Cancer Cell重磅推出 | 多达10种平台的组学数据，全面完善了胶质母细胞瘤（GBM）亚类分型，并挖掘出新的潜在药物治疗靶点，为精准治疗奠定了基础

候选标题：

教你如何玩转10个组学数据，绘制胶质母细胞瘤（GBM）亚类分型精确图谱，挖掘新的潜在药物治疗靶点，为精准治疗奠定了基础

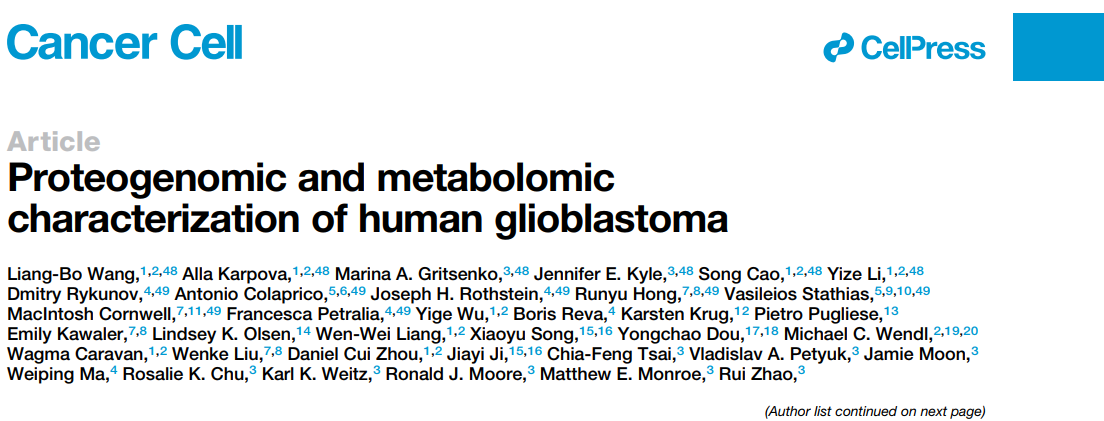
10个组学数据该怎么玩？不知道，不晓得，看着有点懵？那就往下看（C:\Users\Pioneer\AppData\Local\Temp\SGPicFaceTpBq\19360\607119CA.png）

10个组学数据！（C:\Users\Pioneer\AppData\Local\Temp\SGPicFaceTpBq\19360\607119CA.png）怎么，有点懵？一文带您玩转多组学，看研究人员如何全面解析胶质母细胞瘤（GBM）亚类分型精确图谱？

前言

2021年2月华盛顿大学医学院圣路易斯太平洋西北国家实验室Karin D. Rodland团队领导，并和国家癌症研究所（NCI）的临床蛋白质组肿瘤分析联合会（CPTAC）等多个研究机构或实验室合作，在Cancer Cell发表题为“Proteogenomic and metabolomic

characterization of human glioblastoma”的研究成果，通过整合来自多达10种平台的蛋白基因组学和代谢组学数据，揭示了在驱动胶质母细胞瘤（GBM）中起关键作用的基因、蛋白质、代谢物、浸润细胞以及信号通路的详细图谱，并利用不同组学层面信息对GBM进行全面分型。该项研究的研究材料是由CPTAC 收集的99位GBM未治疗的原始病例。



