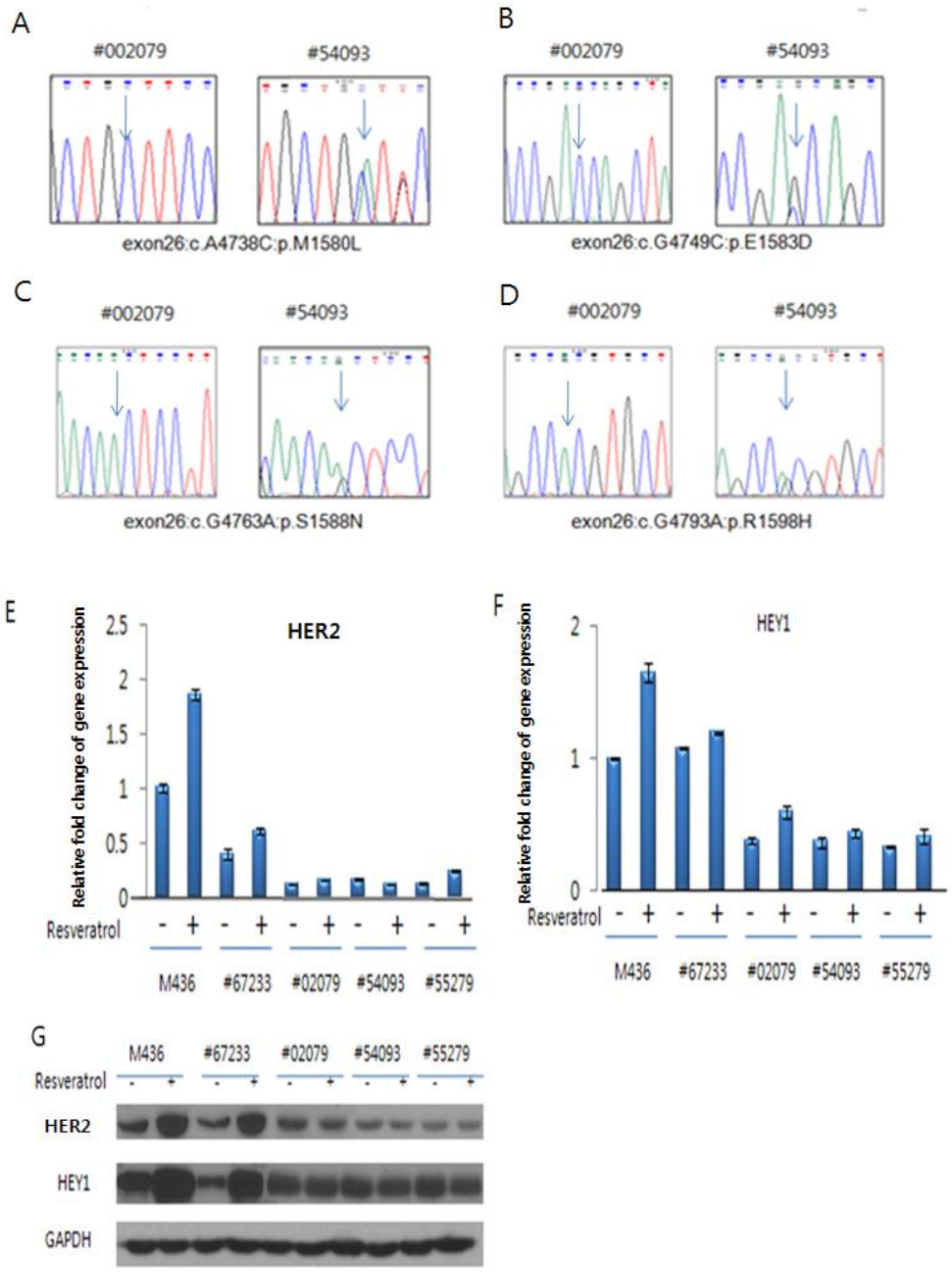


Figure 2-4. Confirmation of novel missense variation found in the NOTCH1 gene by Sanger sequencing and expression level of NOTCH1 target genes. (A-D) An electropherogram showing a missense mutation in the NOCTH1 gene, M1580L (A), E1583D (B), S1588N (C), and R1598H (D). Arrows indicate the position of nucleotide changes. (E) and (F) Graphs showing the expression of Notch signaling target genes including HER2 (E) and HEY1 (F), in NOTCH1 with (MDA-MB-436, #67233) or NOTCH1 mutant (the other cells). (G) Western blotting analysis of HER2 and HEY1 as NOTCH1 target gene.

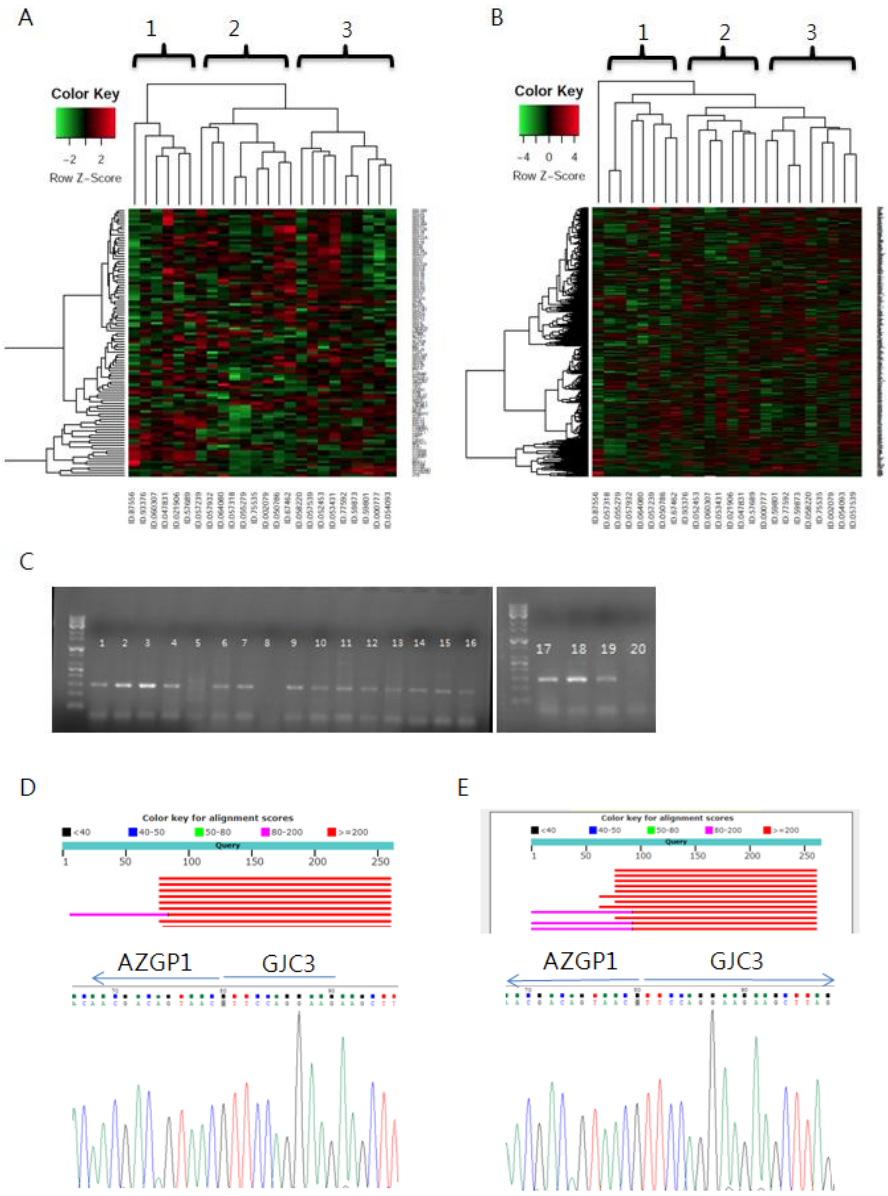


Fig. 2-5. **RNA-seq analysis data and representative gene fusion** (A) Heat map clustering of RNA-seq data. Top 100 genes that show the highest variability among the 24 PDXs were applied to generate the heatmap. According to the clustering, we observed three major subgroups (marked as 1,2, or 3 at the top). (By Macrogen) (B) Similar analysis was performed on the top 500 variable genes, which also show three major clusters. (By Macrogen) (C) RT-PCR results on the AZGP1-GJC3 gene fusion products, in 20 PDXs, showing 300nt specific bands. (By Macrogen) (D) and (E) Representative examples of sequencing alignment (upper panel) and electropherogram data (lower panel) for the AZGP1-GJC3 gene fusion

among the three subgroups, indicating molecular diversity of the TNBCs. Supplementary Tables 2-11 and 2-12 present the top 10 most frequent, frameshift, or stopgain variants. In addition, we found stoploss variants (Supplementary Table 2-13). Of note, we found a frameshift variant of NOTCH2, a stopless variant of ZEB2, and a stopgain variant of TOP2A frequently (see Discussion).

Compared to the number of variants, relatively small numbers of in-frame gene fusions were identified (Supplementary Table 2-14). We found the AZGP1-GJC3 fusion to be the most frequent, showing up in 16 cases out of 24 (Table 2-2). In addition, gene fusions in genes CDK11A, PDCD1LG2, and NUBPL were common in PDXs. Among these identified gene fusions, we could confirm the AZGP1-GJC3 gene fusion. The RT-PCR results in Figure 2-5C show the expression of fusion transcripts in 17 out of the 19 tested PDX samples. Three of these 17 samples were confirmed by Sanger sequencing (Figure 2-5D and 2-5E, Supplementary Figure 2-4).

7) Western blot analysis of major growth signaling proteins confirms their variable expression or activation levels among the PDXs

Even though the CCP and RNA-seq data provide a lot of information on genomic and transcriptomic features of a PDX, it is not always fully translated into a protein amount or activity. Therefore, we examined several proteins involved in cancer signaling by western blotting. Especially, because the PDXs were generated from TNBC specimens, we focused on the expression level of other receptors that may functionally replace ER, PR, and HER2. In addition, we examined the expression of PI3K-AKT and MEK-ERK pathways, which regulate cell proliferation[68]. The results in Figure 2-6 reveal that the EGFR expression level is like the GnRHR level, whereas the Androgen receptor (AR) expression showed a distinct pattern. In some PDXs (#2, 6, 8, and 9) the expression levels of AR and EGFR seem to compensate each other (i.e., inversely correlate). In contrast, some other samples such as #3, 5, and 10 showed low expression of all the three receptors, suggesting the presence of another growth factor receptor in these PDXs. On the other hand, we found the expression patterns of p-AKT and ERK to be similar, whereas p-MEK expression showed a distinct pattern. Interestingly, the p-ERK level did not highly correlate with ERK expression or the p-

MEK amount, implying that there is an alternative regulatory mechanism controlling ERK activation. Lastly we found 50% (5 out of 10) PDX lost the expression of p53, confirming previous results[69]. Altogether, these results showed that TNBC PDXs contain heterogeneous activation of growth signaling mediators.

Table 2-2. Top 10 most frequent gene fusions identified by RNA-seq

Gene 1	Gene 2	No. of cases
AZGP1	GJC3	16
PDCD1LG2	KIAA1432	12
CDK11B	CDK11A	10
PPP1R3F	NUBPL	10
DUS3L	NRTN	6
TTLL12	TTLL12	6
ZNF732	ZNF876P	6
UFC1	DEDD	5
AMDHD1	MRPL15	4
CLPSL1	CLPS	4

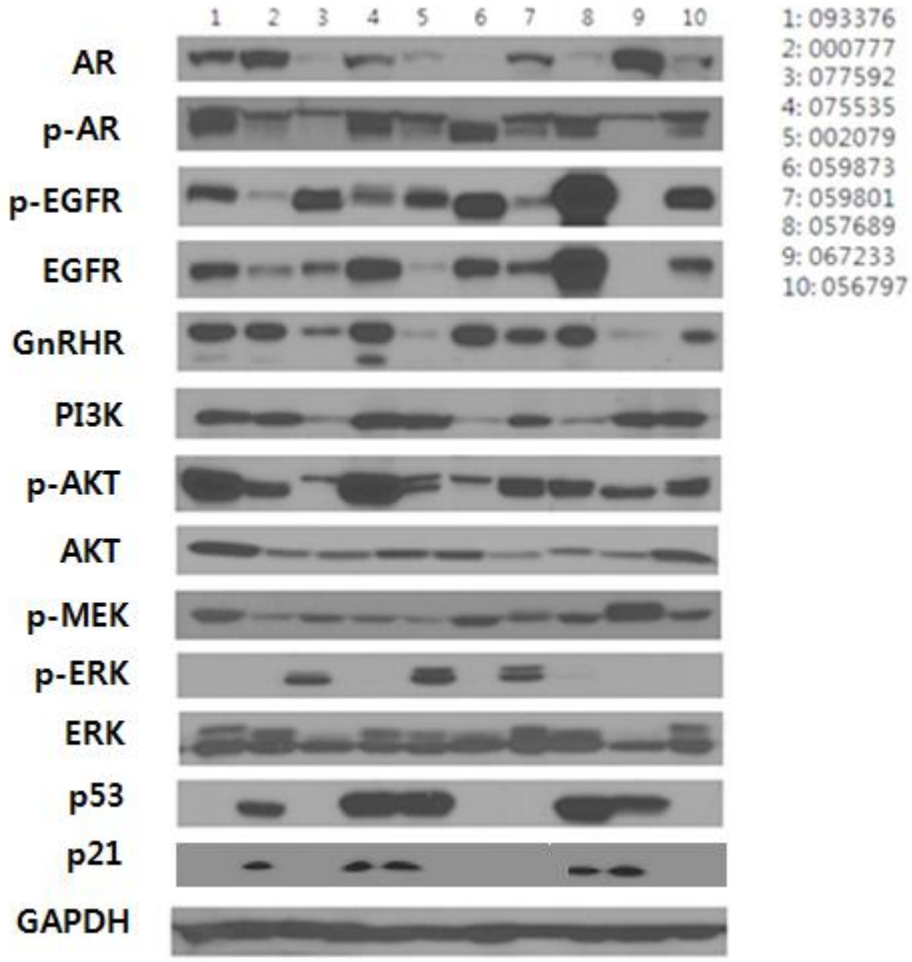


Fig. 2-6. Western blot analysis of key signaling proteins in 10 PDXs. AR; androgen receptor, p-AR; phospho-AR, p-MEK; phospho-MEK (mitogen-activated protein kinase kinase), PI3K; phosphatidyl inositol 3-kinase, EGFR; epidermal growth factor receptor, pEGFR; phospho-EGFR, GnRHR; gonadotropin-releasing hormone receptor, PI3K; Phosphoinositide-3 Kinase, AKT; Protein Kinase B (PKB), p-AKT; phospho-AKT, ERK: Extracellular signal–Regulated Kinases, p-ERK: phospho-ERK, p53; Transformation-Related Protein 53, GAPDH was served as a loading control.

Chapter 3. The first example of preclinical study using PDX model : The study of molecular mechanism for GnRH agonist in breast cancer PDX model

1) Preliminary study of GnRH agonist(zoladex)

Before the main study, we performed preliminary study with a small number of animals. Firstly we measured GnRHR level in 10 PDX (Fig. 3-1A) by Real-time PCR. Depending on sample, the expression level of GnRHR was various. Among them, we selected the highest (#75535), middle(#57689) and the lowest number (#87556) and constructed PDX using two mice, respectively (one for control and one for drug treatment). It was considered that the higher expression of GnRHR, the better the effect of GnRH agonist. When tumor size reached 150mm^3, we treated PBS(the control) or AC(doxorubicin and cyclophosphamide) with zoladex(GnRH agonist) every other day. In the group whose GnRHR level was the highest, tumor growth dramatically decreased, on the contrary, tumor grew continuously in the GnRHR-low group (Fig 3-1B).When tumor size increased enough, we harvest them and perform Annexin V/PI staining to investigate the rate of apoptosis. The results revealed that apoptosis was increased in GnRHR-high group than GnRHR-low group (Fig 3-1C). To find molecular mechanism for these results, we analyzed samples with western blotting. It demonstrated that B-Raf, p-ERK, oncogenic pathway, were significantly decreased in #75535 than other samples (Fig 3-1D). However, these protins were also changed in #87556, giving a room to find another pathway or mechanism.

2) The treatment of GnRH agonist shows profound antitumor effect on tumor growth

To certain the effect of GnRH agonist as a neoadjuvant, we chose #75535(GnRHR high) and #77592(GnRHR low) and implanted these samples into 10 NSG mice respectively. 10 mice were divided into 3 groups (Table3-1). Tumor growths were dramatically decreased in groups treated with AC regardless of zoladex comparing to control group (PBS). In the case of #75535, tumor volumes were significantly reduced more in the group treated with zoladex together than the group treated with only AC. However, in the group implanted #77592, there was no significant difference by administration of zoladex (Fig 3-2A). We also carried out western blotting analysis to examine molecular characteristic. The samples of #75535

showed that B-raf, p-ER,K, p-MEK were decreased and cleaved PARP was increased in AC with zoldadex group markedly(Fig 3-3A), whereas the difference was not distinct in the samples of #77592(Fig 3-3B)

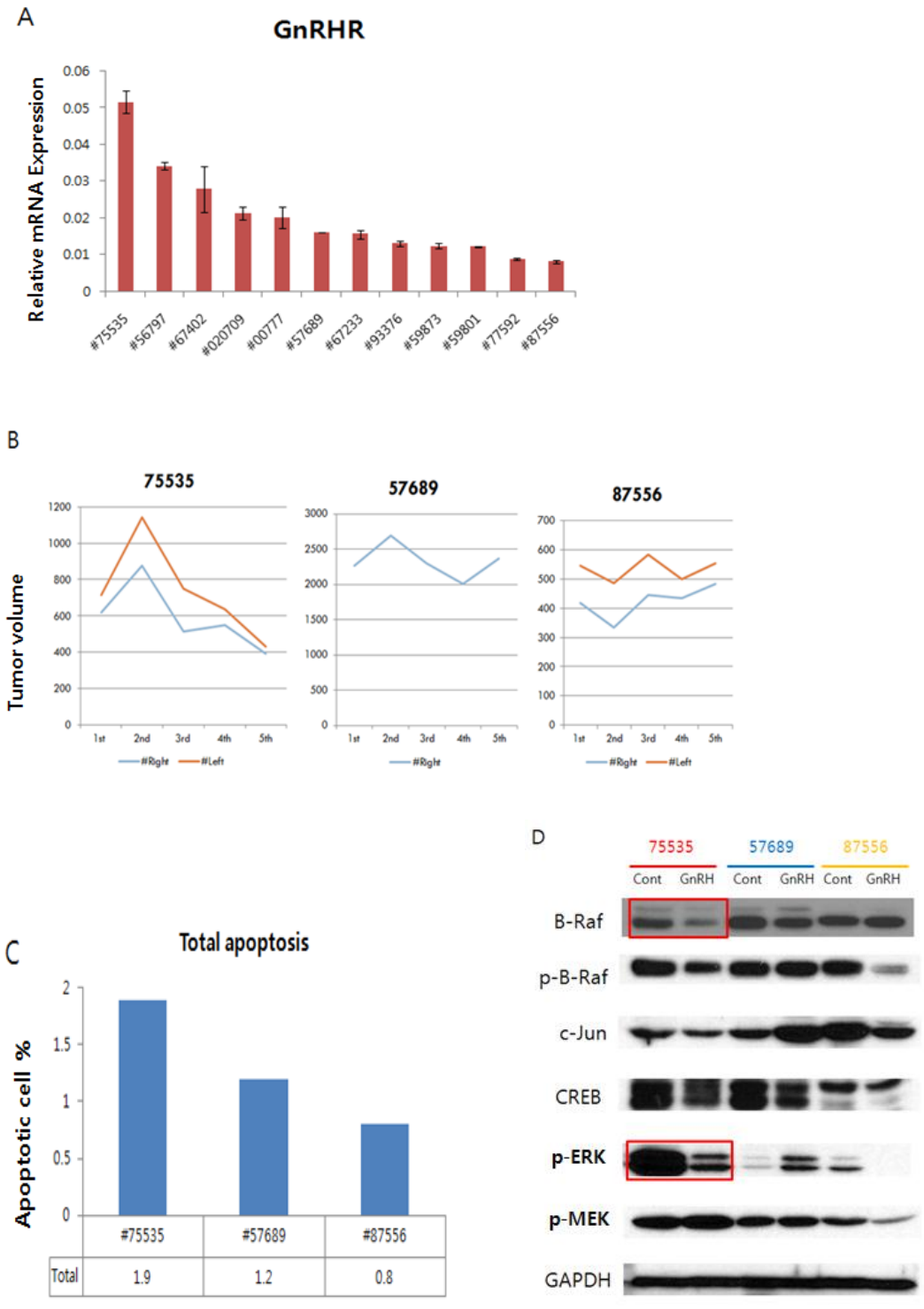


Figure 3-1. Preliminary study result of GnRH agonist (A)qPCR result of GnRH Receptor level in PDX samples (B) Tumor volumes after treatment of AC and zoladex (c)Annexin V/PI staining result showing the rate of apoptotic cells. (D) western blotting analysis of proteins targeted by GnRH agonist.

Table 3-1. The treatment group information for preclinical study

Group	Treatment drug	Drug Dose	Interval of administration	The number of Animals
Group 1	Normal saline(N/S)	-	2day	3
Group 2	Cyclophosphamide doxorubicin	150mg/kg 5gmg/kg	2day 4day	3
Group 3	Cyclophosphamide doxorubicin zoladex	150mg/kg 5mg/kg 8mg/kg	2day 4day 2day	4

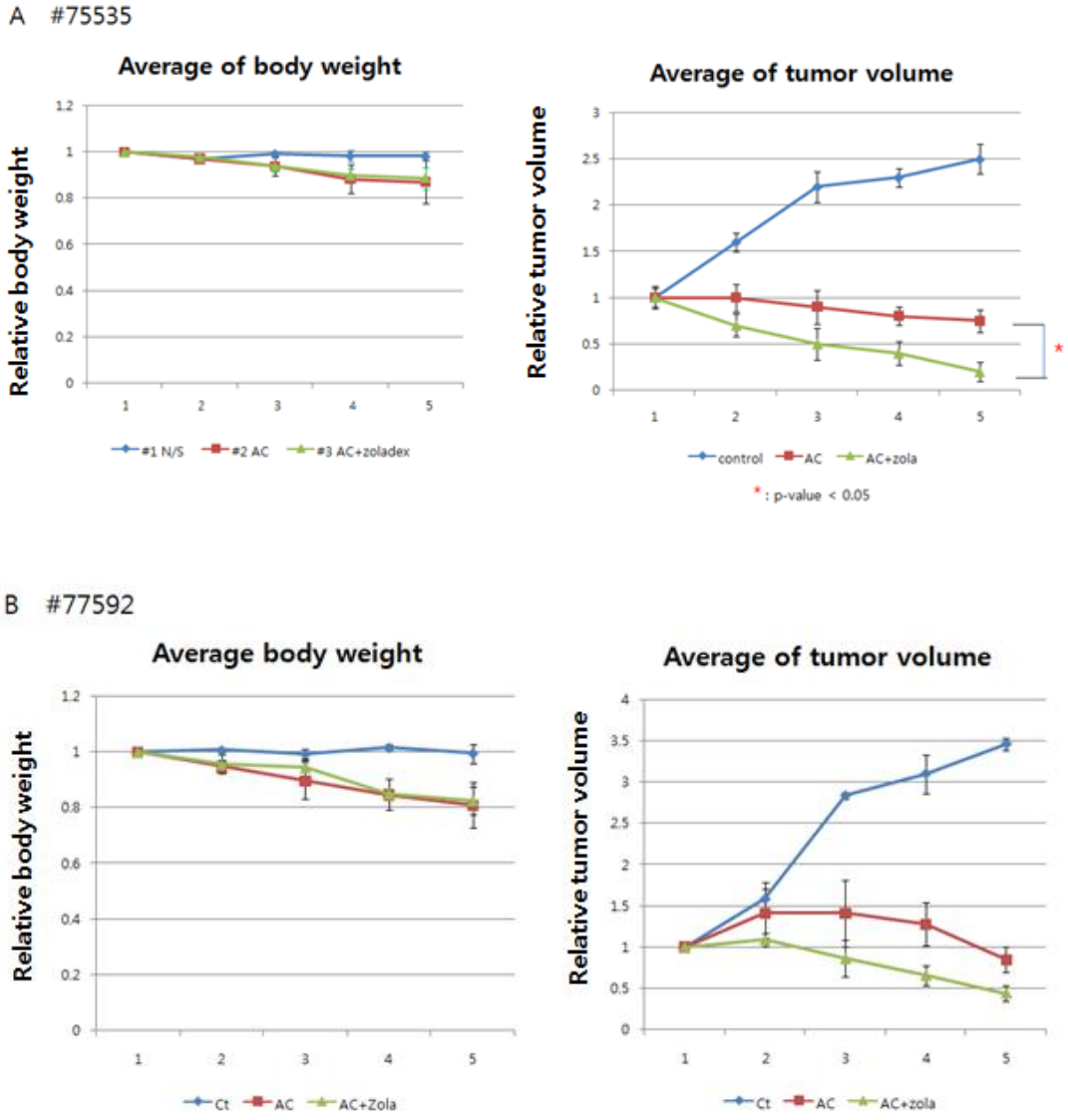


Figure 3-2. Average body weight and tumor volume of PDX mice at each time point. (A)Average of body weight and tumor volume in mice implanted #75535 sample. (B)Average of body weight and tumor volume in mice implanted #77592 sample. In the case of #75535, tumor volumes were significantly reduced more in the group treated with zoladex together than the group treated with only AC. However, in the group implanted #77592, there was no significant difference by administration of zoladex

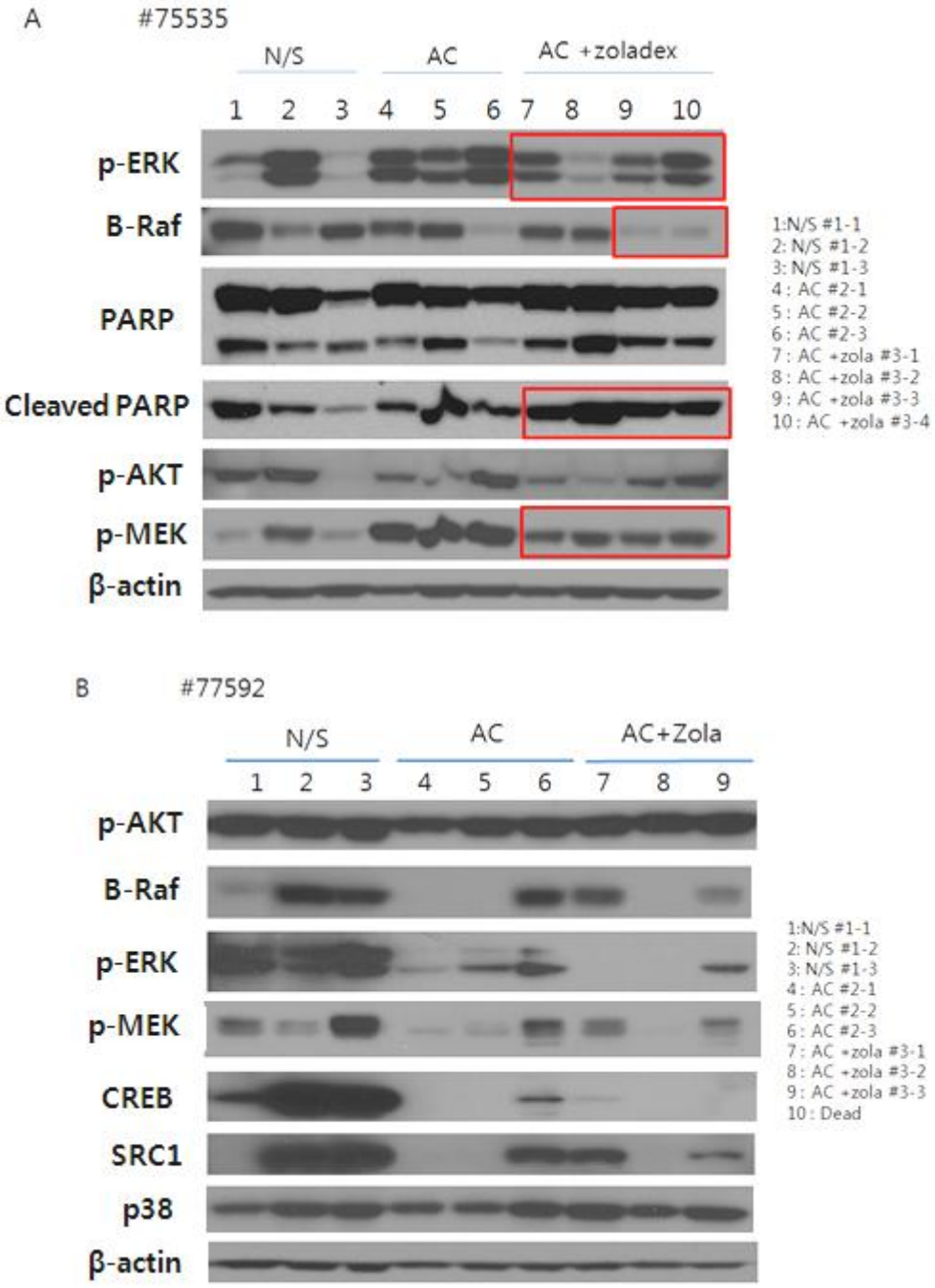


Figure 3-3. Western blotting results of tumor harvested from PDX mice. (A) western blotting analysis of #75535 tumor (B) western blotting analysis of #77592 tumor It showed that B-Raf, p-ERK, p-MEK were decreased and cleaved PARP was increased in AC with zoladex group markedly(Fig 3-3A), whereas the difference was not distinct in the samples of #77592 (Fig 3-3B)

Chapter 4. The second example of preclinical study using PDX model : A crosstalk between miR-155 and YAP in breast cancer

1) RPPA analysis with 30 human TNBC samples

We obtained 80 human TNBC samples from Asan Medical Center and measured the expression of miR-155 in each sample. Among them, we chose 15 samples whose miR-155 expression were high and 15 samples whose miR-155 expression were low (Fig 4-1A). Therefore, we divided these samples into 2 groups (miR-155 high and miR-155 low) and conducted reverse phase protein array (RPPA). The results of RPPA analysis showed that the expressions of 10 proteins increased and the expression of 4 proteins were lowered significantly as miR-155 level increased(Fig 4-1B, 4-1C). We focused that phosphorylated YAP(S127) was the most significantly down-regulated in miR-155 high group, which lead us to investigate the relationship between YAP and miR-155 in breast cancer.

