

# 난소암 치료를 위한 바이오마커 발굴 및 치료 표적 유전자 발굴

바이오공학과  
20190910 이현지

## Introduction

### 난소암 소개

우리 몸의 세포가 여러 원인으로 인해 정상적으로 분화하지 않고, 계속 성장하는 것을 종양(tumor)이라 하며, 이중 악성종양을 암(cancer)라고 한다. 정상적으로 세포는 세포 내 조절기능에 의해 분열하며 성장하고 죽어 없어지기도 하여 세포수의 균형을 유지한다. 그러나 세포의 유전자에 변화가 일어나면 비정상적으로 세포가 변하여 불완전하게 성숙하고, 과다하게 증식하게 되는데, 이를 암이라고 정의한다. 암세포는 세포가 정상적으로 분화되지 않고 성장이 조절되지 않으며, 주위 조직으로 침투(침윤)하고, 멀리 떨어져 있는 조직으로 퍼져나가 성장(전이)하는 것이 특징이다.

난소는 자궁 양쪽에 위치하여 난자를 생산하고 여성호르몬을 만드는 생식기관이며, 난소암은 난소에 생긴 암을 의미한다. 원인은 정확하게 밝혀지지 않았지만 BRCA 유전자 변이를 가지고 있을 때 난소암에 걸릴 가능성이 높다.

난소암은 초기에 별다른 증상이 나타나지 않고, 상당히 진행되더라도 증상이 매우 경미하여 난소암 환자들이 말기에 이르러 병원을 찾는 경우가 많다.

### 난소암 치료법

난소암은 병기에 상관없이 수술적 치료를 시행한 후 항암제를 투여하는 방식으로 치료를 진행한다. 난소암은 재발률이 70% 이상으로 매우 높다는 특성 때문에 재발 자체를 억제하는 것이 중요하다. 이를 위해 2세대 항암제인 표적항암제를 이용한 유지요법을 진행하는 것이 일반적이다.

난소암을 잘 일으키는 BRCA 유전자 변이를 갖고 있는 난소암 환자에게는 BRCA 유전자 변이를 타겟으로하는 PARP inhibitor 라는 표적치료제를 사용한다.

또한 베바시주맵과 PARP 억제제를 함께 사용해 BRCA 혹은 이와 유사한 유전자 변이가 없는 난소암 환자의 치료에도 도움을 줄 수 있다. 이외에도 면역관문억제제를 비롯한 다양한 기전을 가진 신약들도 계속해서 개발되고 있다.

## **PARP inhibitor (Poly ADP-Ribose Polymerase)**

위의 내용과 같이 난소암 환자에게는 BRCA 유전자 변이를 타겟으로 하는 PARP inhibitor 라는 표적치료제를 사용한다. PARP 효소는 DNA 손상을 복구하는 과정에 중요한 역할을 한다. 유전자 변이가 있는 암세포는 다른 DNA 의 손상 복구 경로에 의존하는 경향이 있다. PARP inhibitor 는 PARP 효소의 기능을 차단하여 암세포가 복구 경로에 의존할 수 없도록 한다. 이에 따라 암세포는 생존 능력이 저하되고, 사멸에 이른다.

## **난소암 바이오마커**

바이오마커란, 일반적으로 단백질이나 DNA, RNA 대사물질 등을 이용해 몸 안의 변화를 알아낼 수 있는 지표를 말한다. 특정한 질환을 진단하거나 치료법을 결정하는데 특이적으로 반응하는 물질을 총칭하는 개념이다. 바이오마커는 암의 진단, 약물 반응성 검사에 활용된다.

난소암의 대표적인 바이오마커로는 CA-125(cancer antigen 125)가 있다. 이는 혈액 속에 존재하는 고분자 당단백으로, 난소암이 이를 분비하기 때문에 혈액 중의 CA-125 수치를 측정해 혈액감사로 난소암을 추정한다. 또한, 난소암과 연관된 BRCA1, BRCA2 유전자 변이도 중요한 바이오마커 중 하나이다.

## **난소암 치료법의 한계, PARP inhibitor 저항성**

난소암 치료에서 중요한 부분을 차지하는 PARP inhibitor 는 몇 가지 한계점이 존재한다. 그 중 가장 중요한 문제는 일부 환자들에게서 PARP inhibitor 에 대한 저항성이 나타났다는 것이다. 치료 중에 저항성이 생길 수도 있으나, 대부분 유전자 돌연변이에 의해 저항성을 보인다. 따라서 PARP inhibitor 사용 시 저항성 문제를 극복하기 위해서 본 연구를 진행하게 되었다. PARP inhibitor 에 저항성이 있는 세포에서 치료의 표적이 될 수 있는 유전자를 찾는 것이 연구의 목적이다.

