

# 다형성신경교아종(Glioblastoma Multiforme)의 분자 아형 연구: 애질런트 GeneSpring 및 Mass Profiler Professional Multi-Omics 소프트웨어를 이용한 통합 분석

## 응용 자료

### 저자

Durairaj Renu, Pritha Aggarwal,  
Sunil C. Cherukuri, Pramila Tata  
Strand Life Sciences  
Bangalore, India

Vadiraja Bhat, Carolina B. Livi,  
Michael Rosenberg  
Agilent Technologies Inc.  
Santa Clara, California

Mona Al-Gizawiy, Kathleen  
Schmainda, Shama P. Mirza  
Medical College of Wisconsin  
Milwaukee, Wisconsin

### 개요

중앙 신경계 종양 중 가장 일반적이고 악성도가 높은 종양인 다형성신경교아종(Glioblastoma Multiforme, GBM) 관련 mRNA 발현, miRNA 발현, 단백질 발현 및 암 복제수 이상(copy number aberration, CNA)의 통합 멀티 오믹스 분석을 수행하기 위해, 애질런트 GeneSpring 및 Mass Profiler Professional (MPP) 소프트웨어를 사용하였습니다. 유전체 데이터는 TCGA(The Cancer Genome Atlas) 프로젝트로부터 획득하였으며, TCGA에 의해 정의된 GBM의 분자 아형은 mRNA 발현을 기반으로 하였습니다. 본 연구에서 GBM 아형과 유전체 CNA 사이의 관련성을 탐색하기 위해 GeneSpring 메타데이터 분석 프레임워크를 적용하였습니다. 신경발생(neurogenesis)에 관여하는 mRNA와 그들의 검증된 miRNA 간의 멀티 오믹스 상관 분석을 통해 p53 신호 전달을 억제하는 전신경(Proneural) 종양의 분자 이벤트를 중점으로 소개하였습니다. 또한, TCGA에서 얻은 mRNA 발현 데이터와 독립적인 label-free GBM 단백질체학 연구를 결합하여 분석한 결과, 54개의 시그니처 유전자(signature gene) 및 단백질의 핵심 하위 집합을 밝혀내 GeneSpring/MPP 소프트웨어의 통합 생물학 분석 능력을 입증하였습니다. 핵심 유전체 시그니처를 통해 TCGA 코호트에서 알려진 종양 아형을 명확히 구분할 수 있었습니다. 흥미롭게도 단백질체학 시그니처로 이 코호트를 2 또는 3개 그룹으로 나눌 수 있었습니다. 이는 다른 TCGA 아형이 존재할 수 있음을 시사합니다.



Agilent Technologies

## 소개

신경교종(Gliomas)은 성인에게 가장 흔히 발생하는 중앙 신경계(CNS) 종양입니다. 신경교종은 조직학적으로 성상세포종(astrocytomas), 희돌기세포교종(oligodendroliomas), 상의세포종(ependymomas)으로 나뉘어 지며, 성상세포종은 모든 신경교종 진단사례 중 2/3를 차지하며 세계보건기구(WHO)의 규정에 따라 1 - 4급으로 분류됩니다<sup>1</sup>. 다형성신경교아종(Glioblastoma Multiforme, GBM, WHO 4급 성상세포종)은 이러한 종양 중에서 가장 일반적이고 악성도가 높으며, 전 세계적으로 연간 발병률이 100,000명당 약 3-4명꼴로 알려져 있습니다(<http://www.ncbi.nlm.nih.gov/cancer/4938/glioblastoma/epidemiology>). GBM 환자의 전반적인 중간생존기간(median survival)은 15개월에 불과합니다<sup>2,3</sup>.

앞서 TCGA 프로젝트에서는 GBM 종양 집합에 대한 DNA 염기 서열, 복제수, DNA 메틸화 및 유전자 발현의 유전적 변화와 환자 임상 정보를 밝혀냈습니다<sup>4</sup>. 이 연구 결과는 인간 GBM 종양의 유전체, 전사체, 단백질체의 변화를 종합하고, 특징짓는 것에 대한 큰 관심을 불러일으켜 비정상적으로 기능하는 분자 경로와 종양 아형을 확인할 수 있었습니다. Phillips *et al.*<sup>5</sup> 및 Verhaak *et al.*<sup>6</sup>은 TCGA mRNA 발현 데이터 세트를 군집화 하여 원발 GBM 종양의 아형을 식별하였습니다. Phillips *et al.*<sup>5</sup>은 GBM을 3종의 아형, 즉 전신경(Proneural), 증식성(Proliferative), 중간엽(Mesenchymal)으로 분류하였으며, Verhaak *et al.*<sup>6</sup>은 다중(micro)array 플랫폼으로부터 mRNA 프로파일을 통합하여 전신경(Proneural), 고전(Classical), 중간엽(Mesenchymal) 및 신경(Neural)이라고 하는 4가지 발현 기반의 아형을 정의했습니다. Noushmehr *et al.*<sup>7</sup>은 GBM 종양의 프로모터 DNA 메틸화 변이를 프로파일링하여 전신경(Proneural) 아형에서 우선적으로 집중되는 신경교종-CpG 섬 메틸화 표현형(glioma-CpG island methylator phenotype, G-CIMP)을 정의하였습니다.

본 연구에서는 GBM 아형과 관련된 분자 이벤트를 이해하기 위해 GeneSpring/MPP 멀티 오믹스 소프트웨어의

메타데이터 및 상관 분석 프레임워크를 사용하였습니다.

## 요점

- mRNA 발현 기반의 TCGA 연구<sup>6</sup>로 정의된 GMB의 분자 아형에 대한 이해를 한층 전개하기 위해, 메타데이터 분석 프레임워크를 이용해 아형과 유전체 암 복제수 이상(Copy Number of Aberration, CNA) 사이의 관계를 탐구하였습니다.
- TCGA에서 얻은 mRNA와 miRNA 데이터 세트의 상관 분석을 통해, 알려진 전신경 시그니처 유전자 OLIG2, NKX2-2 및 SOX11를 조절하는 잠재적 후보 miRNA를 식별하였습니다. p53 신호 전달 억제를 유발하는 전신경 종양 아형의 분자 이벤트를 중점적으로 소개하였습니다.
- Label-free 단백질체학적 접근을 이용하여 측정한 GBM 및 대조 환자 시료(뇌전증 수술에서 얻은 뇌 절편)의 독립적 발굴 단백질체학 연구에 대한 분석을 수행하였습니다. 시료 간 상관 분석과 PCA 분석 결과는 대조 시료와 종양 시료를 명확히 분리할 수 있었으며, 또한 종양 시료를 2~3개 하위 집단으로 분리하였습니다.
- TCGA 연구에서 얻은 mRNA 발현 데이터와 label-free 연구에서 얻은 단백질체학 데이터를 결합 분석하여 GBM 아형 시그니처의 하위 집합을 식별함으로써 대규모 TCGA 코호트에서 알려진 종양 아형을 명확히 구분하였습니다.

## 데이터 세트 및 분석법

### TCGA 시료 세부정보

[https://tcga-data.nci.nih.gov/docs/publications/gbm\\_exp/](https://tcga-data.nci.nih.gov/docs/publications/gbm_exp/)에서 다중 array 플랫폼으로부터 얻어진 173개 GBM 종양 시료에 대한 840 개 mRNA 프로파일을 통합 발현값(unified expression value) 형태로 정리한 mRNA 발현 데이터 세트를 획득하였습니다. Agilent 8 × 15K array

를 통해 연구한 534개 miRNA의 발현 데이터는 TCGA 포털에서 다운로드하였습니다. mRNA, miRNA 및 단백질 발현 데이터 세트는 GeneSpring 13.0 (<http://genespring-support.com/>)에서 Generic Single-Color Expression Experiments로 불러오기 하였습니다.

암 복제수 이상 데이터(CNA)는 Brennan *et al.*이 작성한 환자 중심표 ([https://tcga-data.nci.nih.gov/docs/publications/gbm\\_2013/](https://tcga-data.nci.nih.gov/docs/publications/gbm_2013/))에서 얻었습니다. 이 표는 TCGA 코호트에서 시료에 대한 gene-wise CNA를 제공하였습니다. CNA는 측정된 암 복제수(정상 정배수성 유전체 대비 log2 비율로 표현)와 유전체의 변화 정도 측정을 통해 규명됩니다. Brennan *et al.*<sup>8</sup>에 의해 결정된 이상 항목 임계값은 다음과 같았습니다.