2) miR-155 regulates the level of phosphor-YAP

Firstly, to confirm RPPA study result, we conducted western blot in PDX samples. In 3 samples whose expression of miR-155 was high, the expression of pYAP was markedly down-regulated, while miR-155 low samples expressed pYAP high (Fig 4-2A). Then, we made stable cells in which the expression of miR-155 is inhibited (called mirZIP cell), using PDC (patient-derived cancer cell) (Fig 4-2B) and several cancer cell lines including MDA-MB-231 and MDA-MB-436. PDC was cultured from the tumor tissue of PDX mice in which human tissue was implanted. Especially #93376 and #67233 cells, among these cells, we found that inhibition of miR-155 elevated the expression of pYAP, but total YAP level was not largely different between samples (Fig 4-2C). Furthermore, we performed western blot analysis of Hippo pathway proteins (MST, pMST, LATS1, pLATS1, Sav1, pMOB1) with these cells. The results indicated that the expression level of pLATS was increased in #93376, #57689 and #67233 miRZIP cells as compared with each control cell, however MST and pMST and LATS did not differ or rather decresed in their expression(Fig 4-2D)(See discussion). In MDA-MB-231 and MDA-MB-436, pLATS and pYAP were increased in

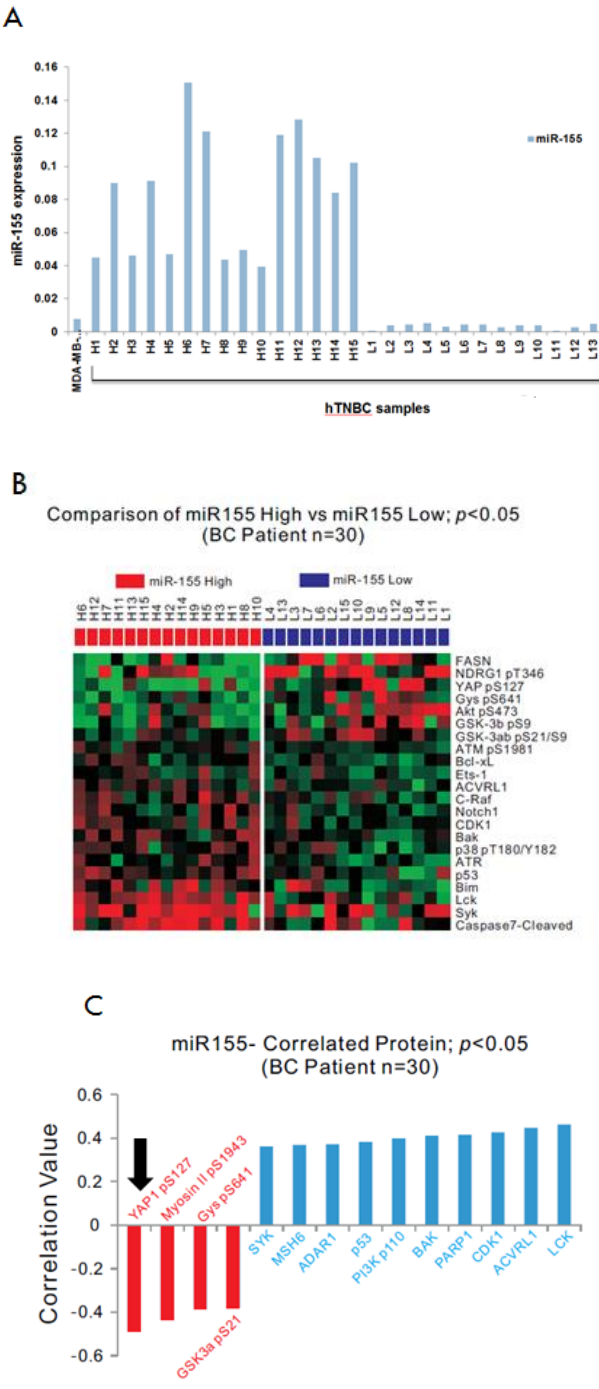
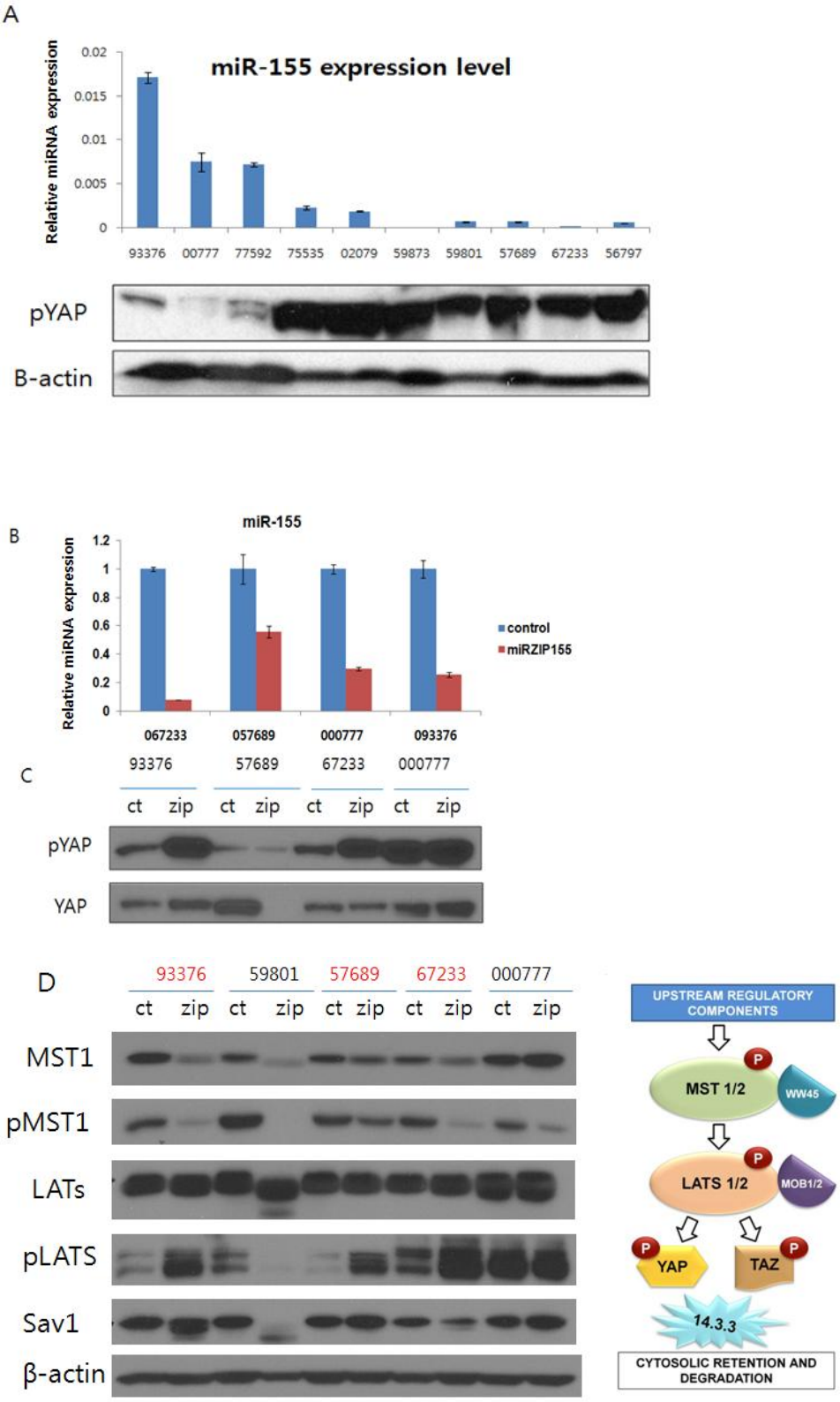


Figure 4-1. RPPA results from 30 TNBC samples. (A) qPCR results of miR-155 in 30 TNBC human samples(By Sinae Kim) (B) Heatmap representing RPPA results in 2 group(miR-155 high and miR-155 low group) (c) miR-155 correlated proteins whose p-value is lower than 0.05 (By Yun-yong Park)



Reference : (71)

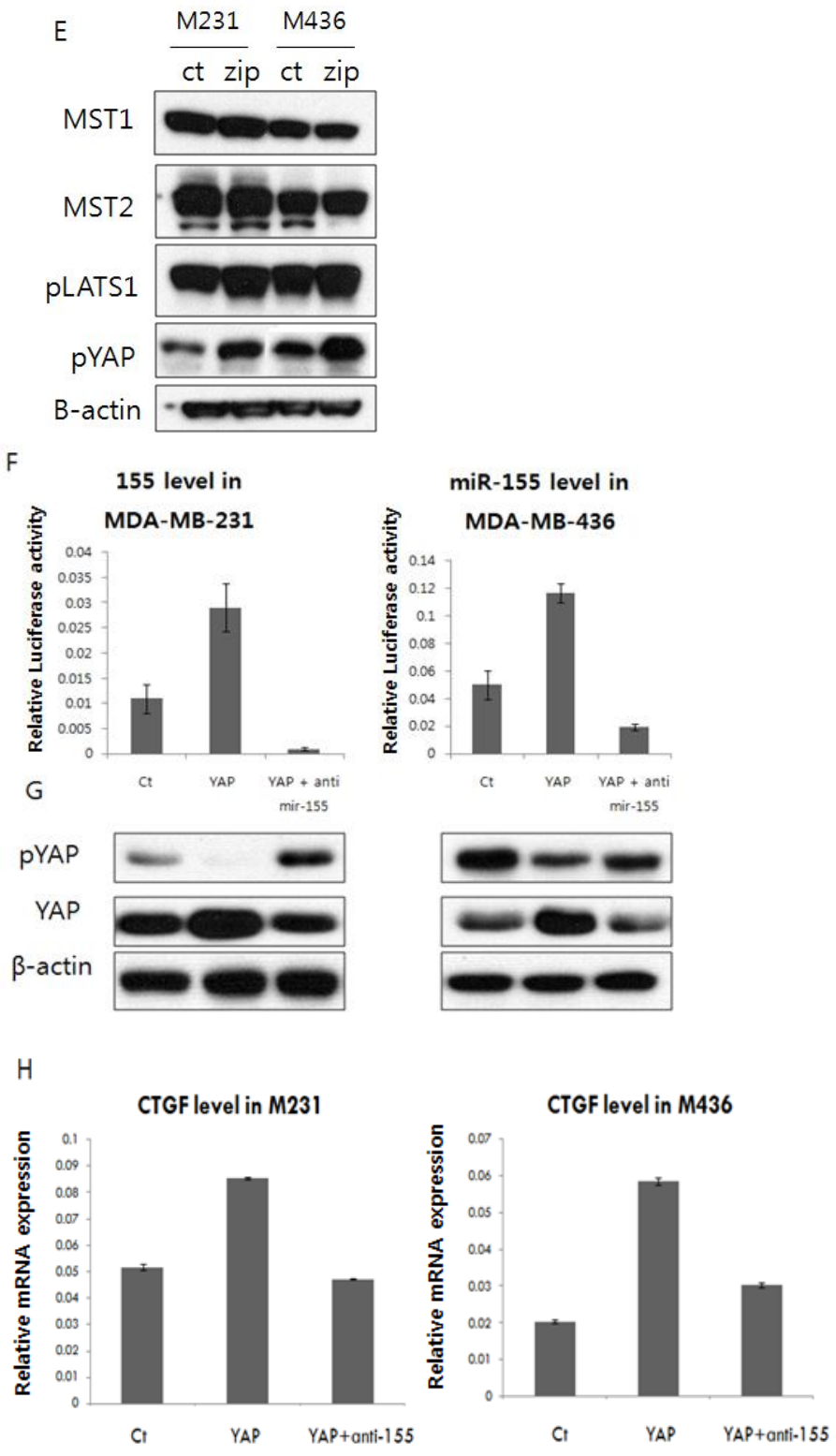


Figure 4-2 western blotting analysis and qPCR results showing miR-155 regulates the level of pYAP (A) Western blotting analysis in PDX sample (B) The expression level of miR-155 in miRZIP cells comparing to the control. (C) Western blotting analysis of pYAP in

miRZIP cell and the control cell. (D) The expression level of several proteins related to hippo pathway in PDC. (E) The expression level of several proteins related to hippo pathway in MDA-MB-231 and MDA-MB-436. (F) qPCR results showing expression level of miR-155 (G) western blotting result of pYAP and YAP in (F) samples. (H) qPCR results of CTGF as YAP target gene.

in miRZIP cells, but other proteins in hippo pathway were not changed (Fig 4-2E). Also, we knock-downed miR-155 with siRNA and investigated the change of pYAP by western blotting.

Before western blotting, we checked whether miR-155 is knock-downed well by quantitative PCR and consequence showed that miR-155 was knock-downed very much(Fig 4-2F). Also, YAP was overexpressed by transfection of YAP plasmid and siRNA of miR-155 was co-transfected with YAP in another cell. Interestingly, overexpression of YAP reduced pYAP level, whereas this effect was compromised by siRNA of miR-155(Fig 4-2G). We also checked the expression of CTGF, a representative YAP target gene, in these cells by conducting qPCR, which confirmed that YAP target genes were also increased as YAP is overexpressed. However, the treatment of miR-155 eliminated this effect (Fig 4-2H). All results of these experiment showed that miR-155 regulates level of pYAP plus its target gene.

3) SVIL is a candidate mediator for the regulation of pYAP by miR-155.

Firstly, we researched literature and information about hippo pathway to find the mediator that regulates pYAP by miR-155. The proteins which are related to hippo pathway were researched whether it can be a target of miR-155 or not using microrna.org website. Kibra, MST1, MST2, TAOK3, HIPK3, MMP5, PATJ, CSGN and MAP4K3 were our candidates and their expression level was screened in control and miRZIP cells.(Fig4-3A) If it is target gene of miR-155, their expression level should be increased in miRZIP cells because miR-155 is knock-downed in miRZIP cells. However, any candidate protein mentioned above did not show expected changes. Therefore, FPKM values from RNA-Seq was used to select new candidates by comparing FPKM values between miR-155 high PDX samples and miR-155 low PDX samples. The genes having different expression value significantly (p-value<0.05) between two groups were listed and the possibility of being target of miR-155 was calculated with their 3'UTR sequence.(miRSVR score from microRNA.org). By selecting genes whose miRSVR scores is lower than -0.5, we could narrowed down to 5 genes from approximately 250 genes(Table 4-1). The new candidates were HIPK3, RAC1, BTBD3, SVIL and UBE2D3. Although HIPK3 is not considered miR-155 at microRNA.org site, it was included as a candidate because it had the lowest p-value between 2 groups. It is well known that YAP is

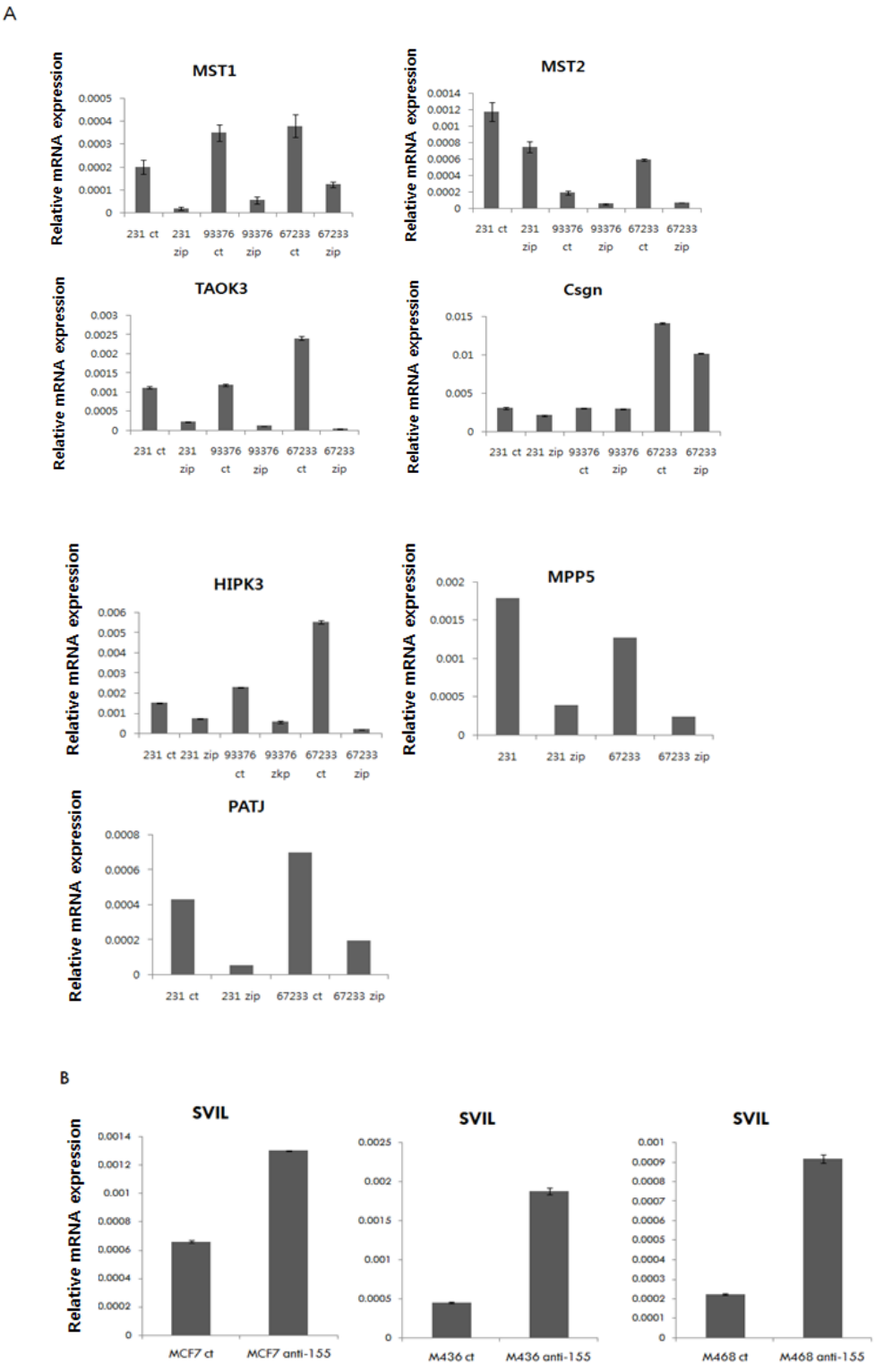


Figure 4-3. qPCR results of candidate gene of the target by miR- 155 (A) qPCR results of 7 proteins in miRZIP cell and the control cell. (B) qPCR results of SVIL expression level after treatment of antagomir of miR-155

Table 4-1. 5 proteins selected by FPKM value from RNA seq and mirSVR value at microrna.org website

gene	P-Value	mirSVR
HIPK3	0.000143	Not target
RAC1	0.000473	-1.1519
BTBD3	0.006831	-0.9091
SVIL	0.017411	-0.6034
UBE2D3	0.01856	-0.8281

important to cell-cell adhesion and cell junction. Among those 5 genes, only RAC1 and SVIL are associated with cytoskeleton and mechanotransduction of cells. SVIL is a gelsolin family of actin-binding proteins, which cap, nucleate, and/or sever actin filaments[70]. Therefore, we evaluated SVIL expression in control and miR-ZIP cell and found the expression was elevated in miRZIP cells (Fig 4-3B), suggesting that SVIL could be a target of miR-155 and mediator to decrease pYAP by miR-155. The hypothesis is that elevated miR-155 inhibit SVIL and this make actin filament stable in the cells, which leads to activate YAP and decrease pYAP in cytoplasm[102]. This possibility can be examined by inhibiting synthesis of actin filament by drug in the next experiment.

4) YAP regulates the expression of miR-155 level directly

While miR-155 regulates phosphorylation of YAP, conversely, it can be possible that YAP also regulate the expression of miR-155. Therefore, we conducted luciferase promoter assay with luciferase vector (pGL3 Enhancer vector) including miR-155 promoter region. Luciferase expression level indicates the amount of the expression of miR-155 indirectly. To observe the effect of YAP on miR-155, luciferase value was measured after overexpression of YAP. It showed luciferase value increased by YAP overexpression, indicating YAP raised the expression of miR-155(Fig4-4A). In addition, luciferase value was decreased when YAP inhibitor, verteporfin, was treated to cells(Fig4-4B). Also, we down-regulated YAP by siRNA in MDA-MB-231 and HS578T cells and performed real-time PCR, which presented the expression of miR-155 dropped in both cells(Fig 4-4C). To see larger effect, we treated both siYAP and siTAZ because TAZ is close paralog of YAP and hippo pathway effector. The treatment of siYAP and siTAZ decreased miR-155 more than treatment of only siYAP in cancer cell line and PDC(Fig4-4D).

To assess whether YAP regulates the expression of miR-155 directly, we searched TEAD binding sequence or similar sequence in the promoter region of miR-155. There were four sites considered that TEAD could bind. Consequently, we performed CHIP assay and conducted quantitative PCR for those sites. Because YAP plasmid contains Flag tag, we immunoprecipitated with Flag Bead and compared with IgG after conducting PCR.

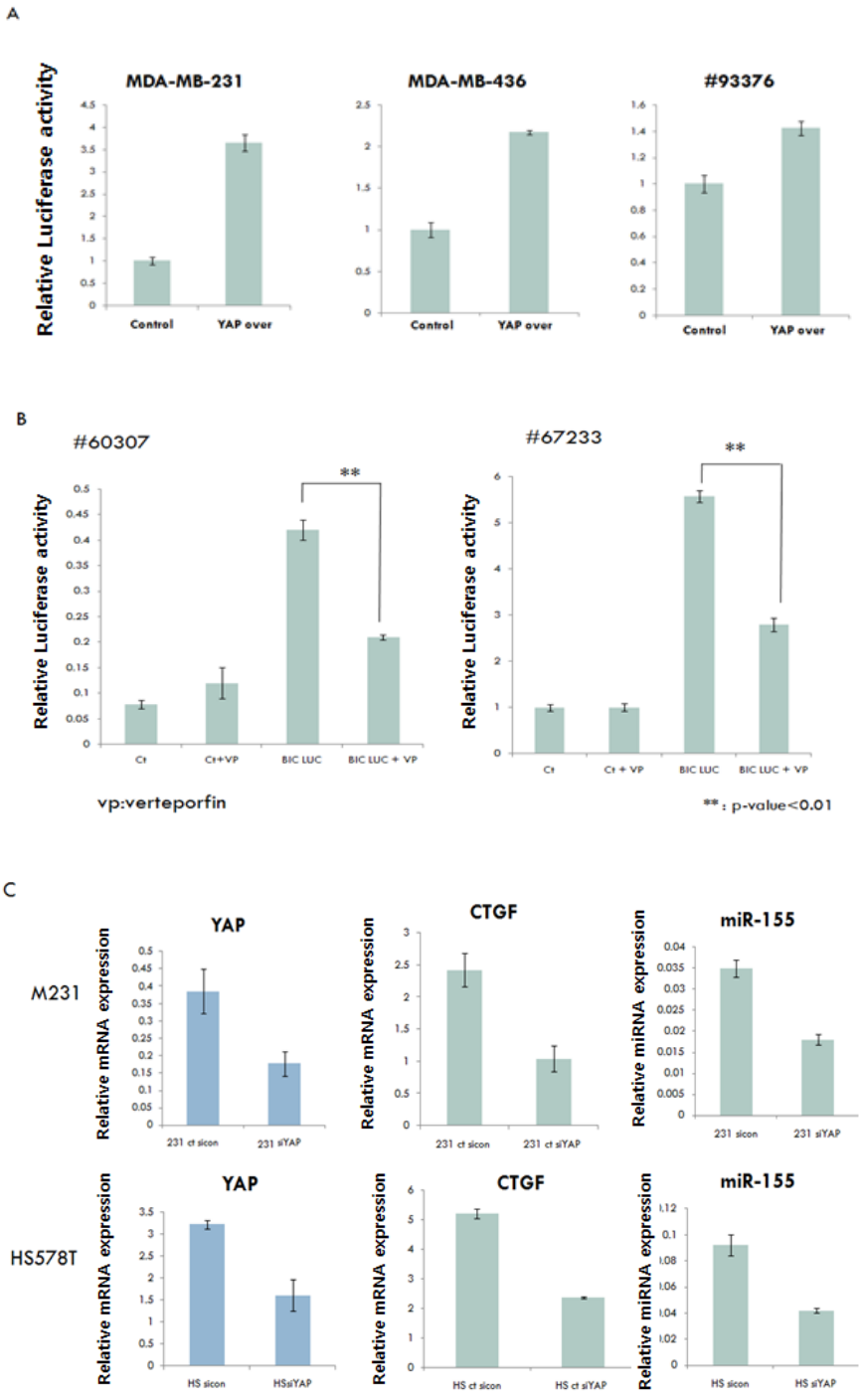
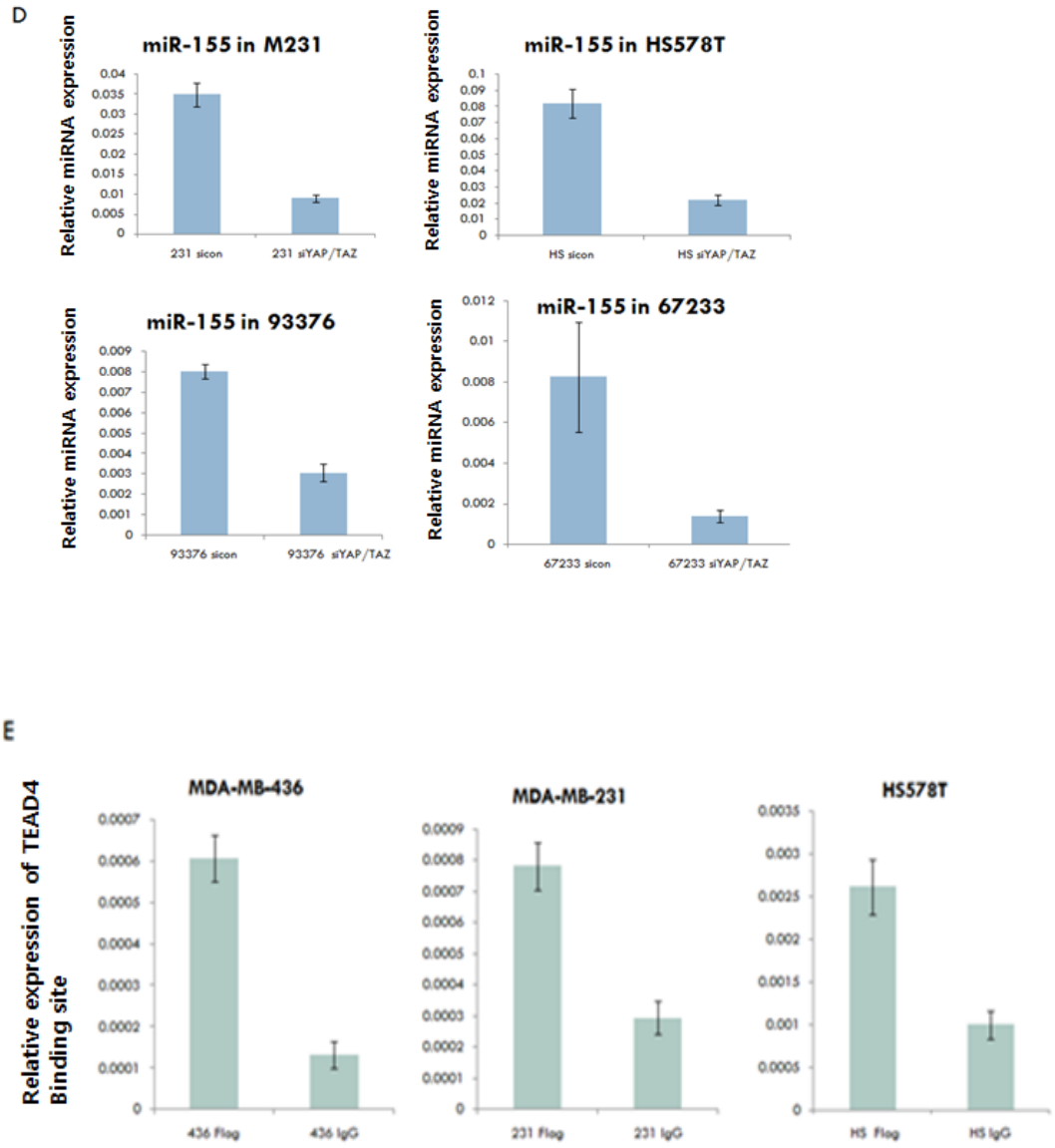
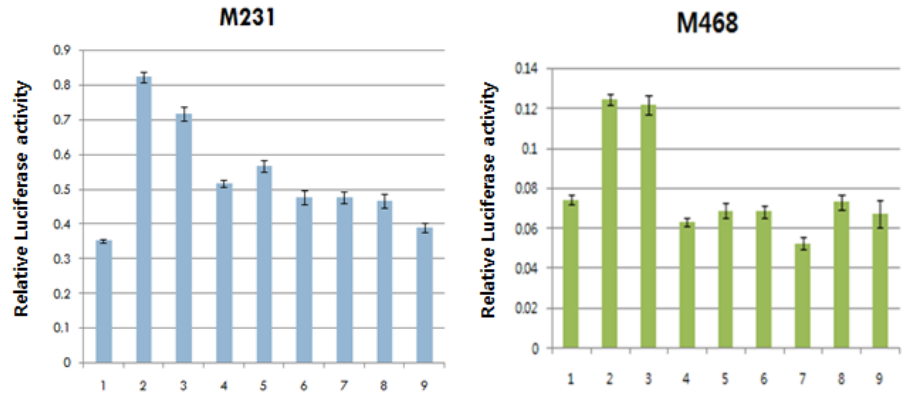


