RESEARCH ARTICLE · Open Access

Sequencing and curation strategies for identifying candidate glioblastoma treatments

Mayu O. Frank[1,2†], Takahiko Koyama[3†], Kahn Rhrissorrakrai[3†], Nicolas Robine[1], Filippo Utro[3], Anne-Katrin Emde[1], Bo-Juen Chen[1,12], Kanika Arora[1], Minita Shah[1], Heather Geiger[1], Vanessa Felice[1], Esra Dikoglu[1,13], Sadia Rahman[1], Alice Fang[1], Vladimir Vacic[1,14], Ewa A. Bergmann[1,15], Julia L. Moore Vogel[1,2,16], Catherine Reeves[1], Depinder Khaira[1], Anthony Calabro[1,17], Duyang Kim[1], Michelle F. Lamendola-Essel[1,18], Cecilia Esteves[1,19], Phaedra Agius[1], Christian Stolte[1], John Boockvar[4], Alexis Demopoulos[5], Dimitris G. Placantonakis[6], John G. Golfinos[6], Cameron Brennan[7], Jeffrey Bruce[8], Andrew B. Lassman[8], Peter Canoll[8], Christian Grommes[7], Mariza Daras[7], Eli Diamond[7], Antonio Omuro[7,20], Elena Pentsova[7], Dana E. Orange[2,9], Stephen J. Harvey[10], Jerome B. Posner[7], Vanessa V. Michelini[10], Vaidehi Jobanputra[1,8], Michael C. Zody[1], John Kelly[3], Laxmi Parida[3], Kazimierz O. Wrzeszczynski[1], Ajay K. Royyuru[3] and Robert B. Darnell[1,2,11*]

Abstract

Background: Prompted by the revolution in high-throughput sequencing and its potential impact for treating cancer patients, we initiated a clinical research study to compare the ability of different sequencing assays and analysis methods to analyze glioblastoma tumors and generate real-time potential treatment options for physicians.

Methods: A consortium of seven institutions in New York City enrolled 30 patients with glioblastoma and performed tumor whole genome sequencing (WGS) and RNA sequencing (RNA-seq; collectively WGS/RNA-seq); 20 of these patients were also analyzed with independent targeted panel sequencing. We also compared results of expert manual annotations with those from an automated annotation system, Watson Genomic Analysis (WGA), to assess the reliability and time required to identify potentially relevant pharmacologic interventions.

Results: WGS/RNAseq identified more potentially actionable clinical results than targeted panels in 90% of cases, with an average of 16-fold more unique potentially actionable variants identified per individual; 84 clinically actionable calls were made using WGS/RNA-seq that were not identified by panels. Expert annotation and WGA had good agreement on identifying variants [mean sensitivity = 0.71, SD = 0.18 and positive predictive value (PPV) = 0.80, SD = 0.20] and drug targets when the same variants were called (mean sensitivity = 0.74, SD = 0.34 and PPV = 0.79, SD = 0.23) across patients. Clinicians used the information to modify their treatment plan 10% of the time.

Conclusion: These results present the first comprehensive comparison of technical and machine augmented analysis of targeted panel and WGS/RNA-seq to identify potential cancer treatments.

* Correspondence: darnelr@rockefeller.edu
†Mayu O. Frank, Takahiko Koyama and Kahn Rhrissorrakrai contributed equally to this work.
[1]New York Genome Center, 101 Avenue of the Americas, New York, NY 10013, USA
[2]Laboratory of Molecular Neuro-Oncology, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA
Full list of author information is available at the end of the article

Frank *et al. BMC Medical Genomics*   (2019) 12:56

Page 2 of 16

Background

Oncology is one of the first areas where next-generation sequencing is being applied [1–3]. Sequencing is used to identify genetic variants that could be pharmacologically targeted, allowing identification of drug effects in stratified populations that may otherwise be missed [2, 4]. Panel-based sequencing, using hybridization and capture of specific regions of key genes or of all genes (whole exomes; WES), versus whole genome sequencing (WGS) are different technologies with different costs that have not previously been directly compared. 10,000 cancer patients sequenced with the MSK-IMPACT panel identified potentially clinically actionable calls in 36.7% of individuals sequenced [5]. Deep sequencing coverage increases sensitivity for rare variants in heterogeneous tumors. WGS, however, does not rely on hybridization and capture, a source of potential bias, and is able to identify non-coding variants such as enhancer bindings sites [6] and increases sensitivity for small copy number variants (CNVs) and missense mutations, indels, [7] intronic variants, [8] and gene fusions [9]. The relative impact of of these technologies on making clinically actionable variant calls is unknown.

Panel-based sequencing is used to identify treatment targets in tumors including glioblastoma (GBM) [10–13], the most common adult brain malignancy with a median survival of 14.2 months [14]. The Cancer Genome Atlas (TCGA) analyzed GBM and established four molecular subtypes [15] defined by IDH1 mutation and methylation status, [16] and more recently three subtypes which takes into consideration tumor purity and heterogeneity, [17] but has not yet led to new therapies. Panel sequencing and WGS/RNA-sequencing (WGS/RNA-seq) provide logical paths forward to identify variants. However, the process of sequence analysis and prioritizing variants is laborious, requiring highly trained experts, particularly in WGS/RNA-seq; this prompted prompting assessment of automated analyses.

The New York Glioblastoma Genome Consortium (NY-GGC) was organized in 2013 at the Rockefeller University and New York Genome Center (NYGC) to conduct a feasibility study of using WGS/RNA-seq to identify tumor-specific variants and potential drug targets, to compare WGS/RNA-seq to panels, and to assess the reliability of automated versus manual analyses. Here, we describe an integrated analysis of 30 GBM patients recruited through seven participating institutions.