- 결손( $x \leq -1$ )
- 부분 결손( $-1 < x \leq -0.7$ )
- 손실( $x < -0.2$ )
- 획득( $0.2 \leq x < 1$ )
- 중간 증폭( $1 \leq x < 2$ )
- 증폭( $x \geq 2$ )

### 단백질체학 시료 세부사항

GBM 및 뇌전증(epilepsy) 표본은 뇌의 측두부, 전두부, 후두부에서 채취하였습니다. GBM 종양 조직과 뇌전증 뇌 조직 표본은 조직 절제 시 바로 취하여 30분 이내에 액체 질소에서 급속 동결시켰습니다. 시료는 질량 분석(MS)이 필요할 때까지 Medical College of Wisconsin(MCW)의 Brain and Spine Tissue Bank에  $-80^{\circ}\text{C}$ 로 보관하였습니다.

MS를 이용한 고효율 label-free 정량법으로 GBM의 종양 조직 표본 10개와 뇌전증 대조 시료 10개에 대해 포괄적 분석을 실시하였습니다. 모든 조직 시료는 이전에 설명한 것과 같이 처리하고 트립신으로 분해하였습니다<sup>9</sup>. 균질화된 종양 조직 단백질체는 SDS-PAGE로 분리하였습니다. PAGE 중에 내부 DNA 마커를 추가하여 겔 기반 분리를 일관성 있게 진행하였습니다. 겔에서 트립신 분해한 펩타이드는 Agilent 6550 iFunnel Q-TOF MS와 연결된 Agilent HPLC-Chip/MS

시스템의 데이터 의존(data-dependent) 모드에서 3회 반복 분석하였습니다. Q-TOF 데이터는 먼저 Agilent Spectrum Mill 소프트웨어에서 UniProt 인간 데이터베이스로 처리 하였으며 검색 결과는 1% FDR(False Discovery Rate)에서 자동 검증되었습니다. Spectrum Mill은 데이터베이스 검색 결과에서 얻은 단백질 정량데이터를 GeneSpring 소프트웨어의 MPP 모듈로 내보내 통계 분석을 가능케 하는 독보적인 기능을 갖추고 있습니다.

## 데이터 분석

이 연구에서 보고된 모든 분석은 GeneSpring/MPP 13.0 멀티 오믹스 생물정보학 소프트웨어로 수행하였습니다. GBM 아형은 유클리드 거리 지표와 Ward 연결법에 따라 840 개 mRNA 프로파일에 대해 계층적 군집 분석을 수행하여 재현하였습니다. TCGA에서 얻은 복제수 값과 기타 시료 파라미터의 유전체 이상 정보는 메타데이터 형식으로 GeneSpring에 보내졌습니다. GeneSpring 메타데이터 프레임워크는 유전자 발현 분석의 일반적인 간접 요인으로 알려진 배치 효과 및 시료 품질 파라미터를 평가하는데 사용하였으며, 또한 복제수 변화와 발현 아형 간의 상관관계를 시각화하는데 사용하였습니다.

TCGA 시료 데이터에서 GBM 아형을 확인하고 단백질체학 연구에서 시료 그룹을 식별하기 위해 주성분 분석(PCA)을 수행하였습니다. 데이터는 PCA를 수행하기 전에 평균중심화 및 척도화 과정을 거쳤습니다.

Gene ontology (GO) 분석은 GO Consortium(<http://geneontology.org>)에서 제공하는 GO terms(생물학적 기능의 논리구조)에 근거하여 수행하였습니다. 계층적 군집 분석을 통해 얻은 군집은 GenSpring의 애질런트 전장 유전체 array 실험으로 변환시켜 유의한 GO terms를 식별하는데 사용하였습니다. 군집에 집중된 GO terms를 식별하기 위해, 모든 알려진 해독 유전자에 대하여 초기하분포를 적용하였습니다.

GeneSpring 상관 분석 프레임워크는 단일 오믹스 플랫폼에서 연구한 분자 간의 상관관계 또는 2가지 오믹스 플랫폼 간의 상관관계뿐만 아니라, 단일 오믹스 기술로 측정한 시료 간의 상관관계도 지원합니다. 이 연구에서는 모든 시료와 개체(entity) 간 상관 분석에 Pearson 유사성 지표를 사용하였습니다.

mRNA와 miRNA 측정값이 모두 제공되는 159개 GBM 시료를 사용하여 GBM의 mRNA 프로파일과 miRNA 프로파일 사이의 상관관계를 분석하였습니다. 신경발생에 관여하는 유전자의 표적 miRNA는 miRWalk 데이터베이스(<http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/>)에서 조회하여 식별했습니다.

## 결과 및 토의

### GBM의 mRNA 발현 기반 분자 아형

173개 GBM 시료에 대한 840개 mRNA 프로파일의 통합 발현값은 그림 1A와 같습니다. 예상한 바와 같이 계층적 군집 분석을 통해 GBM 아형과 일치하는 4가지 주요 군집을 밝혀냈습니다. 뿐만 아니라, 개체의 군집 분석 결과 9개 군집을 얻었고 이들은 GBM 아형에서의 발현 특징에 의해 정의되었습니다(그림 1A). 구체적인 토의 내용은 miRNA 섹션을 참조하십시오.

관찰된 종양 아형이 배치 효과의 영향을 받을 가능성을 배제하기 위해, GeneSpring의 메타데이터 프레임워크를 사용하여 시료 품질 관리 파라미터(예: 배치, 시료 순도(% 종양 핵), 시료 품질(% 세포괴사))와 시료 간의 상관관계를 시각화 하였습니다. 그림 1B는 173개 GBM 시료의 군집 트리에 맞추어 시료 품질 파라미터를 열지도로 요약한 것입니다.

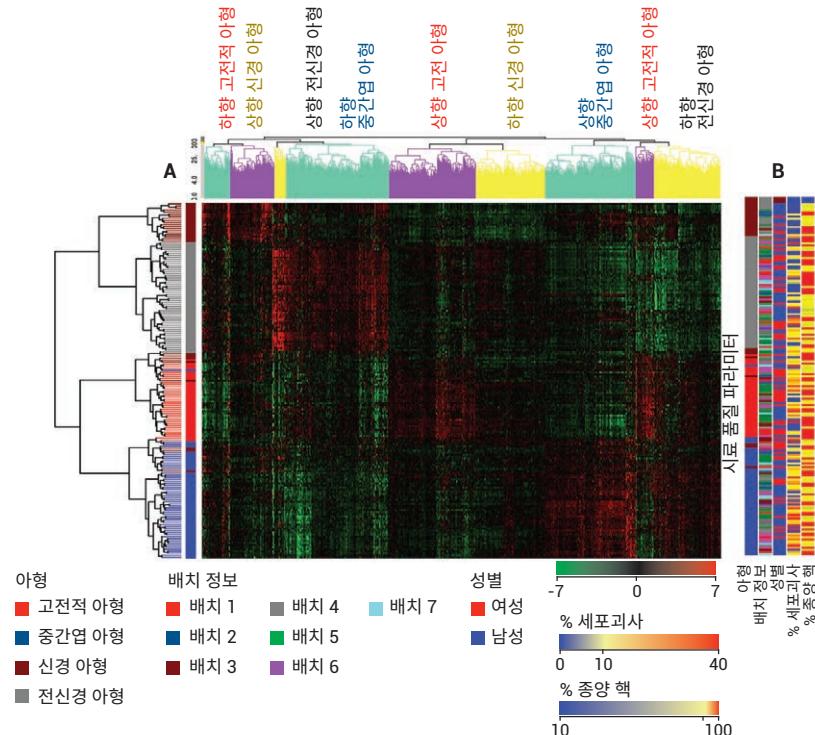


그림 1. A) Verhaak et al.<sup>6</sup>이 정의하고 Agilent GeneSpring으로 검증한 GBM 분자 아형. 이 개체는 9가지 서로 다른 군집으로 그룹화 될 수 있었음. B) 열지도로 시각화한 시료 품질 메타데이터

고효율 실험에서 배치 효과는 고질적인 문제입니다. 어떤 경우에는 생물학적 현상을 가리기도 합니다. 이 GBM 시료에서는 배치 효과 또는 성별 편향이 관찰되지 않았습니다. 또한 시료 순도와 GBM에 정의된 분자 아형 사이에 어떠한 일치도 보이지 않았습니다. 메타데이터 분석을 통해, GBM의 mRNA 발현을 기준으로 정의된 분자 아형이 유전체 분석에서 일반적으로 나타나는 실험적 인공 산물에 기인한 것이 아니라 사실상 완전히 생물학적이라는 것을 확인하였습니다.