Fig 4-4 Luciferase promoter assay and qPCR results showing that YAP regulates the expression of miR-155 level directly (A) Luciferase promoter assay of miR-155 when YAP is overexpressed. (B) Luciferase promoter assay of miR-155 after YAP overexpression and inhibition by verteporfin. (C) qPCR results of YAP, CTGF and miR-155 after knock-down of YAP



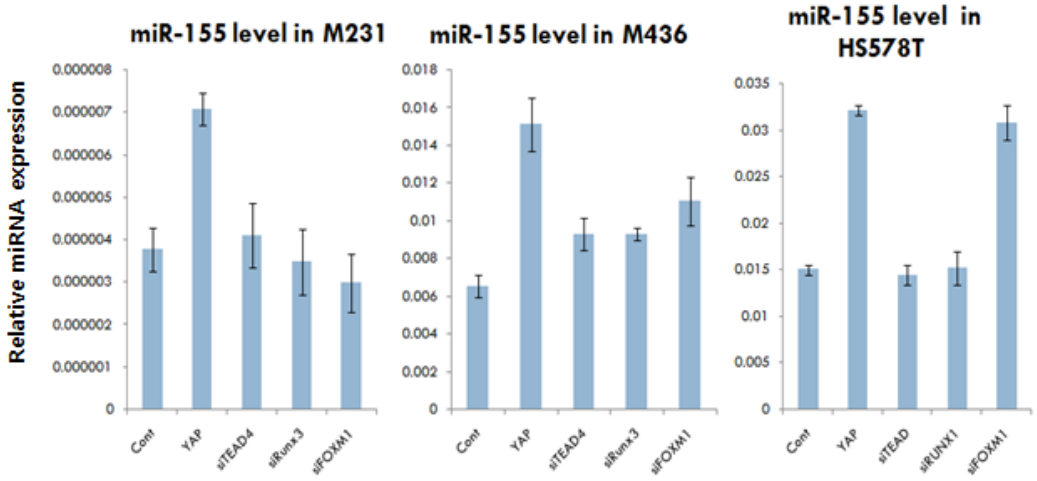
(D) qPCR results of miR-155 after knockdown of YAP and TAZ by siRNA (E) PCR result of TEAD4 binding site after CHIP assay with Flag antibody

F



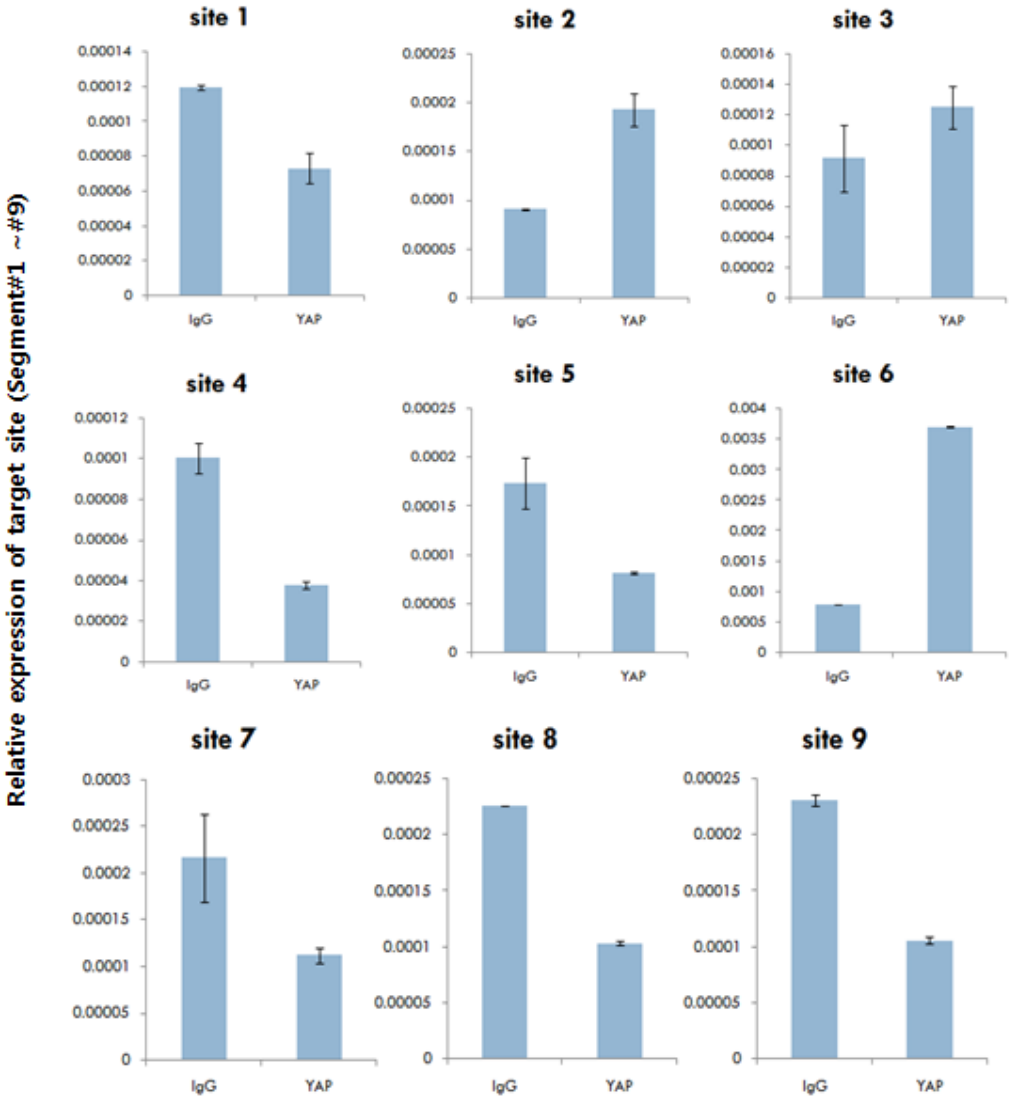
1 : BIC Luc
2 : BIC Luc + YAP
3 : BIC Luc + YAP + VP
4 : BIC Luc + YAP + siRunx3
5: BIC Luc + YAP + siFOXM1
6: BIC Luc + YAP + VP + siRunx3
7 : BIC Luc + YAP + VP + siFOXM1
8 : BIC Luc + YAP + siRunx3 + siFOXM1
9 : BIC Luc + YAP + VP + siRunx3 + siFOXM1

G



(F) Luciferase promoter assay results of miR-155 after several siRNA treatment with or without VP(Verteporfin). siRNA were siRUNX3, FOXM1 which could bind to the promoter region of miR-155. (G) qPCR results of miR-155 after siRNA treatment

H



(H) PCR results after CHIP assay to find binding sites of transcription factor on the promoter of miR-155. The promoter region of miR-155(2kb) was segmented into 9 sites and CHIP assay was performed for each site.

Among 4 sites, only one site was amplified by PCR compared to IgG(Fig4-4E). This site has TEAD4 binding sequence. However, there remained possibility that YAP binds another transcription factor instead of TEAD, such as FOXM1 and RUNX3, and regulates miR-155 by them. Interestingly, there are many binding sites of FOXM1 and RUNX3 in the promoter region of miR-155. Therefore we checked if the expression of miR-155 is affected by FOXM1 and RUNX3. After treatment of siFOXM1 or siRUNX3, we measured mir-155 and discovered that both siFOXM1 and siRUNX3 also lowered expression of mir-155(Fig4-4F and Fig4-4G).

In summary, it is considered that FOXM1, RUNX3 and TEAD4 regulate the expression of miR-155 altogether with YAP. To find exact binding site of these factor, we segmented the 2kb upstream of mir-155 gene into 9 sites, and performed ChIP and qPCR with each primer. (Supplementary Table 5-2) Among them, the 6^{th} site was amplified by PCR, indicating that it is binding site of those factors (Fig 4-4H).

5) The relationship between miR-155 and YAP is maintained in physiological condition

To confirm the relationship between miR-155 and YAP, we established diverse physiological condition and observed the change of the level of miR-155 and YAP. According to the literature, serum starvation induces phosphorylation of YAP and inhibits YAP signaling(72). Therefore, we cultured cells with or without serum and evaluated pYAP(S127) by western blotting(Fig 4-5A). pYAP level was increased in the cells growing without serum. Moreover, whether mir-155 was also changed by serum or not was examined by luciferase assay. As a result, serum starvation induced a reduction of luciferase value, indicating the decrease in miR-155 expression(Fig 4-5B). However, overexpression of YAP recovered this reduction, meaning that YAP stimulates the expression of miR-155 Based on these outcomes, we concluded that relationship between YAP and miR-155 is maintained in physiological condition.

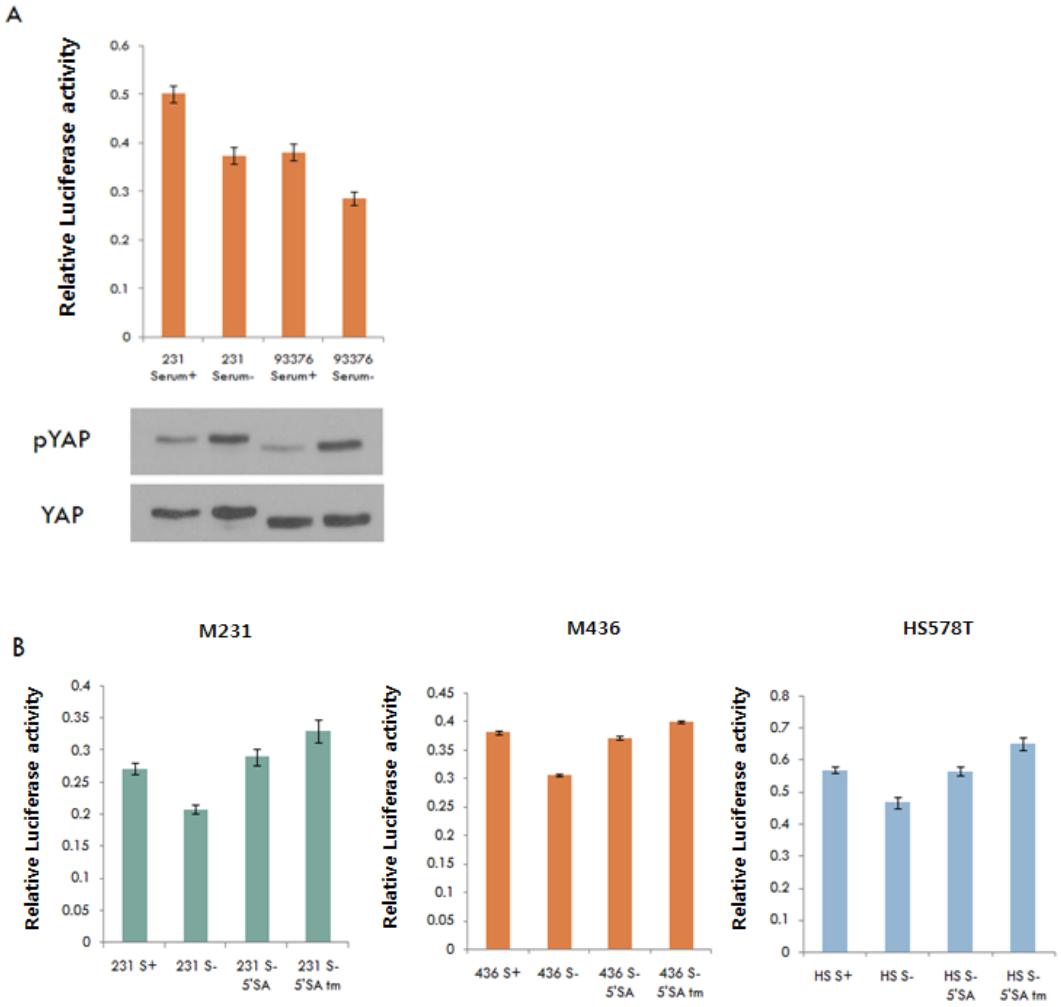


Figure 4-5. Luciferase promoter assay of miR-155, western blotting of pYAP in serum depleted media shows the relationship between YAP and miR-155 in physiological condition. (A) luciferase promoter assay of miR-155 in cells cultured in serum +/- media and western blotting results of pYAP and YAP in those samples. (B) luciferase promoter assay of miR-155 after overexpression of YAP in serum free media. Serum starvation induced a reduction of luciferase value, indicating the decrease in miR-155 expression, however, overexpression of YAP recovered this reduction.

DICUSSION

1. Firstly, we successfully produced 20 PDXs, representing a 72.4% success rate. According to the recent report, the success rate of pancreatic cancer PDX is about 40-50%[102]. Therefore, our success rate is very high comparing to previous reports. The reasons are thought to have minute differences in our procedure. For example, pathologist chooses the most aggressive part of cancer and it is kept freshness until surgery. Also appropriate size of fragment and technique to implant into mice might affect success rate.

Xenograft transplantation of human PDAC cells or tissues was first performed in the late 1990s[73], with subsequent studies reporting a high degree of similarity between the PDXs and primary cancer cells, and passage dependent genetic changes[15]. A recent study using 96 PDAC patient samples estimated the frequency of mutations in a panel of 22 cancer predisposition genes, which led to the identification of 14 pathogenic mutations in 13 patients (13.5%) [74]. Other studies on pancreatic cancer xenografts have analyzed gene expression and/or copy number variations, but have discovered only small numbers of genetic variants[75, 76]. Therefore, our report provides more information about the potentially deleterious variants to pancreatic cancer research field. In our SNP analysis, we initially found 762 deleterious (as predicted by SIFT and Polyphen2) variants in HPDE. Since we included HPDE as a noncancer cell control, we subtracted these variants from the total variants to obtain the number of cancer-specific deleterious variants. However, we cannot exclude the possibility that this subtraction might have missed some variants that is functional in cancer cells. The list of topranked genes (which shows frequent SNPs in multiple PDXs) included a number of promising candidates for functional analysis. For example, LAMB3, the topranked gene, produces lamininb3, which is one of the major components of the extracellular matrix (ECM) of pancreatic cancer[77], and these variants generate diverse types of missense mutations, whose function needs to be further analyzed. By contrast, CD101, which has been reported as a potential risk-associated variant for PDAC[78], plays a role as an inhibitor of CD3-induced T-cell proliferation[79], and so the variants of this gene may have the immuno-modulatory effect on cancer cells Further molecular study will reveal the exact function of these variants in pancreatic cancer.

Because different number of samples were analyzed in SNP array and CCP (18 PDXs in SNP, 8 primary tumor-PDX pairs in CCP), a direct comparison of the clustering result from the two analysis was not possible. However, we were able to compare several samples analyzed in both platforms. For example, we could see the #5/#18 and #6/#13 pairs are closely related in the Group 2 of SNP data (Figure 1-3A) but only #5/#18 pair is closely related in the CCP analysis (Figure 1-4A). Therefore, we think the results of the two techniques are only partly matching. The possible reason of this result might be the difference of analytical platform. Specifically, the SNP array covers about 20,000 exome SNPs throughout the genome but the CCP covers only around 400 cancer genes. Taken together, our findings indicate that the PDX model can provide a faithful representation of patient tumors. For future experiment, it needs to follow specific SNP or variants that seem to be important and research their functional role in pancreatic cancer. Furthermore, these PDXs retained the heterologous nature of pancreatic cancer cells, enabling us to use this model for preclinical research, as well as the basic study of this disease.

2. our previous study of a pancreatic cancer PDX, we found that tumor size is a significant factor of the success of a PDX[6)]. By contrast, in the study of TNBC PDX, we found that the N stage is a clinical factor that affects PDX success. Considering that the TNBC PDXs involved an orthotopic transplant (i.e., a transplant into a mammary fat pad), whereas pancreatic PDX was subcutaneous, the above discrepancy may be caused by the difference in the graft sites. Or simply the difference in the tumor type can explain it. Indeed, in case of non–small cell lung carcinoma, disease-free survival was found to be a significant factor for the graft success[80)]. According to previously reported data, the success rate of orthotopic breast cancer PDX was between 19 and 46 percent [103)]. However, our success rate was 56.4%, which is higher success rate than other reports. The main reasons making higher success rate are thought that we selected sample from under 40 years old woman or from patient having BRCA1 mutation. Also, pathologist selected the part looked like the msot cancerous and aggressive.

Although we found that most of the PDX tumors showed histological characteristics like those of their original primary tumor, we found some cases where there was no match.

For example, Supplementary Figure 2-1F (70212) shows an ER-positive primary tumor, but its PDX was ER negative. Likewise, Supplementary Figure 2-1R (58220) shows a HER2-positive primary tumor, but its PDX has lost the HER2 positivity. These observations can be largely explained by intratumoral heterogeneity and selection during the PDX formation[81]. This kinds of mismatch could be overcome by implanting tissues from the same primary tumor to more than one mouse and analyzing multiple samples.

The c.802-3dupT variant in PTEN is predicted to be benign (ClinVar Miner website) whereas the c.2719A>T variant in the COL1A1 gene was not reported previously. A recent paper showed that the c.3746T>C missense variant has a frequency of ~9% among Korean patients[82]. Considering that c.2719A>T creates a premature stop codon, the impact of the variant needs to be studied further.

Also, in these CCP data of TNBC samples, we observed that variants within 10 primary tumor–PDX pairs had a high rate of matching, ranging from 85% to 96.9% of identity, However, there were also mismatched genes between primary and PDX tissue. Although mice have high degree of genetic similarity to human beings, murine tumors differ from human tumors at both genetic and molecular levels. Especially, genomic and transcriptional variations can be induced by murine microenvironment selection[83]. When primary tumor is grafted, the rapid replacement with murine stroma[84] may results in changes to paracrine regulation of the tumor and its biological properties[85], due to species-specific differences in ligands and receptors. In our study, immunohistochemistry revealed that the receptor status was changed in two samples. Although primary tumors were ER positive, tumors grown in mice were ER negative. It is considered that the loss of ER led to get more aggressiveness and make benefit to tumor growth because absence of receptor for hormone in TNBC can make tumor more aggressive. Also, we compared mismatched genes between primary and PDX tumors. The most different genes were BRAF, FGFR3a and MYH11(Supplementary Table 2-8). BRAF is known as oncogene which phosphorylate MAP2K1, and thereby contrigutes to the MAP kinase signal transduction pathway[86]. FGFR3 is tyrosine-protein kinase that acts as cell surface receptor for fibroblast growth factor[87] and plays an essential role in the regulation of cell proliferation, differentiation and apoptosis[88]. Especially, variants of FGFR3 and MYH11 were not found in primary tissue, on the contrary, variants

were found in all PDX tumor tissue. It is possible that these changes make primary tumor tissue to develop easily in mice by affecting tumor microenvironment. Also, according to the analysis by String (https://string-db.org/) (Supplementary Fig 2-2D), three proteins have strong interactions with HRAS, NRAS and ACTA2. These interaction could provide selection pressure because HRAS and NRAS are well known as oncogene, which can stimulate oncogenic pathway in tumors.

Normally, Notch1 signaling positively contributes to proliferation and survival of cells[89], and many reports lead to the same conclusion in breast cancer as well[90-93]. The Notch pathway is an evolutionarily conserved signaling pathway that regulates stem cell maintenance, cell fate specification, differentiation, proliferation, motility, and survival. In mammalian signal-sending cells, the Notch pathway consists of ligands (Delta-like proteins 1/3/4, Jagged1/2) and receptors (Notch1/2/3/4)[65]. Once bound to its ligand on cells, the Notch receptor is activated, undergoing critical proteolytic steps[84]. In this process, ligand-receptor binding triggers a cleavage of the Notch receptor followed by the gamma-secretase complex. Then, the NICD is free to attach itself to specific DNA-binding proteins located in the nucleus, where it associates with the CSL (CBF1/Su(H)/Lag-1) transcription factor complex, resulting in subsequent activation of the canonical Notch target genes including myc, p21, and the HES-family members[48]. It is known that Notch receptors and ligands were all found to be highly and widely expressed in human breast cancers, compared with normal breast tissues from the margin of tumor section, leading to abnormal growth of breast cancer cells. However, recent researches reveal that Notch can in fact be either oncogenic or tumour suppressive depending on the tissue and cellular context[19]. Notch signalling promotes expression of HES proteins that subsequently bind to target loci and prevent their transcription. Thus, the effect of Notch signaling on gene transcription can be largely repressive as a consequence of HES activity. Transcriptional repression by Notch may also occur by mechanisms that are independent of HES and HEY factors. For example, the ICN has been shown to repress differentiation genes in neural progenitors by both HES-dependent and independent mechanisms[86]. The other target genes of Notch1 are cyclin D1, p21, p27, NF-kB and so on, which regulates cell cycle, proliferation, and transcription. Especially, p21 and p27 are CDK inhibitors, abnormal cell proliferation could occur if their expressions are

low.

The mechanism of Notch1 as tumor suppressor can be explained and summarized by two mechanisms which are cell-autonomous effect and non-cell-autonomous tumor-suppressor activity[19]. The capacity of Notch1 to operate as a tumor suppressor may be a consequence of its cell-intrinsic role in promoting cell cycle exit and differentiation, thus eliminating potential cancer stem cells and tumor-initiating cells[68, 94]. Notch1 prevents the expansion of premalignant stem and progenitor cells that may be carrying oncogenic mutations by promoting their differentiation. However, loss of Notch signaling impairs differentiation and enables the expansion of mutant stem or progenitor cells. Those cells can subsequently acquire secondary oncogenic hits and initiate tumor development. This process is related to pro-differentiation factors such as retinoic acid (RA) signaling[63] and interferon regulatory factor 6 (IRF6)[95]. In addition, as mentioned above, induction of HES and HEY transcriptional repressors may further promote differentiation by preventing the expression of factors that inhibit differentiation or promote cell cycle progression. Notch may also induce differentiation by promoting detachment of stem and/or progenitor cells from the basement membrane via inhibition of Rho-associated kinase 1 (ROCK1), ROCK2 and myotonic dystrophy kinase-related CDC42-binding kinase α (MRCKα)[96]. Notch also actively represses proliferation by promoting the expression of cell cycle inhibitors such as p21 and by antagonizing p63, which promotes the proliferation of immature stem and progenitor cells[19].

The second mechanism is that Notch suppresses tumor development is prevention of the induction of a pro-tumorigenic stroma. Premalignant cells that carry oncogenic mutations fail to initiate tumorigenesis in the absence of a pro-tumorigenic stroma, and intact Notch signaling prevents stromal remodeling by negatively regulating processes such as inflammation[19]. Following loss of Notch signaling, the negative regulation of the inflammatory response is impaired, and inflammation can be initiated by oncogene activity or injury[69, 97]. This results in the induction of a pro-tumorigenic stroma that can foster tumor growth by activating signaling cascades such as the WNT–β-catenin pathway[63]. As an example of Notch1 as tumor suppressor a recent report indicates that weak inhibition of Notch1 signaling increases MDA-MB-231 cell invasion, suggesting that NOTCH1 variants

play a role in metastasis[98]. On the other hand, a loss of function of NOTCH1 has been found in squamous cell carcinoma, pointing to a tumor-suppressive function of NOTCH1[86].

Therefore, the stopgain mutation of NOTCH1 can be oncogenic by itself. Notch1 variants found in our data could be one of the reasons which tumor initiated and proliferated in PDX model. Also, this variant could help cancer growth partly by several mechanisms which are explained above. Those samples having Notch1 mutation can be useful tool for therapeutic agent experiment targeting Notch1 pathway. Also, more research using the genome-editing technology will clarify the physiological meaning of the NOTCH1 variant.

The AZGP1-GJC3 fusion in prostate cancer was reported previously[99], but its function is unclear. AZGP1, α-2-glycoprotein 1, is a zinc-binding glycoprotein with a transporter activity. It can interact with fatty acids and stimulates their degradation, thereby causing a severe fat loss in cancer patients[100]. GJC3,i.e., gap junction protein γ3, is a gap junction protein providing direct connections between cells, and its mutation has been found in hearing loss[94]. Fusion AGZP1-GJC3 therefore can generate a novel membrane protein with a function of a transporter.

To further speculate on the possible impact of the AZGP1-GJC3 fusion, the domains in the individual proteins and domains retained in the fusion protein should be analyzed. As mentioned above, AZGP1 protein is composed with signal peptide(position 1-20) and zinc-alpha-2-glycoprotein(position 21-293). Considering breaking point of AZGP1 (position 112), translation of zinc-alpha-2-glycoprotein would stop in the middle. Therefore, it would lost Ig-like C1-type domain(position 207-292), which could cause loss of normal function. AZGP1 was found in significantly lower levels in invasive breast tissues as compared with adjacent normal tissue[84]. In the case of GJC3, breaking position of GJC3 is 260, so only 261-279 position would be translated. It is thought that only final part of this protein would remain and function of protein would disappear. However, this protein expresses in CNS specific[49], so the effect in breast cancer could not be much influential. In conclusion, impact of AZGP1-GJC3 fusion could be from malfunction of AZGP1, and it is possible to make breast cancer more invasive and aggressive.

The RNA-seq results shown in Figure 2-4A or 2-4B indicate that there is remarkable diversity among TNBCs, even though they can be subdivided into three groups.

In line with this result, WB analysis of the key growth signaling molecules in TNBC PDXs showed its diversity too. This result is consistent with previous findings from primary TNBC tumor studies[97, 101)] and support the usefulness of TNBC PDX as a representative model for the diversity. This heterologous feature of each preast cancer enhances the needs of appropriate therapy for each patient and personalized medicine.

Altogether, our results presented here demonstrate that TNBC PDX is a reliable preclinical model that represents characteristics of individual tumors. Because of this uniqueness, proper selection of a PDX model for a specific application (e.g.,a targeted drug efficacy test or biomarker development) may be a critical factor. Besides, known limitations—such as molecular changes caused by passaging in vivo[65,96)] or variable tumor growth due to intratumoral heterogeneity—need to be carefully handled to maximize the suitability of this model for translational research. Moreover, functional study of each mutation should be performed in the future to make our results more valuable.

3. As a example of using PDX model in cancer study, we performed preclinical study of Zoladex as an adjuvant treatment in TNBC. According to the literature, administration of goserelin(GnRH agonist) with chemotherapy appeared to protect against ovarian failure, reducing the risk of early menopause and improving prospects for fertility in breast cancer patient[104)]. In addition, it is reported that GnRH antagonist repress cell invasion and proliferation *in vitro* [105)]. Therefore we hypothesized that chemotherapy with GnRH agonist as adjuvant therapy can be a good therapeutics for TNBC patient and tumor expressing GnRHR high could be affected by this therapy.

When the GnRH agonist (Zoladex) was introduced, the tumor volume of high-level expression of GnRH receptor group was remarkably reduced compared to the low-level expression of GnRH receptor group. We also found that apoptosis was increased in tumor highly expressing GnRH receptor, by performing Annexin V/PI staining. After that, western blotting analysis of several protein was conducted to reveal the mechanism of GnRH agonist, which showed that the level of B-Raf, p-B-Raf, c-Jun, CREB, p-ERK1/2, p-MEK1/2 were changed by zoladex treatment. In preliminary study, #87556 also showed similar change by treatment of zoladex, even though its GnRHR level was low. This might be due to difference

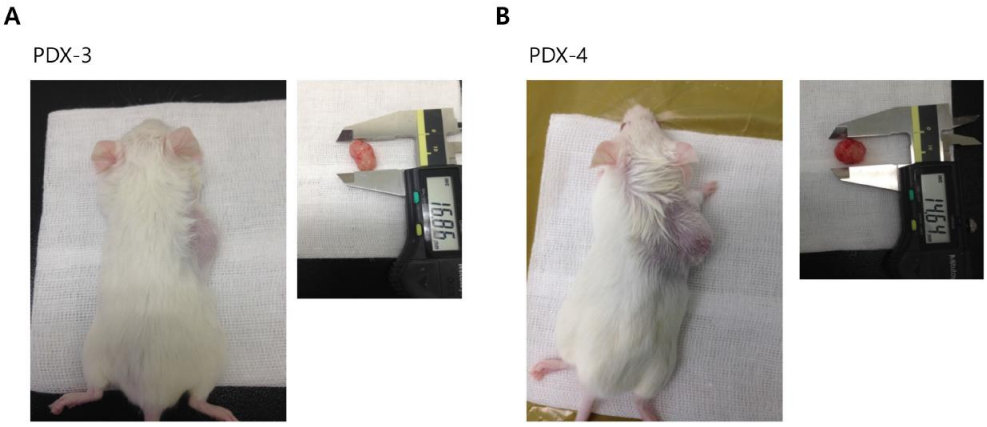
between individual entities or which parts of tumor were implanted, because a tumor is compromised of heterogeous groups. This problem could be improved by increasing number of animals in each group. The difference was very slight between only AC treatment group and the group treated with AC and zoladex together. We assumes that AC is very strong therapeutics, therefore it's effect obscure the effect of zoladex. For this reason, we will reduce the dose of AC and use NOD-SCID instead of NSG to see long term effect because NSG is so weak to see long-term treatment. In future experiment, we will also use 2 samples (GnRHR high sample and low sample) and compare the effect of zoladex depending on the different expression of GnRHR. After termination of animal experiment, we will perform in vitro experiment including western blot, Annexin V staining and treatment siRNA of GnRHR to reveal the mechanism of GnRH agonist therapy. We expect our research can provide the basis of GnRH treatment in the molecular level and suggest improved therapeutic method for TNBC.

As another example of using PDX in cancer research, we conducted experiment the relationship of miR-155 and YAP. Although it is well known that mir-155 and YAP are very oncogenic in various cancer, their relationship has not been studied yet. In human sampel and PDX tissue, miR-155 showed reverse relationship with pYAP. Luciferase assay and real-time PCR after treatment antagomir of mir-155 showed that miR-155 reduced YAP expression. To find mediator causing this effect, we performed real-time PCR about many candidates in cotrol and miRZIP cells. From this, we could assume that SVIL mediates this effect. On the other hand, YAP also regulates expression of miR-155 directly with diverse transcription factor such as TEAD4, FOXM1 and RUNX3. However, we have to identify exact binding site and the most effective transcription factor which affect the expression of miR-155. Anyway, PDX model and PDC were considered as a very useful tool to research this kind of study.