## Results

### PARP inhibitor 저항성과 연관성이 있는 유전자 발굴

난소암에서 PARP inhibitor 저항성 극복을 위한 표적 유전자 발굴을 수행했다. 5종 PARP inhibitor(Niraparib, Olaparib, Veliparib, Rucaparib, Talazoparib)에 대해 유전자 변이에 따른 drug sensitivity 분석을 실행했다. 분석 대상 유전자는 PARP inhibitor에 대해 저항성을 보이는 유전자를 발굴하여 분석 대상으로 선정했다. BRCA1, SMAD4, NF1, MLL2, SMARCA4 총 5개의 유전자가 선정되었다.

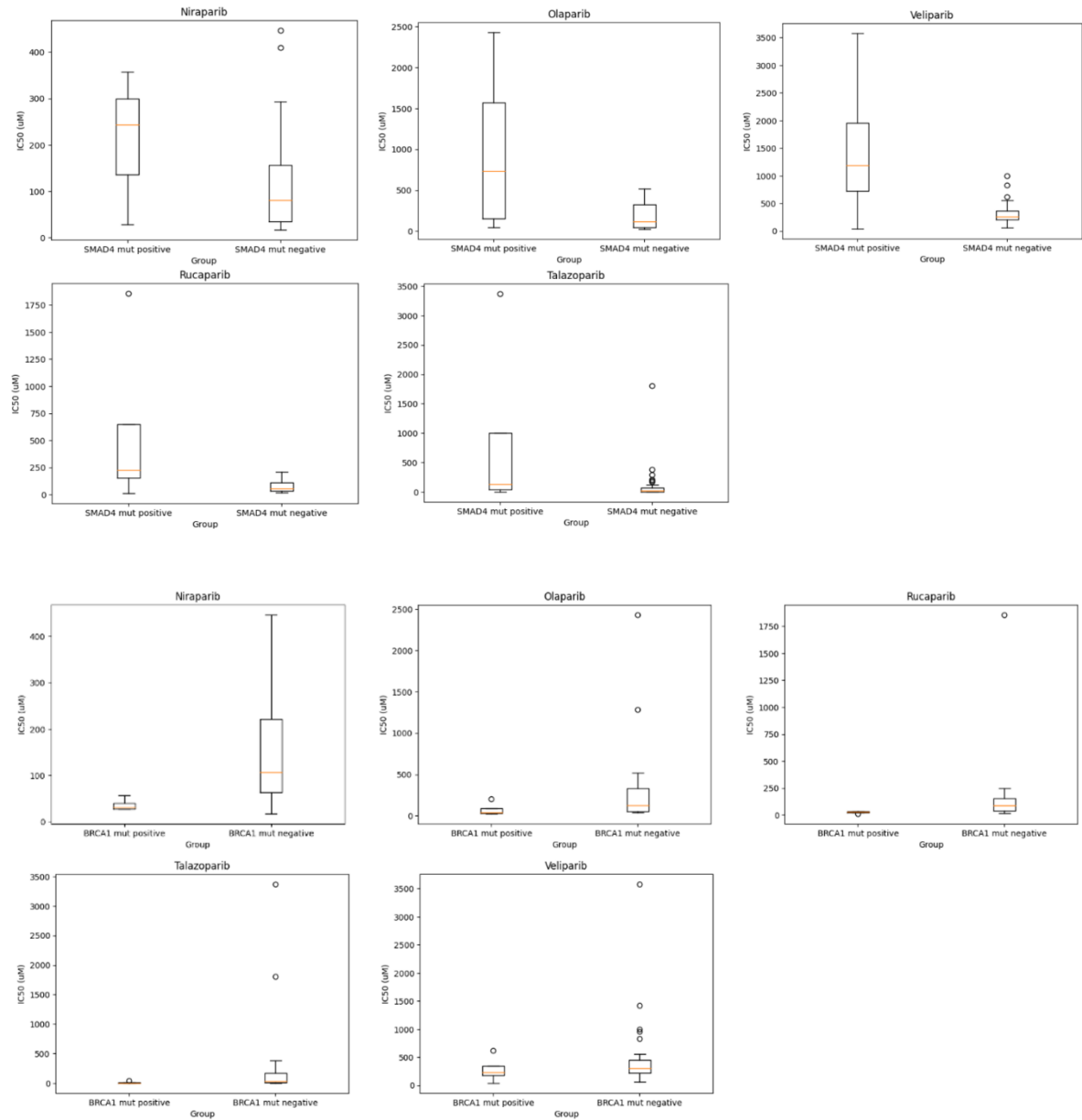
이어서 난소암 세포주들의 변이 정보를 포함하고 있는 데이터에서 특정 유전자에 돌연변이가 있는 세포주, 돌연변이가 없는 세포주로 그룹을 나누어 약물 저항성 분석을 실행했다.