동일한 mRNA 발현 프로파일을 사용하여 173개 GBM 시료를 대상으로 Pearson 상관 분석을 실시하여, 군집 분석에서 관찰된 아형을 확인하였습니다. 시료 상관 분석은 시료 그룹 간에 유사성을 보이는 경우 하나의 그룹 및 조건 내 시료의 발현에도 유사성이 있음을 나타냅니다. 그림 2A 와 같이, 동일한 아형에 속하는 대부분 시료는 발현 프로파일의 유사성을 나타내며 서로 양의(positively)

상관관계를 보였습니다. 시료 간 변동성은 전신경 아형에서 가장 낮게 나타났습니다(그림 1A 및 그림 2). 전신경 아형은 신경세포분화의 발현을 모방하고 증식 속도가 느린 것으로 알려져 있습니다<sup>5</sup>. 이것은 전신경 시료의 군집이 다른 아형에 비해 더욱 밀집해 있는 이유를 설명합니다. 또한 전신경 종양의 증식 속도는 중간엽 및 고전적 아형에 비해 느리다고 알려져 있습니다. 이 사실과 일치하게 거의 모든 전신경 시료가 중간엽 및 고전적 시료와 발현 프로파일에서의 극명한 대비를 이루며 음의(negatively) 상관관계가 있는 것으로 관측되었습니다. 시료 간 변동성이 가장 큰 것은 신경 아형이었습니다.

840 개의 mRNA 프로파일에 대한 PCA를 수행하여 GBM 아형의 존재를 추가로 확인하였습니다. PCA의 점수표는 그림 2B와 같습니다. PCA에서 주성분 1(PC1)은 데이터 세트에서 주된 분산을 점유하는 고유 벡터입니다. 이 연구에서는 PC1을 기준으로 시료

그룹을 분류하기는 했지만 PC1에 걸친 최대 분산은 전신경과 중간엽 종양 사이에서 관측되었습니다. PC2는 고전 및 신경 아형 시료 그룹 사이의 분산이 차지했습니다.

### GBM 분자 아형과 상관관계를 갖는 복제수 이벤트

발현 아형을 구별되게 하는 유전체 이벤트에 대한 통찰을 얻기 위해 유전체 CNA와의 관계를 분석하였습니다. 그림 3A는 아형 열지도와 함께, EGFR, CDKN2A, PDGFRA, TP53, NF1, PTEN 및 FGFR2 유전자의 복제수 증폭 및 결손을 보여주는 열지도를 나타냅니다. 분자 아형에 따라 시료를 정렬한 결과, 매우 흥미로운 관계가 드러났습니다. 예를 들어, 고전적 아형의 GBM 시료는 TP53 또는 NF1의 복제수 변화가 없었습니다. 대신 CDKN2A의 결손과 EGFR의 증폭이 더욱 두드러졌습니다. 이 결과는 Verhaak *et al*<sup>6</sup>이 보고한 관측 결과를 확인시켜 줍니다.

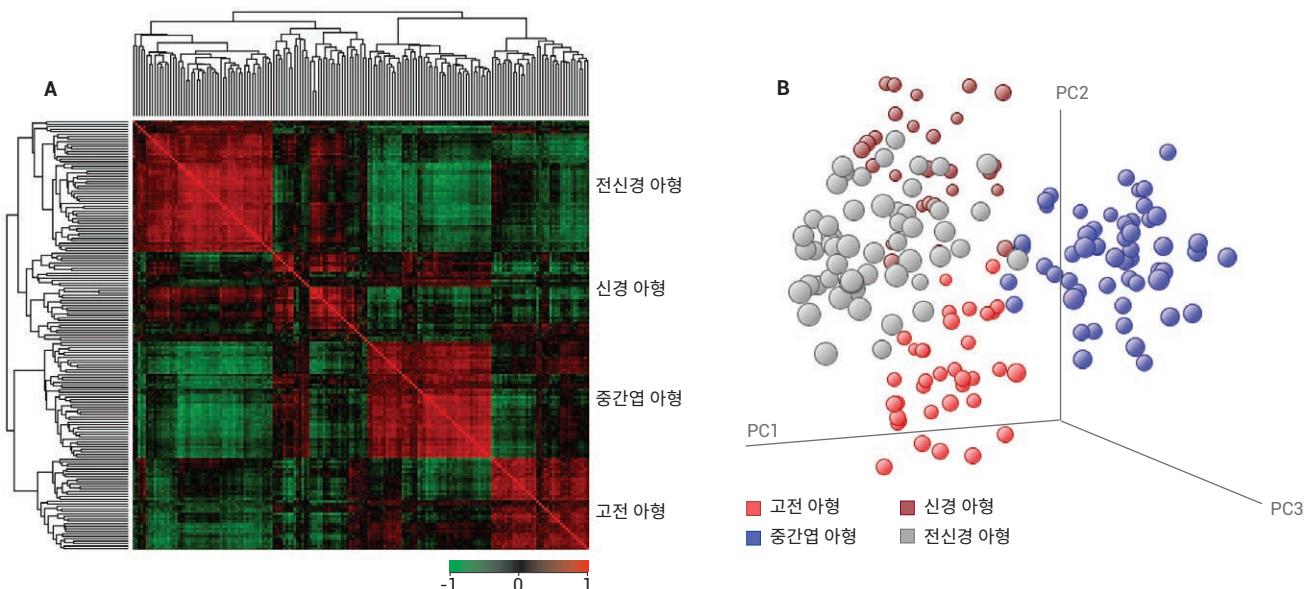


그림 2. Agilent GeneSpring에서 173개 GBM 시료에 대한 시료 간 상관 분석 및 PCA를 통해 TCGA 코호트 중 4가지 아형의 존재를 확인. (A) GBM 시료 간 Pearson 상관관계 및 (B) PCA 결과

주요 마커 유전자의 복제수 값으로 열지도를 정렬하면 GBM 아형 간의 분포를 더욱 명확히 알 수 있습니다. 그림 3B(NF1에 대해 정렬한 열지도)를 보면 고전 및 신경 아형뿐만 아니라 전신경 아형 종양에서도 대개 NF1

돌연변이를 일으키지 않는다는 것을 알 수 있습니다. 이와 대조적으로, Verhaak *et al.*<sup>6</sup>에서 보고된 바와 같이 NF1 반접합체 결손은 중간엽 아형의 특색입니다.

GeneSpring 메타데이터 프레임워크는 GBM 아형에서 발생하는 다른 주목할 만한 변화를 밝혀냈습니다.

- EGFR과 CDKN2A 돌연변이는 다수의 GBM 환자에게서 가장 흔히 나타납니다. 이러한 돌연변이는 고전적 아형에서 두드러지지만 어느 한 아형에 국한되는 것이 아니라 모든 아형에 걸쳐 발생합니다(그림 3A).
- 보고된 바와 같이, TP53 결손은 고전적 아형에서 눈에 띄게 없고, 전신경 아형에서 주로 관찰됩니다.
- PDGFRA가 높은 수준으로 증폭하는 것은 대부분 경우 전신경 아형에 국한되는 것으로 보고되었지만<sup>6</sup>, 다른 아형에서도 중간 수준의 PDGFRA 증폭이 나타나는 것을 관찰하였습니다 (데이터 미제공, PDGFRA 열지도로 정렬 시 관찰 가능).