Methods

Study design

The NY-GGC was formed in 2013 through a collaboration initiated at Rockefeller University, and included Memorial Sloan Kettering Cancer Center, New York University Medical School, Northwell Health (Lenox Hill Hospital and North Shore University Hospital), Columbia University Medical Center and New York Genome Center, sharing tumor and blood samples for sequencing, relevant clinical histories, and raw sequence data (BAM files). Variant call files (VCFs) were analyzed by NY-GGC and by WGA.

Patients

Entry criteria for this study were: minimum age of three, histologically confirmed GBM at referring institution with no requirement for central pathology review, Karnofsky score of at least 60, life expectancy of at least 6 months, and potential interest in further treatment. Clinical data was collected including results of any other clinical sequencing.

All participants provided written informed consent. Protocols were approved by local or central Institutional Review Boards at: Rockefeller University, Biomedical Research Alliance of New York (on behalf of Northwell Health), Memorial Sloan Kettering Cancer Center, New York University School of Medicine, and Weill Cornell Medicine.

WGS and RNA-seq

Paired tumor and normal (blood) samples were sequenced from each individual and analyzed by WGS at 80X tumor and 40X normal coverage as previously described [18]. Ploidy values were used to estimate chromosome, gene, and allele copy number. We analyzed TERT promoter variants, intronic splice site variants (annotated by SnpEff), and exonic variants. Single nucleotide variants (SNVs) were classified by Tiers. Tier 1 variants are defined as variants with known clinical significance in GBM as defined by CIViC (v.alpha, 3/2015). Tier 2 variants are known to be clinically significant in another tumor type as defined by CIViC. Tier 3 are variants of unknown significance (VUS) in known actionable cancer genes with associated drugs. Tier 4 are VUS mutations in Cosmic Cancer Census Genes (v.75) [18].

Where RNA was available, RNA-seq was performed as previously described, [18] with the addition that when the RNA integrity number (RIN) score [19] was less than 7, we used the KAPA Stranded RNA-seq with RiboErase (P/N: KK8483) with Agilent SureSelectXT v6 + Cosmic (P/N: 5190–9308). We used RNA-seq data to annotate variants discovered in WGS according to the level of expression of each variant, estimated by identifying the number of supporting reads and allelic fraction. Gene expression in the tumor samples was assessed and compared with 169 GBM samples from TCGA and a modified z-score was calculated as previously described [18]. Modified z-scores of RNA-seq normalized expression data per gene was used as proxy for differential gene expression. Modified z-score per gene was calculated by subtracting the median transcripts per million (TPM) value (over the TCGA GBM cohort) from each sample's TPM and dividing by the TCGA median absolute

deviation. The z-score therefore represents the number of standard deviations each sample is from the median expression value of a specific gene of the TCGA GBM cancer cohort. For tumor-normal comparisons of splice site variants, percent spliced in (PSI) was calculated as the number of reads supporting the unannotated alternative splicing event divided by the number of canonical reads supporting the annotated event [20] and fusion transcript discovery was performed using RNA-seq data and FusionCatcher [21] as previously described [9].

When associating variants with potential therapies, we prioritized variants of high copy number focal gains and two copy homozygous loss over lower copy number whole arm gains or heterozygous losses. However, at times, lower copy number changes (< 5 copy gains or heterozygous losses) were also reported as others have done [22].

Comparison with Watson Genomic Analytics (WGA) and targeted tumor panels

WGA is an IBM Research proof-of-concept environment of Watson for Genomics described previously, [18, 23, 24] with the addition of algorithmic updates that include basic processing of structural variants (SVs) (version 11/2016). WGA is a cloud-based cognitive system capable of analyzing mutation, gene expression, copy number alteration (CNA) and SV data provided as VCFs by leveraging 20+ structured and unstructured data sources. WGA first performs a Molecular Profile Analysis (MPA) to identify possible driver mutations and drug response biomarkers in a disease-specific manner. MPA evaluates mutations using data from structured databases such as COSMIC and ExAC to search for known variants and remove possible benign germline mutations. It also uses evidence extracted from literature using both machine-based and expert-based manual curation. These data sources are used to create a system capable of categorizing alterations as pathogenic, benign, or VUS in a disease specific manner. WGA performs a similar analysis for CNAs and gene expression changes and takes into the account the functional annotation of a gene when assessing the relevance of a CNA or differential expression. WGA also processes SVs using DNA breakpoints data output by the structural event detection program Delly [25].

From the MPA results, pathogenic and likely pathogenic alterations are assessed for potential direct and indirect (i.e. via pathway mechanisms) therapeutic options. WGA's pathway and drug analysis identifies which therapies are most applicable and categorizes therapies by different levels of evidence from strongest, which includes FDA-recognized marker or mutations predictive of response, to the weakest, which can represent a normally appropriate therapy with a clinically supported resistance marker present in the patient. For each potential treatment option that WGA identifies, it provides all supporting evidence for the use of that therapy, mechanism of action and information on eligible clinical trials if available. WGA currently limits available treatments to molecularly targeted therapies.

To access unstructured data, particularly from literature, clinical trial information and drug label data, WGA applies Natural Language Processing (NLP). In brief, NLP requires a training phase on a high-quality corpus of papers and text that have been manually annotated by subject matter experts for terms and relationships relevant for WGA to understand, including genes, proteins, mutations, drugs, and effect. After this training phase, the NLP model is applied to the full unstructured data set, and after validation by automated methods and subject matter experts, the information is integrated into WGA.

NY-GGC provided VCFs, including CNV and gene expression files as input to WGA. Over the course of this study, IBM and NY-GGC periodically reviewed results and discussed differences between NY-GGC and WGA. Insights and feedback from these sessions along with similar sessions for other studies [24] were used to improve the process of automated curation, for example by identifying additional data sources for consideration and adjusting parameter weights. As new GBM samples became available, the entire cohort to that date was reanalyzed by the then current version of WGA.

Using NY-GGC calls as a truth set, we compared mean sensitivity and positive predictive value (PPV) across patients to determine similarity between manual and automated systems for reported variants and drug targets based on VCFs.