### 약물 저항성 분석 결과

위에서 선정한 5개의 유전자와 5종의 PARP inhibitor에 대해서 실험을 진행했다. 각 유전자에 대해 5종의 PARP inhibitor에 대한 drug sensitivity 실험을 진행했는데, 유전자별로 특정 유전자의 변이가 있는 세포주와 변이가 없는 세포주를 그룹으로 나누어 PARP inhibitor를 처리했을 때 약물 sensitivity를 분석했다.

분석에는 IC50 값을 활용했다. IC50 값은 약물의 효과를 측정하는 데 사용되는 지표로, Inhibitory Concentration 50의 약자이다. 특정 약물이 특정 생물학적 활성에 대해 50%의 억제 효과를 나타내는 예상 농도이며, 낮은 IC50값은 해당 약물의 농도가 적은 상황에서도 높은 효과를 나타낼 수 있음을 의미한다. 즉, IC50 값이 낮을수록 약물의 농도가 낮아도 활성을 억제하는 데 효과적이라는 의미이다.

아래 그래프는 타겟 유전자에 대해 돌연변이가 있는 세포주에 PARP inhibitor를 처리했을 때 세포주들에서 얻어진 IC50 값과 돌연변이가 없는 세포주에 PARP inhibitor를 처리했을 때 세포주들에서 얻어진 IC50값 두 가지를 그래프로 시각화하여 분석했다. 아래 그림은 각각 SMAD4와 BRCA1의 분석 결과 그래프이다.