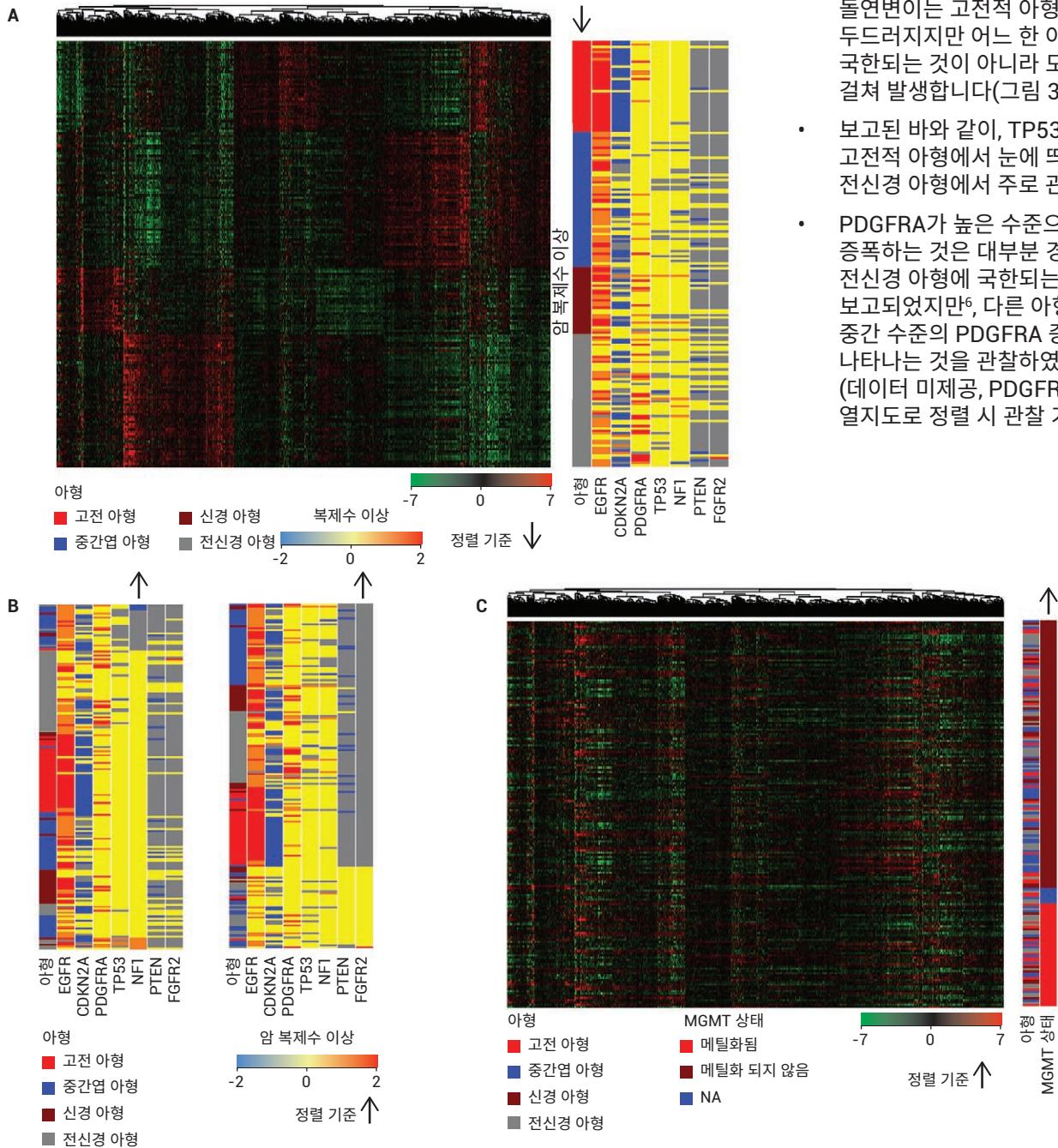


그림 3. A) 분자 아형에 따라 정렬한 GBM 시료와 관련 주요 유전자(EGFR, CDKN2A, PDGFRA, TP53, NF1, PTEN 및 FGFR2)의 암 복제수 이상(CNA). B) NF1과 FGFR2에 따라 각각 정렬한 유전체 암 복제수 이상, 화살표로 표시. (C) MGMT 유전자 메틸화 상태에 따라 정렬한 GBM 시료. 그림 3A에서 군집 패턴의 총 손실이 관찰됨

- MGMT 메틸화는 고전적 아형 치료 반응의 예측 임상 마커입니다<sup>8</sup>. 그러나 MGMT 메틸화와 GBM 아형을 연결할 만한 패턴은 찾지 못했으며, 이는 아직 규명되지 않은 요소가 존재함을 시사합니다 (그림 3C).
- FGFR2는 신경교종에서 종양 억제 유전자로서 기능하는 것으로 알려져 왔습니다<sup>10,11</sup>. 그러나 PathCards는 신경교종이나 GBM 경로에 FGFR2를 포함하고 있지 않습니다 (<http://pathcards.genecards.org/pathway/2538>).

- 그림 3B(FGFR2로 정렬한 열지도)에서는 환자 중 약 1/3에서 FGFR2 결손이 있다는 것을 보여줍니다. 결손 패턴은 PTEN 결손과 일치합니다.
- 그림 3B의 두 정렬된 열지도를 비교한 결과, 상당 수의 환자에서 FGFR2 결손 패턴이 PTEN 결손 패턴을 밀접하게 반영하였습니다. 이러한 환자들은 각각 10q23.3 과 10q26.1에 위치한 PTEN과 FGFR2의 순서로 이어지는 단일 결손 이벤트를 포함하고 있을 가능성을 시사합니다.

- Klinz et al.<sup>12</sup>은 GBM 아형의 복제수 이벤트 빈도를 요약하였습니다(표 1). GeneSpring의 메타데이터 프레임워크를 사용하여 분석한 내용은 Klinz et al<sup>12</sup>이 요약한 복제수 이벤트와 대체로 유사했습니다.

#### 아형 특이적 발현에서의 miRNA 조절

그림 4는 계층적 군집 분석(그림 1A)에서 확인된 9개 군집의 mRNA 발현 패턴을 나타냅니다. 군집에 대해 GO functional enrichment 분석을

표 1. 교모세포종(Glioblastoma) 아형에서의 암 복제수 이상 빈도(Klinz et al.<sup>12</sup>)

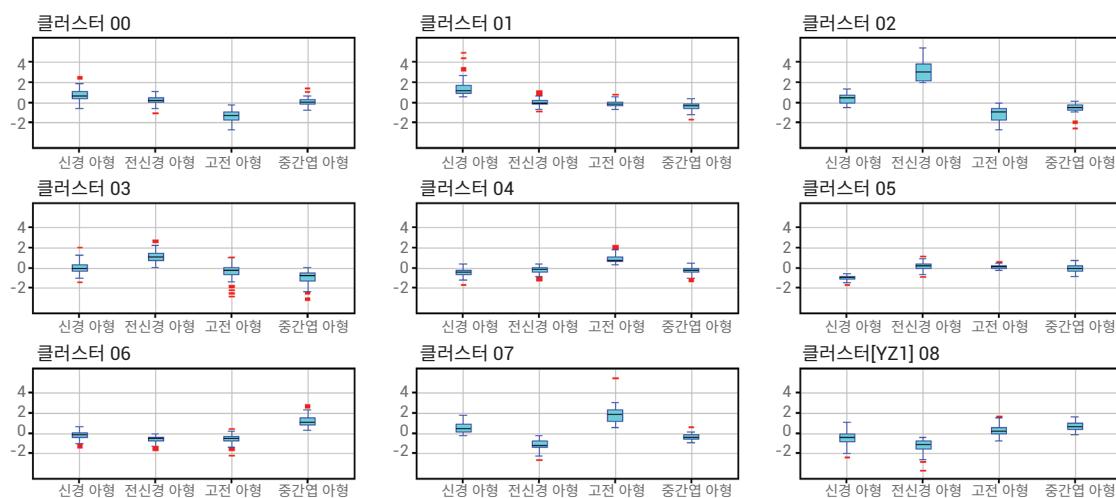
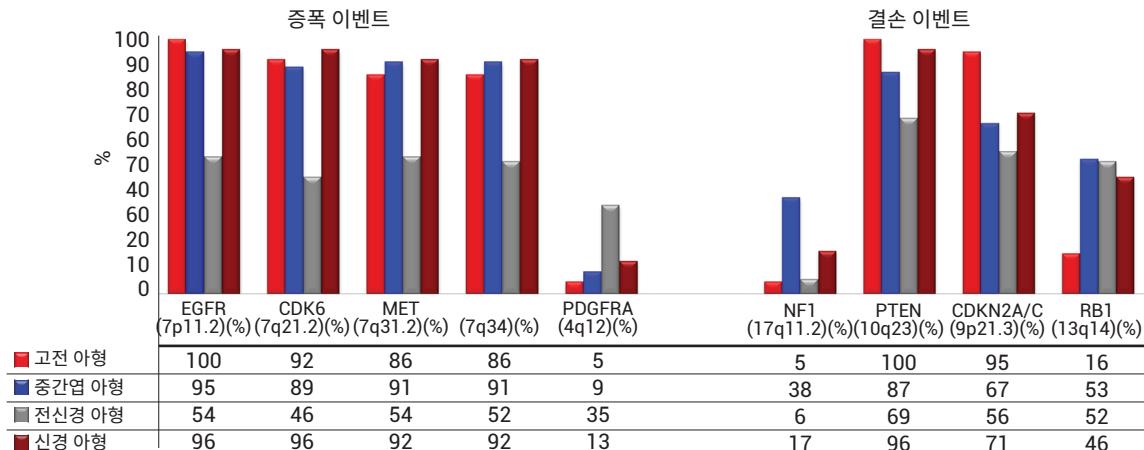


그림 4. 아형 특이적 발현 패턴을 나타내는 9개 군집의 상자 도표

실시한 결과, 세포외기질 조직, 상처 치유 반응, 스트레스 반응, 신경계 발달, 신경아교세포 분화, 액틴세포골격 조직, 염색질 조직 등의 세포생리학 이벤트에 주로 관여하는 840개의 시그니처 유전자를 식별하였습니다. GeneSpring에서 추가 분석을 수행하여 GBM에 영향을 미치는 경로에 참여하는 것으로 알려진 유전자(예: RTK, AKT, RB, P53 및 PI3K)와 문자 아형을 정의하는 유전자 시그니처(데이터 미제공) 사이에 유의한 중첩이 없다는 것을 확인하였습니다.