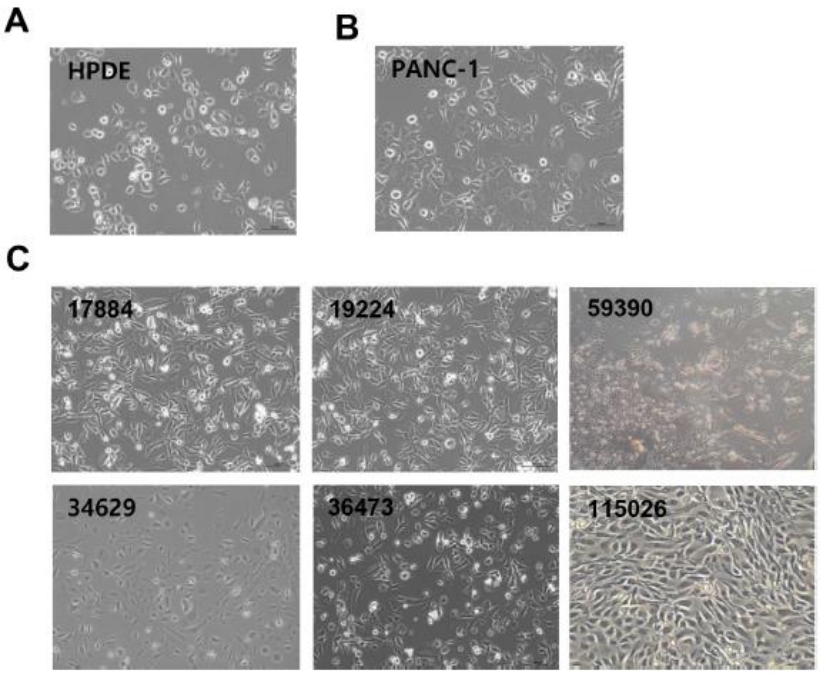
CONCLUSION

PDX model has a lot of advantages to research cancer study because it represents the unique features of each cancer patient and retain diverse characteristics of cancer. We developed PDX model for pancreatic ductal adenocarcinoma(PDAC) and triple negative breast cancer(TNBC). It is thought to be worthwhile because PDAC is the most challenging type of cancer to treat and TNBC has high recurrent, metastatic potential but still lacks appropriate targeted drug. We analyzed these tumor at molecular and genomic levels by several tools including bioinformatics, which revealed various variants in PDAC and TNBC giving new insights for cancer. Moreover, this paper presents two preclinical study to show that PDX models are useful for drug screening and preclinical evaluation. We investigated the efficacy of GnRH agonist for TNBC and studied the relationship of miR-155 and YAP in breast cancer. Taken altogether, our study shows a good example which PDX model can be applied in cancer study.

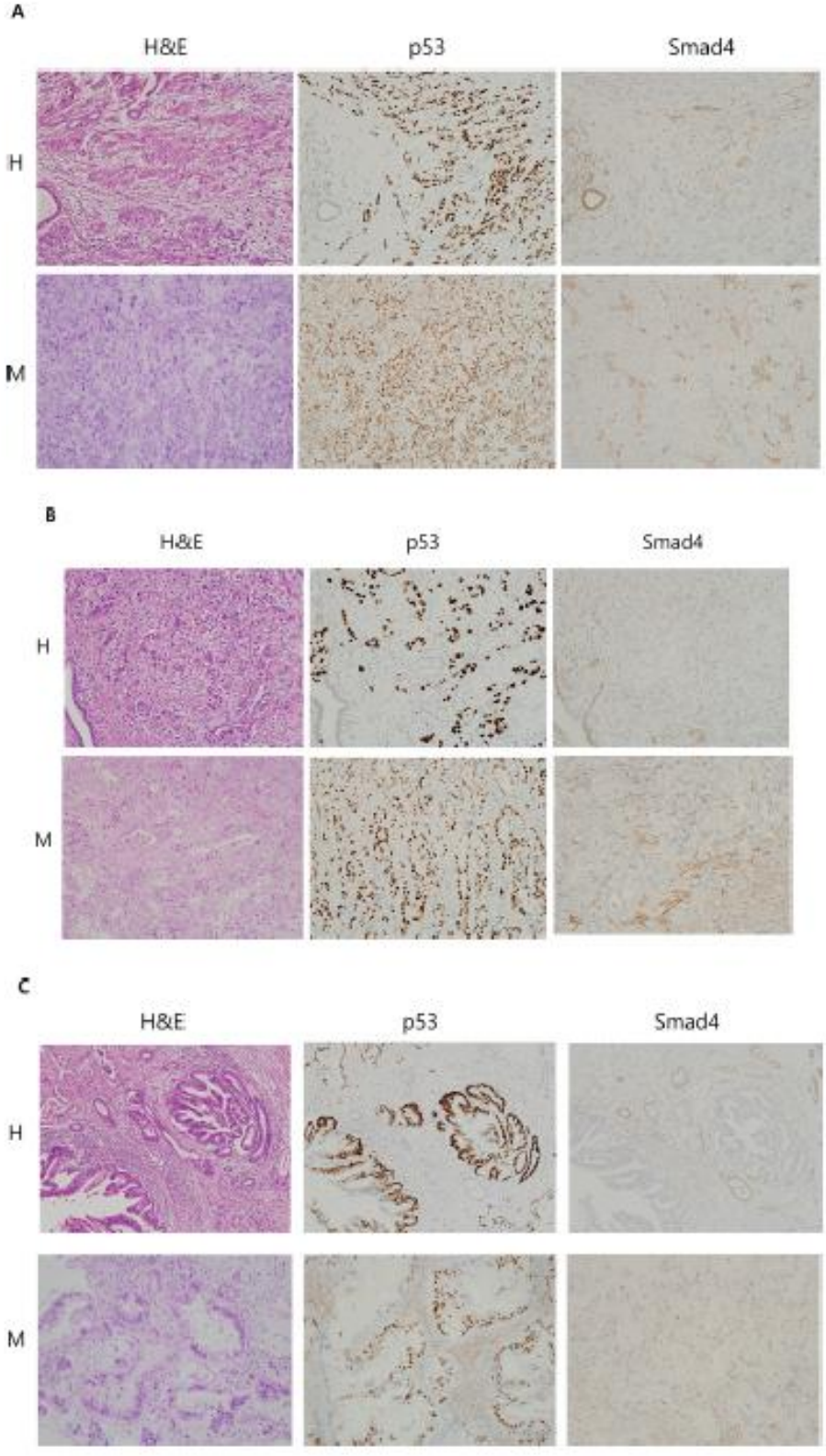
Supplementary Data

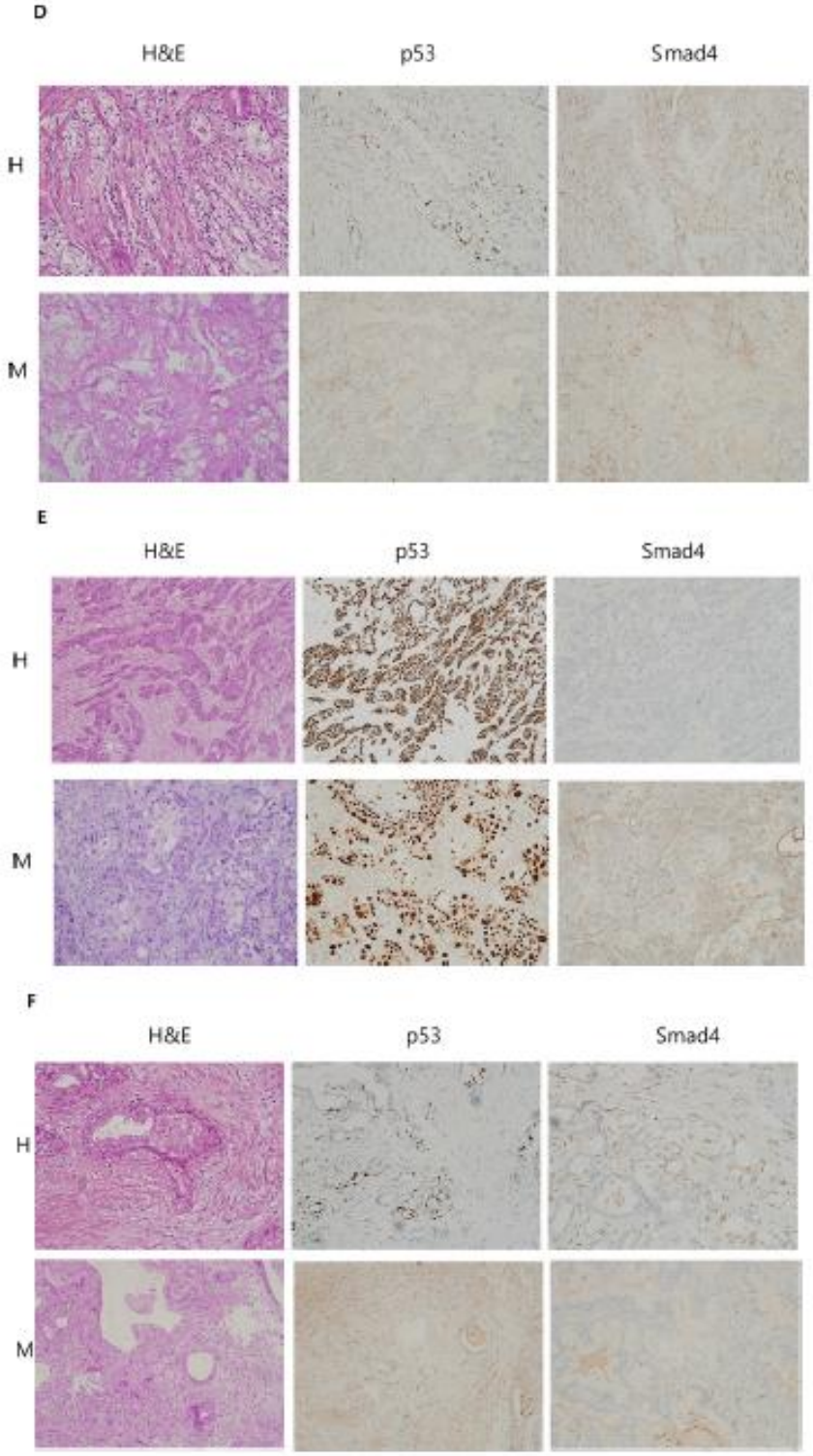


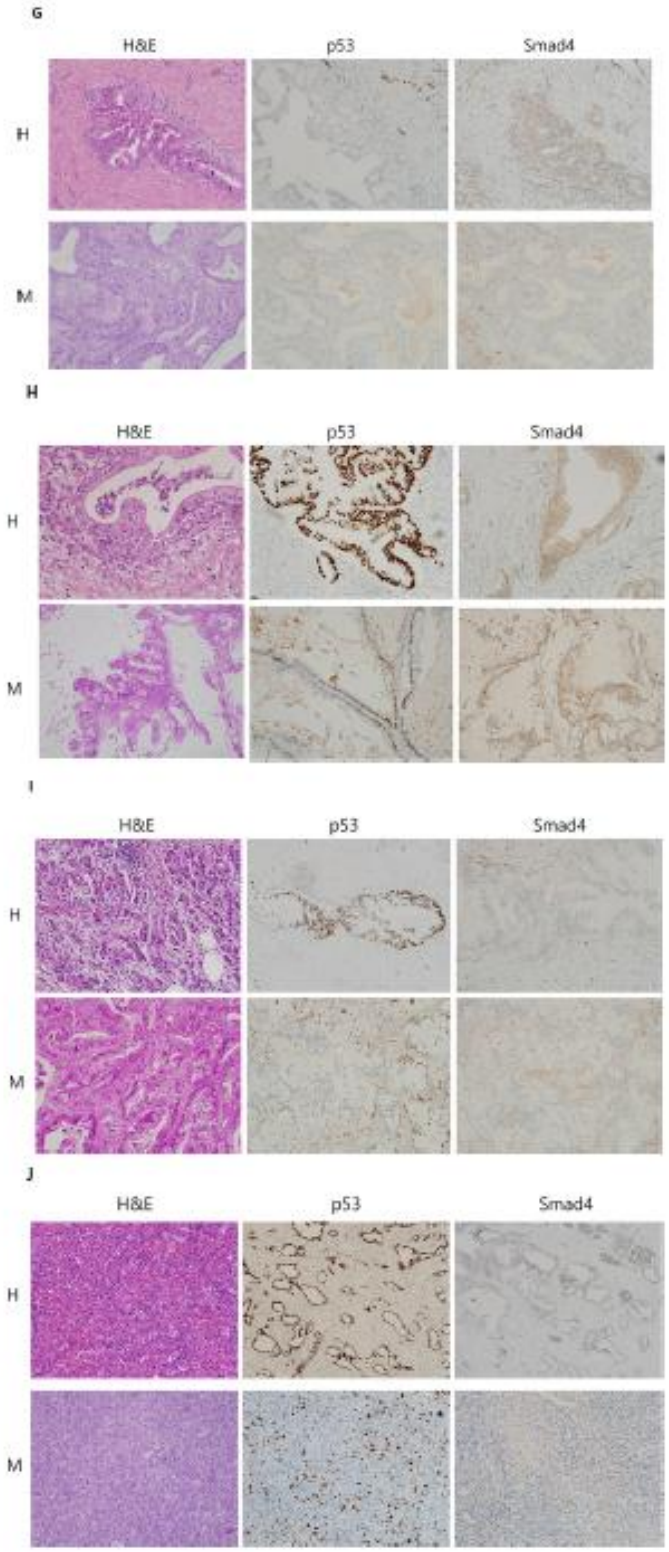
Supplementary Figure S1-1: Examples of pancreatic patient-derived xenografts (PDXs) generated in NOD-SCID mice. Left: Before dissection; right: tumor in calipers after dissection.

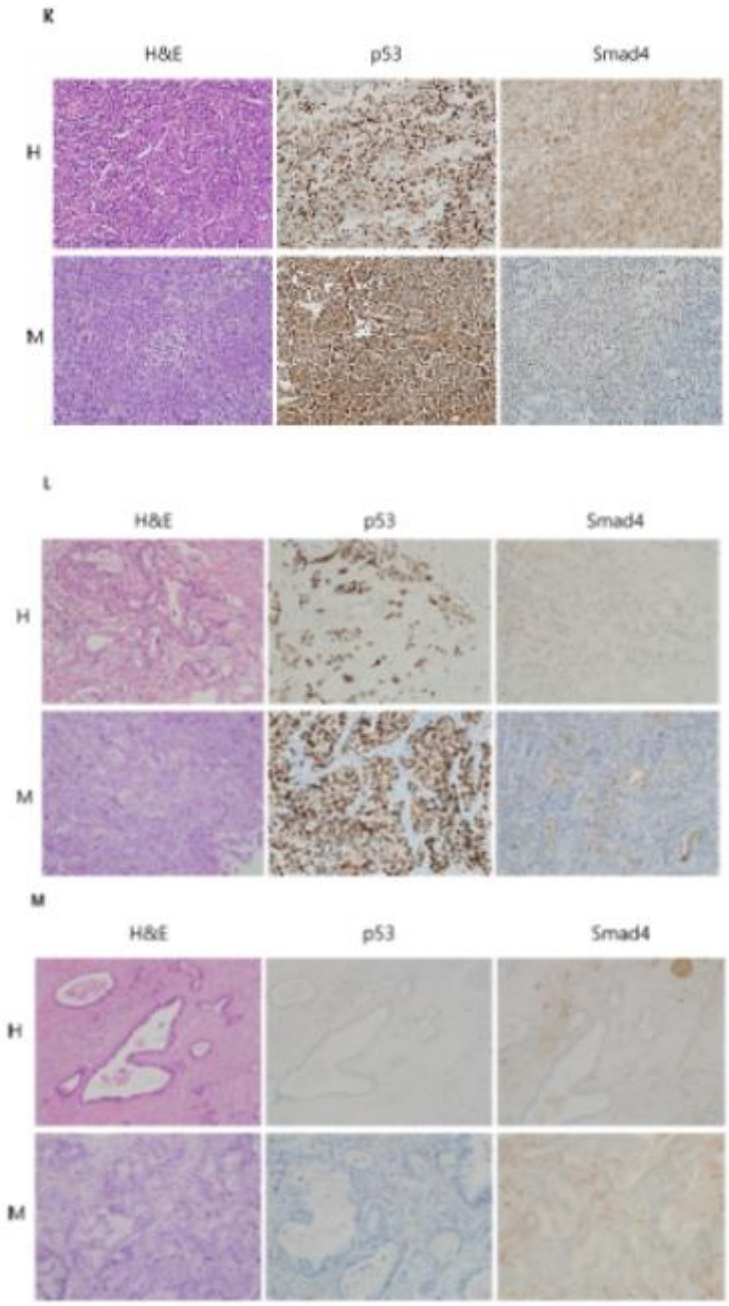


Supplementary Figure S1-2: Representative picture of primary cancer cells derived from patient-derived xenografts (PDXs) and cell lines. A. Human pancreatic ductal epithelial cell (as a normal control). B. Panc1 pancreatic ductal adenocarcinoma cell (as a cancer control). C. Human pancreatic primary cancer cells from PDX tissues(x200) (by Eun Ji Lee)

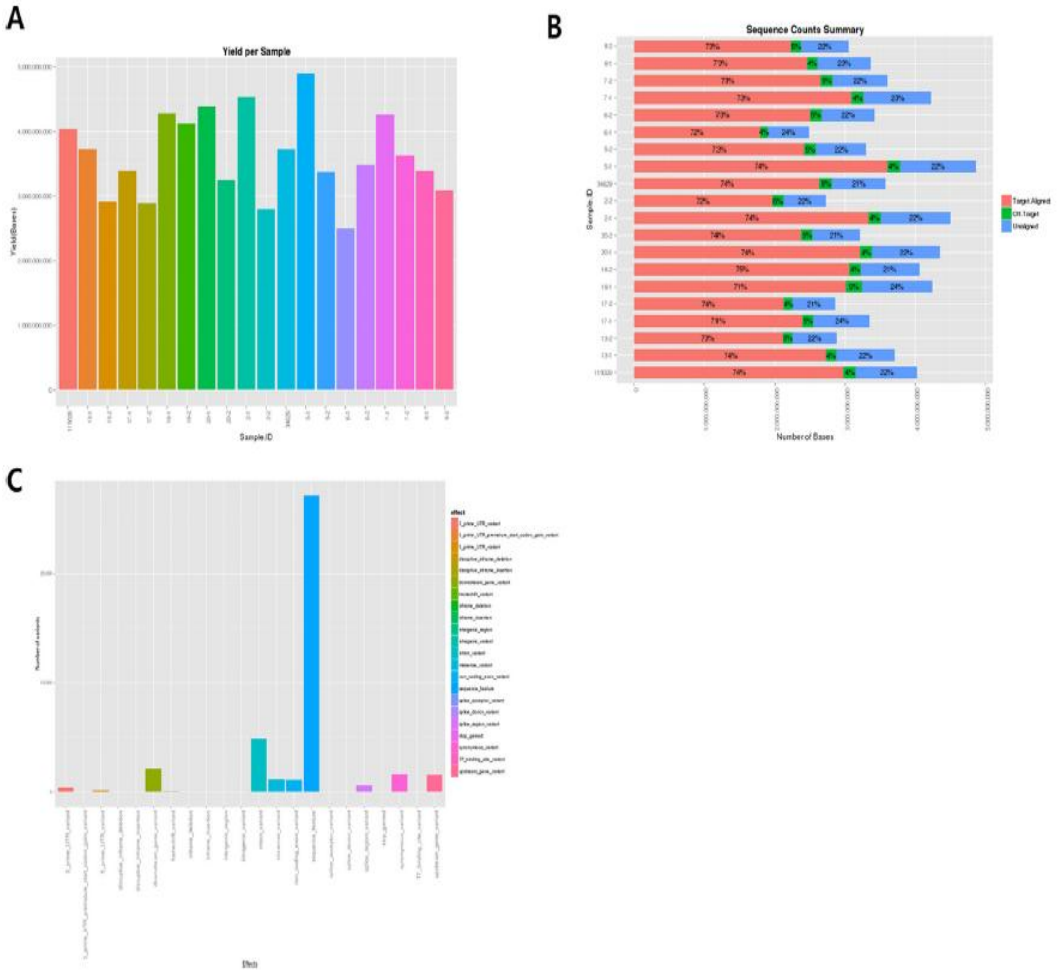

A
H&E
p53
Smad4
H
M
B
H&E
p53
Smad4
H
M
C
H&E
p53
Smad4
H
M


D
H&E
p53
Smad4
H
M
E
H&E
p53
Smad4
H
M
F
H&E
p53
Smad4
H
M

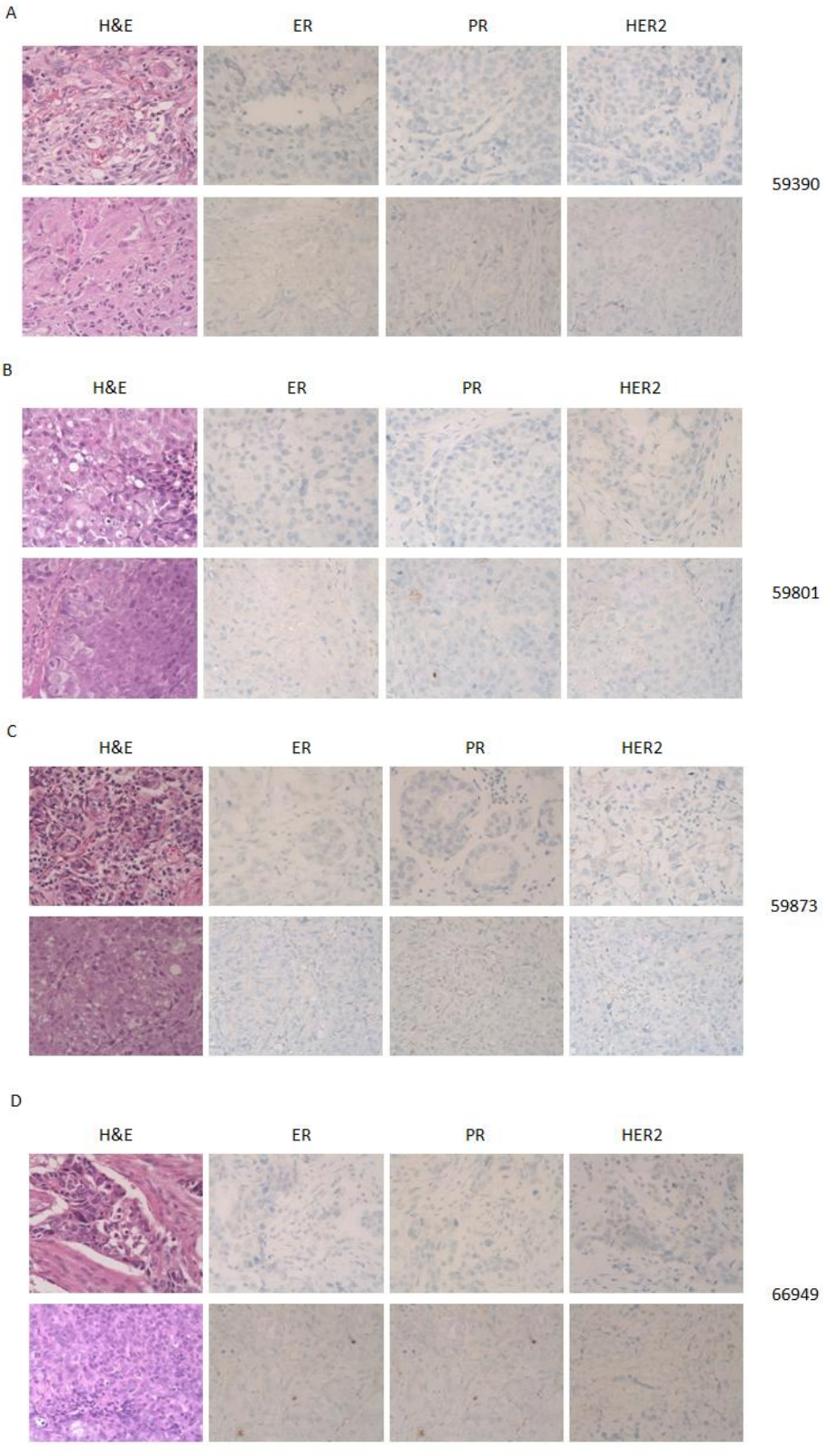

G
H&E
p53
Smad4
H
M
H
H&E
p53
Smad4
H
M
I
H&E
p53
Smad4
H
M
J
H&E
p53
Smad4
H
M



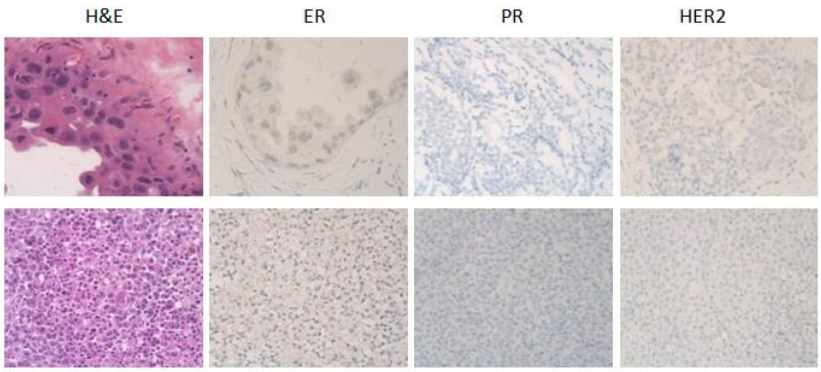
Supplementary Figure S1-3: A to M. Immunohistochemistry (IHC) or hematoxylin and eosin (H&E) histological data from the primary tumors AMC004 to AMC017 (upper panels; marked as H(Human)) or a patient-derived xenograft (PDX) tumor (lower panels; marked as M(Mouse)). Representative pictures of H&E staining, and p53 and SMAD4 IHC are shown.