Comparison with targeted panels

Prior targeted NGS panel testing results were available for 20 patients and were compared with WGS results. The percentage of variants that were in common, that were uniquely called by WGS, or that were uniquely called by the targeted panel were calculated and shown in a heatmap. Targeted NGS panel testing was done by Memorial Sloan Kettering Cancer Center IMPACT (Panel 1), FoundationOne (Panel 2), New York University Next Generation Sequencing Tumor 50 Panel (Panel 3), Weill Cornell Medical Center Precision Medicine's whole exome sequencing assay (Panel 4), University of San Francisco's 500 Cancer Gene Panel (Panel 5) and Caris Molecular Intel (Panel 6).

Therapeutic targets and drug recommendations

After annotating the tumor-specific gene variants, relative to normal germline DNA, based on SNV, CNV, SV, and RNA-seq data, [18] variants were associated with drugs in the NYGC database.

NYGC database was assembled by manual curation of publically available data from the National Comprehensive Cancer Network, (https://www.nccn.org/), US Food and

Frank *et al. BMC Medical Genomics* (2019) 12:56

Page 4 of 16

Drug Administration (https://www.fda.gov/Drugs/InformationOnDrugs/ApprovedDrugs), CIViC - Clinical Interpretations of Variants in Cancer (civic.genome.wustl.edu), Precision Cancer Therapy-MD Anderson (https://pct.mdanderson.org/), OncoKB (oncokb.org), canSar (https://cansar.icr.ac.uk), Pharmacogenomics Knowledgebase - PharmGKB (www.pharmgkb.org), Clinical Trials.gov (clinicaltrials.gov) and from directed literature searches. The current NYGC drug to gene database contains 260 genes associated with a least one drug.

Prioritization and rationale of drug recommendations was based on further manual assessment by the NY-GGC including but not limited to: variant Tier, quality of data supporting variant call, interpretation of the consequences of VUS in light of literature research (structural and functional analysis of protein interactions, prior knowledge of analogous variants), including analysis of X-Ray crystallographic structures, drug FDA approval status, drug identification in a current GBM trial, and record of drug success in the treatment of GBM and/or other cancer types specific to the variant.

Each individual's results from NY-GGC were first discussed in an internal tumor board comprised of those involved in development of the analytic pipeline, bioinformaticians, project managers, pathway analysts, and clinical experts, and subsequently at a NY-GGC tumor board with that same team together with referring physicians and collaborating physicians and scientists. Further clinical use of the results provided were at the sole discretion of the referring physician.

Statistical analysis

Mean tumor purity and ploidy were calculated with standard deviations (SD). Median number of SNVs are reported with interquartile range (IQR). Correlation between RNA-seq and exonic SNV variant allele frequency (VAF) was assessed using Pearson's correlation coefficient. Mean number of alternative splice variants was calculated with SD. Mean sensitivity and PPV were calculated to compare agreement between the number of calls made between WGS and WGA.

Results

Patients and study process

Between March 2015 and July 2016, 36 patients were screened and 30 were enrolled (Table 1). Four were excluded due to final pathology indicating diagnosis other than GBM and 2 others died before sequencing began. Three participants had two separate tumor samples sequenced; one from tumors resected from two distinct brain areas, another from an enhancing and non-enhancing region on MRI, and a third from samples

Table 1 Patient characteristics

	$n = 30$
Age, median (range)	63 (25–81)
Female, no. (%)	12 (40%)
Resections, no. (%)	
Initial	19 (63%)
Biopsy only	2 (7%)
Re-resection	7 (23%)
Re-resection of subtotal	2 (7%)
Prior treatment of 7 recurrent tumors, no.	
RT/TMZ	6
Bevacizumab	3
Cetuximab	2
RT/Nivolimab	1
RT/Rindopepimut	1
CCNU	1
Gamma knife	1
Prior cranial RT, unrelated to GBM	1
Days from initial resection to sample submission, median (IQR)	67 (266)
Sample preservation, no. (%)	$n = 33$
Fresh frozen	13 (40%)
OCT embedded	12 (36%)
Formalin-fixed paraffin-embedded	8 (24%)
Tumor biomarkers, no./no. assessed	
EGFR amplification	8/20
MGMT methylation	12/30
IDH1 R132H mutation	2/30

All tumors were histologically confirmed as WHO grade IV gliomas. However, one had focal sarcomatous features, another was initially reviewed as pleomorphic xanthoastrocytoma with anaplastic features then as an epitheliod GBM upon re-review, and a third had a PNET-like component. SD = standard deviation, IQR = interquartile range, RT = radiation therapy, TMZ = temozolomide, CCNU = Lomustine

representing different histological characteristics. Throughout the study, we refined our processes for sample collection, squencing, analysis, interpretation, and dissemination of information to referring physicians. The average time from sample receipt to the tumor board meetings was 4.5 months (SD = 2.1) and the average time from post bioinformatic pipeline analysis to the tumor board meetings was 1.9 months (SD = 1.1).

Tumor purity and ploidy

Tumor purity ranged from 15 to 95% (mean = 71%, SD = 16%) for all samples. Two samples had tumor purity < 20% (15 and 19%). The average estimated ploidy was 2.03 (range 1.59 to 4.06, SD = 1.235). Four samples were hyperploid and three were hypoploid.

WGS analysis

Samples from only three patients contained a Tier 1 SNV, two with the *IDH1* R132H variant and one with *BRAF* V600E, similar to the incidence of these variants in 591 GBM patient samples in MSKCC's cBioPortal (2 and 1% respectively). Samples from three patients contained one Tier 2 variant: *TP53* R175H, *TP53* R273C, and *PIK3CA* E542K. In contrast, there were 163 Tier 3 variants in 90 genes across 29 samples. There were 143 variants in Tier 4 genes, but none were deemed targetable. Only one sample had no SNVs in targetable genes. In total, 44 genes and 32 samples contained targetable SNVs, the vast majority of which were Tier 3 variants.