GO 분석에서는 신경계 발달(즉, 신경발생)을 전신경 아형에서 상향 조절되고 중간엽 아형 종양에서 하향 조절되는 유전자 군집에 집중된 주요 생물학적 프로세스로서 식별하였습니다 (그림 1A). 군집에서 25개의 유전자가 신경계 발달에 관여하는 것을 기능적으로 해독하였습니다(그림 5). 이 25개 유전자를 대표 전신경 아형 유전자로 선택했습니다. 선택된 유전자는 앞서 보고한 전신경 시그니처 유전자 OLIG2, NFKX2-2, SOX11, ASCL1을 포함하였습니다<sup>5,6</sup>.

MicroRNA는 신경교종의 종양 발생 및 종양 억제 경로를 조절함으로써 종양 형성에 관여합니다<sup>13,14</sup>. miRWalk 데이터베이스를 조회하여 신경계 발달에 관여하는 표적 유전자로 보고되어왔던 miRNA를 식별하였습니다<sup>15</sup>. miRWalk 데이터베이스에서 검증된 것으로 표시된 miRNA는 하향 분석을 고려하였습니다.

GeneSpring의 멀티 오믹스 상관 분석 프레임워크를 사용하여 모든 아형을 대상으로 신경계 발달에 관여하는 25개 mRNA와 검증된 38개 miRNA의 Pearson 쌍별 상관관계를 분석하였습니다. 분석을 위해 mRNA와 miRNA 측정값이 제공되는 TCGA의 159개 GBM 시료 그룹을 사용하였습니다. 전신경과 중간엽 아형 사이에서 발현에 유의한 차이를 보인 38개 miRNA 중( $t$ -검정  $p \leq 0.05$ 에서 조정), miR-222, miR-221, miR-223, miR-34a, miR-29a, miR-27a, miR-155 및 miR-21는 25개 mRNA에 대해 음의 상관관계를 보였습니다(그림 6).

중간엽 아형에서 전신경 아형에 비해 1.5배 상향 조절된 miRNA는 신경계 발달에 관여하는 유전자와 음의 상관관계를 보이는 군집에 유의하게 집중되어 있습니다. 그림 6에서는 집중된 miRNA를 빨간색 막대로 표시했습니다. 이 군집의 miRNA-mRNA 쌍은 대부분  $-0.3 \sim -0.5$  범위에서 음의 상관관계를 가집니다.

해당 GBM 아형에서 mRNA-miRNA 조절 요소를 더욱 자세히 이해하기 위해, 그림 6에서 음의 상관관계를 보였던 miRNA와 보고된 전신경 아형 마커 4개 (OLIG2, NFKX2-2, SOX11 및 ASCL1)를 선택하였습니다. 이렇게 선택된 mRNA와 miRNA 사이의 상관관계는 각각의 전신경 및 중간엽 아형 시료에서 발현 유사성을 기준으로 분석하였습니다. 상관관계 분석 결과는 그림 7과 표 2로 나타났습니다.

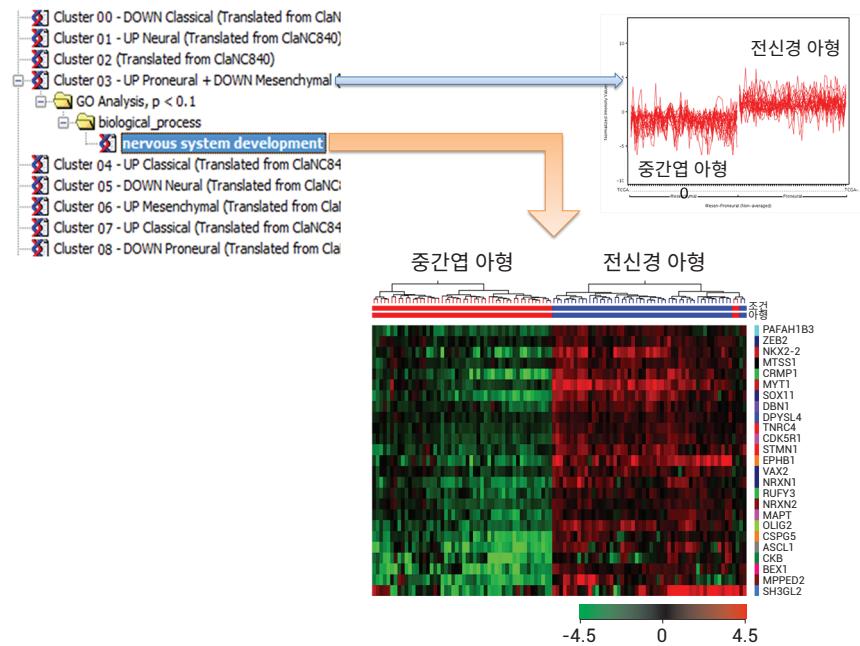


그림 5. 전체 유전자 집합 대비 전신경 아형에서 상향 조절되는 유전자 및 중간엽 종양 아형에서 하향 조절되는 유전자에 대한 GO 분석. 신경계 발달에 관여하는 25개 유전자를 규명함

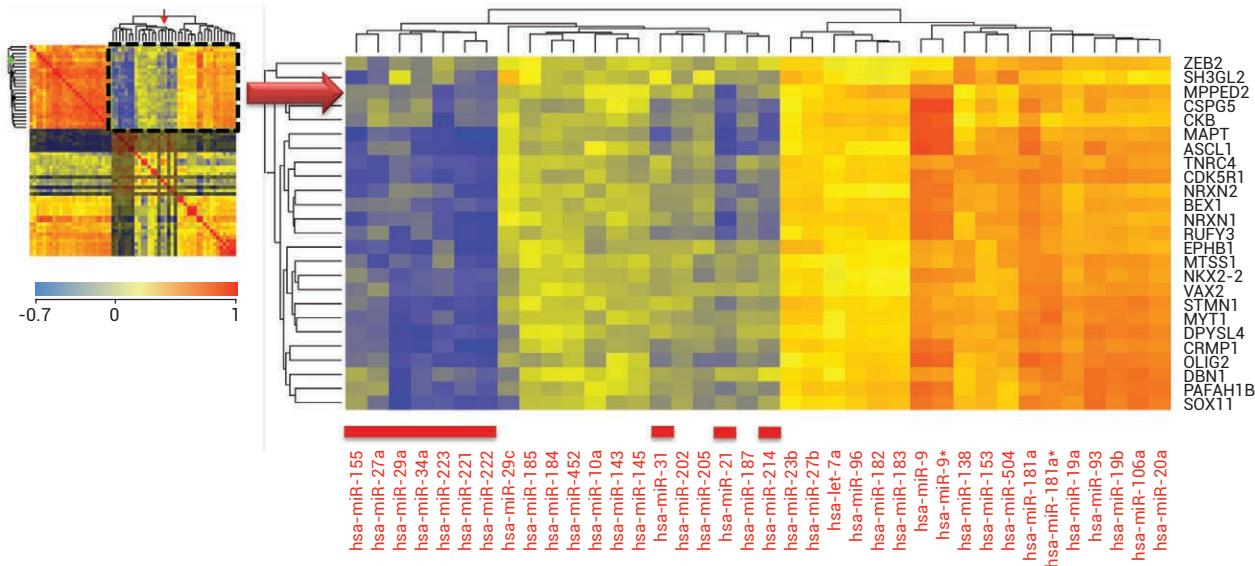


그림 6. 신경계 발달에 관여하는 유전자 25개와 전신경 아형 및 중간엽 아형에서 유의한 차이를 보이는 발현(t-검정  $p \leq 0.05$ 에서 조정)을 가진 38개 miRNA 사이의 관계를 Pearson 쌍별 상관 분석한 결과. 확대 보기는 x축에 miRNA, y축에 mRNA를 배치한 두 번째 사분면을 나타냄. 유사성은 4개 아형에 걸친 mRNA 프로파일과 miRNA 프로파일을 기준으로 계산

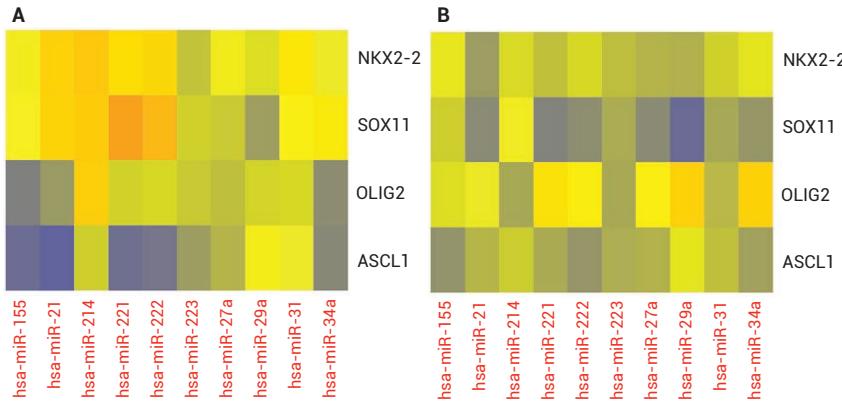


그림 7. OLIG2, NKX2-2, SOX11 및 ASCL1과 중간엽 종양 아형에서 상향 조절을 나타낸 miRNA의 상관관계. 전신경 아형 시료(A) 및 중간엽 아형 시료(B)에서의 상관관계