Supplementary Figure S1-4: Genetic analysis of patient-derived xenograft (PDX) primary tumors using Comprehensive Cancer Panel. A. Graph showing the yield per sample in the Next-Generation Sequencing (NGS) reads. B. Sequence Counts Summary showing aligned, off target, and unaligned sequence counts for each of the tumor samples. C. Regional annotation of the variants identified in the panel. Note the high number of the sequence_feature (marked as blue) variant type, which is in the protein coding region, due to the data being obtained from the exome sequencing of 402 cancer-related genes. (By Macrogen)

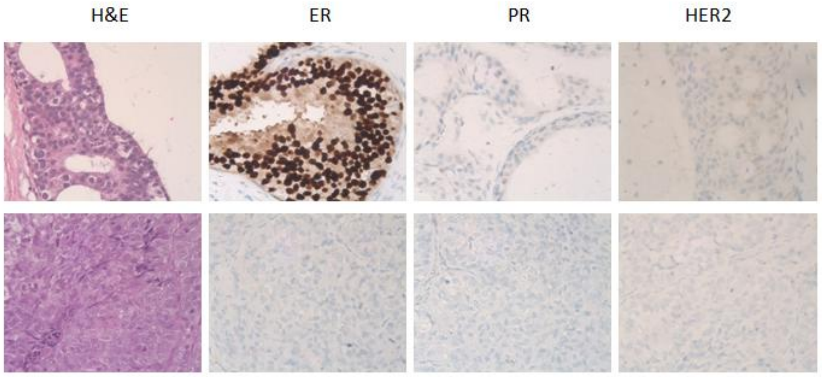

A
H&E
ER
PR
HER2
59390
B
H&E
ER
PR
HER2
59801
C
H&E
ER
PR
HER2
59873
D
H&E
ER
PR
HER2
66949

E



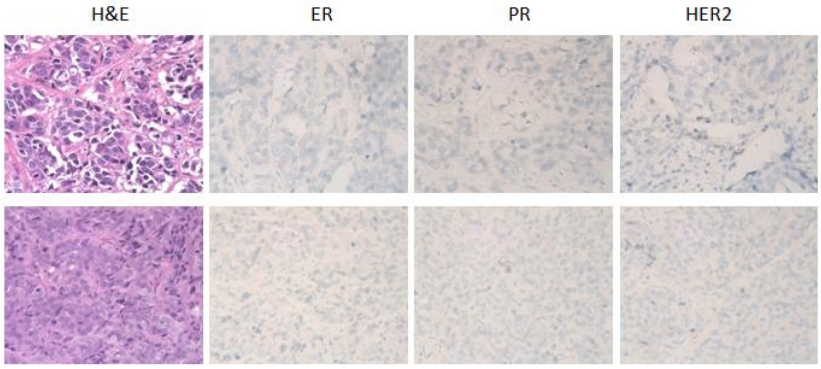
70151 (mouse tumor 는 lymphoma)

F



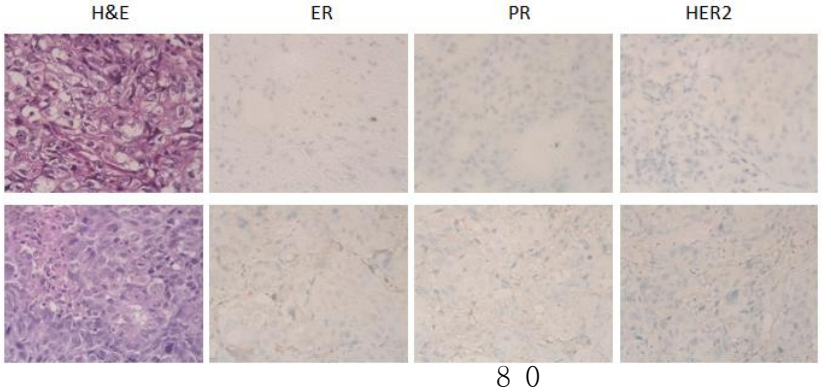
70212 (ER match 안됨)

G



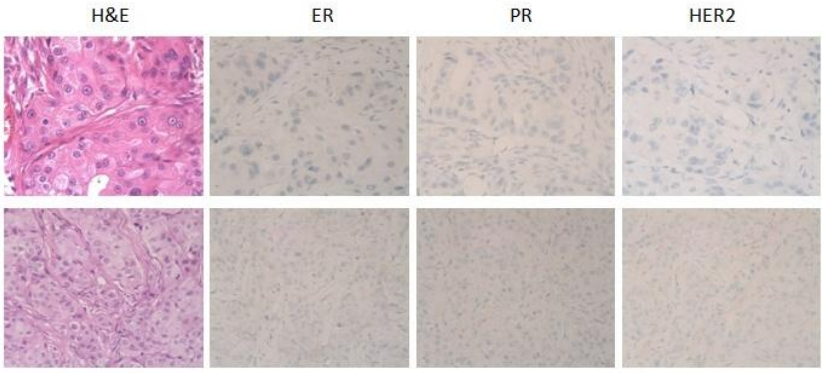
75535

H



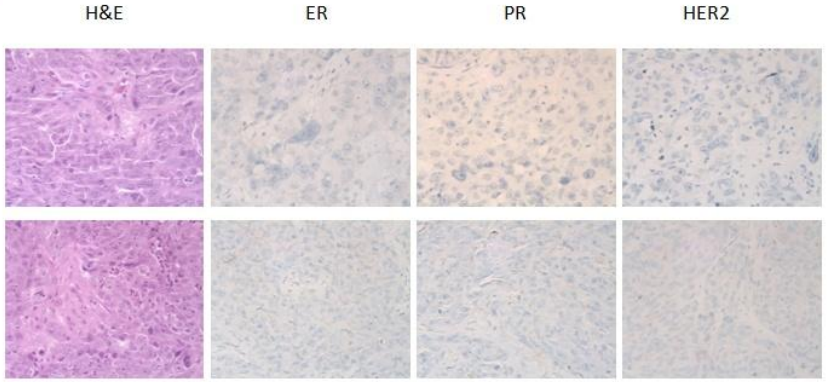
77592

I



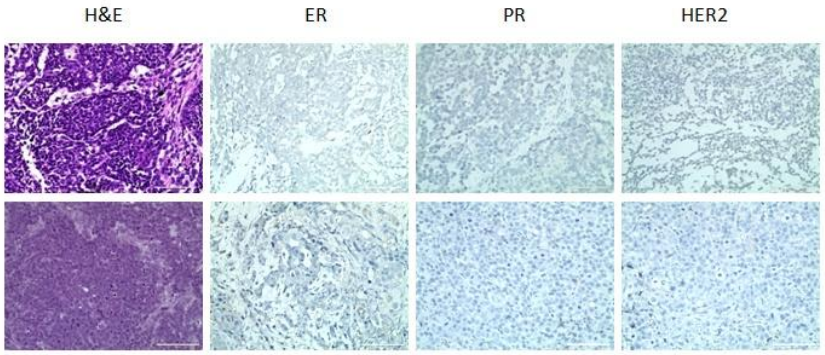
87556

J



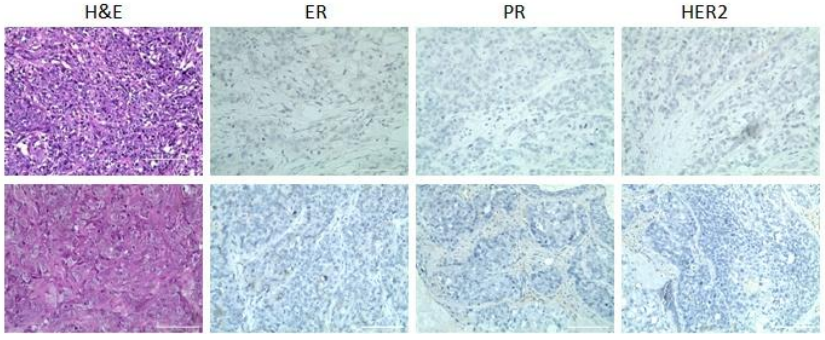
93376

K



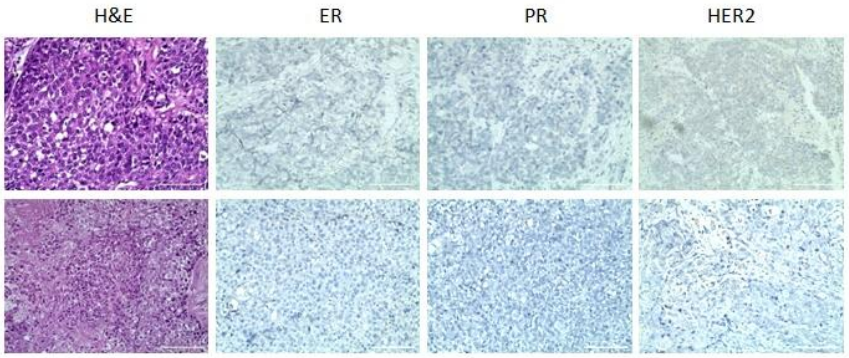
50786

L



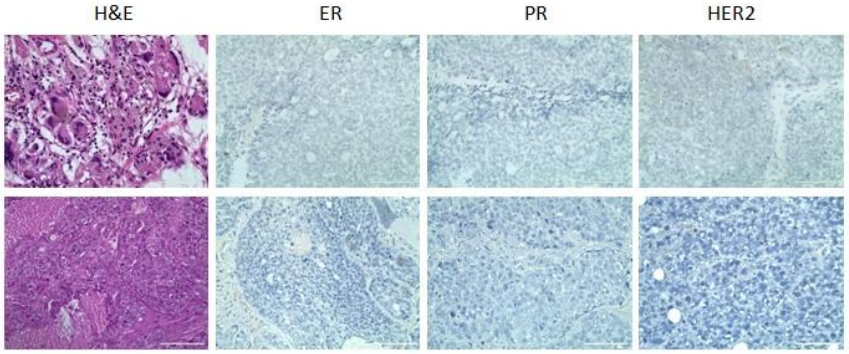
52433

M



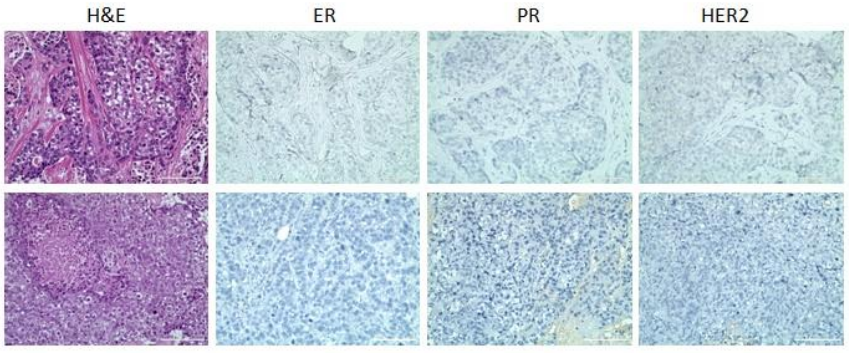
55279

N



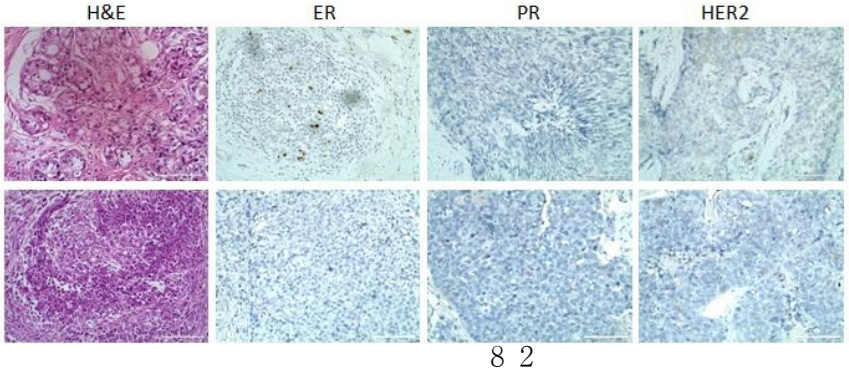
55431

O

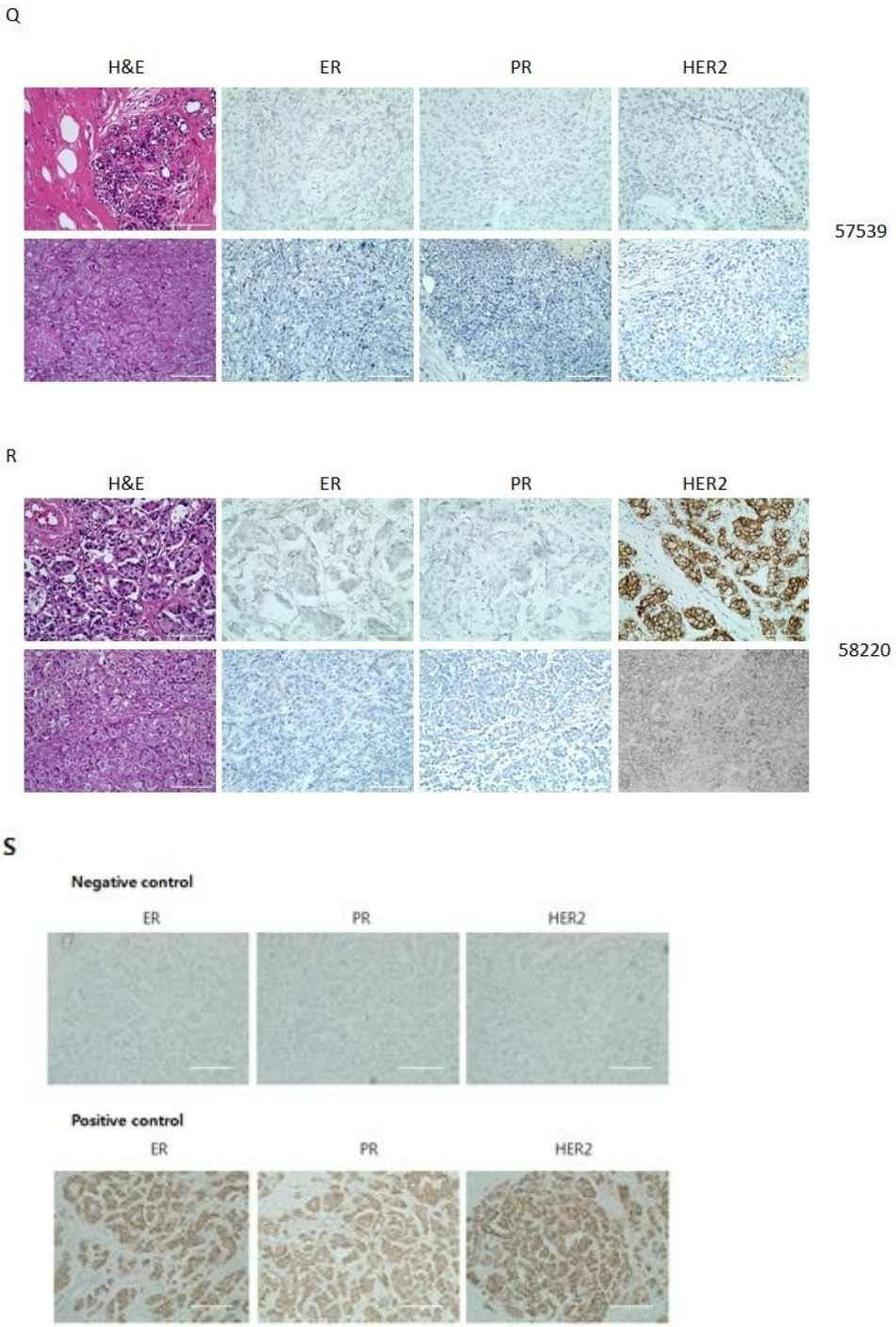


57239

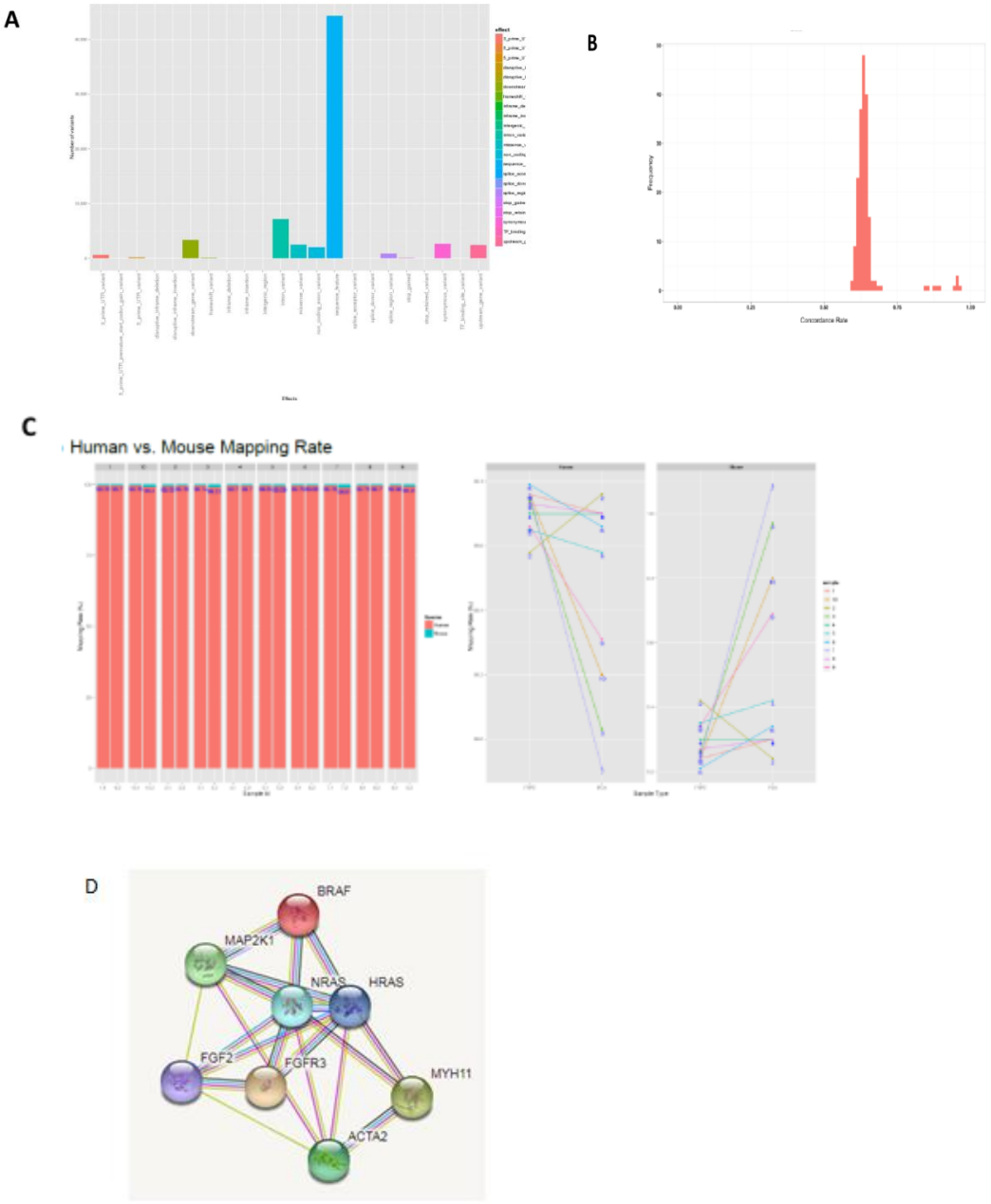
P



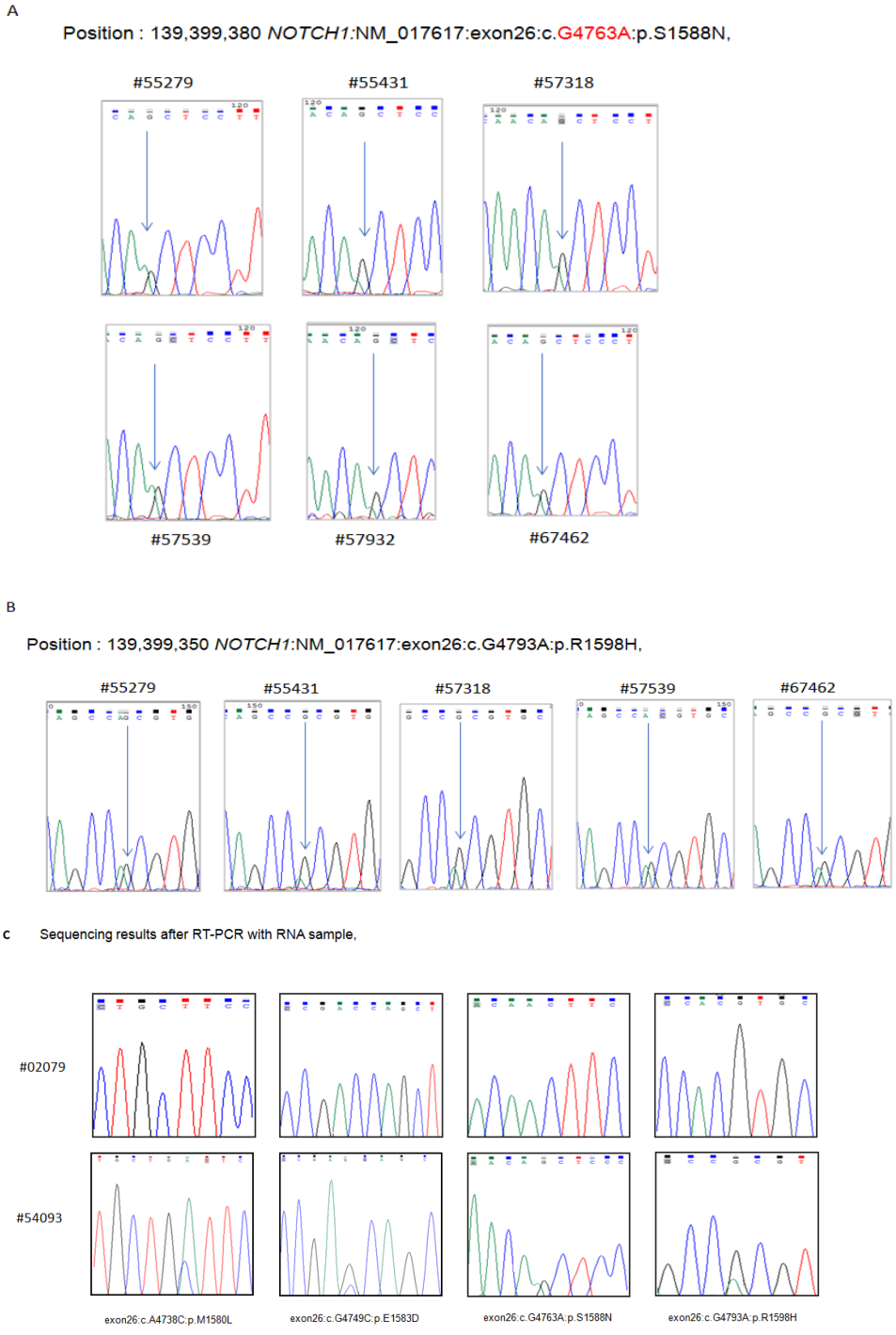
57318



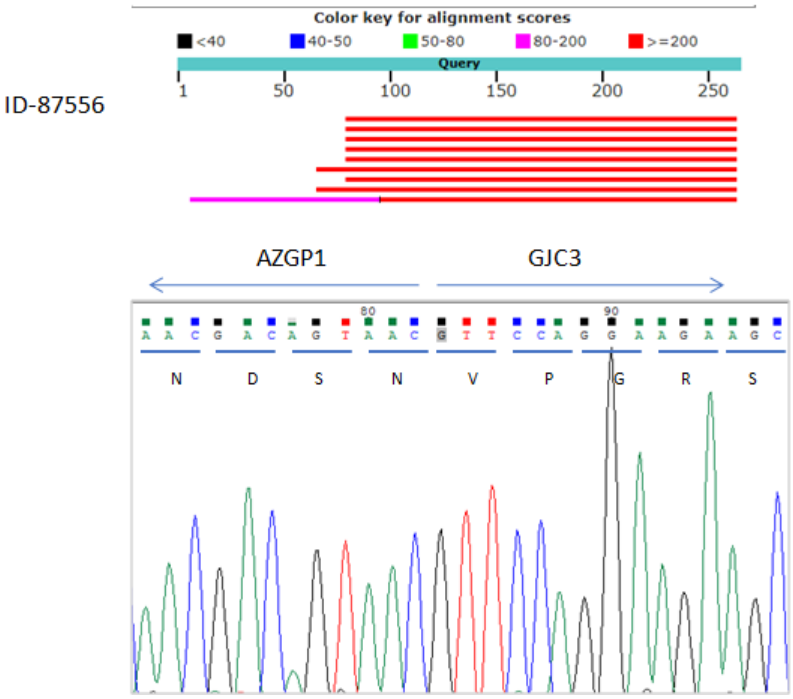
Supplementary Figure 2-1. IHC data for each primary tumor–PDX tumor pair. 59390 (A), 59801 (B), 59873 (C), 66949 (D), 70151 (E), 70212 (F), 75535 (G), 77592 (H), 87556 (I), 93376 (J), 50786 (K), 52433 (L), 55279 (M), 55431 (N), 57239 (O), 57318 (P), 57539 (Q), and 58220 (R). The upper panels are from a human primary tumor, and the lower panels are from PDX tumors; 400× magnification. (S) positive and negative control.



Supplementary Figure 2-2. A. Graph showing the numbers of variants classified by type. B. The concordance histogram of the CCP analysis. C. A graph showing the human vs mouse mapping rate. Note that 0.2–1.1% of mouse sequences possibly resulted from infiltrating stromal cells. D. Interaction network between top 3 differential genes by drawing in STRING website(https://string-db.org/)



Supplementary Figure 2-3. A. Electropherograms for a Notch1 variant (G4763A) found in six PDXs. B. Electropherograms for a Notch1 variant (G4793A) found in five PDXs. C. Sequencing results of cDNA from RT-PCR with RNA sample of PDX tumor tissue.



Supplementary Figure 2-4. Sequencing alignment (upper panel) and electropherogram data (lower panel) for the AZGP1-GJC3 gene fusion in PDX 87556.

Supplementary Table S1-1: Clinical parameters of pancreatic cancer primary cells (by Eun Sung Jun)

Primary cell ID	Gender	Age	Pathological parameters				
			Stage	pT	pN	M	Histology_type
17884	F	48	IIB	3	1	0	Pancreatic ductal adenocarcinoma
19224	M	61	IIB	3	1	0	Pancreatic ductal adenocarcinoma
34629	F	64	IIA	3	0	0	Pancreatic ductal adenocarcinoma
36473	F	59	IIB	3	1	0	Pancreatic sarcomatoid adenocarcinoma*

Five-digit numbers were assigned for identification purposes and the clinical features of each primary cell type are listed.

Supplementary Table S1-2. List of the genes with single nucleotide polymorphisms (SNPs) that were predicted as being functionally important (By Eunji Kim, Hyun Cue Lee and Buhm Han)

Gene name										
DNAAF2	*C8orf31*	*C1orf65*	*DOPEY1*	*GTPBP3*	*CCRL2*	*P4HA3*	*SARDH*	*SBF2*	*CPN2*	*SVEP1*
TESPA1	*TCF7L2*	*SCRIB*	*GBP3*	*ONECUT2*	*RGS22*	*KRT76*	*NGFR*	*OR4X1*	*CHMP1A*	*NXPE1*
PKD1L2	*MYOM2*	*FAM71F2*	*TTC9*	*MOCOS*	*TCF7*	*C15orf26*	*TMEM176B*	*EGFLAM*	*HIPK4*	*FGFBP3*
MYCBPAP	*OR7D2*	*GOLGA1*	*FAM129C*	*ACOX1*	*FAM194B*	*DUSP6*	*PRSS21*	*KLHL38*	*TBX6*	*C10orf71*
KIAA0753	*OR4C46*	*NFATC3*	*CDC20B*	*PAK6*	*MYO15A*	*CCDC8*	*OR2F1*	*KCNU1*	*C14orf159*	*MPDU1*
OR5H6	*MYO10*	*PRPH*	*OR52L1*	*MIR548H4*	*FAM198A*	*EVC*	*PMEL*	*ANKRD6*	*RIC8A*	*OR51B2*
BTBD16	*OR10G9*	*TRAPPC11*	*MYOM3*	*NOX5*	*ALDH1L1*	*PITRM1*	*TIAM1*	*ADAMTS16*	*CASKIN2*	*OR51G1*
RGS11	*LOC100505666*	*TAS2R7*	*PRIM2*	*SPESP1*	*SERPINA9*	*QRICH2*	*HCLS1*	*MROH2B*	*CARD14*	*HELT*
KRTAP11-1	*ADAM15*	*KRT3*	*ITGA1*	*RIN3*	*KRT83*	*KIAA1524*	*TIPIN*	*TAF1B*	*RAI1*	*HUNK*
TPPP2	*OR1A1*	*NIN*	*TDG*	*TSKU*	*NAV1*	*OXER1*	*CYP4A22*	*TTN*	*WDR90*	*WNK2*
TRIOBP	*PUS1*	*OR2K2*	*VN1R5*	*PALD1*	*RABL6*	*DMBT1*	*CYP4F8*	*STK36*	*PLA2G4D*	*ARSB*
LMO7	*ZNF30*	*C3orf70*	*FAM53B*	*TTF1*	*HEATR1*	*AQP11*	*TNXB*	*07-Mar*	*OR6S1*	*CFH*
CCDC169-SOHLH2	*TBRG1*	*USP6*	*TEX13B*	*NIPSNAP3B*	*FTSJ3*	*MRPL37*	*AKAP3*	*OR5H2*	*TEKT1*	*MTCH2*
CCDC169	*ADNP*	*AZI1*	*USP45*	*ZNF484*	*ZNF761*	*NIPSNAP3A*	*TTLL6*	*RGL3*	*LRRN2*	*C7*
KCNK4	*OR5D14*	*ACTBL2*	*OR1J2*	*MTUS1*	*ACAT2*	*ABCC4*	*FMO2*	*TAS2R42*	*ICOSLG*	*KIF20B*

Supplementary Table S1-2(Continued)

VARS2	*SGSH*	*F5*	*AFAP1-AS1*	*RRS1*	*RNF43*	*GALNT2*	*PASK*	*OR4D6*	*SP110*	*CBR3-AS1*
PLCD1	*IGSF10*	*FAM118A*	*AFAP1*	*LRRFIP1*	*USP35*	*TADA1*	*SIGLEC6*	*ZNF229*	*CCDC137*	*CBR3*
TIAM2	*KRT36*	*FEZ1*	*BAG3*	*PRUNE*	*MMP27*	*PADI3*	*USP31*	*LOC100506134*	*LTBP3*	*TCF3*
MCPH1	*MROH5*	*TRIP11*	*CRNN*	*TRIM45*	*OR52J3*	*MGAM*	*MRC1*	*TTC21B*	*GGH*	*CPVL*
COL11A2	*IL1RL2*	*GPR112*	*PRICKLE1*	*COL9A2*	*PPP2R4*	*ADARB2*	*ACIN1*	*SHBG*	*MYEOV*	*PLA2G7*
TRPA1	*DOCK5*	*ILVBL*	*PLEKHH1*	*PHF21A*	*OR8D2*	*KDM6B*	*IL27*	*SAT2*	*CAPG*	*BTD*
HKDC1	*AP1G2*	*COL6A6*	*IQGAP2*	*PPARGC1A*	*PDLIM5*	*CPT1B*	*OR10H2*	*CDC6*	*VARS*	*GPC4*
AP1G1	*CCDC149*	*F2*	*IFI44L*	*ZMIZ2*	*SLC22A14*	*CHKB-CPT1B*	*RPGR*	*CCNJL*	*SLC2A7*	*FGF2*
SUCLG2	*NMBR*	*RAB11FIP2*	*KIAA1751*	*TRIM31*	*SLC22A20*	*CHIA*	*ETFA*	*APOB*	*ACO2*	*NUDT6*
OCEL1	*PLOD3*	*RALGPS1*	*MICA*	*IMPG1*	*CD320*	*MATN2*	*IRS1*	*WBSCR27*	*POLR3H*	*FUT2*
SURF6	*AMBN*	*BPI*	*MYBPC1*	*BANK1*	*PDLIM7*	*TRAK2*	*CCR5*	*TRIML1*	*CRTAP*	*PLEKHG2*
CHI3L1	*OR56B1*	*PCDHGA9*	*EML2*	*EFCC1*	*GRID1*	*PYCRL*	*PPP1R3A*	*TET1*	*OR2G2*	*C6orf10*
NRAP	*ARSD*	*PCDHGB6*	*DCAF4*	*GTF2E1*	*C4orf26*	*ZNRF4*	*XRCC1*	*INHA*	*MEFV*	*LCTL*
FPR1	*OR1I1*	*PCDHGA10*	*BOC*	*CX3CR1*	*SVOPL*	*GRIN3B*	*MMP7*	*FAM188B*	*PLEKHG5*	*NEK4*
FEZ2	*TAPBPL*	*PCDHGB7*	*FAM171A1*	*MPHOSPH10*	*C19orf57*	*COBL*	*FUK*	*NFXL1*	*SLC15A2*	*SAMD9*
NBR1	*C1orf127*	*PCDHGA11*	*C1R*	*PRR5*	*LTBP2*	*C2CD2L*	*ZKSCAN3*	*MYH15*	*PPAP2C*	*FBXO34*
PGLYRP3	*FAM149B1*	*ERI2*	*HHIPL2*	*PRR5-ARHGAP8*	*SLC52A1*	*KCTD20*	*CTU1*	*VPS13C*	*AKAP11*	*IDO2*