Among clinically actionable SNVs, *PIK3R1* and *RB1* were mutually exclusive, consistent with previous observations [26, 27]. Ten SNVs were identified in *PIK3CA* and *PIK3R1*. Of the three variants identified in *PIK3CA*, one was assigned to Tier 2 (E542, within the catalytic subunit), and the other two were assigned to Tier 3. Seven Tier 3 variants were identified in *PIK3R1*, all of which were in the iSH2 regulatory domain.

All but one sample had CNVs that were considered targetable. We observed 24 arm-scale and nine focal losses of chromosome 10 that included *PTEN*, consistent with previous observations in GBM [28–30]. Of those 24, we observed 10 cases containing a secondary SNV. EGFR focal amplifications were observed containing the *EGFRvIII*, *EGFRvV* and A289V variants. Six patients were previously treated with temozolomide; two had copy number losses of CDKN2A, and one also had a copy number loss of RB1 [26].

TERT promoter mutations, which can have prognostic consequences depending on factors such as MGMT promoter methylation status, were found in 19 of 33 samples, consistent with prior studies. [31–34] There were 35 to 5881 (median = 157, IQR = 111) exonic SNVs per sample, of which two were exceptionally high (1954 and 5881), one previously treated with temozolomide and one not. Missense mutations accounted for 43% of calls that resulted in drug associations, and CNV and SV data were used in 23% of therapeutic associations (Fig. 1a).

RNA analysis

Of 30 samples, sufficient RNA quantity for library preparation was extracted from 27 and high quality sequences with sufficient coverage were obtained from 26. There was good correlation in the VAF between RNA-seq and exonic SNVs identified by WGS (Pearson's correlation coefficient r = 0.622, *p*-value = 1.899e-13, Fig. 2). RNA-seq identified 113 of 155 (73%) variants identified by WGS. We observed more similarity in allelic frequencies in DNA and RNA data for genes with higher read count in a specific sample (i.e. overamplified DNA and higher RNA transcript count).
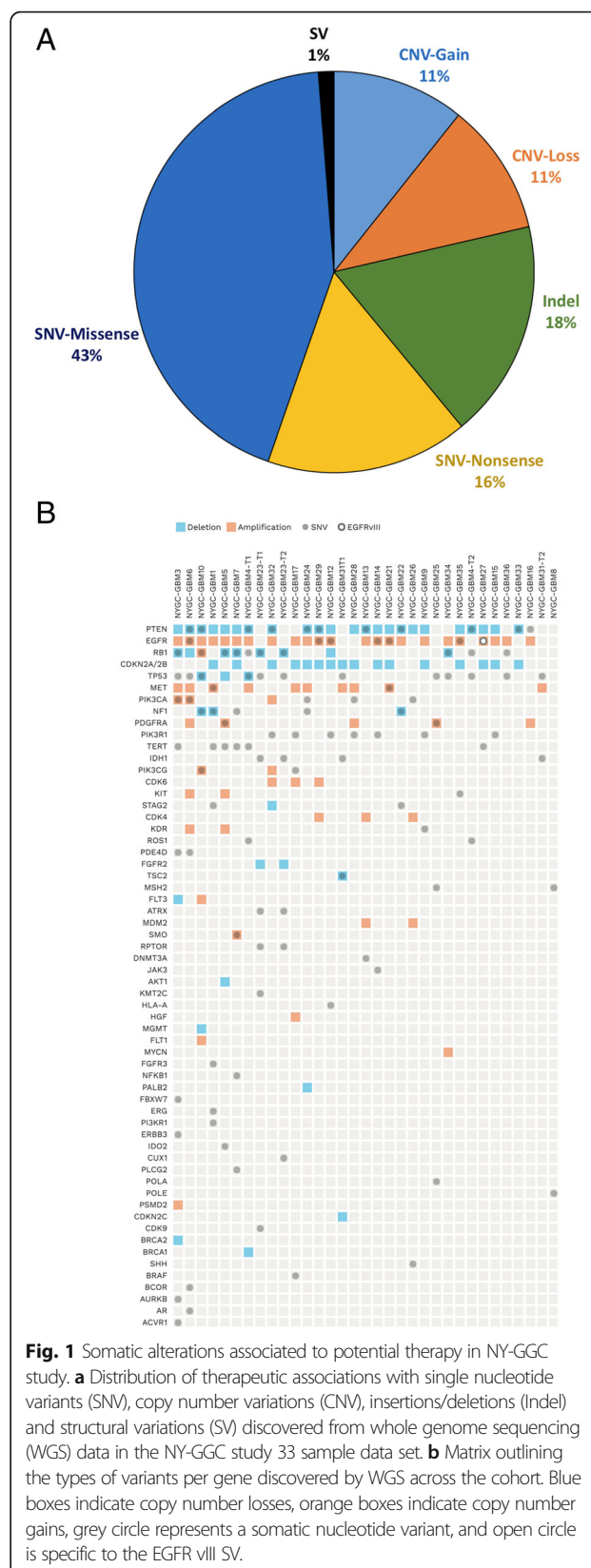