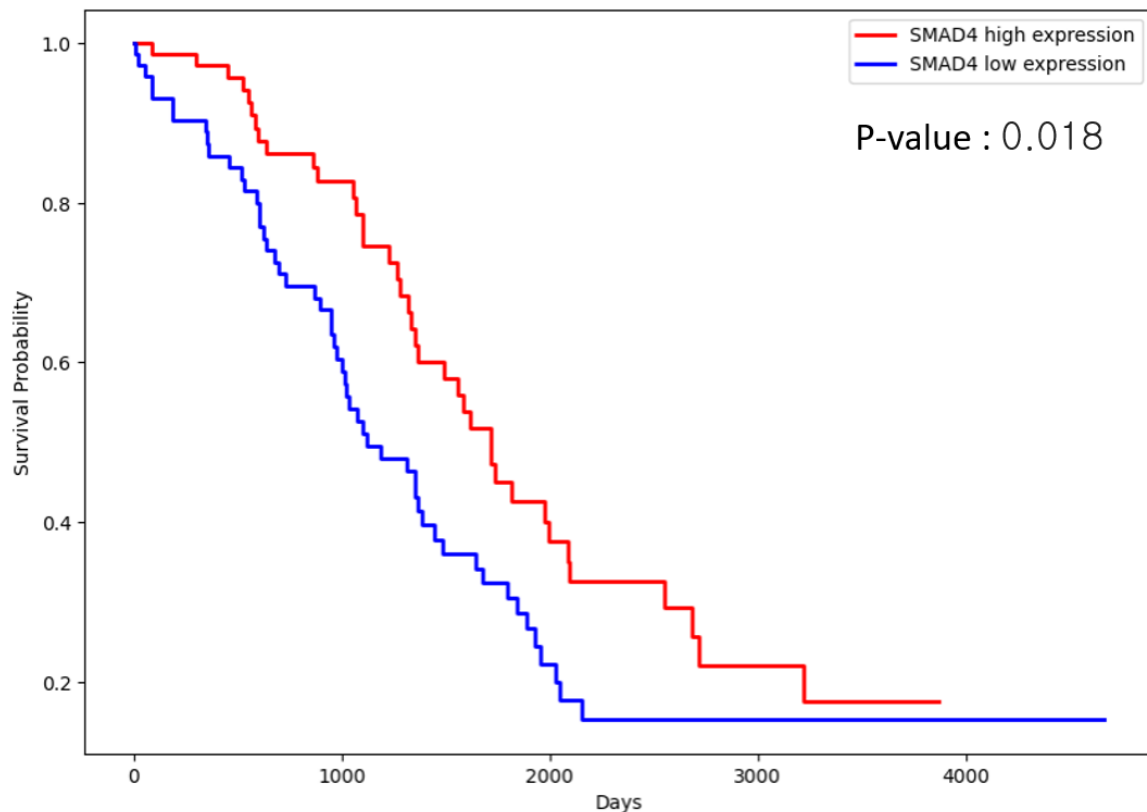


5개의 PARP inhibitor에 대해 SMAD4를 제외한 4개의 유전자는 모두 비슷한 경향을 보였다. 유전자 돌연변이가 있는 세포주에서 IC50 값이 낮았다. PARP inhibitor의 활성이 좋으며, 돌연변이가 있을 때 세포들이 잘 죽는다는 것을 의미한다.

SMAD4는 그 반대의 경향성을 보였다. SMAD4 유전자에 돌연변이가 존재할 경우 PARP inhibitor에 대한 drug sensitivity를 보이는 결과를 관찰했다. 본 연구의 목적과 일치하는 것은 SMAD4였기에, 추가 분석은 SMAD4만을 이용해 진행했다.

## 신약개발 표적물질 생존분석

약물 저항성 분석 결과에 따라 이후 추가적인 분석은 SMAD4에 대해 진행했다. 생존분석은 Kaplan - Meier 방식으로 진행했다. 생존 데이터는 특정 사건(주로 환자 사망)이 발생할 때까지 지속 시간을 의미한다. Kaplan - Meier 방식은 생존 확률을 추정하기 위해 각 구간의 생존율을 구하여 곱하는 방식으로 누적 생존율을 추정하는 방식이다.



위 그림은 SMAD4의 생존 분석 결과 그래프이다. 분석 결과, SMAD4 high expression group 과 SMAD4 low expression group 의 생존율이 크게 차이 나는 것을 확인할 수 있었다. 이는 SMAD4를 바이오마커로 선정할 수 있는 근거가 된다.

SMAD4 low expression group 은 SMAD4 변이가 있는 그룹과 유사하다. SMAD4 low expression group 에서 생존율이 낮은 결과를 통해, SMAD4의 발현이 낮은 그룹에서는 PARP inhibitor 가 잘 작용하지 않아 (저항성이 있음) 생존율이 떨어진다는 가설 설정이 가능하다. (단, TCGA 환자에서 SMAD4의 변이는 관찰되지 않음) 따라서, SMAD4 low expression group 환자들을 위해 대안이 되는 표적 유전자 발굴이 필요하다.

## Gene Dependency Score 분석 결과

표적 유전자 발굴은 dependency map database 를 활용해 진행했다. 이 database 는 여러 세포주에 각각의 유전자를 하나씩 knock out 하여 세포의 생존, 성장 등의 특성에 미치는 영향에 대한 정보를 제공한다. CRISPR 기술을 사용해 수집된 데이터를 활용했다. SMAD4 변이 유무, 발현 정도 차이로 그룹을 나눈 후 gene dependency score 를 비교했다.

Gene dependency score 가 높을수록 해당 유전자가 세포 생존이나 기능에 중요한 역할을 한다는 것을 의미한다. 음수 값의 경우, 음의 방향으로 값이 커질수록 특정 유전자를 knock out 했을 때 세포가 쉽게 죽는다는 것을 의미한다.