표 2. 전신경 아형 마커 NKX2-2, SOX11 및 OLIG2와 전신경 아형 시료에서 하향 조절된 일부 miRNA 사이의 상관 계수 관측값

	OLIG2		NKX2-2		SOX11	
	전신경 아형	중간엽 아형	전신경 아형	중간엽 아형	전신경 아형	중간엽 아형
hsa-miR-21	-0.278	-0.036	0.18	-0.272	0.168	-0.322
hsa-miR-214	0.193	-0.238	0.218	-0.106	0.2	-0.014
hsa-miR-221	-0.124	0.12	0.139	-0.17	0.36	-0.34
hsa-miR-222	-0.112	0.078	0.155	-0.113	0.276	-0.31
hsa-miR-223	-0.149	-0.236	-0.162	-0.178	-0.127	-0.23
hsa-miR-27a	-0.175	0.072	-0.007	-0.208	-0.145	-0.317
hsa-miR-29a	-0.12	0.18	-0.093	-0.216	-0.267	-0.41
hsa-miR-31	-0.113	-0.195	0.101	-0.126	0.067	-0.239
hsa-miR-34a	-0.312	0.177	-0.024	-0.069	0.091	-0.288

증식과 침습성을 유도하는 것으로 알려진 miR-21<sup>16</sup>은 전신경 아형 시료의 전사 활성자인 NKX2-2 및 SOX11와 양의 상관관계를 보였으나, 중간엽 아형 시료에서는 음의 상관관계를 보였습니다. miR-21<sup>17</sup>의 표적 관련 문헌과 miRWalk 데이터베이스 조회에서는 NKX2-2와 SOX11가 miR-21의 표적으로 나타나지 않았습니다. 본 연구에서는 GBM 종양에서 miR-21이 잠재적으로 NKX2-2와 SOX11을 조절할 가능성이 있는 것으로 나타났습니다.

희돌기교세포 발달 조절자 OLIG2는 전신경 아형 시료에서 miR-34a, miR-21 및 miR-155와 음의 상관관계를 보였으나, 중간엽 아형 시료에서는 양의 상관관계를 보이거나 아예 상관관계를 나타내지 않았습니다. 이전에는 TP53 돌연변이와 LOH 이벤트가 전신경 아형의 특징인 것으로 규명하여 보고되었습니다. 이 연구에서는 전신경 코호트에서 희돌기교세포 발달 조절자 OLIG2의 발현은 증가하고 miR-34a의 발현은 감소하는 것을 관측하였습니다 (그림 8). OLIG2는 세포 운명을 결정하고 세포 증식을 제어하는 데 중추적인 역할을 하며, p53 종양-억제자 경로 작용인자인 p21을 억제하는 것으로 보고되었습니다<sup>18</sup>. miR-34a는 p53 신호 전달 경로를 조절하는 종양 억제자입니다<sup>19,20</sup>. 이는 전신경 아형 시료에서 TP53이 기능적으로 손실되는 원인이 유전자의 결손, 또는 OLIG2의 증가 및 miR-34a의 감소에 의한 억제일 가능성�이 있다는 것을 시사합니다.

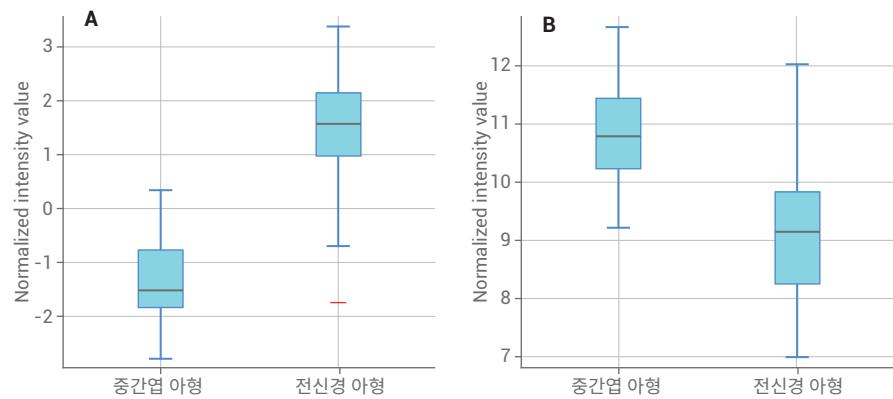


그림 8. 전신경 아형 시료와 중간엽 아형 시료에서 (A) OLIG2와 (B) miR-34a 발현을 나타낸 상자 도표

### 단백질체학 분석

mRNA 발현이 항상 직접적인 단백질 정량데이터로 번역되는 것은 아닙니다. 단백질 수준에서 GBM의 유전체 관찰 결과를 검증하기 위해 독립적인 단백질체학 데이터 세트를 생성하였습니다. RNA와 단백질 수준에서 유의한 GBM 시그니처를 식별하고 독립적 기술을 사용하여 독립 환자 코호트에서 TCGA 관찰 결과를 확인하는 것이 목적이었습니다.

### Label-free 단백질체학 분석

MPP 소프트웨어를 이용하여 GBM 종양 시료 10개와 대조 시료(뇌전증 뇌 조직) 10개의 label-free 단백질 데이터 세트에 대해 초기 프로파일링 분석을 수행하였습니다. 단백질 데이터베이스 검색 결과를 통해 전체 데이터 세트 중 고유 단백질 그룹 14,187개를

식별하였습니다. 발현 프로파일링에서 하나의 종양 시료가 대조군 시료와 유사함을 보여, 해당 시료를 이후의 분석에서 제외하였습니다(데이터 미제공).

9개 GBM과 10개 대조 시료를 통계 분석한 결과, 다르게 발현된 단백질 587개를 식별하였습니다(T-검정, 보정된  $p$ -값  $\leq 0.05$  및  $FC \geq 2.0$ ). 이 587개 단백질은 대조 시료와 종양 시료를 명확히 구분하였습니다. 이는 시료 간 상관관계 및 PCA 분석으로 확인하였습니다(그림 9A 및 9B). PCA와 상관 분석 결과는 대조 시료와 종양 시료의 단백질 발현 프로파일 간에 명확한 차이를 나타냈습니다. 또한, 연구 대상이었던 제한적 수량의 종양 시료에서 종양 하위 그룹의 잠재적 패턴을 식별하였습니다.

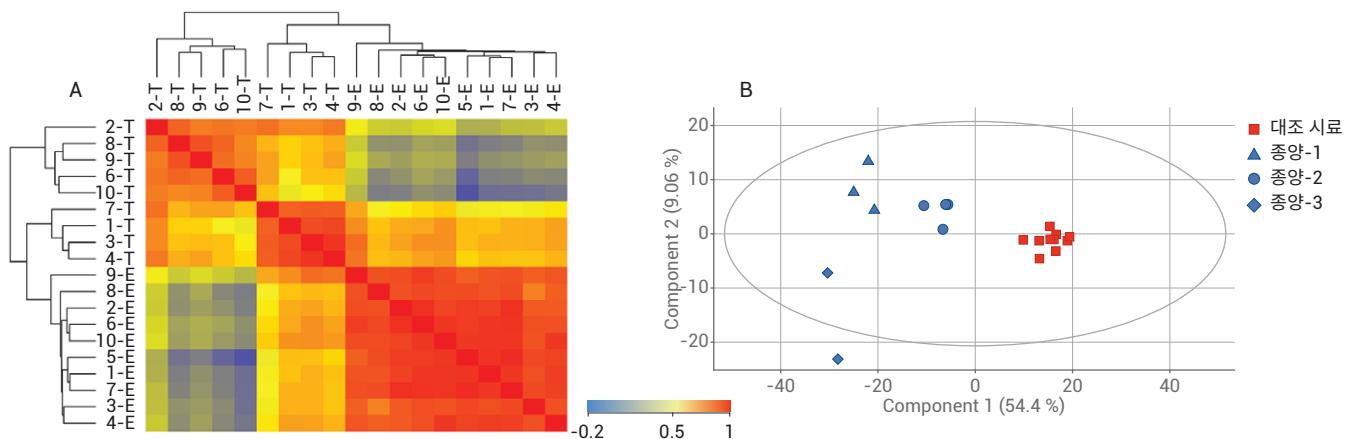


그림 9. 대조 시료(-E)와 종양 시료(-T)를 대조 시료와 종양 시료에서 다르게 발현되는 587개 단백질을 기준으로 분리. A) 시료 간 상관관계. B) PCA

GeneSpring 메타데이터 프레임워크를 사용하여 관찰된 하위 그룹 패턴을 종양 시료를 위해 제공되는 메타데이터로 설명할 수 있는지 평가하였습니다. 그림 10은 나이, 성별, 종양 위치, MGMT(GBM에 대한 알려진 임상 예후 마커)의 메틸화 상태와 같은 메타데이터로 정렬된 종양 하위 그룹을 보여줍니다. 이 메타데이터는 다르게 발현되는 단백질에 근거한 종양 내에서 관찰된 하위 집단을 명확히 설명하지 못했습니다.