Fig. 1 Somatic alterations associated to potential therapy in NY-GGC study. **a** Distribution of therapeutic associations with single nucleotide variants (SNV), copy number variations (CNV), insertions/deletions (Indel) and structural variations (SV) discovered from whole genome sequencing (WGS) data in the NY-GGC study 33 sample data set. **b** Matrix outlining the types of variants per gene discovered by WGS across the cohort. Blue boxes indicate copy number losses, orange boxes indicate copy number gains, grey circle represents a somatic nucleotide variant, and open circle is specific to the EGFR vIII SV.
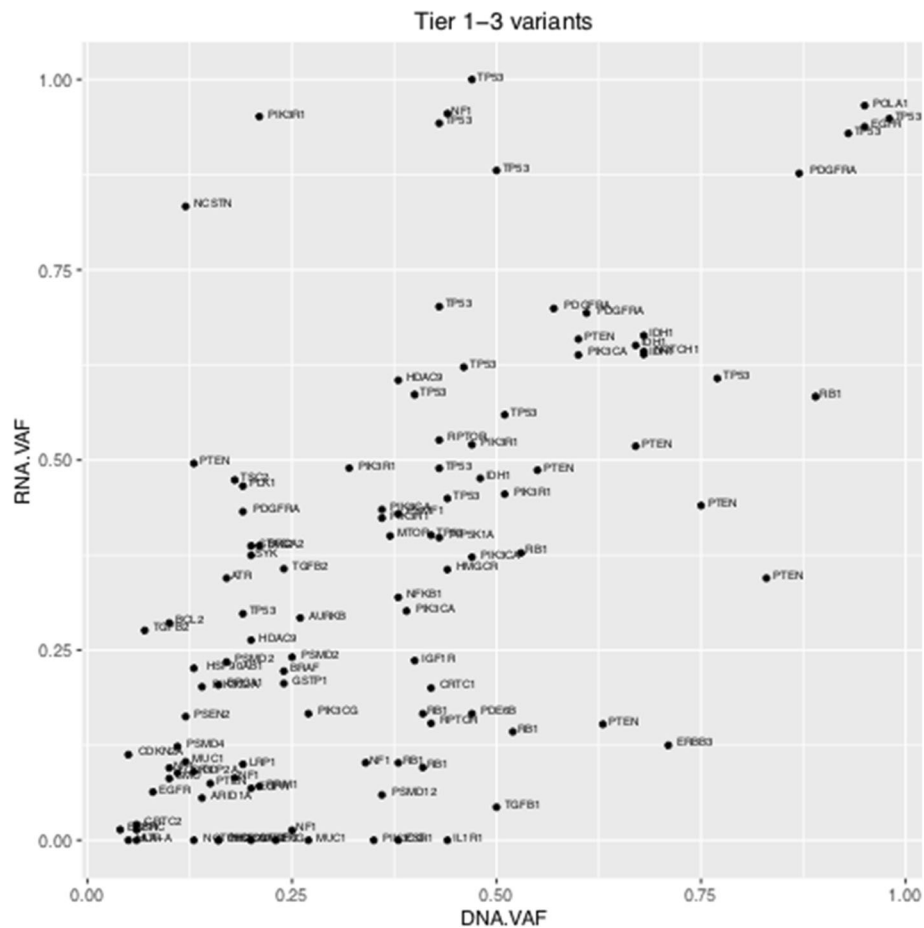
Fig. 2 DNA variant allele frequency correlation with RNA variant allele frequency. Correlation of variant allele frequency (VAF) in WGS vs RNA-seq plotted for all Tier 1–3 variants with sufficient (≥5 total reads) coverage in RNA

RNA-seq revealed an average of seven alternative splice variants (PSI > 10%) in Cancer Census genes per individual (SD = 5.02, range 1 to 26). One variant (c.2419-274_2443del) in NYGC-GBM1 disrupted a splicing acceptor site and generated an exon 11 skipping event in the MET transcript. [18] Additionally, both *EGFR*vIII and *EGFR*vV exon deletions were confirmed in RNA-seq data (Fig. 3). We also discovered a *CHST11-PKP2* fusion (NYGC-GBM13, Fig. 4). One of the breakpoints disrupted *CHST11* and the other was in the intergenic region just upstream of *PKP2*. The *CHST11* promoter may drive a high expression of this fusion transcript (modified z-score of RNA-seq normalized expression =3.01). *PKP2* is associated with *EGFR* regulation and this fusion may result in activation of *EGFR* signaling pathway [35].

Identification and utilization of potentially clinically actionable information

We identified one or more potential treatment targets in all 30 tumor samples. WGS/RNA-seq identified 61 disrupted genes, including both SNVs and CNVs, which were

associated with 87 targeted treatment options including 62 clinical trials for 39 therapies (Table 2). The most commonly identified therapeutic targets were *PTEN, EGFR, CDKN2A,* and *MET* (Fig. 1b). Two samples with high mutation burden were found to have Tier 4 variants in *MSH2*, a mismatch repair (MMR) gene for which checkpoint inhibitors were identified as having potential utility to increase immune response to tumors with high neoantigen load [36].