### 1. 변이 유무에 따른 분석

다음은 SMAD4의 변이 유무에 따른 분석 결과를 정리한 표이다.

| Gene     | P-value         | 변이 있는 세포주 그룹  | 변이 없는 세포주 그룹  |
|----------|-----------------|---------------|---------------|
| MAPK1    | 0.0003583434271 | -0.9643021247 | -0.1283703622 |
| INTS6    | 0.0006985277639 | -1.772328447  | -0.6205858322 |
| TTC7A    | 0.0008126539174 | -0.5755603926 | -0.1819756184 |
| MTA2     | 0.001756739246  | -0.8224920598 | -0.2729927419 |
| SOX2     | 0.003437355674  | -0.5007732163 | -0.170472423  |
| MAP2K1   | 0.005255507143  | -0.5004093316 | -0.1117956298 |
| PPP2R2A  | 0.005998120153  | -0.5252384892 | -0.1287423348 |
| RPL28    | 0.008260177838  | -1.347527386  | -0.8032781127 |
| TYRO3    | 0.009253613484  | -0.6591326018 | -0.3882366906 |
| KAT2A    | 0.009308273368  | -0.5809385855 | -0.2898911877 |
| ARHGEF5  | 0.009470140687  | -0.6281156533 | -0.36185643   |
| DOCK5    | 0.009805565447  | -0.5624650112 | -0.113929235  |
| HSPA8    | 0.01092913017   | -1.493095405  | -0.6527274343 |
| HIC2     | 0.01208204322   | -0.5541738259 | -0.0774067286 |
| PKN2     | 0.01244936864   | -0.8480871283 | -0.444565498  |
| SCARF2   | 0.01452646178   | -0.5807476792 | -0.1585978437 |
| MED15    | 0.01586286388   | -0.6710967393 | -0.1594184846 |
| ARF5     | 0.01706965237   | -0.6358629963 | -0.2516370367 |
| CDK4     | 0.01861236022   | -1.480016323  | -0.6004179077 |
| TIMM29   | 0.01966532771   | -0.8811358739 | -0.5806537005 |
| HEXIM1   | 0.02181049931   | -0.6848315467 | -0.3736097017 |
| YAP1     | 0.02605114663   | -1.262273746  | -0.5317037888 |
| ASH2L    | 0.02841221661   | -0.6809619208 | -0.3158388382 |
| TUBD1    | 0.02825946383   | -0.7340870443 | -0.401224784  |
| H2AW     | 0.02985952897   | -0.5787383739 | -0.3423502248 |
| EIF1     | 0.03571549451   | -1.619206336  | -1.065396872  |
| PPP2CA   | 0.04519724551   | -1.507283832  | -0.7478467715 |
| PRPF4B   | 0.04545577078   | -0.9152391968 | -0.603594451  |
| ACACA    | 0.04615960385   | -0.6231343039 | -0.3010762063 |
| SAP130   | 0.04649156598   | -0.7302210246 | -0.4636548051 |
| KRTAP9-6 | 0.04682546292   | -0.608252887  | -0.3868084038 |
| ILK      | 0.04776365357   | -0.878820545  | -0.4457324311 |
| FASN     | 0.04965966143   | -0.517742628  | -0.2311320193 |

Gene dependency score 가 높을수록 해당 유전자가 세포 생존이나 기능에 중요한 역할을 한다는 것을 의미한다. 음수 값의 경우, 음의 방향으로 값이 커질수록 특정 유전자를 knock out 했을 때 세포가 쉽게 죽는다는 것을 의미한다.

본 연구의 목적에 맞는 후보 유전자는 SMAD4 변이가 있을 때 세포가 쉽게 죽는 유전자이다. 따라서, 변이가 있는 그룹에서 gene dependency score 가 -0.5보다 낮은 것을 기준으로 했고, 변이가 없는 세포주에서 변이가 있는 세포주보다 gene dependency score 가 약 1.5배 높은 유전자를 선정했다.

## 2. 발현 정도에 따른 분석

다음은 SMAD4의 발현 정도에 따른 분석 결과를 정리한 표이다.

| Gene     | P-value        | Low expression group | High expression group |
|----------|----------------|----------------------|-----------------------|
| SMARCA4  | 0.001905306902 | -0.633678624         | -0.2505815985         |
| WNK1     | 0.002486252276 | -0.7792962134        | -0.4228625029         |
| RAB6A    | 0.004362271922 | -0.5009528457        | -0.1044452672         |
| VPS4A    | 0.004687503779 | -0.7432819345        | -0.34249359           |
| GFPT1    | 0.01243022551  | -0.717103201         | -0.4088484084         |
| ARHGAP29 | 0.01285016637  | -0.5394828657        | -0.1668539466         |
| RIC1     | 0.01346252594  | -0.7341095754        | -0.3076168595         |
| TUBB4B   | 0.01472029281  | -0.5456385834        | -0.3077800753         |
| WWTR1    | 0.01516471373  | -0.5619011649        | -0.3042706833         |
| KLF5     | 0.01537158338  | -0.616327931         | -0.2498946977         |
| CAND1    | 0.01565026151  | -0.5319110086        | -0.220943715          |
| TRPM7    | 0.01987069467  | -0.7870177672        | -0.4314869551         |
| SMARCB1  | 0.0239712164   | -0.8086554707        | -0.4964206952         |
| YRDC     | 0.03400554018  | -2.246116731         | -1.486850828          |
| TFRC     | 0.03420404147  | -0.9310879627        | -0.5984811158         |
| EFR3A    | 0.0422406507   | -0.6048393495        | -0.3122176712         |
| SLC2A1   | 0.04576406222  | -0.5958086471        | -0.3069922865         |
| USP9X    | 0.04909943503  | -0.6929039523        | -0.3982794832         |