中文标题：人类恶性胶质瘤的蛋白基因组学和代谢组学图谱

研究对象：人胶质母细胞癌

发表期刊：Cancer Cell

影响因子：26.602

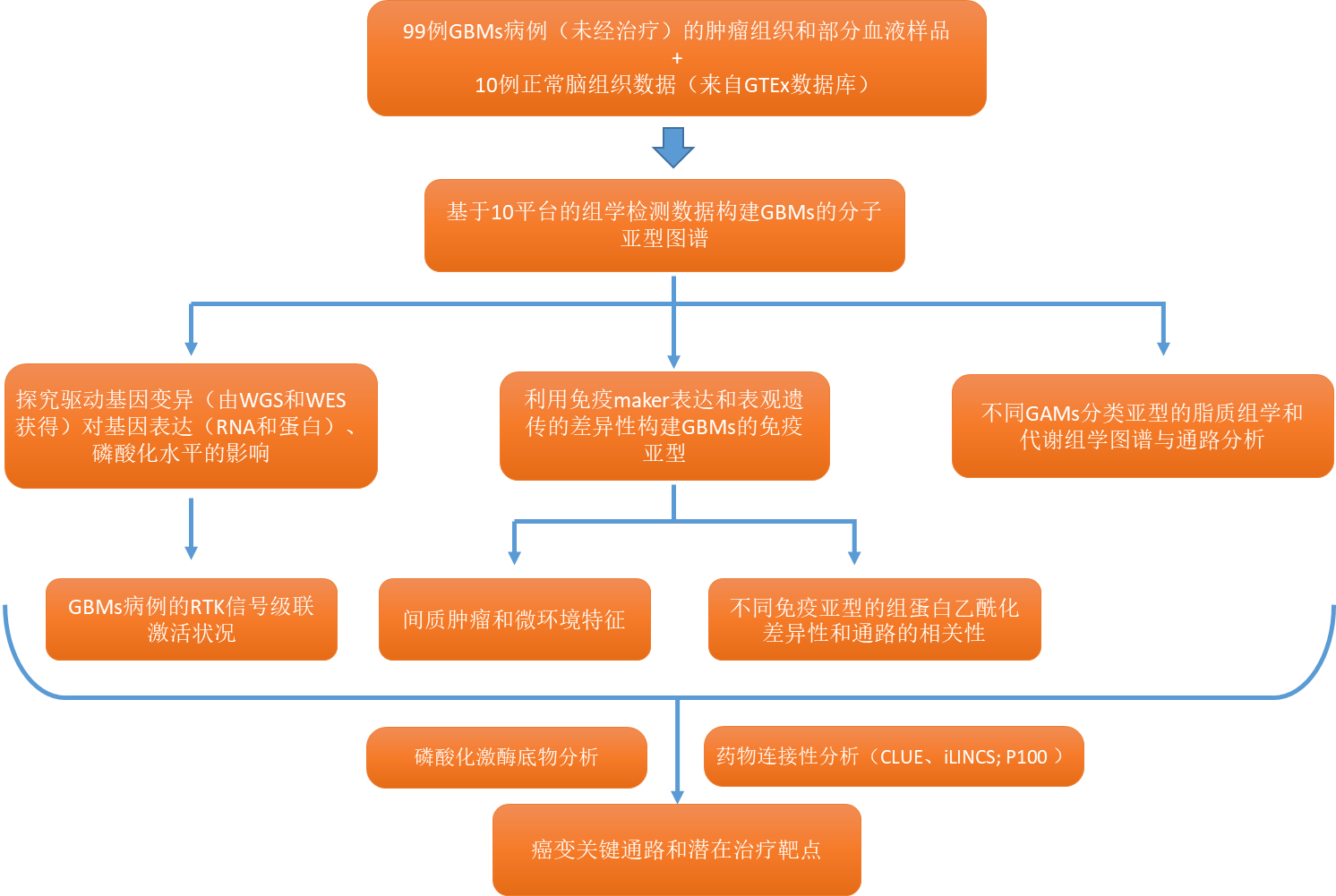
发表时间：2021年2月

运用生物技术：全基因组测序（WGS）、全外显子测序（WES）、全转录组测序（RNA-seq）、小RNA测序（miRNA-seq）、单细胞核测序（snRNA-seq）、DNA甲基化芯片、蛋白质组学、修饰（磷酸化，乙酰化）蛋白质组学、代谢组学、脂质组学等

研究背景

胶质母细胞瘤（GBM）是最具浸润性和破坏性的神经系统癌症之一，美国每年大约有12,000例新病例，中位生存期不到2年。在GBM亚型分类方面，已有一些研究进展，如基于驱动基因突变或甲基化进行分类，但适用面还不够。在治疗方面，手术切除、化学疗法和放射疗法仍然是治疗的标准，近来也提出了些有希望的免疫疗法，如免疫检查点抑制剂、疫苗、嵌合抗原受体T细胞（CAR-T）疗法和病毒疗法，但尚无明确的III期临床试验效果。尽管MGMT启动子甲基化的患者对替莫唑胺（Temozolomide）的反应很好，但对以转录组学方法分类的特定亚型患者，仍没有一种有效的治疗方法。

研究思路



修改见ppt

研究结果

1. 研究概况

本文运用10种组学技术对99例GBMs病例的癌症组织和部分患者的血液样本进行检测，同时结合GTEx数据库的10例正常人脑样品数据进行整合分析后（图1A），获得了基于免疫的新亚型，发现了在TP53突变型肿瘤中DNA修复途径的激活是通过提高了DNA修复基因的磷酸化位点水平，在受体酪氨酸激酶（RTK）改变的肿瘤中出现明显的磷酸化信号受限，而在经典样GBM肿瘤中出现组蛋白H2B乙酰化的富集和巨噬细胞含量较低的现象。另外，研究人员利用单细胞核测序技术探究了各类型细胞对肿瘤特征形成的贡献，并分析了间充质亚型，进而获悉了肿瘤和浸润性免疫细胞中的上皮-间质转化（EMT）特征。