We identified potentially synergistic combination therapy options such as those targeting multiple mutations occurring in the same pathway or targeting multiple arms of the same pathway in seven patients. Three variant types, *EGFR* amplification, *PTEN* loss, and PIK3R1 SNVs were most frequently associated with potential combination targets. In NYGC-GBM24, *EGFR* gain and *PTEN* loss were identified, which could independently activate the PI3K/AKT pathway. Cetuximab and everolimus were suggested to inhibit activation of EGFR and mTOR downstream of AKT, respectively [37]. In NYGC-GBM1, *PTEN* loss, *MET* gain and *PI3KR1* variants
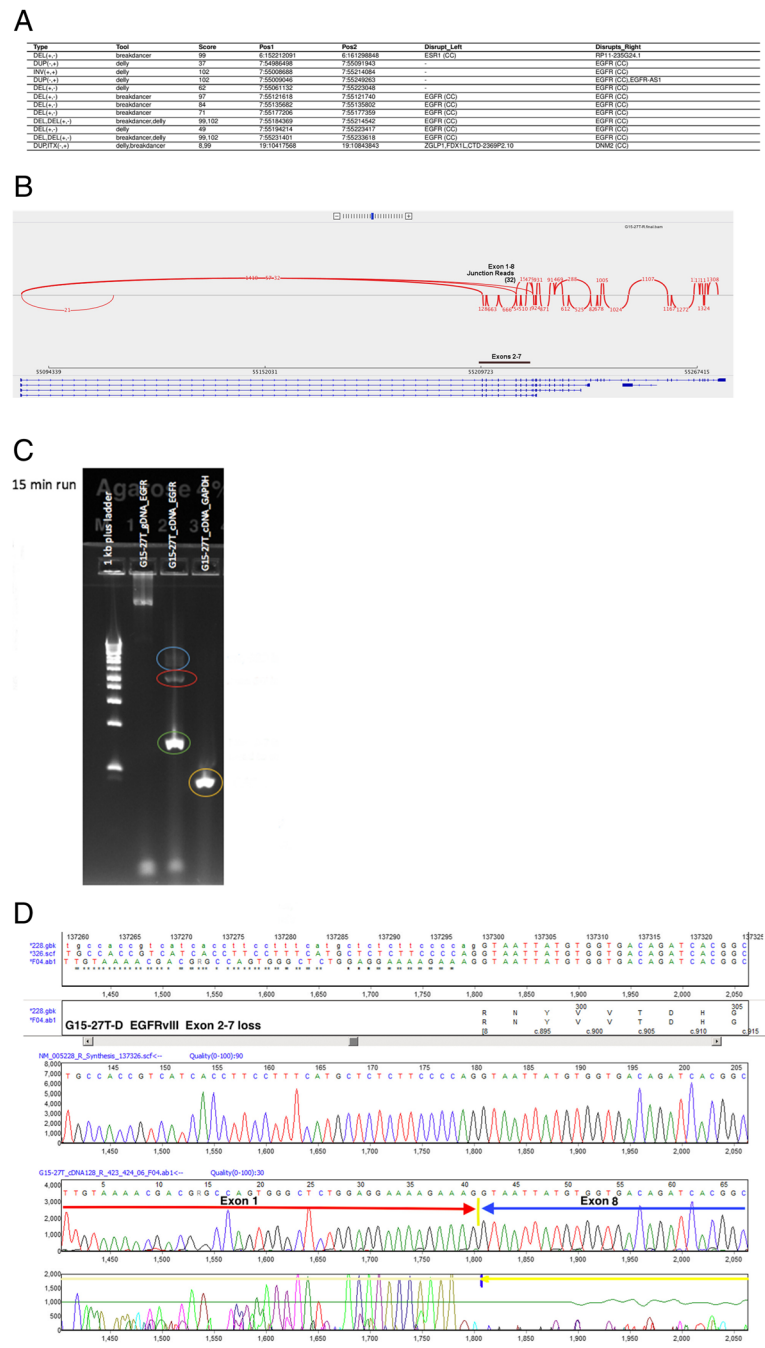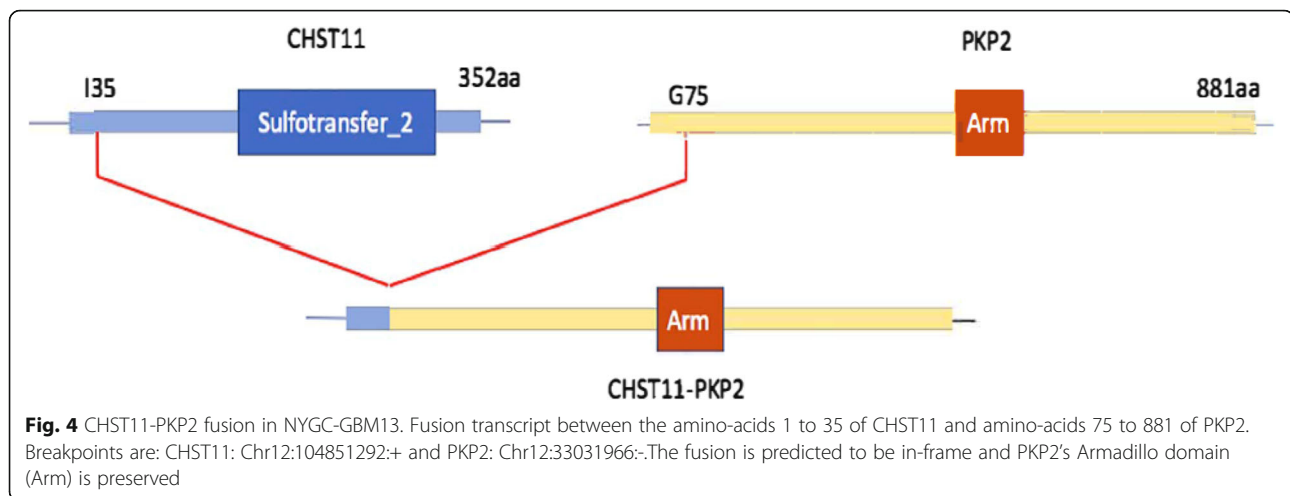
Fig. 3 EGFR vIII detection in NYGC-GBM 27. **a** Table of EGFR SVs called by two SV callers (Delly, Breakdancer) in WGS data. **b** Sashimi plot from RNA-Seq data indicating 32 Exon1–8 junctions reads (exon 2–7 del). **c** RT-PCR confirmation of EGFRvIII. Blue circle indicates wild-type EGFR (exons 1–8) 929 bp fragment, green circle indicates EGFRvIII exons 2–7 loss 128 bp, red circle indicates EGFR potential exon 6–7 loss, yellow circle indicates GAPDH 87 bp control. **d** Sanger sequencing of gel band confirming presence of EGFR vIII with Glycine codon insertion as previously reported, (Chr7: 55,087,058-55,223–523; c.335_1135delinsggt, p.V30_R297delinsG)

were present, and associated with a PI3K inhibitor (BKM120) and a MET inhibitor (INC280). [18, 38, 39] NYGC-GBM9 had a PIK3R1 variant and *PTEN* loss and a PI3K inhibitor (BKM120) and everolimus were suggested [40]. In NYGC-GBM17 with *BRAF* V600E, *PIK3R1* loss, and *MET* and *EGFR* gain, the combination of vemurafenib and INC280 were identified, aiming to avoid resistance to inhibitors targeting *BRAF* V600E [41–44].