### 단백질체학 및 유전자 발현의 종합적 분석

GBM 종양에서 관찰된 하위 그룹을 설명하기 위해, 단백질체학 연구를 통해 식별된 다르게 발현되는 단백질과 TCGA에서 식별된 다르게 발현되는 단백질을 통합적으로 분석하였습니다. 다르게 발현된 단백질 587개(이전 섹션 참조)와 Verhaak *et al.*<sup>6</sup>이 보고한 아형 특이적 유전자 840개 사이의 교집합에는 54개 유전자와 이들이 인코딩하는 54개의 단백질이 포함되었습니다(그림 11). 이 54개 문자를 사용하여 두 연구의 주요 발견을 확인할 수 있었습니다.

이전 섹션에서는 종양 시료와 정상 시료를 분리하고 587개의 다른게 발현된 단백질의 존재비를 기반으로 종양 하위 그룹을 식별하는 내용을 서술하였습니다. 그림 12A와 12B는 54개 단백질의 존재비를 기준으로 한 시료 간 상관계 열지도와 PCA 도표를 나타냅니다. 이 그림에서는 종양 하위 그룹은 물론이고 종양 시료와 정상 시료가 명확히 분리되었음을 확인할 수 있습니다.

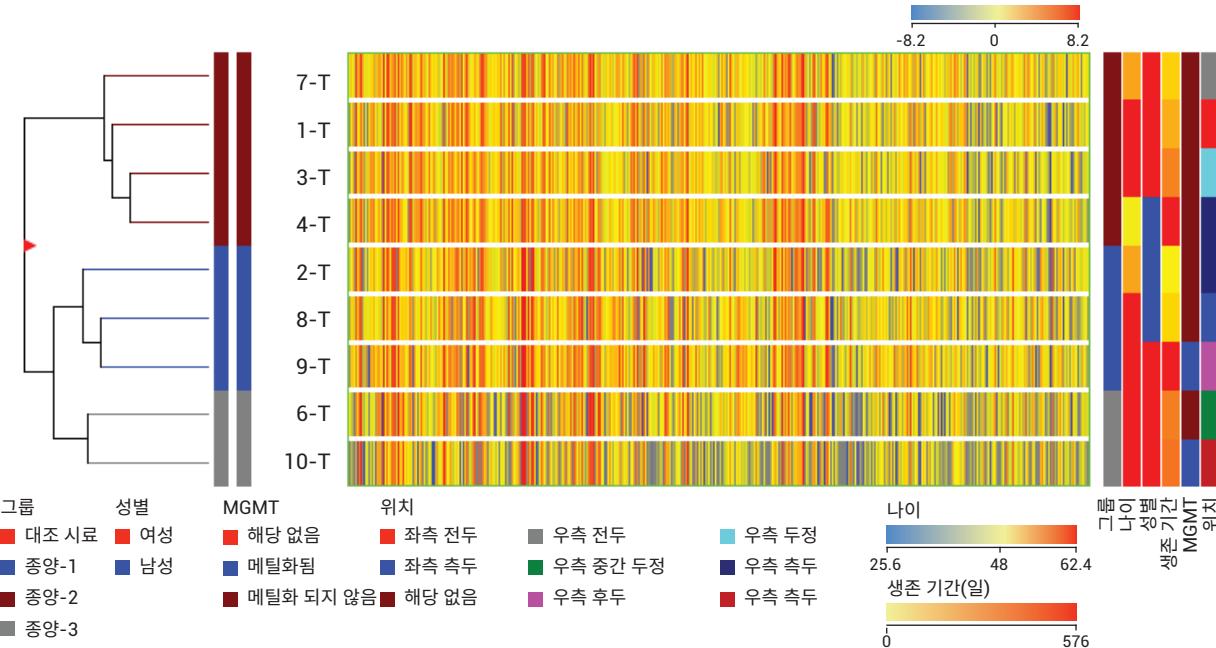


그림 10. 시료 메타데이터(나이, 성별, 생존 기간, MGMT 유전자 메틸화 상태, 뇌 위치)로 정렬한 종양 시료 군집 분석으로 관찰된 시료 그룹

프로파일링 분석으로부터  
다르게 발현된 587개 단백질  
T-검정 p-값  $\leq 0.05$  및 FC  
 $\geq 2.0$

단백질체학 실험을 통해  
측정된 TCGA의 아형 특이적  
유전자 840개 중 655개

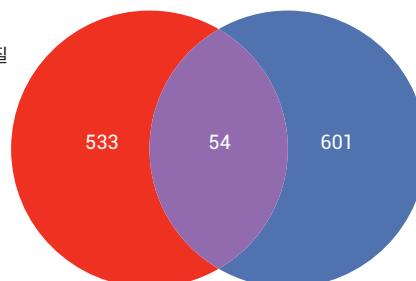


그림 11. 발굴 차이 분석의 587개 단백질과 TCGA의 840개 아형 특이적 유전자 사이의 중첩 부분을 보여주는 벤 다이어그램. 840개 유전자 중 655개는 단백질체학 실험을 통해 측정

이 단백질체학 연구에서 관측된 하위 그룹이 GBM에서 보고된 분자 아형과 관련이 있는지를 검증하기 위해 54개 단백질에 대응하는 유전자를 TCGA mRNA 발현 데이터에 기반하여 군집화 하였습니다. 그림 13은 54개 유전자의 mRNA 발현 프로파일에 대해 실시한 유전자 시료의 군집 분석 결과입니다. 그림 13은 54개 유전자의 발현이 GBM 종양들을 4개의 보고된 분자 아형으로 높은 수준의 확실성에서 분류할 수 있다는 것을 확인해줍니다.

## 결론

GeneSpring/MPP 소프트웨어는 복잡한 생물학적 데이터 세트를 철저히 조사할 수 있는 강력한 데이터 분석 도구 모음을 제공합니다. 이 사례 연구에서는 기존 및 새로 작성된 유전체와 단백질체 데이터 세트를 활용하여 GBM 종양에 대한 다차원 분석을 수행하였습니다. GeneSpring/MPP 소프트웨어에서 현재 제공하는 혁신적 메타데이터와 상관관계 프레임워크를 사용하여 대규모 TCGA 연구의 주요 발견을 재현하고 확장하였습니다. 예를 들어, 단일 GeneSpring 실험을 통해 TCGA에서 보고한 GBM 종양의 유전체 아형 분류를 확인할 수 있었습니다. 뿐만 아니라, TCGA에서 GBM 주요 유전자의 체세포 결손을 이미 보고한 바 있지만, 본 연구에서는 GeneSpring 메타데이터 프레임워크를 통해 동일한 환자에게서 예상보다 높은 비율의 FGFR2 및 PTEN 동시 결손을 밝혀냈습니다. 이 중요한 발견은 2개의 암 관련 유전자에서 이상이 발생한 것이 두 개의 독립적 결손 이벤트 때문이 아니라 단일의 유전체적 재배열 때문일 수 있다는 가능성을 시사합니다.

여러 가지 오믹스 기술에서 얻은 데이터를 비교한 결과를 통해 GBM 종양에 관한 알려지지 않은 생물학적 정보를 발견할 수 있습니다. 예를 들어, TCGA 프로젝트에서는 GBM 종양을 4개의 알려진 아형으로 분류할 수 있는 유전자 840개를 식별하였습니다. 이 연구의 단백질체 데이터는 정상 조직 대비 종양 조직에서 다르게 발현되는 587개 단백질을 식별했으며, 단백질 발현만으로 종양을 분류할 수 있음을 나타냈습니다. GeneSpring/MPP 상관관계 프레임워크에서 유전체와 단백질체 데이터 세트를 비교하여 오직 54개 유전자와 단백질로 구성된 핵심 시그니처를 식별하였습니다.