SMAD4 low expression group 은 SMAD4 변이가 있는 그룹과 유사하기에, 발현 정도에 따른 분석 결과에서는 위 과정과 반대되는 경우를 적용하여 후보 유전자를 선정했다.

SMAD4의 변이, 발현 정도에 따른 후보 표적 유전자 수는 51개였다.

## Discussion

SMAD4의 변이, 발현 정도에 따라 발굴한 후보 표적 유전자 51개에 대해 약물 분석과 문헌 분석을 진행했다. 문헌 분석은 chatGPT API 와 NCBI 를 활용했고, 난소암과의 연관성 논문이 존재하는지 확인했다..

후보 표적 유전자에 대해 난소암과의 연관성 논문이 존재하는 유전자는 21개였으며, SMAD4 변이가 있는 세포주에서 저해제로 항암 실험을 수행한 유전자는 CDK4, SMARCA4, ACACA로, 총 3개였다.

다음은 약물 분석 결과이다.

#### CDK4

|         | IC50           |                   |                    |                   |
|---------|----------------|-------------------|--------------------|-------------------|
| 약물이름 -> | <b>AT-7519</b> | <b>CGP-082996</b> | <b>Palbociclib</b> | <b>Ribociclib</b> |
| IGROV-1 | 0.306962       | N/A               | 6.950464           | 70.615288         |
| OVKATE  | 8.733215       | N/A               | 412.090583         | 275.831214        |
| SW626   | 1.274812       | N/A               | 141.302247         | 53.470191         |
| OV-90   | 1.169768       | N/A               | 514.244451         | 108.484591        |

CDK4의 경우, 항암 활성이 검증되었으나, 물질 개발이 되어 있지 않은 후보 유전자가 존재한다. 따라서 siRNA 등 실험을 통해 표적 유전자들을 검증할 필요가 있다.

#### SMARCA4

|         | IC50         |
|---------|--------------|
| 약물이름 -> | <b>PFI-3</b> |
| IGROV-1 | 154.382072   |
| OVKATE  | 데이터 없음       |
| SW626   | 154.780211   |
| OV-90   | 87.811512    |

#### ACACA

|          | IC50            |
|----------|-----------------|
| 약물 이름 -> | <b>BAY ACCi</b> |
| IGROV-1  | 105.898509      |
| UWB1.289 | 178.203806      |
| OVCA433  | 58.013312       |
| DOV13    | N/A             |

그 외 다른 유전자에 대한 약물 분석 결과는 없었다.

난소암과의 연관성 논문이 존재하는 유전자 중 하나를 소개하자면, ASH2L이 있다. ASH2L의 원형 RNA(circRNA)와 난소암 발달, 전이 및 예후 사이의 밀접한 관계가 있음이 밝혀졌다. CircASH2L은 난소암 세포의 침입과 증식을 향상시키고, miR-665에 결합하여 VEGFA 발현을 촉진한다. CircASH2L이 VEGFA 의존적 방식으로 난소암 세포의 성장과 침윤에 영향을 가하므로



CircASH2L을 억제하면 난소암 마우스 모델에서 종양의 성장, 림프관 신생 및 혈관 신생이 억제된다. CircASH2L의 녹다운은 생체 내에서 종양 진행을 현저하게 억제 하고 sh-circASH2L이 VEGFA 매개 림프관 신생 및 혈관 신생을 억제하여 종양 형성을 조절하는 것으로 보인다는 것을 나타낸다. CircASHL은 난소암 조직 및 세포주에서 과발현된다는 것이 입증되었다