1. 基于蛋白质基因组学和代谢组学特征的GBMs分子分型

研究人员对这10种组学数据以及前人研究数据进行了全面整合分析，并对不同单一组学和多组学数据进行聚类分型，构建出了GBMs的精细图谱（图1C）。

首先研究人员对基因组数据进行分析，发现了一类新亚型的GBMs病例--非热点位突变的*IDH1*突变亚型（传统的是*IDH1* R132H突变型和*IDH1*野生型）。与TCGA的GBMs队列基因组数据对比分析后，发现在癌基因（如*EGFR*和*PDGFRA*）以及肿瘤抑制因子*PTEN*和*NF1*中存在许多结构变异（SV）。WES和WGS结果显示，*TERT*启动子（*TERT*p）突变的等位基因变异频率（VAF）> 5％（图1B）。同时，拷贝数分析也鉴定到了常见的焦点和臂级拷贝数变异（CNV）。

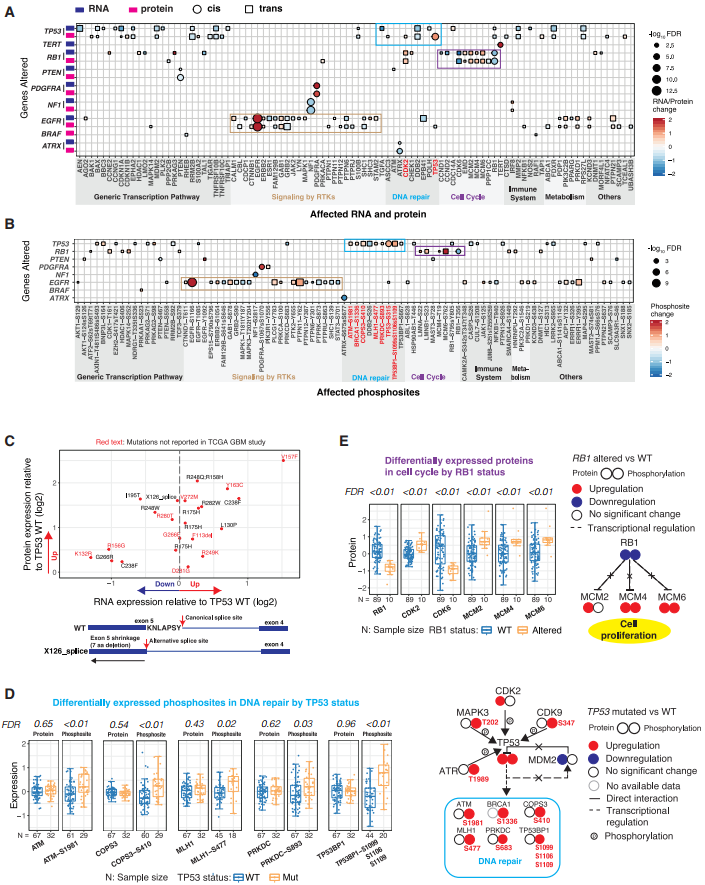
研究人员将RNA、蛋白质和磷酸化位点的表达数据与前人研究的基因表达数据（来自TCGA）进行对比分析后，发现具有相似的表达趋势，并将*IDH1*野生型聚类分型为3个亚型，而且这3个亚型具有明显差异独特的富集通路（图1C）。