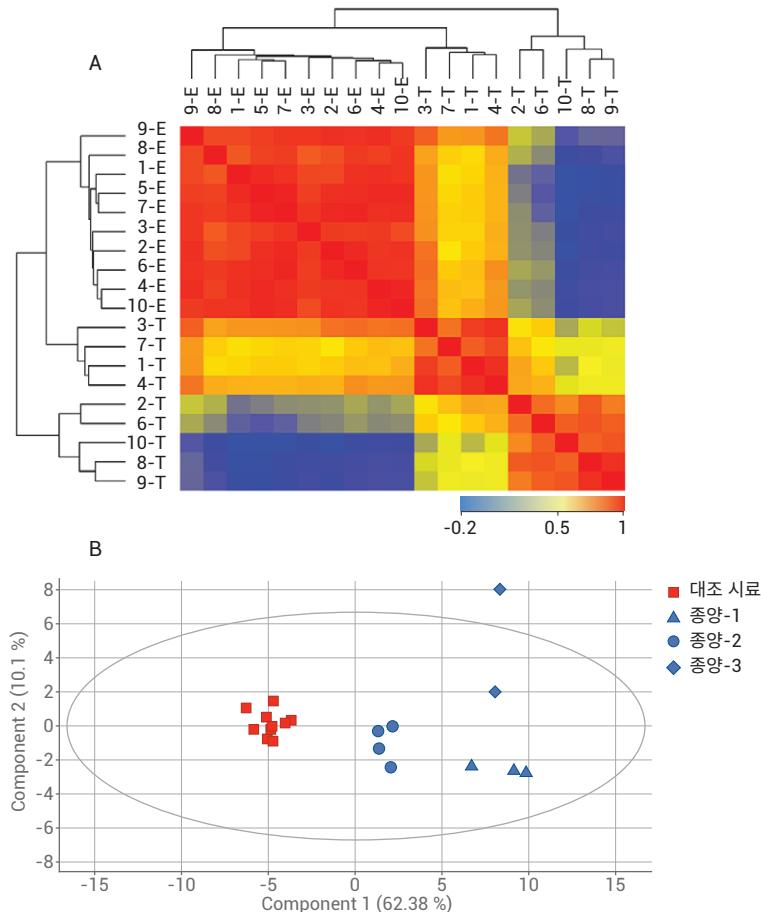


그림 12. Agilent GeneSpring의 TCGA mRNA 발현 실험으로부터 유전자 시그니처를 번역하여 식별한 54개 단백질을 기초로 한 대조 시료(-E)와 종양 시료(-T)의 분리. A) 시료 간 상관관계, B) PCA

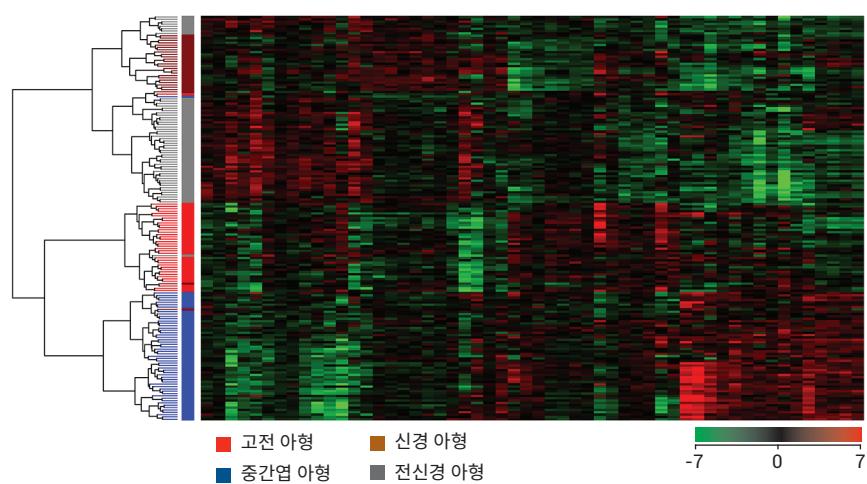


그림 13. 단백질체학 연구에서 식별한 54개 mRNA 시그니처의 발현을 기초로 173개 TCGA 종양 시료를 군집하여 GBM의 보고된 분자 아형을 규명

핵심 시그니처는 840개 유전자 및 587개 단백질에 필적하는 분류 성능을 입증하였고, GeneSpring의 멀티 오믹스 데이터 분석 능력을 잘 나타내 주었습니다. 관측 결과는 잠재적 진단 및 임상적 의의를 지닌 후속 연구의 가능성을 열어 두고, GBM 종양의 이질성은 시그니처 유전자의 mRNA와 단백질 정량데이터로 가장 잘 규정할 수 있음을 시사하였습니다.

## 감사의 글

Agilent GeneSpring 메타데이터 및 상관관계 프레임워크 개발 중 데이터, 피드백, 유익한 토론을 공유해준 Human Toxome Mapping Consortium 회원들에게 감사드립니다. MR은 NIH 보조금 1R01ES020750-01 "MAPPING THE HUMAN TOXOME BY SYSTEMS TOXICOLOGY"(PI T. Hartung)에서 일부 지원을 받았습니다. SPM은 James C. Benjamin Fund for Brain Tumor Research, Froedtert Foundation Grant, Advancing a Healthier Wisconsin - Clinical and Translational Science Institute Pilot Research Grant에서 지원을 받았습니다.

## 참고문헌

1. Kleihues, P; Sobin, L..H; World Health Organization classification of tumors. **2000, Volume 88, Issue 12**, p 2887.
2. Brandes, A. A; et al. Temozolomide as a second-line systemic regimen in recurrent high-grade glioma: A phase II study. *Ann. Oncol.* **2001**, 12, pp 255~257.
3. Martinez, R; Rohde, V; Schackert, G. Different molecular patterns in glioblastoma multiforme subtypes upon recurrence. *J. Neurooncol.* **2010** 96, pp 321~329.
4. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **2008**, 455, pp 1061~1068.
5. Phillips, H. S; et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* **2006**, 9, pp 157~173.
6. Verhaak, R. G; et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* **2010**, 17, pp 98~110.
7. Noushmehr, H; et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* **2010**, 17, pp 510~522.
8. Brennan, C. W; et al. The Somatic Genomic Landscape of Glioblastoma. *Cell Oct 10* **2013**, 155(2), pp 462~77.
9. Heroux, M. S; et al. Comprehensive characterization of glioblastoma tumor tissues for biomarker identification using mass spectrometry-based label-free quantitative proteomics. *Physiological Genomics* **2014**, 46, pp 467~481.
10. Toedt, G; et al. Molecular signatures classify astrocytic gliomas by IDH1 mutation status. *Int. J. Cancer.* **2011**, 128(5), pp 1095~1103.
11. PhosphoSitePlus: <http://www.phosphosite.org/proteinAction.do?id=1262&showAllSites=false>
12. Klinz, F. J; et al. In *Genetic Diversity of Glioblastoma Multiforme: Impact on Future Therapies, Glioma - Exploring Its Biology and Practical Relevance*; Dr. Anirban Ghosh, Ed.; **2011**; ISBN: 978-953-307-379-8.
13. Ventura, A; Jacks, T. MicroRNAs and cancer: short RNAs go a long way. *Cell* **2009**, 136(4), pp 586~591.
14. Calin, G. A; Croce, C. M. MicroRNA signatures in human cancers. *Nat. Rev. Cancer* **2006**, 6(11), pp 857~866.
15. Dweep, H; et al. miRWALK-database: prediction of possible miRNA binding sites by "walking" the genes of 3 genomes. *Journal of Biomedical Informatics* **2011**, 44, pp 839~47.
16. Møller, H. G; et al. A Systematic Review of MicroRNA in Glioblastoma Multiforme: Micro-modulators in the Mesenchymal Mode of Migration and Invasion. *Mol. Neurobiol.* **2013**, 47, pp 131~144.
17. Buscaglia, L. E; Li, Y. Apoptosis and the target genes of microRNA-21. *Chin. J. Cancer.* Jun **2011**, 30(6), pp 371~380.
18. Ligon, K. L; et al. Olig2-regulated lineage-restricted pathway controls replication competence in neural stem cells and malignant glioma. *Neuron* **2007**, 53, pp 503~517.
19. Hermeking, H. p53 Enters the MicroRNA World. **2007, Volume 12, Issue 5**, pp 414~418.
20. Silber, J; et al. miR-34a repression in proneural malignant gliomas upregulates expression of its target PDGFRA and promotes tumorigenesis. *PLoS One* **2012**, 7(3):e33844.

[www.agilent.com/chem](http://www.agilent.com/chem)

이 정보는 사전 고지 없이 변경될 수 있습니다.

© Agilent Technologies, Inc., 2015  
2015년 4월 2일, 한국에서 발행  
5991-5505KO

서울시 용산구 한남대로 98, 일신빌딩 4층 우)04418  
한국애질런트테크놀로지스(주) 생명과학/화학분석 사업부  
고객지원센터 080-004-5090 [www.agilent.co.kr](http://www.agilent.co.kr)



Agilent Technologies