Supplementary Table S1-2(Continued)

IGF2R	*CCP110*	*ACSM3*	*MOV10L1*	*AK1*	*CEP120*	*SNAP47*	*PIK3CG*	*CLDN23*	*KLRB1*	*AQPEP*
ALDH1A1	*CWC22*	*C15orf55*	*ALPK2*	*TBC1D1*	*TTLL5*	*RBBP8NL*	*KRT39*	*DFNA5*	*DLG1*	*SLC7A9*
SLC28A1	*ZNF195*	*MYBBP1A*	*HSF5*	*AOC2*	*LNPEP*	*NKAIN4*	*DSC1*	*PARP15*	*DDX51*	
NELL1	*C9orf171*	*C3orf30*	*TNS1*	*ENPEP*	*SPINK5*	*IRF6*	*MAP4K5*	*IL17RA*	*C19orf45*	
ABO	*FAM71F1*	*BIVM-ERCC5*	*02-Mar*	*NKTR*	*EHHADH*	*PNMA2*	*MAPT-AS1*	*TMC6*	*CAMKK2*	
DNAJB11	*RSPH4A*	*ERCC5*	*ACTRT2*	*PDZRN4*	*C9orf50*	*HEMK1*	*SPPL2C*	*PGAM5*	*CPOX*	
IGSF5	*LOC100652739*	*KHDC3L*	*USP29*	*CDH11*	*RPGRIP1L*	*PRMT6*	*C2orf71*	*LRRC55*	*NQO1*	
ZNF519	*RAET1E*	*NME8*	*GORASP2*	*CHURC1-FNTB*	*BST1*	*CCT3*	*WDR72*	*TMPRSS2*	*HDHD1*	
C14orf80	*RNU6-28*	*MAST3*	*ALDH5A1*	*RAB15*	*EZH2*	*PEPD*	*SH3TC2*	*CDON*	*DDX27*	
OR6K2	*TP53BP1*	*EIF1B-AS1*	*DHDH*	*MUSK*	*MINK1*	*CTSS*	*NOTCH4*	*HEXA*	*PDCD6IP*	
OR6C65	*PCDHGA1*	*MYRIP*	*INADL*	*KIAA0141*	*SLC14A1*	*ZNF140*	*PLCE1*	*CCDC116*	*GFAP*	
C11orf82	*PCDHGA2*	*CHRNA9*	*KIAA1210*	*NUAK2*	*YDJC*	*GLI1*	*H6PD*	*LRRC71*	*IQGAP3*	
OR5AN1	*PCDHGA3*	*GPC5*	*PRODH2*	*LRP10*	*C9orf131*	*PYGB*	*FSIP2*	*OR10J1*	*GDPD4*	
NUGGC	*PCDHGB1*	*LMF1*	*SCARF1*	*COL16A1*	*ITGA8*	*NT5DC1*	*ZNF276*	*SPATA3*	*GRID2IP*	
OR5D13	*PCDHGA4*	*TPSG1*	*SPG11*	*NEK5*	*COL15A1*	*COL10A1*	*FANCA*	*CGN*	*OR10Q1*	
MPEG1	*PCDHGB2*	*MTA1*	*CD101*	*SPEF2*	*ANKRD16*	*HIVEP1*	*SPAG17*	*PON2*	*KIAA1377*	
NHLRC2	*PCDHGA5*	*TPO*	*EMR2*	*C1QTNF3-AMACR*	*SWT1*	*PDCD1*	*VWDE*	*ANKRD30A*	*LRIT1*	

Supplementary Table S1-2(Continued)

MTMR6	*PCDHGB3*	*OR2B11*	*LOC100131691*	*AMACR*	*CHMP4A*	*SULT1C3*	*CERS4*	*KDELR3*	*SLCO1C1*
OR8U8	*PCDHGA6*	*AHSG*	*MZF1*	*GUCY2D*	*HHAT*	*PCDH15*	*DLGAP5*	*C7orf31*	*COL24A1*
OR5AR1	*PCDHGA7*	*SYT8*	*MTHFSD*	*TNFRSF13B*	*TOP1MT*	*PNLIPRP3*	*ATP5S*	*PPARGC1B*	*RNASET2*
MRM1	*PCDHGB4*	*MAGEB16*	*CACNA1H*	*COL4A3*	*RP1*	*ZNF546*	*KIF4B*	*GPR101*	*TLR10*
FCN2	*PCDHGA8*	*ACTL9*	*SCUBE2*	*UROC1*	*IQCE*	*CYP4F2*	*DKK2*	*CASC5*	*MAGEC3*
TMCC1	*PCDHGB5*	*AGT*	*DSPP*	*SPNS3*	*KRT71*	*APOL5*	*CAPN13*	*TMPRSS4*	*B4GALNT3*
ARAP2	*C3orf20*	*CDH23*	*TNK2*	*SHROOM3*	*PDIA4*	*LAMB3*	*EGFL6*	*SH3RF3*	*OLR1*
GGA3	*IL17RD*	*KLF10*	*PPAN-P2RY11*	*TNRC6C*	*DKKL1*	*E2F2*	*MYZAP*	*TECPR1*	*OR51B5*
KRT13	*PLB1*	*INMT-FAM188B*	*P2RY11*	*ZSCAN5A*	*BMP2K*	*NOD1*	*GCOM1*	*CELSR2*	*OR51I2*
TTC29	*ALDH2*	*INMT*	*IL4I1*	*FAM170A*	*MTRR*	*LAMB4*	*VPS53*	*FBRSL1*	*IGSF22*
HTR3D	*TTC24*	*ERV3-1*	*NUP62*	*PZP*	*WDR55*	*ABCB5*	*RECK*	*TMC8*	*MKI67*

Supplementary Table S1-3. Functional annotation of the variants

(By Eunji Kim, Hyun Cue Lee and Buhm Han)

Effect	dbSNP	dbSNP(%)	Novel	Novel(%)	Total
***TOTAL	30,796	10000	10,031	10000	40,827
3_prime_UTR_variant	284	92	100	100	384
5_prime_UTR_ premature_start_ codon_gain_variant	14	5	0	0	14
5_prime_UTR_variant	97	31	33	33	130
disruptive_inframe_ deletion	7	2	16	16	23
disruptive_inframe_ insertion	4	1	10	10	14
downstream_gene_ variant	1,621	526	490	488	2,111
frameshift_variant	13	4	15	15	28
inframe_deletion	0	0	7	7	7
inframe_insertion	0	0	16	16	16
intergenic_region	1	0	0	0	1
intragenic_variant	0	0	2	2	2
intron_variant	3,441	1120	1,418	1410	4,859
missense_variant	804	261	342	341	1,146
non_coding_exon_ variant	867	282	228	227	1,095
sequence_feature	20,612	6690	6,543	6520	27,155
splice_acceptor_ variant	22	7	2	2	24
splice_donor_variant	0	0	3	3	3
splice_region_variant	340	110	243	242	583
stop_gained	0	0	15	15	15
synonymous_variant	1,482	481	144	144	1,626
TF_binding_site_ variant	6	2	1	1	7
upstream_gene_variant	1,181	383	403	402	1,584
#of variants in coding and UTR					**1804**

UTR: Untranslated region; TF: transcription factor.

Supplementary Table S1-4. List of the variants that were predicted to be deleterious, as identified using Comprehensive Cancer Panel (CCP) data (By Macrogen and Je-Keun Rhee)

CHROM	POS	ID	REF	ALT	Variant Type	Allele	Effect	Impact	Gene_Name	Transcript_BioType	Rank	HGVS.c	Num Samples Affected	Num Homo Ref	Num Hetero	Num Homo Alt
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.934-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`11/16	c.850-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.883-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.850-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.853-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	nonsense_mediated_decay	`10/15	n.*663-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.850-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.925-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.895-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	nonsense_mediated_decay	`10/15	n.*663-2A>G	2	18	2	0

Supplementary Table 1-S4(Continued)

chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.892-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.883-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	nonsense_mediated_decay	`2/7	n.244-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`1/5	c.112-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	nonsense_mediated_decay	`4/9	n.*179-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`10/15	c.892-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	protein_coding	`5/6	c.466-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	retained_intron	`4/4	n.541-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	retained_intron	`1/2	n.215-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	nonsense_mediated_decay	`8/8	n.*179-2A>G	2	18	2	0
chr1	45797760	rs77542170	T	C	SNP	C	splice_acceptor_variant&intron_variant	HIGH	`MUTYH	nonsense_mediated_decay	`9/9	n.*589-2A>G	2	18	2	0
chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`9/47	c.1140delA	18	2	18	0
chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`9/46	c.927delA	18	2	18	0

Supplementary Table 1-S4(Continued)

chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`6/44	c.729delA	18	2	18	0
chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`6/44	c.729delA	18	2	18	0
chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`9/46	c.1140delA	18	2	18	0
chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`6/22	c.729delA	18	2	18	0
chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`6/23	c.729delA	18	2	18	0
chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`2/19	c.1218delA	18	2	18	0
chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`2/17	c.1218delA	18	2	18	0
chr1	1.45E+08	rs369106087	GT	G	DEL	G	frameshift_variant	HIGH	`PDE4DIP	protein_coding	`2/17	c.90delA	18	2	18	0
chr2	85536533	.	TC	T	DEL		frameshift_variant	HIGH	`TCF7L1	protein_coding	`12/12	c.1718delC	2	18	2	0
chr3	37818889	.	C	T	SNP	T	stop_gained	HIGH	`ITGA9	protein_coding	`24/28	c.2548C>T	2	18	2	0
chr5	38496637	rs3729740	C	T	SNP	T	splice_acceptor_variant&intron_variant	HIGH	`LIFR	nonsense_mediated_decay	`2/6	n.109-1G>A	4	16	2	2
chr5	55243415	.	G	A	SNP	A	stop_gained&splice_region_variant	HIGH	`IL6ST	protein_coding	`13/15	c.1843C>T	13	7	13	0
chr5	55243415	.	G	A	SNP	A	stop_gained&splice_region_variant	HIGH	`IL6ST	protein_coding	`15/17	c.1843C>T	13	7	13	0
chr5	55243415	.	G	A	SNP	A	stop_gained&splice_region_variant	HIGH	`IL6ST	protein_coding	`14/16	c.1843C>T	13	7	13	0
chr5	55243415	.	G	A	SNP	A	stop_gained&splice_region_variant	HIGH	`IL6ST	protein_coding	`12/14	c.1660C>T	13	7	13	0

Supplementary Table S1-4(Continued)

chr6	93956653	.	G	T	SNP	T	stop_gained	HIGH	`EPHA7	protein_coding	`15/17	c.2583C>A	14	5	12	2
chr8	48805816	rs11411516	A	AG	INS	AG	frameshift_variant&splice_region_variant	HIGH	`PRKDC	protein_coding	`31/87	c.3729dupC	20	0	0	20
chr8	48805816	rs11411516	A	AG	INS	AG	frameshift_variant&splice_region_variant	HIGH	`PRKDC	protein_coding	`31/86	c.3729dupC	20	0	0	20
chr8	1.46E+08	.	A	C	SNP	C	splice_donor_variant&intron_variant	HIGH	`RECQL4	protein_coding	`15/21	c.2463+2T>G	7	13	7	0
chr8	1.46E+08	.	A	C	SNP	C	splice_donor_variant&intron_variant	HIGH	`RECQL4	processed_transcript	`13/19	n.2590+2T>G	7	13	7	0
chr8	1.46E+08	.	A	C	SNP	C	splice_donor_variant&intron_variant	HIGH	`RECQL4	processed_transcript	`14/20	n.2506+2T>G	7	13	7	0
chr8	1.46E+08	rs11342077	CG	C	DEL	C	frameshift_variant&splice_region_variant	HIGH	`RECQL4	protein_coding	`15/22	c.2296delC	20	0	0	20
chr9	1.34E+08	.	A	AC	INS	AC	frameshift_variant	HIGH	`NUP214	protein_coding	`29/36	c.4484_4485insC	18	2	8	10
chr9	1.34E+08	.	A	AC	INS	AC	frameshift_variant	HIGH	`NUP214	protein_coding	`29/36	c.4481_4482insC	18	2	8	10
chr9	1.34E+08	.	A	AC	INS	AC	frameshift_variant	HIGH	`NUP214	protein_coding	`29/36	c.4451_4452insC	18	2	8	10
chr9	1.34E+08	.	A	AC	INS	AC	frameshift_variant	HIGH	`NUP214	protein_coding	`4/11	c.959_960insC	18	2	8	10
chr9	1.34E+08	.	A	AC	INS	AC	frameshift_variant	HIGH	`NUP214	protein_coding	`2/2	c.812_813insC	18	2	8	10
chr10	89720633	.	CT	C,CTT	DEL,INS	T	splice_acceptor_variant&intron_variant	HIGH	`PTEN	protein_coding	`7/8	c.802-3dupT	20	0	20	0
chr10	89720633	.	CT	C,CTT	DEL,INS	T	splice_acceptor_variant&intron_variant	HIGH	`PTEN	processed_transcript	`2/2	n.229-3dupT	20	0	20	0
chr17	7579470	.	C	CG	INS	CG	frameshift_variant	HIGH	`TP53	protein_coding	`4/11	c.216dupC	8	12	7	1

Supplementary Table S1-4(Continued)

chr17	7579470	.	C	CG	INS	CG	frameshift_variant	HIGH	`TP53	protein_coding	`3/7	c.216dupC	8	12	7	1
chr17	7579470	.	C	CG	INS	CG	frameshift_variant	HIGH	`TP53	protein_coding	`3/9	c.216dupC	8	12	7	1
chr17	7579470	.	C	CG	INS	CG	frameshift_variant	HIGH	`TP53	protein_coding	`4/12	c.216dupC	8	12	7	1
chr17	7579470	.	C	CG	INS	CG	frameshift_variant	HIGH	`TP53	protein_coding	`4/12	c.216dupC	8	12	7	1
chr17	7579470	.	C	CG	INS	CG	frameshift_variant	HIGH	`TP53	protein_coding	`4/11	c.216dupC	8	12	7	1
chr17	7579470	.	C	CG	INS	CG	frameshift_variant	HIGH	`TP53	protein_coding	`4/5	c.216dupC	8	12	7	1
chr17	7579470	.	C	CG	INS	CG	frameshift_variant	HIGH	`TP53	protein_coding	`4/5	c.216dupC	8	12	7	1
chr17	7579470	.	C	CG	INS	CG	frameshift_variant	HIGH	`TP53	protein_coding	`5/6	c.216dupC	8	12	7	1
chr18	48573537	.	G	T	SNP	T	stop_gained	HIGH	`SMAD4	protein_coding	`2/12	c.121G>T	1	19	1	0
chr18	48573537	.	G	T	SNP	T	stop_gained	HIGH	`SMAD4	protein_coding	`3/9	c.121G>T	1	19	1	0
chr18	48573537	.	G	T	SNP	T	stop_gained	HIGH	`SMAD4	protein_coding	`2/3	c.121G>T	1	19	1	0
chr18	48573537	.	G	T	SNP	T	stop_gained	HIGH	`SMAD4	protein_coding	`2/12	c.121G>T	1	19	1	0
chr18	48573537	.	G	T	SNP	T	stop_gained	HIGH	`SMAD4	protein_coding	`2/3	c.121G>T	1	19	1	0
chr18	48573537	.	G	T	SNP	T	stop_gained	HIGH	`SMAD4	protein_coding	`2/3	c.121G>T	1	19	1	0
chr18	48573537	.	G	T	SNP	T	stop_gained	HIGH	`SMAD4	protein_coding	`2/4	c.121G>T	1	19	1	0
chr18	48573537	.	G	T	SNP	T	stop_gained	HIGH	`SMAD4	protein_coding	`2/3	c.121G>T	1	19	1	0

Supplementary Table S1-5. The number of known variants identified in each of the tumors by Comprehensive Cancer Panel (CCP) analysis (By Macrogen)

Sample	nSNPs	nInsertions	nDeletions	nComplex	nTotal	nHets	nHomVar	nSingletons	heterozygosity	hetHomRatio
115026	616	30	50	38	734	456	278	23	1.00E-07	1.64
13-1	586	34	51	33	704	398	306	0	1.00E-07	1.3
13-2	584	33	52	32	701	397	304	1	1.00E-07	1.31
17-1	606	39	57	31	733	434	299	6	1.00E-07	1.45
17-2	581	36	52	27	696	330	366	0	1.00E-07	0.9
18-1	554	31	51	37	673	356	317	2	1.00E-07	1.12
18-2	552	33	46	37	668	347	321	0	1.00E-07	1.08
2-1	587	33	50	33	703	398	305	3	1.00E-07	1.3
2-2	561	34	48	33	676	330	346	0	1.00E-07	0.95
20-1	563	33	55	34	685	400	285	16	1.00E-07	1.4
20-2	538	35	50	36	659	308	351	16	1.00E-07	0.88
34629	591	29	52	34	706	389	317	18	1.00E-07	1.23
5-1	581	34	51	35	701	416	285	0	1.00E-07	1.46
5-2	582	34	51	33	700	418	282	1	1.00E-07	1.48
6-1	592	36	54	33	715	411	304	2	1.00E-07	1.35
6-2	541	34	49	36	660	331	329	0	1.00E-07	1.01
7-1	596	34	51	38	719	393	326	0	1.00E-07	1.21
7-2	598	34	46	37	715	385	330	1	1.00E-07	1.17
8-1	635	40	52	35	762	483	279	8	2.00E-07	1.73
8-2	563	40	45	33	681	326	355	0	1.00E-07	0.92

Supplementary Table S1-6: The number of novel variants identified in each of the tumors by Comprehensive Cancer Panel (CCP) analysis (By Macrogen)

Sample	nSNPs	nInsertions	nDeletions	nComplex	nTotal	nHets	nHomVar	nSingletons	heterozygosity	hetHomRatio
115026	42	1	15	2	60	57	3	13	0	19
13-1	40	3	15	3	61	57	4	2	0	14.2
13-2	35	3	13	3	54	53	1	3	0	53
17-1	34	2	15	2	53	49	4	0	0	12.2
17-2	33	3	15	4	55	52	3	0	0	17.3
18-1	42	2	18	3	65	63	2	3	0	31.5
18-2	36	2	14	2	54	53	1	0	0	53
2-1	42	4	13	4	63	58	5	4	0	11.6
2-2	29	3	13	3	48	42	6	0	0	7
20-1	36	2	14	2	54	48	6	11	0	8
20-2	31	2	13	3	49	43	6	5	0	7.17
34629	28	1	13	4	46	45	1	8	0	45
5-1	34	4	14	3	55	51	4	0	0	12.8
5-2	35	4	17	3	59	58	1	1	0	58
6-1	32	3	11	3	49	44	5	3	0	8.8
6-2	33	3	16	3	55	49	6	4	0	8.17
7-1	37	2	13	4	56	54	2	3	0	27
7-2	24	2	14	4	44	42	2	0	0	21
8-1	44	1	17	4	66	58	8	1	0	7.25
8-2	42	1	16	2	61	58	3	4	0	19.3

Supplementary Table S1-7: Coverage of the Comprehensive Cancer Panel (CCP) analysis (By Macrogen)

Sample	×1(%)	×5(%)	×10(%)	×15(%)	×20(%)	×25(%)	×30(%)	×50(%)	total_bases_ aligned_bases	mean_ coverage_ depth
115026	99.8	99.6	99.5	99.3	99.3	99.2	99.1	98.9	2,978,186,056	1760
13-1	99.8	99.6	99.5	99.4	99.3	99.2	99.2	98.9	2,723,116,470	1610
13-2	99.9	99.7	99.5	99.4	99.3	99.2	99.1	98.9	2,109,245,791	1250
17-1	99.8	99.6	99.5	99.4	99.3	99.2	99.1	98.9	2,389,805,922	1420
17-2	99.8	99.6	99.5	99.3	99.3	99.2	99.1	98.8	2,121,760,748	1260
18-1	99.9	99.7	99.6	99.5	99.4	99.3	99.2	99	3,010,370,819	1780
18-2	99.9	99.8	99.6	99.6	99.5	99.4	99.3	99.1	3,051,246,526	1810
2-1	99.8	99.6	99.5	99.4	99.4	99.3	99.3	99.1	3,330,983,419	1970
2-2	99.8	99.6	99.3	99.1	98.9	98.7	98.6	98.1	1,964,565,898	1160
20-1	99.8	99.6	99.5	99.4	99.3	99.1	99	98.7	3,212,141,634	1900
20-2	99.7	99.5	99.3	99.3	99.2	99.1	99	98.8	2,369,919,932	1400
34629	99.9	99.7	99.6	99.5	99.4	99.3	99.2	99	2,632,180,839	1560
5-1	99.8	99.6	99.5	99.4	99.4	99.3	99.2	99	3,594,801,890	2130
5-2	99.9	99.7	99.6	99.5	99.5	99.4	99.3	99.1	2,416,077,435	1430
6-1	99.8	99.5	99.4	99.2	99	98.9	98.8	98.5	1,787,704,037	1060
6-2	99.9	99.6	99.5	99.4	99.3	99.2	99.2	98.9	2,494,578,771	1480
7-1	99.8	99.6	99.5	99.3	99.3	99.1	99.1	98.8	3,088,396,179	1830
7-2	99.9	99.7	99.6	99.5	99.4	99.3	99.2	99	2,645,211,631	1570
8-1	99.8	99.6	99.4	99.4	99.3	99.2	99.1	98.8	2,457,867,852	1460
8-2	99.8	99.5	99.3	99.1	99	98.9	98.9	98.6	2,224,617,317	1320

Each column indicates the depths, ranging from ×1 to ×50, and each cell represents the proportion of genes with a higher depth than indicated.

Supplementary Table 2-1. Duration of PDX growth in the first xenograft

Sample No.	Implantation date	Harvest date	Period of 1st graft
056797D	2014-06-25	2014-10-26	123
057689E	2014-06-27	2014-10-07	102
059390	2014-07-03	2014-11-06	126
059801E	2014-07-03	2014-10-07	96
059873C	2014-07-04	2014-10-30	118
067233	2014-07-24	2015-02-12	203
066949F	2014-07-24	2015-03-05	224
067462G	2014-07-25	2015-03-31	249
070151E	2014-08-01	2014-10-29	89
070212E	2014-08-01	2014-12-26	147
075535E	2014-08-19	2015-03-05	198
077592I	2014-08-25	2014-10-30	66
087556	2014-09-25	2015-03-04	160
093376F	2014-10-14	2015-01-29	107
000777	2015-01-05	2015-03-10	64
002079J	2015-01-08	2015-03-11	62
021906	2015-03-10	2015-12-03	268
047831F	2015-05-22	2015-06-23	32
050786I	2015-05-29	2015-07-20	52
052433I	2015-06-05	2016-03-31	300
054093E	2015-06-12	2016-01-12	214
055279F	2015-06-19	2015-11-06	140
055431	2015-06-22	2015-10-23	123
057239F	2015-06-30	2015-08-13	44
057318D	2015-06-30	2015-11-06	129
057409F	2015-06-30	2016-01-11	195
057539F	2015-06-30	2015-09-18	80
057932G	2015-07-02	2015-10-02	92
058220	2015-07-03	2015-11-06	126
060307E	2015-07-10	2016-01-12	186
064080D	2015-07-21	2015-12-23	155
068083A	2015-08-03	2015-12-23	142

Supplementary Table 2-2. Duration of PDX growth in thesecond xenograft

Sample No.	2nd Implantation date	Harvest date	Period of growth
057689E	2014-10-07	2015-01-19	104
059801E	2014-10-07	2014-12-23	77
059801E	2014-10-07	2014-12-26	80
056797D	2014-10-16	2015-02-16	123
070151E	2014-10-29	2014-12-08	40
077592I	2014-10-30	2015-03-04	125
059873C	2014-10-30	2015-01-13	75
059873C	2014-10-30	2014-12-12	43
059390	2014-11-06	2015-01-29	84
093376F	2015-01-29	2015-07-10	162
087556	2015-03-04	2015-07-10	128
067233	2015-02-12	2015-07-10	148
066949F	2015-03-05	2015-07-10	127
002079J	2015-03-11	2015-05-18	68
057239F	2015-08-13	2015-09-18	36
050786I	2015-09-18	2015-11-06	49
057932G	2015-10-02	2015-12-02	61
055431	2015-10-23	2015-12-23	61
057318D	2015-11-06	2015-12-31	55
15S-068083	2015-12-23	2016-03-17	85
15S 064080	2015-12-23	2016-03-17	85
15S-060307E	2016-01-11	2016-03-17	66

Supplementary Table 2- 3. Clinical Information on the TNBCs that successfully formed PDX

No	Diagnosis	Age	ER	PR	HER2	HER2 SISH	Sub type	size (cm)	LN_+	LN_ total	Recur	death	T	N	M	1st period	2nd period
1	RECURRENT INVASIVE DUCTAL CARCINOMA	29	0	0	0	x	TN	6	3	3	1	S	3	3	0	268	-
2	RESIDUAL INVASIVE DUCTAL CARCINOMA	41	0	0	0	x	TN	6.5	15	18	1	S	3	3	1	64	-
3	RESIDUAL INVASIVE DUCTAL CARCINOMA	48	0	0	0	x	TN	4.8	1	10	0	S	3	1	0	62	68
4	RESIDUAL INVASIVE DUCTAL CARCINOMAS (X2)	53	0	0	0	x	TN	4	0	0	0	S	3	0	0	32	-
5	RESIDUAL METAPLASTIC CARCINOMA	42	0	0	0	x	TN	1.5	0	3	0	S	2	0	0	52	49
6	INVASIVE DUCTAL CARCINOMA	34	0	0	2	o	HER2	3.5	1	21	x	S	2	1	0	300	
7	RESIDUAL INVASIVE DUCTAL CARCINOMA	59	0	0	0	x	TN	2.3	3	11	0	S	2	1	0	214	-
8	RESIDUAL INVASIVE DUCTAL CARCINOMA	35	0	0	0	x	TN	7.7	3	23	0	S	3	3	0	140	-
9	RECURRENT ADENOID CYSTIC CARCINOMA SOLID AND CRIBRIFORM TYPE	55	0	0	0	x	TN	2.1	0	3	1	S	1	0	0	123	61
10	RESIDUAL INVASIVE DUCTAL CARCINOMA	31	2	0	1	x	TN	4	3	6	1	D	2	1	1	44	36
11	RESIDUAL INVASIVE DUCTAL CARCINOMA	34	0	0	0	x	TN	2.2	1	3	0	S	2	1	0	129	55
12	RESIDUAL INVASIVE DUCTAL CARCINOMA	43	0	0	0	x	TN	2.5	2	4	1	S	3	1	1	80	-
13	INVASIVE DUCTAL CARCINOMA, NUCLEAR	28	0	0	0	x	TN	1.8	0	2	0	S	1	0	0	102	104
14	RESIDUAL INVASIVE DUCTAL CARCINOMA	40	0	0	0	x	TN	9	12	17	1	N/A	3	3	0	92	61
15	RESIDUAL INVASIVE DUCTAL CARCINOMA	42	4	0	3	x	ER/HER2	8	21	22	0	N/A	2	3	1	126	-
16	RESIDUAL INVASIVE DUCTAL CARCINOMAS, multifocal	39	0	0	0	x	TN	2.6	0	6	1	S	2	0	0	96	77/80
17	INVASIVE MICROPAPILLARY CARCINOMA	57	0	0	0	x	TN	2	2	13	0	S	1	1	0	118	75/43
18	RESIDUAL INVASIVE DUCTAL CARCINOMA	42	3	3	2	o	ER/PR/HER2	4.8	2	8	0	S	2	1	0	186	66
19	RESIDUAL INVASIVE DUCTAL CARCINOMA	48	0	0	0	x	TN	5.6	0	3	0	S	2	0	0	155	85
20	RESIDUAL INVASIVE DUCTAL CARCINOMA	32	0	0	0	x	TN	2.1	1	10	1	S	2	1	0	249	-
21	RESIDUAL INVASIVE DUCTAL CARCINOMAS, MULTIPLE	37	0	0	3	x	HER2	5.2	4	18	0	S	3	3	0	142	85
22	RESIDUAL INVASIVE DUCTAL CARCINOMA	43	0	0	0	x	TN	2.4	2	14	0	S	2	1	0	198	-
23	RESIDUAL INVASIVE DUCTAL CARCINOMA	36	0	0	0	x	TN	2.5	0	5	0	S	2	0	0	66	125
24	RESIDUAL INVASIVE DUCTAL CARCINOMA	41	0	0	0	x	TN	7	15	25	0	S	3	1	1	160	128
25	INVASIVE DUCTAL CARCINOMA	24	0	0	0	x	TN	2.1	0	3	1	S	3	0	0	107	162

Supplementary Table 2-4. A summary of comparative IHC analysis of primary and PDX tumors (By Pathology Department)

No	Dx	Primary_Histology	op_ER	op_PR	op_HER2	op_Ki67	Xenograft_Histology
56797	IDC	small & large cluster	3+1=4	0	2+	60-70	diffuse sheet
57689	metaplastic	large cluster	0	0	0	80-90	diffuse sheet
59390	metaplastic	large cluster	0	0	0	80-90	large cluster
59801	IDC	small & large cluster	0	0	0	70-80	small & large cluster
59873	IDC	small cluster	0	0	0	>90	small cluster
77592	IDC	large cluster	0	0	0	80-90	large cluster
70212	IDC	small cluster	3+5	0	2		Large cluster
93376	IDC	large cluster	0	0	1		(m,s) Large cluster
67233	IDC	small cluster	3+3	0	3		(m) Large cluster
87556	IDC	small cluster	0	0	0		(s) small cluster
66949	IDC	small & Large cluster	0	0	0		(m) Large cluster
75535	IDC	Large cluster	0	0	0		(m) Large cluster
70151	IDC	small cluster	0	0	3+	30-40	diffuse sheet

Supplementary Table 2-5. The list of rare variants with a high impact, as determined by CCP analysis (By Macrogen and Suhwan Chang)