图一 GBMs队列的蛋白基因组图谱总览

1. 驱动基因变异对癌基因蛋白表达丰度和磷酸化的影响

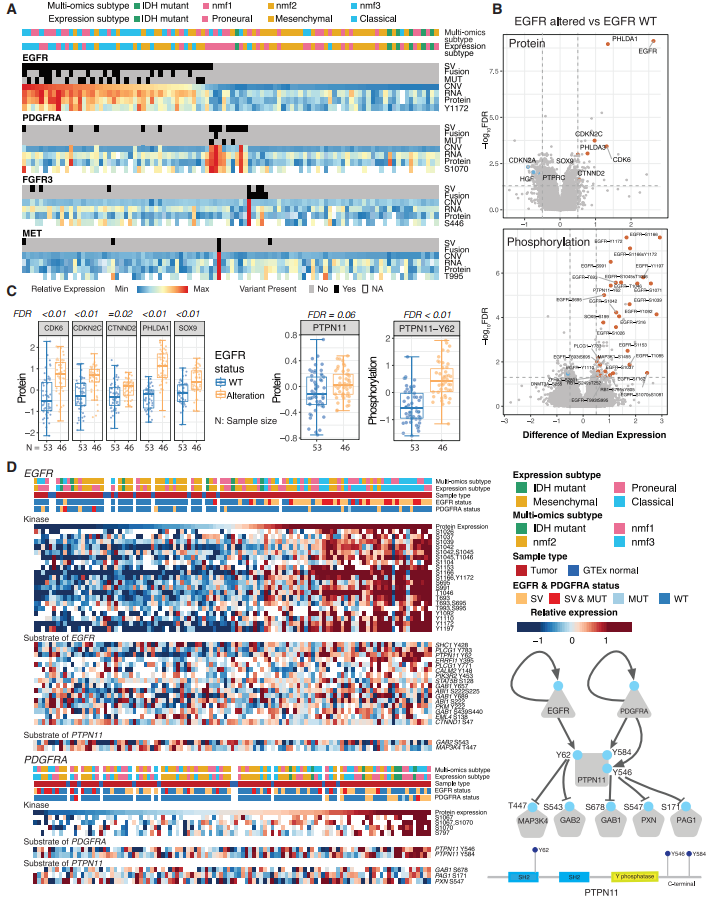
为了探究上游驱动基因变异对基因的表达（RNA水平）和下游蛋白、蛋白磷酸化水平的影响，研究人员将特异驱动基因的变异（突变，CNVs，基因融合，SVs）与RNA、蛋白质表达与磷酸化水平进行了联合分析，并发现95个顺反式的磷酸化事件（图2A, 2B）。如，基因*EGFR* 和*PDGFRA*的突变对其mRNA、蛋白质的表达和在S1166 和 S1067/S1070位点的磷酸化有较强的顺式作用；而在反式中，基因*EGFR*变异则导致CTNNB1 (β-catenin)的mRNA水平降低，但蛋白水平却是上升的，同时也增强了PTPN11 (Shp2) 在Y62处和 PLCG1 在 Y783的磷酸化水平。另外，*TP53*变异在其DNA结合结构域有大量的无义突变（图2C），同时也观察到一些蛋白或蛋白磷酸化水平（ATR, MAPK3, CDK2, 和 CDK9）上升 。在*TP53*突变个体中，一些DNA修复基因却只有磷酸化水平上升，蛋白表达并未增加，这可能显示出特殊的磷酸化调节（图2D）。在*RB1*变异患者中则表现出明显的RB1 下调和 MCM2、MCM4、 MCM6 蛋白的上调（图2E）。



图二 *TP53*基因变异对DNA修复基因和*RB1*基因变异对细胞周期基因的顺反式效应

1. RTK级联信号在GBM中被激活

前人研究表明，在GBMs患者中与RTK相关的基因座（如*EGFR*，*PDGFRA*和*MET*）经常会扩增。在研究队列中，研究人员鉴定到45例存在*EGFR* SVs，所有的都有基因拷贝数扩增，这表明SV和CNV具有高度的一致性（图3A）。在*EGFR*变异患者中，存在高度的EGFR自磷酸化，同时伴随着普列克汀同源样域家族A成员蛋白（PHLDA1和PHLDA3），转录因子SOX9，细胞粘附蛋白CTNND2（δ-catenin）以及细胞周期蛋白CDK6和CDKN2C的表达和磷酸化水平的增加（图3B和3C）。研究人员还对EGFR和PDGFRA进行了激酶底物研究，确定了GAB1在Y689 和 Y657位点存在高水平的磷酸化，并伴随着EGFR的高表达。另外， PTPN11的Y546和Y584磷酸位点与PDGFRA表达强相关（图3D），另在肺癌中还伴随着*ALK*基因融合。