Fig. 4 CHST11-PKP2 fusion in NYGC-GBM13. Fusion transcript between the amino-acids 1 to 35 of CHST11 and amino-acids 75 to 881 of PKP2. Breakpoints are: CHST11: Chr12:104851292:+ and PKP2: Chr12:33031966:-.The fusion is predicted to be in-frame and PKP2's Armadillo domain (Arm) is preserved

We considered levels of expression in pathways downstream of potentially targetable variants. NYGC-GBM7 contained an *EGFR* gain and a nonsense SNV in the gene Smoothened (*SMO*), which is directly targetable by vismodegib. The *SMO* variant (p.Arg421*/c.1261C > T) is of unknown function. It was confirmed by RNA-seq and Sanger sequencing, however no significant alteration in *SMO* transcript level (z-score = − 0.93) was identified. SMO activation leads to release of GLI1 from SUFU-mediated cytoplasmic sequestration and to nuclear translocation and transcription activation [45]. Nonsense mutations are often not directly associated with RNA decay, however RNA-seq analysis showed significant upregulation of downstream transcripts in the SMO-SUFU-GLI pathway, including those encoding *MDM2* and *IGFBP6,* which are associated with tumor proliferation [46]. Vismodegib, was well tolerated, but not efficacious in recurrent GBM [47] but given that this study did not stratify enrollees based on genomic variants, it was considered potential therapeutic option.

Of 30, there were four patients in whom the WGS/RNA-seq done in this study were used to make treatment decisions. NYGC-GBM17 was a 33 year-old female whose initial pathology was reviewed as a low-grade pleimorphic xanthroastrocytoma (PXA), later re-reviewed as an epitheliod glioblastoma. She was treated with radiation, temozolomide, and multiple procedures for cyst drainage. NY-GGC identified a *BRAF* V600E variant, also found by FoundationOne, which has been reported to occur in both PXA and epitheliod GBM tumors, and this finding supported a recent case report suggesting that epithelioid GBM may arise from a PXA with a *BRAF* V600E mutation [48]. After presentation at the NY-GCC tumor board, the patient was treated with drugs targeting BRAF (dabrafenib and trametinib) for 12 months. At the time of this writing, she was 18 months post initial diagnosis, free of focal neurologic deficits with stable disease on MRI.

NYGC-GBM25 was a 24 year-old man with a partially resected GBM that recurred and was re-resected 1 month later. He was treated with chemoradiation and intra-arterial Bevacizumab on a clinical trial (NCT 01811498). NY-GGC identified a high mutational burden, which was likely due to a truncating mutation in a MMR gene, *MSH2* [49–51] and a missense VUS in POLA1, which combined with MMR defects can lead to higher rates of mutation than MMR defects alone [52]. At NY-GGC tumor board, checkpoint inhibitors were identified as possible targeted treatments, based on evidence that they are effective against MMR-defective tumors and that there may be potential benefit of concurrent therapy [53, 54]. The referring physician stopped adjuvant temozolomide, in light of evidence that MSH6 protein loss could lead to progression on temozolomide in GBM, [55, 56] and started nivolumab. Shortly thereafter, the patient underwent re-resection and this sample was submitted for resequencing in another study. This sample, collected 7 months after the initial sample, had a similarly high mutation burden with many of the same potentially targetable variants. The hypermutation phenotype was again chosen by the physician as the focus of treatment, discontinuing nivolumab and starting pembrolizumab. The patient later suffered multiple central nervous system infections and seizures and died 6 months after the third resection.

NYGC-GBM12 was a 64 year-old male treated with temozolomide, bevacizumab, intra-arterial cetuximab (based on FISH evidence of EGFR amplification) and multiple resections. NY-GGC identified an extreme focal EGFR amplification, and suggested erlotinib despite pre-treatment with cetuximab. He received 3 cycles of erlotinib but did not have a significant clinical response and died 5 months later.

Overall, NY-GGC tumor boards occured at an average of 4.5 months (SD = 2.1 months) after sample submission and follow up indicated that results were utilized in the