CHR	POS	ID	REF	ALT	Variant Type	QUAL	DP	Allele	Effect	Impact	Gene_Name	HGVS.c	HGVS.p
1	144852390	rs61804988	C	T	SNP	1419772.14	158396	T	stop_gained	HIGH	`PDE4DIP	c.7053G>A	p.Trp2351*
1	145075683	rs2762779	C	T	SNP	1685391.14	197288	T	stop_gained	HIGH	`PDE4DIP	c.180G>A	p.Trp60*
2	190670539	.	TA	TAA,T	INS,DEL	169985.94	37272		frameshift_variant	HIGH	`PMS1	c.488delA	p.Lys163fs
2	190670539	.	TA	TAA,T	INS,DEL	169985.94	37272	A	frameshift_variant	HIGH	`PMS1	c.488dupA	p.Leu164fs
2	190719499	.	G	T	SNP	443.8	7309	T	stop_gained	HIGH	`PMS1	c.1501G>T	p.Gly501*
2	216299548	.	C	G	SNP	209.07	43	G	splice_acceptor_variant&intron_variant	HIGH	`FN1	c.149-1G>C	
2	223158885	.	C	G	SNP	4508.41	16748	G	splice_donor_variant&intron_variant	HIGH	`PAX3	c.586+1G>C	
3	52439899	.	A	AG	INS	888414.1	94154	AG	frameshift_variant	HIGH	`BAP1	c.812_813insC	p.Gln272fs
4	55973937	.	C	CA	INS	58613.38	57698	CA	frameshift_variant	HIGH	`KDR	c.1378dupT	p.Trp460fs
5	38496637	rs3729740	C	T	SNP	201115.55	44828	T	splice_acceptor_variant&intron_variant	HIGH	`LIFR	n.109-1G>A	
5	67591007	.	C	T	SNP	91391.41	73436	T	stop_gained	HIGH	`PIK3R1	c.1600C>T	p.Arg534*
5	176722037	.	A	AGAT	INS	929777.1	53839	ATCG	frameshift_variant	HIGH	`NSD1	c.7669_767 0insATCG	p.Gly2557fs
6	39484	.	C	T	SNP	11061.32	39976	T	stop_gained	HIGH	`IRF4	c.244C>T	p.Arg82*
6	93956653	.	G	T	SNP	2304.71	13897	T	stop_gained	HIGH	`EPHA	c.2583C>A	p.Cys861*

												7		
												`SAMD		
7	92734381	.	G	A	SNP	14019.32	45590	A	stop_gained	HIGH	9	c.1030C>T	p.Arg344*	
		rs15007300							frameshift_variant&st		`KMT2			
7	151945071	7	G	GT	INS	552670.1	108343	GT	op_gained	HIGH	C	c.2447dupA	p.Tyr816fs	
									frameshift_variant&s		`PRKD			
8	48805816	rs11411516	A	AG	INS	762428.28	15903	AG	plice_region_variant	HIGH	C	c.3729dupC	p.Arg1244fs	
									splice_acceptor_varia		`CSMD	c.3311-		
8	113651141	.	C	T	SNP	2708.32	30099	T	nt&intron_variant	HIGH	3	1G>A		
									frameshift_variant&s		`RECQ			
8	145738767	rs11342077	CG	C	DEL	1556373.28	38482	C	plice_region_variant	HIGH	L4	c.2296delC	p.Arg766fs	
											`NUP2	c.4484_448		
9	134073362	.	A	AC	INS	2866.5	91	AC	frameshift_variant	HIGH	14	5insC	p.Glu1495fs	
											`NOTC			
9	139390966	.	G	A	SNP	8701.43	11462	A	stop_gained	HIGH	H1	c.7225C>T	p.Gln2409*	
											`NOTC			
9	139397637	.	G	A	SNP	15977.82	24075	A	stop_gained	HIGH	H1	c.5164C>T	p.Gln1722*	
									splice_donor_variant			c.634+1G>		
10	89712017	.	G	A	SNP	10487.32	31291	A	&intron_variant	HIGH	`PTEN	A		
									splice_acceptor_varia			c.802-		
10	89720633	.	CT	CTT,C	INS,DEL	31287.95	5813	T	nt&intron_variant	HIGH	`PTEN	3dupT		
		rs14665027	GTA									c.955_958d		
10	89720798	3	CT	G	DEL	18871.27	34765		frameshift_variant	HIGH	`PTEN	elACTT	p.Thr319fs	
											`CYP2			
10	96540410	rs4986893	G	A	SNP	37717.43	40263	A	stop_gained	HIGH	C19	c.636G>A	p.Trp212*	
												c.488_507d		
												elTCTGGG		
			TG~									ACCCCCC		
11	533301	.	A	T	DEL	29446.38	7973	T	frameshift_variant	HIGH	`HRAS	GGGACCC	p.Leu163fs	

11	47260369	.	G	A	SNP	6667.43	16491	A	stop_gained	HIGH	`DDB2	c.1253G>A	p.Trp418*
											`CCND		
12	4383368	.	C	G	SNP	442.43	17695	G	stop_gained	HIGH	2	c.162C>G	p.Tyr54*
									frameshift_variant&s			c.265-	
									plice_acceptor_varian			12_271delT	
									t&splice_region_varia			TTTGTTC	
			AT~						nt&splice_region_vari			CCAGGGA	
13	48916720	.	G	A	DEL	8841.35	9928		ant&intron_variant	HIGH	`RB1	GGTT	p.Gly89fs
									splice_donor_variant		`BUB1	n.452+1G>	
15	40462773	.	G	A	SNP	1135.42	24224	A	&intron_variant	HIGH	B	A	
									stop_gained&splice_r				
16	2120577	rs45480292	C	T	SNP	10177.43	47892	T	egion_variant	HIGH	`TSC2	c.1837C>T	p.Gln613*
									splice_donor_variant		`FANC	c.426+1G>	
16	89877336	.	C	T	SNP	42393.43	41726	T	&intron_variant	HIGH	A	A	
17	7577136	.	TC	T	DEL	61104.38	45663	T	frameshift_variant	HIGH	`TP53	c.801delG	p.Arg267fs
												c.454_466d	
			CG~									elCCGCCC	
17	7578463	.	GG	C	DEL	181274.38	29054	C	frameshift_variant	HIGH	`TP53	GGCACCC	p.Pro152fs
			GTC									c.412_418d	
			TTG									elGCCAAG	
17	7578511	.	GC	G	DEL	173085.38	29140	G	frameshift_variant	HIGH	`TP53	A	p.Ala138fs
17	7579414	.	C	T	SNP	3527.39	7112	T	stop_gained	HIGH	`TP53	c.273G>A	p.Trp91*
17	7579462	.	AG	A	DEL	18398.35	7628	A	frameshift_variant	HIGH	`TP53	c.224delC	p.Pro75fs
17	29683525	.	A	T	SNP	198.47	6027	T	stop_gained	HIGH	`NF1	c.7663A>T	p.Lys2555*
											`COL1		
17	48266848	.	T	A	SNP	30173.35	34904	A	stop_gained	HIGH	A1	c.2719A>T	p.Lys907*
									stop_gained&splice_r				
17	75484811	.	G	A	SNP	633.2	16587	A	egion_variant	HIGH	`SEPT9	c.1127G>A	p.Trp376*
17	78317127	.	T	C	SNP	73412.39	47377	C	splice_donor_variant	HIGH	`RNF21	c.6330+2T>	

									&intron_variant		3	C	
19	1207033	.	A	T	SNP	34390.39	43857	T	stop_gained	HIGH	`STK11	c.121A>T	p.Lys41*
19	41744438	.	G	A	SNP	30046.43	24097	A	stop_gained	HIGH	`AXL	c.1058G>A	p.Trp353*
19	42795107	.	T	G	SNP	362.43	24583	G	stop_gained	HIGH	`CIC	c.4914T>G	p.Tyr1638*
												c.2789_279	
20	57485023	.	CTG	C	DEL	135737.38	63855		frameshift_variant	HIGH	`GNAS	0delTG	p.Val930fs
									splice_acceptor_varia		`CYP2		
22	42525188	.	C	T	SNP	10871.43	30353	T	nt&intron_variant	HIGH	D6	c.353-1G>A	
									splice_acceptor_varia				
X	48117969	.	A	G	SNP	20379.43	23295	G	nt&intron_variant	HIGH	`SSX1	c.185-2A>G	

Supplementary Table 2- 6. The list of the top 10 most frequent rare variants with predicted strong effects as determined by CCP analysis (By Suhwan Chang)

CHR	POS	REF	ALT	Variant Type	Effect	Gene_ Name	HGVS.c	HGVS.p	# of Affected	# of Homo Ref	# of Hetero	# of Homo Alt
10	89720633	CT	CTT,C	INS,DEL	splice_acceptor_variant&intron_variant	`PTEN	c.802-3dupT		19	0	19	0
17	48266848	T	A	SNP	stop_gained	`COL1A1	c.2719A>T	p.Lys907*	19	1	19	0
9	134073362	A	AC	INS	frameshift_variant	`NUP214	c.4484_4485insC	p.Glu1495fs	18	0	9	9
9	139397637	G	A	SNP	stop_gained	`NOTCH1	c.5164C>T	p.Gln1722*	14	6	14	0
2	190670539	TA	TAA,T	INS,DEL	frameshift_variant	`PMS1	c.488delA	p.Lys163fs	12	8	12	0
6	93956653	G	T	SNP	stop_gained	`EPHA7	c.2583C>A	p.Cys861*	10	10	6	4
2	216299548	C	G	SNP	splice_acceptor_variant&intron_variant	`FN1	c.149-1G>C		6	6	4	2
17	75484811	G	A	SNP	stop_gained&splice_region_variant	`SEPT9	c.1127G>A	p.Trp376*	5	15	5	0
2	190719499	G	T	SNP	stop_gained	`PMS1	c.1501G>T	p.Gly501*	4	16	4	0
2	223158885	C	G	SNP	splice_donor_variant&intron_variant	`PAX3	c.586+1G>C		3	17	3	0
15	40462773	G	A	SNP	splice_donor_variant&intron_variant	`BUB1B	n.452+1G>A		3	17	3	0

Supplementary Table 2-7. The list of the top 10 most frequent rare variants with predicted moderate effects according to CCP analysis

CHR	POS	ID	REF	ALT	Gene_ Name	Gene_ID	HGVS.c	HGVS.p	1000G Maf	# of Samples Affected	# of Hetero	# of Homo Alt
5	112154843	rs201478136	A	G	`APC	ENSG00000134982	c.1114A>G	p.Asn372Asp	0.000457875	6	4	2
2	47656972	rs17224367	C	T	`MSH2	ENSG00000095002	c.1168C>T	p.Leu390Phe	0.004578755	6	6	0
3	142274770	rs77208665	T	C	`ATR	ENSG00000175054	c.2290A>G	p.Lys764Glu	0.004578755	6	6	0
4	1806633	rs56240927	C	T	`FGFR3	ENSG00000068078	c.1355C>T	p.Thr452Met	0.004578755	5	4	1
4	62758425	rs141921843	C	T	`LPHN3	ENSG00000150471	c.1532C>T	p.Thr511Ile	0.004120879	3	3	0
4	106157703	rs147836249	T	G	`TET2	ENSG00000168769	c.2667T>G	p.Phe889Leu	0.004120879	3	3	0
2	29416520	rs141010693	A	G	`ALK	ENSG00000171094	c.4433T>C	p.Met1478Thr	0	2	2	0
1	150785811	rs2275237	G	A	`ARNT	ENSG00000143437	c.2117C>T	p.Pro706Leu	0.000457875	2	2	0
1	156846256	rs55892037	T	C	`NTRK1	ENSG00000198400	c.1697T>C	p.Met566Thr	0.000457875	2	2	0
2	234669619	rs35350960	C	A	`UGT1A1	ENSG00000242366	c.686C>A	p.Pro229Gln	0.000457875	2	2	0

Supplementary 2-8. The most mismatched genes between primary and PDX tumor

CHROM	POS	ID	REF	ALT	Variant Type	Allele	Effect	Impact	Gene_Name	Gene_ID	Feature_Type	`mismatched pairs
7	140434597	rs61291709	GAAA	GAAAA,G	INS,DEL	G	intron_variant	MODIFIER	`BRAF	ENSG00000157764	transcript	10
7	140434597	rs61291709	GAAA	GAAAA,G	INS,DEL	GAAAA	intron_variant	MODIFIER	`BRAF	ENSG00000157764	transcript	10
7	140434597	rs61291709	GAAA	GAAAA,G	INS,DEL	G	intron_variant	MODIFIER	`BRAF	ENSG00000157764	transcript	10
7	140434597	rs61291709	GAAA	GAAAA,G	INS,DEL	GAAAA	intron_variant	MODIFIER	`BRAF	ENSG00000157764	transcript	10
7	140434597	rs61291709	GAAA	GAAAA,G	INS,DEL	G	intron_variant	MODIFIER	`BRAF	ENSG00000157764	transcript	10
7	140434597	rs61291709	GAAA	GAAAA,G	INS,DEL	GAAAA	intron_variant	MODIFIER	`BRAF	ENSG00000157764	transcript	10
7	140434597	rs61291709	GAAA	GAAAA,G	INS,DEL	G	intron_variant	MODIFIER	`BRAF	ENSG00000157764	transcript	10
7	140434597	rs61291709	GAAA	GAAAA,G	INS,DEL	GAAAA	intron_variant	MODIFIER	`BRAF	ENSG00000157764	transcript	10
4	1808017	.	G	A	SNP	A	downstream_gene_variant	MODIFIER	`FGFR3	ENSG00000068078	transcript	10
4	1808017	.	G	A	SNP	A	downstream_gene_variant	MODIFIER	`FGFR3	ENSG00000068078	transcript	10
4	1808017	.	G	A	SNP	A	downstream_gene_variant	MODIFIER	`FGFR3	ENSG00000068078	transcript	10
4	1808017	.	G	A	SNP	A	intron_variant	MODIFIER	`FGFR3	ENSG00000068078	transcript	10
16	15853537	.	A	G	SNP	G	non_coding_exon_variant	MODIFIER	`MYH11	ENSG00000133392	transcript	10
4	1808017	.	G	A	SNP	A	missense_variant	MODERATE	`FGFR3	ENSG00000068078	transcript	10
4	1808017	.	G	A	SNP	A	missense_variant	MODERATE	`FGFR3	ENSG00000068078	transcript	10
4	1808017	.	G	A	SNP	A	missense_variant	MODERATE	`FGFR3	ENSG00000068078	transcript	10
4	1808017	.	G	A	SNP	A	missense_variant	MODERATE	`FGFR3	ENSG00000068078	transcript	10
4	1808017	.	G	A	SNP	A	missense_variant	MODERATE	`FGFR3	ENSG00000068078	transcript	10
16	15853537	.	A	G	SNP	G	missense_variant	MODERATE	`MYH11	ENSG00000133392	transcript	10
16	15853537	.	A	G	SNP	G	missense_variant	MODERATE	`MYH11	ENSG00000133392	transcript	10
16	15853537	.	A	G	SNP	G	missense_variant	MODERATE	`MYH11	ENSG00000133392	transcript	10
16	15853537	.	A	G	SNP	G	missense_variant	MODERATE	`MYH11	ENSG00000133392	transcript	10

Supplementary Table 2-9. Frequent *NOTCH1* variants found in TNBC PDXs(by Suhwan Chang)

Sample Name	Chrom.	Position	Allele Call	Frequency	Quality	Allele Source	Description
2079	chr9	139,399,405	Homozygous	99.3	4282.8	Novel	NOTCH1:NM_017617:exon26:c.A4738C:p.M1580L,
54093	chr9	139,399,405	Heterozygous	3.8	15.8	Novel	NOTCH1:NM_017617:exon26:c.A4738C:p.M1580L,
2079	chr9	139,399,403	Homozygous	99.3	4282.8	Novel	NOTCH1:NM_017617:exon26:c.G4740T:p.M1580I,
54093	chr9	139,399,403	Heterozygous	3.8	15.6	Novel	NOTCH1:NM_017617:exon26:c.G4740T:p.M1580I,
2079	chr9	139,399,394	Homozygous	99.3	4301.7	Novel	NOTCH1:NM_017617:exon26:c.G4749C:p.E1583D,
54093	chr9	139,399,394	Heterozygous	4	17.9	Novel	NOTCH1:NM_017617:exon26:c.G4749C:p.E1583D,
2079	chr9	139,399,380	Homozygous	100	31963.1	Novel	NOTCH1:NM_017617:exon26:c.G4763A:p.S1588N,
54093	chr9	139,399,380	Heterozygous	32.9	1851.4	Novel	NOTCH1:NM_017617:exon26:c.G4763A:p.S1588N,
55279	chr9	139,399,380	Heterozygous	9.4	138.6	Novel	NOTCH1:NM_017617:exon26:c.G4763A:p.S1588N,
57318	chr9	139,399,380	Heterozygous	8.9	75.9	Novel	NOTCH1:NM_017617:exon26:c.G4763A:p.S1588N,
57539	chr9	139,399,380	Heterozygous	12.2	396.5	Novel	NOTCH1:NM_017617:exon26:c.G4763A:p.S1588N,
57932	chr9	139,399,380	Heterozygous	6.9	22.7	Novel	NOTCH1:NM_017617:exon26:c.G4763A:p.S1588N,
67462	chr9	139,399,380	Heterozygous	8.2	60.7	Novel	NOTCH1:NM_017617:exon26:c.G4763A:p.S1588N,
2079	chr9	139,399,350	Homozygous	100	31768.4	Novel	NOTCH1:NM_017617:exon26:c.G4793A:p.R1598H,
54093	chr9	139,399,350	Heterozygous	33.2	1888	Novel	NOTCH1:NM_017617:exon26:c.G4793A:p.R1598H,
55279	chr9	139,399,350	Heterozygous	10	156.3	Novel	NOTCH1:NM_017617:exon26:c.G4793A:p.R1598H,
55431	chr9	139,399,350	Heterozygous	3.6	16.6	Novel	NOTCH1:NM_017617:exon26:c.G4793A:p.R1598H,
57318	chr9	139,399,350	Heterozygous	9.3	81.9	Novel	NOTCH1:NM_017617:exon26:c.G4793A:p.R1598H,
57539	chr9	139,399,350	Heterozygous	12.8	434.1	Novel	NOTCH1:NM_017617:exon26:c.G4793A:p.R1598H,
67462	chr9	139,399,350	Heterozygous	8.4	61.4	Novel	NOTCH1:NM_017617:exon26:c.G4793A:p.R1598H,

Supplementary Table 2-10. The number of variants in a coding region or UTR for each of PDX (By Macrogen)

PDX sample	# of mutated genes in coding region	# of variants in coding region	# of mutated genes in UTR	# of variants in UTR
ID-000777	3756	5821	7907	22564
ID-002079	3347	5053	7261	21019
ID-021906	3276	4898	7205	20259
ID-047831	3174	4743	7354	20392
ID-050786	3118	4526	7344	19801
ID-052453	3505	5304	7585	21802
ID-053431	3594	5486	7911	23135
ID-054093	3492	5235	7645	22026
ID-055279	3506	5339	7516	22262
ID-057239	3145	4636	6968	18119
ID-057318	3333	4980	7222	20569
ID-057539	3189	4636	7118	19775
ID-057932	2987	4407	6804	17531
ID-058220	2982	4359	7035	18797
ID-060307	3498	5200	7796	21781
ID-064080	3204	4766	7382	20877
ID-57689_	3468	5223	7726	21564
ID-59801_	3752	5735	7694	22453
ID-59873_	3296	4863	7296	20349
ID-67462_	3128	4631	7248	20120
ID-75535_	3504	5411	7629	23076
ID-77592_	3475	5194	7196	20644
ID-87556_	4262	6672	8310	25806
ID-93376_	3253	4857	7172	20299

Supplementary Table 2-11. The top 10 exonic frameshift variants found in PDXs by RNA-seq analysis
(By Suhwan Chang)

Num of mutated sample	Chr	Start	End	Ref	Alt	Gene	Exonic Function	AA Change
16	chr17	17697102	17697102	G	-	RAI1	frameshift deletion	p.Q280fs
9	chr20	1559029	1559029	C	-	SIRPB1	frameshift deletion	p.D130fs
7	chr22	21841122	21841126	TCTTA	-	PI4KAP2	frameshift deletion	p.D160fs
7	chr1	31732603	31732603	A	-	SNRNP40	frameshift deletion	p.F154fs
7	chr19	52887146	52887146	A	-	ZNF880	frameshift deletion	p.K105fs
7	chr5	175388070	175388070	A	-	THOC3	frameshift deletion	p.I319fs
6	chr15	20648688	20648691	CAGA	-	HERC2P3	frameshift deletion	p.S705fs
6	chr17	26708303	26708304	GT	-	SARM1	frameshift deletion	p.V205fs
6	chr1	120612003	120612004	GG	-	NOTCH2	frameshift deletion	l:p.P6fs
6	chr2	128938487	128938488	TG	-	UGGT1	frameshift deletion	.N1308fs

Supplementary Table 2-12. The top 10 stopgain variants found in PDX by RNA-seq analysis (By Suhwan Chang)

# of mutated samples	Chr	Start	End	Ref	Alt	Function	Gene	Exonic Function
23	chr5	139936760	139936760	G	T	exonic	SRA1	stopgain
22	chr9	139937799	139937799	G	A	exonic	NPDC1	stopgain
22	chr7	64438667	64438667	G	A	exonic	ZNF117	stopgain
21	chr1	144852390	144852390	C	T	exonic	PDE4DIP	stopgain
17	chr9	86327446	86327446	C	T	exonic	AK300656	stopgain
17	chr19	59093464	59093464	C	T	exonic	MGC2752	stopgain
12	chr6	31124849	31124849	C	T	exonic	CCHCR1	stopgain
12	chr17	38556250	38556250	-	A	exonic	TOP2A	stopgain
11	chr17	74077797	74077797	C	T	exonic	ZACN	stopgain
10	chr1	17084304	17084304	C	T	exonic	MST1L	stopgain

Supplementary Table 2-13. The list of the top 10 most frequent, stoploss variants identified by RNA-seq
(By Suhwan Chang)

# of mutated samples	Chr	Start	End	Ref	Alt	Gene	Exonic Function	AA Change
18	chr2	145147020	145147020	A	C	ZEB2	stoploss	p.X1209E
8	chr17	62464464	62464464	A	G	MILR1	stoploss	:p.X106W
7	chr8	38458239	38458239	T	C	RNF5P1	stoploss	:p.X160W
4	chr10	102777342	102777342	T	C	PDZD7	stoploss	:p.X518W
3	chr9	116800	116800	C	G	FOXD4	stoploss	:p.X440Y
3	chr7	1009016	1009016	-	T	COX19	stoploss	1delinsM
2	chr19	13255323	13255323	T	A	STX10	stoploss	:p.X247C
2	chr6	29858259	29858259	A	T	HLA-H	stoploss	:p.X151C
2	chr11	111963920	111963920	A	G	SDHD	stoploss	:p.X144W
1	chr6	29798157	29798157	T	A	HLA-G	stoploss	:p.X344R

Supplementary Table 2-14. The number of in-frame gene fusions for each of PDX

PDX ID	# of gene fusions
ID-000777	18
ID-002079	37
ID-021906	35
ID-047831	50
ID-050786	17
ID-052453	35
ID-053431	16
ID-054093	27
ID-055279	27
ID-057239	34
ID-057318	17
ID-057539	20
ID-057932	46
ID-058220	16
ID-060307	24
ID-064080	27
ID-57689_	36
ID-59801_	23
ID-59873_	55
ID-67462_	23
ID-75535_	30
ID-77592_	59
ID-87556_	22
ID-93376_	26

Supplementary Table 5-1. Primer sequences for qPCR

Gene	Forward primer	Reverse primer
ERBB2	5′-AGC CGC GAG CAC CCA AGT-3′	5′-TTG GTG GGC AGG TAG GTG AGT T-3′
HEY1	5′-TGG ATC ACC TGA AAA TGC TG-3′	5′-CGA AAT CCC AAA CTC CGA TA-3′
SVIL	5'-CTC CAT GTC TGC GAG TCA AA-3'	5'-CAG CAG GCT TAC CAA AGT CC-3'
TEAD4	5'-TGA TGC AGA GGG TGT ATG GA-3’	5'-GAT CAG CTC ATT CCG ACC AT-3'
RUNX3	5'-AGC ACG GAG CAG AGG AAG T-3’	5'-TCG GAA CTG AAC CCA TTC TC-3’
FOXM1	5'-CCA CTC TTC CAA GGG AGG GCT C-3'	5'-TAG GAC TTC TTG GGT CTT GGG GTG-3’
PATJ	5'-CAA GAT CCC TCA CCA TCC AT-3'	5'-GTG CCT CCT GGG AAT AGA CA-3'
MPP5	5'-CTT GCA TCA CCA GAG GAA CA-3’	5'-TTC CTC TCT CCT ACG CCT CA-3’
MST1	5'-GAT GTG GCA GAT GCT GAA GA-3'	5'-GCT GCT CAC GTT GTA GTG GA-3'
MST2	5'-GAT GAG CTG GAT TCC CAC AC-3'	5'-TCC ATC TTC TTC TTC CTC ATC C-3'
TAOK3	5'-AAC ATG GAA CCA GCC TGA AC-3'	5'-ATT GTG CTT TCG TCA TCG TG-3'
HIPK3	5'-AAA CCA ACA TGG GAA ATC CA-3'	5'-GAC GGG CAT AAG AAG GAT GA-3'
CSNK1G2	5'-AAG GAG CGG TAC CAG AAG ATC G-3'	5'-GAA GAG CTT CCG CAG GTA GTC A-3'
CTGF	5'-CAA GGG CCT CTT CTG TGA CT-3’	5'-ACG TGC ACT GGT ACT TGC AG-3’
YAP	5'-CCT TCT TCA AGC CGC CGG AG-3’	5'-CAG TGT CCC AGG AGA AAC AGC-3’
hRPL13a	5'-CGA AGA TGG CGG AGG TGC AG-3'	5'-GGT TTT GTG GGG CAG CAT AC-3'

Supplementary Table 5-2. Primer sequences for ChIP

Site /Segment #	Forward primer	Reverse Primer
Primer for 4 site which have TEAD binding sequence or similar sequence		
Site 1	5'-GCC AGC TCT TAC TTT CTC AAG C-3'	5'-CAC CTG TAG TCC CAG CTA CTC TG-3'
Site 2	5'-AGG ACT CTC TTC CTG GCT TAC AG-3'	5'-GGT CAA AAG GGT GGA TCA CTA AT-3'
Site 3	5'-CGA CAA AAC CAA AAA TTA AAA CG-3'	5'-TGT CTG GTA CCC GCA TTT ATA CT-3'
Site 4	5'-TTG CTG TAG GCT GTA TGC TGT TA-3'	5'-AAC ATC CCA GTG ACC AGA TTA TG-3'
Primer for 9 segment of 2kb promoter region of miR-155		
Segment1	5'-ACG AGC TCT TAC GCG TGC TA-3'	5'-GAC AGA CAG GAG GGC TTC AG-3'
Segment2	5'-TTG CTC ATT CCG AAT CAC TG-3	5'-GCC GTG GAC TTT CCT GAC C-3
Segment3	5'-TCC TGT CTG TCC TGG TGA GG-3'	5'-GCT CCC CTT TCC CTT CCT-3
Segment4	5'-CAT TCG ACT GCG TTT CCT TT-3'	5'-ACC TCC GGC TCG AAA AAC-3'
Segment5	5'-GAA GGG AAA GGG GAG CAC-3	5'-TCG TGA CTC ATA ACC GAC CA-3'
Segment6	5'-AGT CAC GAG CGG GCT ATA AA-3'	5'-CTC TCC CCA TCC GAT CAC-3’
Segment7	5'-TTG GGG AAA CAG AGT TAG GC-3'	5'-AGG GTT AGC CAG GCT TCA TT-3'
Segment8	5'-TGA GCT TCT GTG CCT GTT TG-3	5'-GGT CCA TTC CTG AAA GCT GA-3'
Segment9	5'-CGT GCC AAG GTG AGA TTT TT-3	5'-GCT TTA CAA CAG TAC CGG ATG C-3'

Supplemetary Table 5-3. siRNA primer sequences

YAP1	GAGAAAUUUACUAUAUAAAUCUCUUUAUAUAGUAAAUUUCUCUU
TAZ	CCGAAUUCCUGCGUUUCAAUCUCUUGAAACGCAGGAAUUCGGUU;
FOXM1	CUAUCA AGGAGGAAGAAAUCCAGCCUCUCGGCUGGAUUUCUUCCUCCUUGAUAGUU
RUNX3:	CAGGCAAUGACGAGAACUACUCCGCUCUCGCGGAGUAGUUCUCGUCAUUGCCUGUU
SVIL	5’ GAGAACAAGGGAAUGUUGAUU 3’ and : 5’ UCAACAUUCCCUUGUUCUCUU 3’

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. International journal of cancer. 2015;136(5):E359-86.

2. Torre LA, Siegel RL, Ward EM, Jemal A. Global Cancer Incidence and Mortality Rates and Trends--An Update. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2016;25(1):16-27.

3. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012;483(7391):603-7.

4. van Weerden WM, Romijn JC. Use of nude mouse xenograft models in prostate cancer research. The Prostate. 2000;43(4):263-71.

5. DiMasi JA, Reichert JM, Feldman L, Malins A. Clinical approval success rates for investigational cancer drugs. Clinical pharmacology and therapeutics. 2013;94(3):329-35.

6. Jung J, Lee CH, Seol HS, Choi YS, Kim E, Lee EJ, et al. Generation and molecular characterization of pancreatic cancer patient-derived xenografts reveals their heterologous nature. Oncotarget. 2016;7(38):62533-46.

7. Ko AH. Progress in the treatment of metastatic pancreatic cancer and the search for next opportunities. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(16):1779-86.

8. Kamisawa T, Wood LD, Itoi T, Takaori K. Pancreatic cancer. Lancet (London, England). 2016;388(10039):73-85.

9. Heestand GM, Murphy JD, Lowy AM. Approach to patients with pancreatic cancer without detectable metastases. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2015;33(16):1770-8.