图三 RTKs变异对基因的表达、磷酸化状态和下游靶标的影响

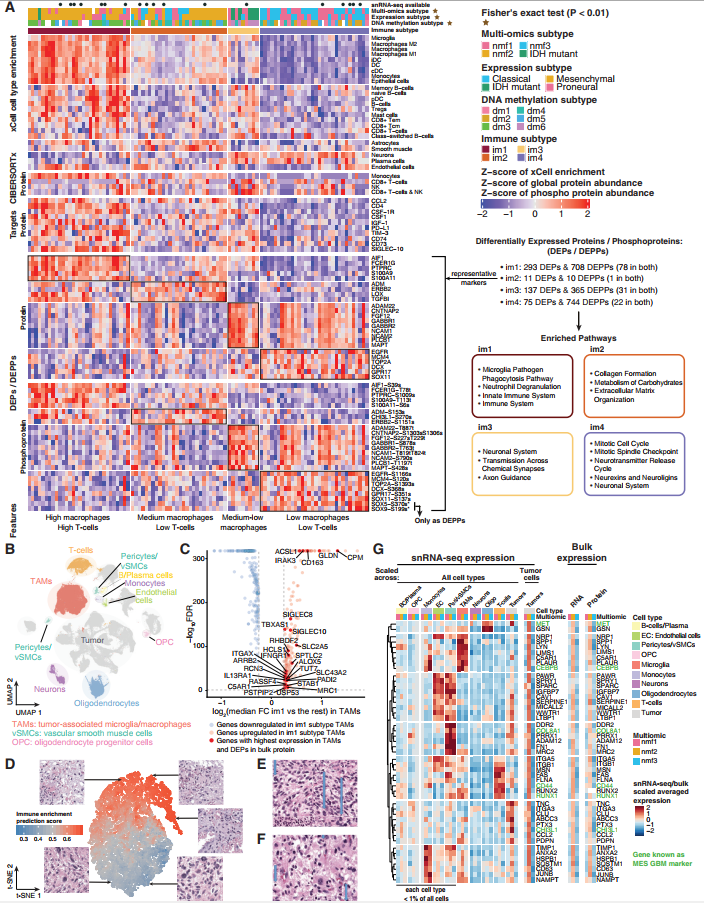
1. GBM免疫亚型具有明显不同的免疫标志表达和表观遗传事件

研究人员首先根据细胞类型的免疫富集得分（用xCell进行单样本GSEA分析得到）将GBMs分成4个亚型（im1，im2，im3，im4），这四个亚型具有独特的通路富集（图4A）。

接着对18个GBM样品（7 im1、5 im2、1 im3、5 im4样品；4A图）进行了snRNA-seq，发现TAMs在GBM TME中构成主要的非肿瘤细胞群（图4B）。与其他免疫亚型（分别为平均7％和19％）相比，im4亚型显示出的较低的T细胞和TAM细胞的浸润比例（分别平均为1.3％和6％）。im1肿瘤在总体水平上具有较高的小胶质细胞和巨噬细胞表达得分，但在snRNA-seq中的im1-3中未观察到此结果，这可能是由TAMs自身的细胞数比例高和基因高表达引起的。而普通转录组和蛋白组数据结果正好也支持这一假设，在im1亚型中的TAMs细胞的RNA和蛋白质表达水平均出现上调（图4C）。

另外，研究人员应用深度学习模型（Deep Learning Model）对不同免疫亚型肿瘤切片的H＆E染色图片进行形态学差异分析发现，与其他亚型（im1-3）相比，im4的维度（t-SNE）降低了（图4D-4F），这显示im4中有大量的大细胞，其中一些是巨型细胞，这在im1-3中很少见。

历来，GBM间充质信号起源一直存在争议，研究人员用snRNA-seq数据对肿瘤和免疫细胞中的间充质特征进行了解释。在nmf2亚型的肿瘤细胞中发现了与EMT相关的基因上调（图4G）。其中，间充质标记*CHI3L1*和*MET*以及其他EMT相关基因（如*TNC*，*ITGA3*和*PDPN*）在肿瘤细胞中的表达量最高。一些EMT相关基因在非肿瘤细胞类型（如TAM、T细胞，周细胞，内皮细胞和少突胶质细胞）中也有表达，但这些标记基因几乎没有亚型特异性的差异。另发现，nmf2亚型的肿瘤组织细胞中EMT相关的上调基因在普通转录组和蛋白质组的表达水平也显著上升（图4G）。除免疫浸润外，蛋白质组学结果也证实了nmf2亚型具有固有的间充质肿瘤细胞特异性。

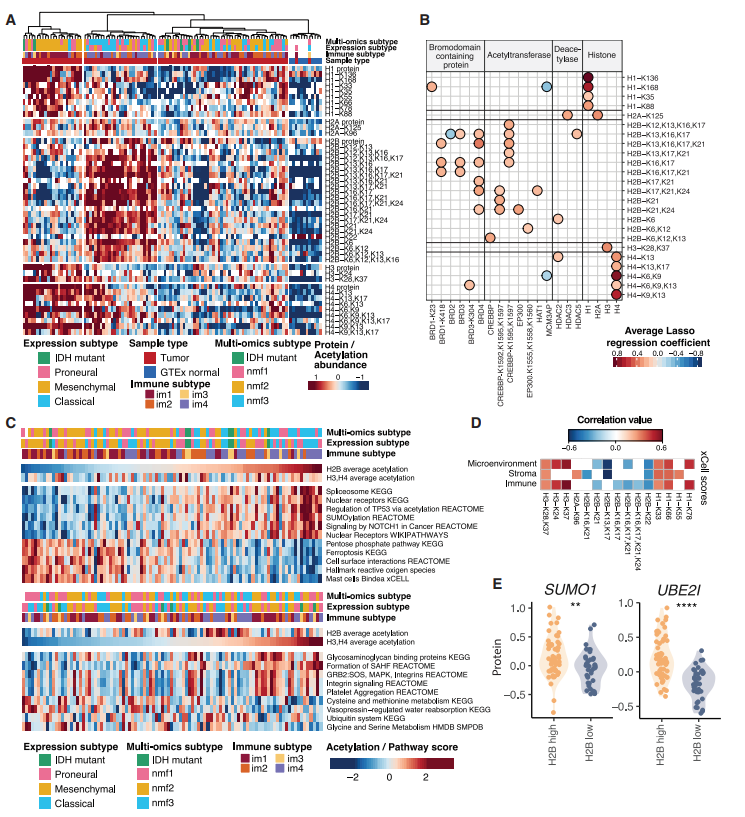


图四 四种免疫亚型的细胞类型富集、免疫标记表达和富集通路

1. 组蛋白的乙酰化差异与特殊亚型和途径的关系

组蛋白乙酰化组结果显示，肿瘤样品中组蛋白（H1、H2A、H2B、H3.3和 H4）有超过30个乙酰化位点。对这些位点进行非监督聚类，发现在肿瘤亚型中具有不同的乙酰化位点的组蛋白（H1、H2B、H3.3和H4）（图5A）。

为了了解HAT、BRD与H2B乙酰化位点之间的潜在联系，研究人员对组蛋白乙酰基转移酶（HATs），含溴结构域的蛋白（BRD）和脱乙酰基酶（HDAC）的乙酰化位点与蛋白表达、乙酰化丰度之间进行了Lasso线性回归。如CREBBP和EP300，它们的蛋白表达和乙酰化水平都显示出了与H2BK12，K13，K16，K17和K21和H2B-K21和K24位点之间的显著关联（图5B）。观察到H2B乙酰化与TME富集得分之间显著负相关（图5D），而其他组蛋白则显示正相关。进一步分析表明，与免疫浸润有关的一些途径，例如肥大症，肥大细胞和活性氧物质途径与H2B乙酰化呈负相关，而剪接体、核受体和SUMOylation途径呈正相关（图5C）。在具有高H2B乙酰化水平的样品中，SUMOylation途径中的两个关键蛋白SUMO1和UBE2I呈上调（图5E）。 有趣的是，UBE2I与细胞周期调节因子CDK6（Pearson r = 0.413）适度相关，后者是在SUMOylation稳定后的UBE2I目标。这些观察结果表明，H2B乙酰化有助于区分免疫细胞和其他细胞类型。