10. Gillen S, Schuster T, Meyer Zum Buschenfelde C, Friess H, Kleeff J. Preoperative/neoadjuvant therapy in pancreatic cancer: a systematic review and meta-analysis of response and resection percentages. PLoS medicine. 2010;7(4):e1000267.

11. Spadi R, Brusa F, Ponzetti A, Chiappino I, Birocco N, Ciuffreda L, et al. Current

therapeutic strategies for advanced pancreatic cancer: A review for clinicians. World journal of clinical oncology. 2016;7(1):27-43.

12. Hingorani SR, Petricoin EF, Maitra A, Rajapakse V, King C, Jacobetz MA, et al. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer cell. 2003;4(6):437-50.

13. Cook N, Olive KP, Frese K, Tuveson DA. K-Ras-driven pancreatic cancer mouse model for anticancer inhibitor analyses. Methods in enzymology. 2008;439:73-85.

14. Colvin EK, Scarlett CJ. A historical perspective of pancreatic cancer mouse models. Seminars in cell & developmental biology. 2014;27:96-105.

15. Mattie M, Christensen A, Chang MS, Yeh W, Said S, Shostak Y, et al. Molecular characterization of patient-derived human pancreatic tumor xenograft models for preclinical and translational development of cancer therapeutics. Neoplasia (New York, NY). 2013;15(10):1138-50.

16. Delitto D, Pham K, Vlada AC, Sarosi GA, Thomas RM, Behrns KE, et al. Patient-derived xenograft models for pancreatic adenocarcinoma demonstrate retention of tumor morphology through incorporation of murine stromal elements. The American journal of pathology. 2015;185(5):1297-303.

17. Melisko ME, Gradishar WJ, Moy B. Issues in Breast Cancer Survivorship: Optimal Care, Bone Health, and Lifestyle Modifications. American Society of Clinical Oncology educational book American Society of Clinical Oncology Meeting. 2016;35:e22-9.

18. Bhattacharya R, Banerjee K, Mukherjee N, Sen M, Mukhopadhyay A. From molecular insight to therapeutic strategy: The holistic approach for treating triple negative breast cancer. Pathology, research and practice. 2017;213(3):177-82.

19. Evans KW, Yuca E, Akcakanat A, Scott SM, Arango NP, Zheng X, et al. A Population of Heterogeneous Breast Cancer Patient-Derived Xenografts Demonstrate Broad Activity of PARP Inhibitor in BRCA1/2 Wild-Type Tumors. Clinical cancer research : an official journal of the American Association for Cancer Research. 2017;23(21):6468-77.

20. Arango NP, Yuca E, Zhao M, Evans KW, Scott S, Kim C, et al. Selinexor (KPT-330) demonstrates anti-tumor efficacy in preclinical models of triple-negative breast cancer. 2017;19(1):93.

21. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. Nature reviews Clinical oncology. 2016;13(11):674-90.

22. Irshad S, Ellis P, Tutt A. Molecular heterogeneity of triple-negative breast cancer and its clinical implications. Current opinion in oncology. 2011;23(6):566-77.

23. Ortmann O, Weiss JM, Diedrich K. Gonadotrophin-releasing hormone (GnRH) and GnRH agonists: mechanisms of action. Reproductive biomedicine online. 2002;5 Suppl 1:1-7.

24. Jayes FC, Britt JH, Esbenshade KL. Role of gonadotropin-releasing hormone pulse frequency in differential regulation of gonadotropins in the gilt. Biology of reproduction. 1997;56(4):1012-9.

25. Caunt CJ, Finch AR, Sedgley KR, McArdle CA. GnRH receptor signalling to ERK: kinetics and compartmentalization. Trends in endocrinology and metabolism: TEM. 2006;17(8):308-13.

26. Wang L, Bogerd J, Choi HS, Seong JY, Soh JM, Chun SY, et al. Three distinct types of GnRH receptor characterized in the bullfrog. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(1):361-6.

27. Fatima P, Hossain MM, Rahman D, Suman GM. Outcome of pregnancies after inadvertent exposure to GnRH agonist in early pregnancy. Mymensingh medical journal : MMJ. 2011;20(2):303-7.

28. Beckers NG, Macklon NS, Eijkemans MJ, Ludwig M, Felberbaum RE, Diedrich K, et al. Nonsupplemented luteal phase characteristics after the administration of recombinant human chorionic gonadotropin, recombinant luteinizing hormone, or gonadotropin-releasing hormone (GnRH) agonist to induce final oocyte maturation in in vitro fertilization patients after ovarian stimulation with recombinant follicle-stimulating hormone and GnRH antagonist cotreatment. The Journal of clinical endocrinology and metabolism. 2003;88(9):4186-92.

29. Seitz S, Buchholz S, Schally AV, Weber F, Klinkhammer-Schalke M, Inwald EC, et al. Triple negative breast cancers express receptors for LHRH and are potential therapeutic targets for cytotoxic LHRH-analogs, AEZS 108 and AEZS 125. BMC cancer. 2014;14:847.

30. Kwok CW, Treeck O, Buchholz S, Seitz S, Ortmann O, Engel JB. Receptors for luteinizing hormone-releasing hormone (GnRH) as therapeutic targets in triple negative breast cancers (TNBC). Targeted oncology. 2015;10(3):365-73.

31. Song Q, Huang R, Li J, Fan J, Zheng S, Zhang B, et al. The diverse distribution of risk factors between breast cancer subtypes of ER, PR and HER2: a 10-year retrospective multi-center study in China. PloS one. 2013;8(8):e72175.

32. Naor Z. Signaling by G-protein-coupled receptor (GPCR): studies on the GnRH receptor. Frontiers in neuroendocrinology. 2009;30(1):10-29.

33. Kono H, Nakamura M, Ohtsuka T, Nagayoshi Y, Mori Y, Takahata S, et al. High expression of microRNA-155 is associated with the aggressive malignant behavior of gallbladder carcinoma. Oncology reports. 2013;30(1):17-24.

34. Kong W, Yang H, He L, Zhao JJ, Coppola D, Dalton WS, et al. MicroRNA-155 is regulated by the transforming growth factor beta/Smad pathway and contributes to epithelial cell plasticity by targeting RhoA. Molecular and cellular biology. 2008;28(22):6773-84.

35. Kong W, He L, Coppola M, Guo J, Esposito NN, Coppola D, et al. MicroRNA-155 regulates cell survival, growth, and chemosensitivity by targeting FOXO3a in breast cancer. The Journal of biological chemistry. 2010;285(23):17869-79.

36. Eis PS, Tam W, Sun L, Chadburn A, Li Z, Gomez MF, et al. Accumulation of miR-155 and BIC RNA in human B cell lymphomas. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(10):3627-32.

37. Gironella M, Seux M, Xie MJ, Cano C, Tomasini R, Gommeaux J, et al. Tumor protein 53-induced nuclear protein 1 expression is repressed by miR-155, and its restoration inhibits pancreatic tumor development. Proceedings of the National Academy of Sciences of the United States of America. 2007;104(41):16170-5.

38. Greither T, Grochola LF, Udelnow A, Lautenschlager C, Wurl P, Taubert H. Elevated expression of microRNAs 155, 203, 210 and 222 in pancreatic tumors is associated with poorer survival. International journal of cancer. 2010;126(1):73-80.

39. Mattiske S, Suetani RJ, Neilsen PM, Callen DF. The oncogenic role of miR-155 in breast cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of

Preventive Oncology. 2012;21(8):1236-43.

40. Feng J, Gou J, Jia J, Yi T, Cui T, Li Z. Verteporfin, a suppressor of YAP-TEAD complex, presents promising antitumor properties on ovarian cancer. OncoTargets and therapy. 2016;9:5371-81.

41. Overholtzer M, Zhang J, Smolen GA, Muir B, Li W, Sgroi DC, et al. Transforming properties of YAP, a candidate oncogene on the chromosome 11q22 amplicon. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(33):12405-10.

42. Shin SH, Kim SC, Hong SM, Kim YH, Song KB, Park KM, et al. Genetic alterations of K-ras, p53, c-erbB-2, and SMAD4 in pancreatic ductal adenocarcinoma and their correlation with patient survival. Pancreas. 2013;42(2):216-22.

43. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic acids research. 2012;40(Web Server issue):W452-7.

44. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Current protocols in human genetics. 2013;Chapter 7:Unit7.20.

45. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods (San Diego, Calif). 2001;25(4):402-8.

46. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature methods. 2012;9(4):357-9.

47. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics. 2011;12:323.

48. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome research. 2012;22(9):1760-74.

49. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England). 2013;29(1):15-21.

50. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al.

The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010;20(9):1297-303.

51. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research. 2010;38(16):e164.

52. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, Sun MG, et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. PLoS computational biology. 2011;7(5):e1001138.

53. Wu TD, Watanabe CK. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics (Oxford, England). 2005;21(9):1859-75.

54. Moore PS, Orlandini S, Zamboni G, Capelli P, Rigaud G, Falconi M, et al. Pancreatic tumours: molecular pathways implicated in ductal cancer are involved in ampullary but not in exocrine nonductal or endocrine tumorigenesis. British journal of cancer. 2001;84(2):253-62.

55. Blackford A, Serrano OK, Wolfgang CL, Parmigiani G, Jones S, Zhang X, et al. SMAD4 gene mutations are associated with poor prognosis in pancreatic cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2009;15(14):4674-9.

56. Li M, Becnel LS, Li W, Fisher WE, Chen C, Yao Q. Signal transduction in human pancreatic cancer: roles of transforming growth factor beta, somatostatin receptors, and other signal intermediates. Archivum immunologiae et therapiae experimentalis. 2005;53(5):381-7.

57. Matthaios D, Zarogoulidis P, Balgouranidou I, Chatzaki E, Kakolyris S. Molecular pathogenesis of pancreatic cancer and clinical perspectives. Oncology. 2011;81(3-4):259-72.

58. Nedaeinia R, Avan A, Manian M, Salehi R, Ghayour-Mobarhan M. EGFR as a potential target for the treatment of pancreatic cancer: dilemma and controversies. Current drug targets. 2014;15(14):1293-301.

59. Heestand GM, Kurzrock R. Molecular landscape of pancreatic cancer: implications for current clinical trials. Oncotarget. 2015;6(7):4553-61.

60. Gysin S, Lee SH, Dean NM, McMahon M. Pharmacologic inhibition of RAF-->MEK-->ERK signaling elicits pancreatic cancer cell cycle arrest through induced expression of p27Kip1. Cancer research. 2005;65(11):4870-80.

61. Sibley K, Stern P, Knowles MA. Frequency of fibroblast growth factor receptor 3 mutations in sporadic tumours. Oncogene. 2001;20(32):4416-8.

62. Muller KE, Marotti JD, de Abreu FB, Peterson JD, Miller TW, Chamberlin MD, et al. Targeted next-generation sequencing detects a high frequency of potentially actionable mutations in metastatic breast cancers. Experimental and molecular pathology. 2016;100(3):421-5.

63. Heikkinen K, Mansikka V, Karppinen SM, Rapakko K, Winqvist R. Mutation analysis of the ATR gene in breast and ovarian cancer families. Breast cancer research : BCR. 2005;7(4):R495-501.

64. Durocher F, Labrie Y, Soucy P, Sinilnikova O, Labuda D, Bessette P, et al. Mutation analysis and characterization of ATR sequence variants in breast cancer cases from high-risk French Canadian breast/ovarian cancer families. BMC cancer. 2006;6:230.

65. Castera L, Krieger S, Rousselin A, Legros A, Baumann JJ, Bruet O, et al. Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes. European journal of human genetics : EJHG. 2014;22(11):1305-13.

66. Yu XM, Jaskula-Sztul R, Ahmed K, Harrison AD, Kunnimalaiyaan M, Chen H. Resveratrol induces differentiation markers expression in anaplastic thyroid carcinoma via activation of Notch1 signaling and suppresses cell growth. Molecular cancer therapeutics. 2013;12(7):1276-87.

67. Osipo C, Patel P, Rizzo P, Clementz AG, Hao L, Golde TE, et al. ErbB-2 inhibition activates Notch-1 and sensitizes breast cancer cells to a gamma-secretase inhibitor. Oncogene. 2008;27(37):5019-32.

68. Craig DW, O'Shaughnessy JA, Kiefer JA, Aldrich J, Sinari S, Moses TM, et al. Genome and transcriptome sequencing in prospective metastatic triple-negative breast cancer uncovers therapeutic vulnerabilities. Molecular cancer therapeutics. 2013;12(1):104-16.

69. Berrada N, Delaloge S, Andre F. Treatment of triple-negative metastatic breast cancer: toward individualized targeted treatments or chemosensitization? Annals of oncology : official journal of the European Society for Medical Oncology. 2010;21 Suppl 7:vii30-5.

70. Archer SK, Claudianos C, Campbell HD. Evolution of the gelsolin family of actin-binding proteins as novel transcriptional coactivators. BioEssays : news and reviews in molecular, cellular and developmental biology. 2005;27(4):388-96.

71. Cacemiro MC, Berzoti-Coelho MG, Cominal JG, Burin SM, Castro FA. Hippo pathway deregulation: implications in the pathogenesis of haematological malignancies. Journal of clinical pathology. 2017;70(1):9-14.

72. Adler JJ, Johnson DE, Heller BL, Bringman LR, Ranahan WP, Conwell MD, et al. Serum deprivation inhibits the transcriptional co-activator YAP and cell growth via phosphorylation of the 130-kDa isoform of Angiomotin by the LATS1/2 protein kinases. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(43):17368-73.

73. Mohammad RM, Dugan MC, Mohamed AN, Almatchy VP, Flake TM, Dergham ST, et al. Establishment of a human pancreatic tumor xenograft model: potential application for preclinical evaluation of novel therapeutic agents. Pancreas. 1998;16(1):19-25.

74. Hu C, Hart SN, Bamlet WR, Moore RM, Nandakumar K, Eckloff BW, et al. Prevalence of Pathogenic Mutations in Cancer Predisposition Genes among Pancreatic Cancer Patients. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2016;25(1):207-11.

75. Walters DM, Stokes JB, Adair SJ, Stelow EB, Borgman CA, Lowrey BT, et al. Clinical, molecular and genetic validation of a murine orthotopic xenograft model of pancreatic adenocarcinoma using fresh human specimens. PloS one. 2013;8(10):e77065.

76. Wennerstrom AB, Lothe IM, Sandhu V, Kure EH, Myklebost O, Munthe E. Generation and characterisation of novel pancreatic adenocarcinoma xenograft models and corresponding primary cell lines. PloS one. 2014;9(8):e103873.

77. Zboralski D, Bockmann M, Zapatka M, Hoppe S, Schoneck A, Hahn SA, et al. Divergent mechanisms underlie SMAD4-mediated positive regulation of the three genes encoding the basement membrane component laminin-332 (laminin-5). BMC cancer. 2008;8:215.

78. Reid-Lombardo KM, Fridley BL, Bamlet WR, Cunningham JM, Sarr MG, Petersen

GM. Inflammation-related gene variants as risk factors for pancreatic cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2011;20(6):1251-4.

79. Soares LR, Tsavaler L, Rivas A, Engleman EG. V7 (CD101) ligation inhibits TCR/CD3-induced IL-2 production by blocking Ca2+ flux and nuclear factor of activated T cell nuclear translocation. Journal of immunology (Baltimore, Md : 1950). 1998;161(1):209-17.

80. Ilie M, Nunes M, Blot L, Hofman V, Long-Mira E, Butori C, et al. Setting up a wide panel of patient-derived tumor xenografts of non-small cell lung cancer by improving the preanalytical steps. Cancer medicine. 2015;4(2):201-11.

81. Morgan KM, Riedlinger GM, Rosenfeld J, Ganesan S, Pine SR. Patient-Derived Xenograft Models of Non-Small Cell Lung Cancer and Their Potential Utility in Personalized Medicine. Frontiers in oncology. 2017;7:2.

82. Jeong HM, Kim RN, Kwon MJ, Oh E, Han J, Lee SK, et al. Targeted exome sequencing of Korean triple-negative breast cancer reveals homozygous deletions associated with poor prognosis of adjuvant chemotherapy-treated patients. Oncotarget. 2017;8(37):61538-50.

83. Wang D, Li JR, Zhang YH, Chen L. Identification of Differentially Expressed Genes between Original Breast Cancer and Xenograft Using Machine Learning Algorithms. 2018;9(3).

84. DeRose YS, Wang G, Lin YC, Bernard PS, Buys SS, Ebbert MT, et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. Nature medicine. 2011;17(11):1514-20.

85. Hidalgo M, Amant F, Biankin AV, Budinska E, Byrne AT, Caldas C, et al. Patient-derived xenograft models: an emerging platform for translational cancer research. Cancer discovery. 2014;4(9):998-1013.

86. Brooks YS, Ostano P, Jo SH, Dai J, Getsios S, Dziunycz P, et al. Multifactorial ERbeta and NOTCH1 control of squamous differentiation and cancer. The Journal of clinical investigation. 2014;124(5):2260-76.

87. Monsonego-Ornan E, Adar R, Feferman T, Segev O, Yayon A. The transmembrane mutation G380R in fibroblast growth factor receptor 3 uncouples ligand-mediated receptor activation from down-regulation. Molecular and cellular biology. 2000;20(2):516-22.

88. Ornitz DM, Itoh N. The Fibroblast Growth Factor signaling pathway. Wiley interdisciplinary reviews Developmental biology. 2015;4(3):215-66.

89. Hales EC, Taub JW, Matherly LH. New insights into Notch1 regulation of the PI3K-AKT-mTOR1 signaling axis: targeted therapy of gamma-secretase inhibitor resistant T-cell acute lymphoblastic leukemia. Cellular signalling. 2014;26(1):149-61.

90. Reedijk M, Odorcic S, Chang L, Zhang H, Miller N, McCready DR, et al. High-level coexpression of JAG1 and NOTCH1 is observed in human breast cancer and is associated with poor overall survival. Cancer research. 2005;65(18):8530-7.

91. Mungamuri SK, Yang X, Thor AD, Somasundaram K. Survival signaling by Notch1: mammalian target of rapamycin (mTOR)-dependent inhibition of p53. Cancer research. 2006;66(9):4715-24.

92. Ling H, Sylvestre JR, Jolicoeur P. Notch1-induced mammary tumor development is cyclin D1-dependent and correlates with expansion of pre-malignant multipotent duct-limited progenitors. Oncogene. 2010;29(32):4543-54.

93. Klinakis A, Szabolcs M, Politi K, Kiaris H, Artavanis-Tsakonas S, Efstratiadis A. Myc is a Notch1 transcriptional target and a requisite for Notch1-induced mammary tumorigenesis in mice. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(24):9262-7.

94. Hong HM, Yang JJ, Su CC, Chang JY, Li TC, Li SY. A novel mutation in the connexin 29 gene may contribute to nonsyndromic hearing loss. Human genetics. 2010;127(2):191-9.

95. Gao H, Korn JM, Ferretti S, Monahan JE, Wang Y, Singh M, et al. High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. Nature medicine. 2015;21(11):1318-25.

96. Ben-David U, Ha G, Tseng YY, Greenwald NF. Patient-derived xenografts undergo mouse-specific tumor evolution. 2017;49(11):1567-75.

97. Brennan DF, Dar AC, Hertz NT, Chao WC, Burlingame AL, Shokat KM, et al. A

Raf-induced allosteric transition of KSR stimulates phosphorylation of MEK. Nature. 2011;472(7343):366-9.

98. Nueda ML, Naranjo AI, Baladron V, Laborda J. Different expression levels of DLK1 inversely modulate the oncogenic potential of human MDA-MB-231 breast cancer cells through inhibition of NOTCH1 signaling. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2017;31(8):3484-96.

99. Nacu S, Yuan W, Kan Z, Bhatt D, Rivers CS, Stinson J, et al. Deep RNA sequencing analysis of readthrough gene fusions in human prostate adenocarcinoma and reference samples. BMC medical genomics. 2011;4:11.

100. Tisdale MJ. Zinc-alpha2-glycoprotein in cachexia and obesity. Current opinion in supportive and palliative care. 2009;3(4):288-93.

101. Mayer IA, Abramson VG, Lehmann BD, Pietenpol JA. New strategies for triple-negative breast cancer--deciphering the heterogeneity. Clinical cancer research : an official journal of the American Association for Cancer Research. 2014;20(4):782-90.

102. Pergolini I, Morales-Oyarvide V, Mino-Kenudson M, Honselmann KC, Rosenbaum MW, Nahar S, et al. Tumor engraftment in patient-derived xenografts of pancreatic ductal adenocarcinoma is associated with adverse clinicopathological features and poor survival. PloS one. 2017;12(8):e0182855.

103. Whittle JR, Lewis MT, Lindeman GJ, Visvader JE. Patient-derived xenograft models of breast cancer and their predictive power. Breast cancer research : BCR. 2015;17:17

104. Moore HC, Unger JM, Phillips KA, Boyle F, Hitre E, Porter D, et al. Goserelin for ovarian protection during breast-cancer adjuvant chemotherapy. The New England journal of medicine. 2015;372(10):923-32.

105. Wang C, Ma Y, Feng S, Liu K, Zhou N. Gonadotropin-releasing hormone receptor-targeted paclitaxel-degarelix conjugate: synthesis and in vitro evaluation. Journal of peptide science : an official publication of the European Peptide Society. 2015;21(7):569-76.

국문 요약

Patient-derived xenograft(PDX) 모델은 지금까지 cancer cell lines 으로 xenograft 를 만들어 실험하였던 것에 비해 많은 장점이 있다. 환자의 암 조직을 수술 시 채취하여 마우스에 바로 심어 종양 형성을 유도함으로써, 종양이 그 환자만의 암 조직의 특성을 유지하게 할 수 있다. 또한 암이 한 종류의 세포가 아닌 다양한 종류의 세포들이 모여있다는 특성을 반영함으로써 비임상 시험을 포함한 암 관련 연구에 효과적인 도구가 될 수 있다.

따라서 우리는 췌장암과 삼중음성유방암에 대한 PDX 모델을 구축한고 그러한 샘플에서 다양한 분석을 시도하였다. 췌장암은 말기에 이르러서까지 거의 증상이 나타나지 않음으로써 조기 진단이 쉽지 않고 5 년 생존률도 10% 미만인 사망률도 높은 암종이다. 또한 삼중음성 유방암은 ER, PR HER2 와 같은 3 가지 수용체가 잘 발현되지 않는 유방암의 종류로 다른 종류의 유방암에 비해 치료하기가 쉽지 않고 결과도 좋지 않은 것으로 보고되고 있다. 따라서 췌장암과 삼중음성유방암을 보다 현실적으로 연구할 수 있는 도구 및 모델의 필요성이 요구되고 있는 상황이다.

췌장암의 경우 우리는 총 20 개의 PDX 모델을 만들었고 성공률은 74%를 보여주었다. 췌장암 조직을 작은 조각으로 잘라 피하에 심는 방식을 취하였다. 성공한 PDX 와 원래 환자의 암의 유사성을 비교하기 위해 면역염색을 통해 p53 과 SMAD4 의 발현양상을 조사하였는데 20 개 모두 일치함이 들어났다. 더불어 an exome single nucleotide polymorphism (SNP) array 와 Comprehensive Cancer Panel(CCP) 를 실시하였는데 암과 관련한 variant 가 PDX 와 원래 샘플간 94% 가 일치함을 알 수 있었다. 그 외에도 Polyphen2 와 Sorting Intolerant from Tolerant (SIFT) 로 예측을 하여 623 개의 variant 를 찾아내고 CCP 에서는 409 개의 발암억제유전자 및 발암촉진 유전자를 찾아냄으로서 PDX 와 본래 암의 gene mutation profile 을 분석할 수 있었고 이들이 매우 다양한 분자적 성질을 갖고 있다는 것을 다시 한번 확인하였다.

삼중음성유방암의 PDX 샘플 24 개에 대해 CCP 와 high-throughput RNA sequencing (RNA-seq) 를 실시하였고, 이를 통해 역시 다양한 variant 및 gene-fusion 을 찾아낼 수 있었다. 이들이 갖는 의미는 고찰 (Discussion) 에서 논의하였

다.

이와 같은 PDX 모델이 실제로 전임상시험 연구에 어떻게 효과적으로 사용될 수 있는지를 보여주기 위해 2 가지 실험의 예를 덧붙였다. 첫번째 실험은 GnRH agonist(zoladex) 약물이 neoadjunvant theraphy 로써 사용 시 종양크기 감소 효과가 있는지 독성을 유발하지는 않는지를 알기 위해 PDX 모델에 기존 항암치료제(AC) 에 GnRH agonist 를 함께 투여하여 살펴보았다. 그 결과 GnRH receptor 발현이 큰 샘플을 심은 마우스에서는 GnRH agonist 를 함께 투여한 실험군이 기존 항암제만 투여한 실험군보다 항암 효과가 좋은 것으로 드러났다.

두번째 예시로서는 miR-155 와 YAP 과의 관계를 유방암에서 연구하는데 PDX 모델 및 PDX 에서 유래한 patient-derived cell (PDC) 가 어떻게 사용될 수 있는지를 보여주었다. 실험 결과 miR-155 는 YAP 의 phosphorylation 을 조절하고 그 매개체로는 SVIL 이 miR-155 의 target 이 됨으로써 이러한 효과가 나타날 수 있을 것이란 가능성이 제시되었다. 더불어 YAP 역시 miR-155 의 발현을 조절한다는 사실을 luciferase promoter assay 와 western blotting, qPCR 및 ChIP assay 실험을 통해 보여주었다.

결론적으로 우리는 암 연구에 유용한 PDX 모델을 췌장암과 삼중음성 유방암에서 구축하였고 다양한 분석을 통해 분자유전학적인 특징을 알아내었다. 더불어 우리가 구축한 PDX 모델을 이용하여 GnRH agonist 효과를 보는 전임상시험과 miR-155 와 YAP 의 관계를 알아보는 암 관련연구를 예시로 제시함으로써 PDX 모델의 유용함을 보여주었다.

Acknowledgement

LIO 실험실에 들어온지가 엊그제 같은데 벌써 봄여름가을겨울이 몇 번씩 지나고 시간은 쏜 살 같이 흘러 어느덧 새로운 봄을 맞이하며 이렇게 졸업을 준비하게 되었네요. 보통 빨리 시간이 지나갈 때 쏜 살 같다는 말을 하는데, 쏜 살은 손을 놓는 순간 쏘아버린 곳을 향해 곧바로 나아가지만 저는 이리저리 부딪히며 겨우 겨우 과녁판에 도착했다는 기분이 듭니다. 또한 그만큼 한 번 쏘아버린 화살이라고 쳐다만 본 것이 아닌, 순간순간 엇나가는 화살이 되지 않도록 끊임없이 노력해왔던 것 같습니다. 이와 같이 부족한 제가 여러가지 어려움을 이겨내고 무사히 과녁판에 도달 할 수 있도록 지속적으로 지도해주신 저희 지도교수님이신 장수환 교수님께 진심으로 감사 드립니다. 특히 과거, 방황하는 저를 믿고 불러주시고 개인적 어려움에 있을 때도 항상 들어주시고 도와주시며 제가 다시 일어서서 열정을 갖고 학문에 임할 수 있게 해주셨기에 교수님을 만나지 못했다면 현재의 저는 없었으리라 생각합니다.

그 외에 이러한 논문을 위해 많은 데이터를 작성해 주신 다른 여러분들 에게도 깊은 감사 말씀 드립니다. 더불어 몇 년 동안 함께 실험하며, 고충을 나누어주었던 모든 우리 LIO 실험실 분들께도 감사 드립니다. 특히 2 년 동안 제 옆에서 실험을 배우고 또 바쁠 때 저를 많이 도와주기도 하였던 정민이에게 고맙다는 말을 해주고 싶습니다. 또한 많은 경험으로 난관에 부딪힐 때마다 도와주신 김시내 박사님, 바쁘신데도 열심히 여러 가지 동물 실험을 함께 해주신 최연숙 박사님, 같은 학생으로서 처음부터 옆에서 항상 공감해주고 슬프고 기쁜 것을 나누었던 민지, 한국인보다 더 한국인 같은 Duc, 항상 조용히 자기 일을 열심히 하는 은아, 타국에 와서도 잘 적응하기 위해 노력하는 bruce, 실험실원이 얼마되지 않은 막내 명지까지 모두 즐거운 시간들을 만들어 주셔서 감사합니다

또한 제가 가장 사랑하는 부모님, 저희 엄마 아빠께도 깊은 감사 드립니다. 저보다 더 제 삶을 사랑해주시기에, 수 많았던 유혹에서도 일어나고 헤쳐나올 수 있었습니다. 두 발로 서기 힘들 때면 저의 두 다리가 되어 주시며 굳건히 서있을 수 있도록 해주시고, 넘어져 있을 때면 다시 일어날 수 있도록 용기를 북돋아 주셨기에 오늘을 맞이 할 수 있었습니다. 저 역시도 그만큼 부모님을 사랑한다는 것을 알아주셨으면 좋겠습니다

끝으로, 이 자리를 빌어 손우찬 교수님께도 진심으로 여러 가지로 죄송하고 감사했다는 말씀 드리고 싶습니다. 예전 교수님과 함께 슬라이드를 보고 많은 것을 가르쳐 주시던 시간들은 이제 제 기억에 행복했던 기억으로 남아 있습니다. 그동안 죄송스런 마음에 차마 찾아 뵙지 못하고 인사 드리지 못했지만 예전에 주신 가르침에 감사 드리며 그 시기에 갖게 해주셨던 초심을 잊지 않고 끝까지 갈 수 있도록 노력하겠습니다.

지난 몇 년간 저의 표면적인 목표는 박사학위를 받는 것이었지만 제 꿈은 박사가 되는 것이 아닙니다. 즉, 박사는 끝이 아니라 시작이란 것을 잘 알고 있습니다. 진정한 연구자가 되는 첫발을 내딛었다는 마음가짐으로 박사 후에도 열심히 연구에 매진하고, 정말 가치있고 인류에 보탬이 될 수 있는 연구를 하는 과학자가 되도록 노력하겠습니다. 다시 한번 모든 분께 감사드립니다