图五 组蛋白乙酰化、免疫亚型和通路的关系

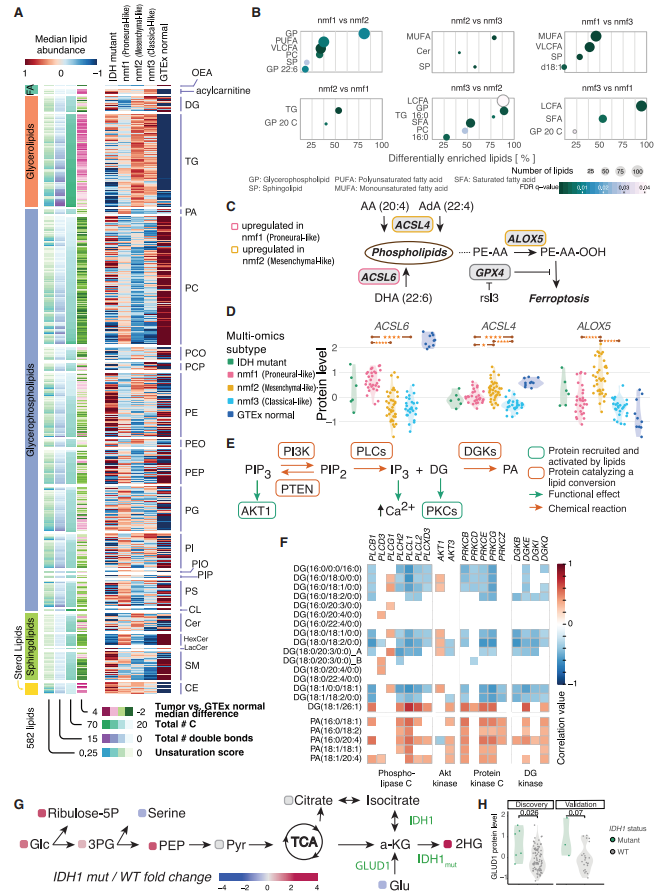
1. 与GBM亚型相关的脂质组学和代谢组学特征

研究人员从75例肿瘤组织和7例GTEx正常组织共定量了582个脂质，而在4个亚型（多组学分型）中鉴定到超过500个差异脂质（Wilcoxon; FDR < 0.05, 图1C）。结果显示，间充质样亚型（nmf2）显示出三酰甘油（TGs）的丰度升高，以及磷脂酰胆碱（PCs）和其他类型的磷脂的水平降低（图6A和6B）。前神经元样亚型（nmf1）富含长链脂肪酸（PU）和长链（LC）的超长链脂肪酸脂质（VLCFA）和甘油磷脂（图6B）。 至于代谢产物，前神经样亚型（nmf1）的肌酸酐和高半胱氨酸水平显著升高，L-半胱氨酸和帕拉金糖醇水平则降低（Wilcoxon; FDR<0.05）。

另外，研究人员探究了含22:4和22:6碳链的脂质含量与其相关的神经保护蛋白的关系，如22:6（如docosahexaenoic acid, DHA）和ACSL6（乙酰辅酶A合成酶）的共调节（图6C）。肿瘤样品的ACSL6蛋白表达大幅降低（图6D），含DHA的磷脂酰甘油（PGs）和TG含量增加，而其他磷脂酰乙醇胺（PE），PC和磷脂酰丝氨酸（PS）的含DHA脂质呈现下调趋势。与间充质样亚型（nmf2）相比，前神经样亚型（nmf1）表现出ACSL6和含DHA的磷脂表达则升高。 就DHA代谢而言，正常组织与前神经元样亚型（nmf1）最相似，而与间充质样亚型（nmf2）最不相似。

除此之外，研究人员还检测二酰基甘油（DG）和与DG合成相关酶的水平（图6E），显示了DG与AKT1，PLCD3和PLCG1蛋白表达之间的显著相关性（图6F）。有趣的是，PLCG1的磷酸化也受到*EGFR*变异的影响（图2B）。

接着研究人员比较了*IDH*-突变型和*IDH*-WT型肿瘤样品的代谢产物丰度。2-HG是*IDH*突变肿瘤中代谢最丰富的代谢产物（中位log2FC = 3.62，FDR <0.05），还有其他差异代谢产物（p <0.05，但没有超过FDR截止值）（图6G）。*IDH*突变型中参与糖酵解的几种代谢物丰度增加，而丝氨酸和谷氨酸水平降低。 谷氨酸可能通过GLUD1和IDH1催化的反应促进α-酮戊二酸和2-HG的生成。 支持这一假设的是在*IDH*-突变型中GLUD1蛋白表达与谷氨酸丰度之间的负相关（Pearson r = 0.29，p <0.01）和GLUD1的显著上调（Wilcoxon p值<0.05）。验证队列也证实了*IDH*突变型肿瘤中GLUD1表达升高（Wilcoxon p值= 0.07，图6H）。



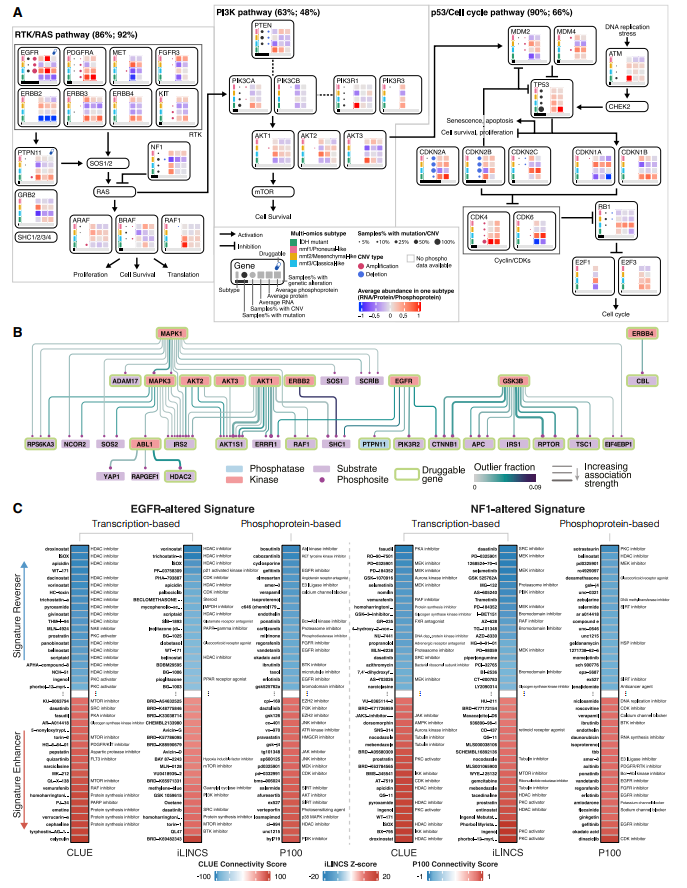
图六 GBMs亚型的脂质组学和代谢组学特征

1. GBM中的关键致癌途径和治疗靶点

研究人员整合了GBM中的遗传变异以及RNA，蛋白质和磷酸位点的数据，获得了三个重要致癌信号通路：RTK / RAS，PI3K / AKT和p53 /cell cycle（图7A）。 而在RTK途径中，表达异常值百分比远高于的基因变异率（图3）。 此外，肿瘤样本在这三种途径中的至少具有一种遗传改变或异常表达。这说明，这三个通路在GBM形成中可能起重要作用，为治疗提供方向。

为了进一步寻求潜在药物治疗靶点，研究人员结合DGIdb和DEPO数据库的可药物性信息，进行了激酶底物和离群值分析，并获得了一些可药物对（图7B）。 结果发现，GSK3B磷酸化与其下游蛋白的磷酸化呈正相关，该下游蛋白参与哺乳动物雷帕霉素靶标（mTOR）信号传导（例如RPTOR和TSC1）和Wnt信号传导（例如CTNNB1和APC）。mTOR信号的另一个参与者AKT1S1与AKT1，AKT2和AKT3激酶有许多重要联系。 还发现EGFR在N端磷酸化CTNNB1 S33，已知其介导CTNNB1蛋白酶体降解。 磷酸化位点异常分析证实了GSK3B，AKT1，MAPK1，MAPK3和EGFR的相互作用。MAP激酶级联反应与包括ABL1激酶在内的多种蛋白质相关，反过来又和HDAC2脱乙酰酶（~5%异常值）具有很强的相关性。而在这种调控中具有大量磷酸化离群值的多种底物的激酶可能是潜在的治疗靶标。

最后，研究人员使用变异特异性转录组学（CLUE和iLINCS）和磷酸化蛋白质组学（P100）特征（突变型肿瘤与WT型肿瘤）进行药物连通性分析，获得了最可能逆转或增强EGFR特征和NF1特征的二十种化合物（药物）及其已知的作用机理（图7C）。这为潜在治疗靶标提供了药物的选择。



图七 GBMs治疗的可选关键通路和潜在药物靶标

研究结论

研究人员通过整合来自10个平台的蛋白基因组学和代谢组学数据，鉴定到了关键的磷酸化事件（如，磷酸化的PTPN11和PLCG1），这些磷酸化蛋白是介导致癌途径激活的潜在开关和*EGFR-*，*TP53-*和*RB1-*变异型肿瘤的潜在靶标。使用普通组学方法发现了GBMs具有不同免疫细胞类型的免疫亚型，并通过snRNA-seq进行验证，而这些亚型又与特异性表达和组蛋白乙酰化模式相关。 经典类和免疫性低的GBMs中的组蛋白H2B乙酰化主要受BRD，CREBBP和EP300的驱动。全面的代谢组学和蛋白质组学数据证实了*IDH*突变型肿瘤中各亚型的特定脂质分布和显著各异的整体代谢变化。这项研究所挖掘聚焦的生物学关系可能有助于GBMs患者分型，并能有效的指导个性化精准治疗。

小鹿推荐

单一组学数据往往只能反映某生物学过程或现象的某方面信息，这就好比盲人摸象一般。这一点在癌症研究中，体现的尤为突出。本文中，研究人员为了更全面解析胶质母细胞瘤（GBMs）致病因、寻求更精准的个性化治疗，动用了多达10个平台的组学技术（蛋白基因组学、代谢组学），并结合公共数据库，对几乎覆盖全组学数据进行了整合分析，不仅对GBMs进行了更为细致全面的亚型分类，同时挖掘出更为可能的致病通路和关键靶点，并更深层次的利用药物连接性分析，获得了可用药物对。这一套流程下来，可谓是一气呵成，也为我们研究肿瘤或其他复杂成因的生物学问题提供了研究策略和方案。