Table 2 Key targetable variants and associated drugs and clinical trials

Sample	Key Therapeutic Variants	Associated Drugs	Clinical Trials
GBM1	PIK3R1 R562del, MET R755fs MET 11, focal gain	BKM120, INC280	NCT01870726 NCT02386826
GBM3	PIK3CA V344G	BKM120	NCT01870726 NCT01349660
GBM4	EGFR gain, PTEN W111C, whole arm loss	Cetuximab, Everolimus	NCT01238237 NCT01870726
GBM5	PDGFRA/KIT/KDR gain	Nilotinib	NCT01140568 NCT01871311
GBM6	PIK3CA G542K, PTEN N59 fs, EGFR whole arm gain	BKM120, Everolimus, Cetuximab	NCT02142803 NCT01349660 NCT01238237
GBM7	SMO R421*, NF1 Q270*	Vismodegib, MEK162	NCT00980343 NCT01885195
GBM8*	POLE P1505S, MSH2 splice site donor c.366 + 1G > A, High mutation burden	Pembrolizumab, Nivolumab	NCT02337686 NCT02017717
GBM9	EGFR gain, KDR R1022*, PIK3R1 E443del	Cetuximab, ABT41, Bevacizumab, BKM120	NCT02573324 NCT01349660
GBM10	MGMT loss, PTEN R130*, whole arm loss NF1 E1722*	Temozolomide, Everolimus, MEK162	NCT01885195
GBM12	EGFR A289V, focal gain, PIK3R1 D560G	BKM120, Afatanib	NCT01349660 NCT01934361
GBM13	MDM2 focal gain, CDK4 focal gain, PTEN R47S	RG7112/RG7388, AMG232, Palbociclib, Everolimus	NCT01877382 NCT02143635 NCT01227434 NCT01870726
GBM14	EGFR A289V, focal gain, PIK3R1 W597G, PTEN whole arm loss and focal deletion	BKM-120, Everolimus	NCT01349660 NCT01934361 NCT01870726
GBM15	PIK3R1 T473P, EGFR whole arm gain, CDKN2A homozygous focal deletion, PTEN whole arm loss	BKM-120, Cetuximab, ABT-414, Afatanib, ABBV-221, Palbociclib, Ribociclib, Everolimus	NCT01339052 NCT02423525 NCT02573324 NCT02365662 NCT01227434 NCT02345824
GBM16	PTEN Y16*, EGFR focal gain, PDGFRA focal gain	Everolimus, Cetuximab, ABT-414, Afatanib, ABBV-221, Nilotinib, Crenolanib	NCT02423525 NCT02573324 NCT02365662 NCT02626364 NCT01140568 NCT01871311
GBM17	BRAF V600E, EGFR gain	Vemurafanib, Cobimetinib, ABT414	NCT02537600 NCT02573324 NCT02423525
GBM21	EGFR R222C, focal gain, MET P791L, focal gain, PTEN whole arm loss, CDKN2A homozygous focal deletion	Cetuximab, ABT-414, Afatanib, ABBV-221 Crizotinib, INC280, Everolimus, Palboci-clib, Ribociclib	NCT02423525 NCT02573324 NCT02365662 NCT02540161 NCT02034981 NCT02386826 NCT01227434
GBM22	PTEN V119F, whole arm loss, STAG2 focal deletion, NF1283fs, focal loss, TP53 R158H	Everolimus, Olaparib, Veliparib, MEK162, Temsirolimus/Docetaxel	NCT01390571 NCT02152892
GBM23	IDH1 R132H, RPTOR A578G	AG-120, AG-881, BAY-1436032, Everoli-mus, INK128	NCT02073994 NCT02481154 NCT02746081 NCT01434602 NCT02142803
GBM24	PIK3CA R93W, EGFR whole arm amplification, MET focal gain, PTEN R335*, T277I, whole arm loss, CDKN2A homozygous focal deletion, PALB2 whole arm loss	BKM-120, Cetuximab, ABT-414, Afatanib, ABBV-221, Crizotinib, INC280, Everolimus, Palbociclib, Ribociclib, Olaparib, Veliparib	NCT01870726 NCT01339052 NCT02423525 NCT02573324 NCT02365662 NCT02386826 NCT01227434 NCT02345824 NCT01390571 NCT02152982
GBM25	POLA1 G1178, MSH2 splice site donor c.366 + 1G > A, TP53 R175H, G245S, PDGFRA Y375H, PDGFRA/KIT/KDR focal gain, High mutation burden	Pembrolizumab, Nivolumab, Paclitaxel, Nilotinib	NCT02337686 NCT02017717 NCT02379416 NCT01140568 NCT01871311
GBM26	PIK3CA R88Q, MDM2 focal gain, CDK4 focal gain, PTEN whole arm loss	BKM-120, RG-7112, AMG-232, Palbociclib, Ribociclib, Everolimus	NCT01249660 NCT01339052 NCT01877282 NCT01723020 NCT01390571 NCT02152982 NCT02255461
GBM27	EGFRvIII, EGFR focal gain, CDKN2A homozygous focal deletion, PTEN whole arm loss	Afatnib, Rindopepimut, CAR-T, Cetuxi-mab, ABT-414, Afatanib, Palbociclib, Ribociclib, Everolimus	NCT01480479 NCT02423525 NCT02664363 NCT02573324 NCT02423525 NCT00703625 NCT01390571 NCT02152982
GBM28	PIK3R1 Q579fs, PIK3CA D939G, MET focal gain, PDGFRA focal gain, CDKN2A homozygous focal deletion, PTEN whole arm loss	BKM-120, Crizotinib, INC280, Nilotinib Palbociclib, Ribociclib, Everolimus	NCT01870726 NCT01339052 NCT01870726 NCT01339052 NCT02365662 NCT01140568 NCT01390571 NCT02152982
GBM29	EGFR A289V, EGFR focal gain, PTEN R130*, whole arm loss, CDKN2A homozygous focal deletion,	Cetuximab, ABT-414, Afatanib, Everolimus, Palbociclib, Ribociclib	NCT02573324 NCT02423525 NCT0070362 NCT01390571 NCT02152982 NCT01390571

Table 2 Key targetable variants and associated drugs and clinical trials *(Continued)*

Sample	Key Therapeutic Variants	Associated Drugs	Clinical Trials
	CDK4 focal gain		NCT02152982
GBM31	IDH1 R132H, TSC2 P1215fs, TP53 R175H, CDKN2A homozygous focal deletion	AG-120, AG-881, BAY-1436032, Everolimus, Temsirolimus, MLN0128, Temsirolimus/Docetaxel, Palbociclib, Ribociclib	NCT02073994 NCT02481154 NCT02746081 NCT002238946 NCT02142803 NCT01390571 NCT02152982
GBM32	PIK3R1 A483P, STAG2 focal deletion, PTEN M198R, whole arm loss	BKM-120, Olaparib, Veliparib, Everolimus	NCT01870726 NCT01390571 NCT02152982 NCT01434602
GBM33	PTEN Q97*, whole arm loss, CDKN2A focal loss	Everolimus, Temsirolimus, Palbociclib, Ribociclib	NCT01390571 NCT02152982
GBM34	EGFR gain, PTEN whole arm loss, TP53 C242S, V143 M, MYCN gain	Cetuximab, ABT-414, Afatanib, Everolimus, Temsirolimus, Docetaxel, CP-0610, MK-8628, GSK2820151	NCT02573324 NCT02423525 NCT00703625 NCT02698176 NCT02630251 NCT01877382
GBM35	EGFR focal gain, PIK3R1 L372dup, CDKN2A homozygous loss, PTEN whole arm loss, KIT A207V	Cetuximab, ABT-414, Afatanib, BKM-120 Palbociclib, Ribociclib, Everolimus, Imatinib, Nilotinib	NCT02573324 NCT02423525 NCT02345824 NCT01390571 NCT02152982
GBM36	NF1 c.1062 + 1 Splice Site Donor, TP53 T211I	MEK162, Temsirolimus, Docetaxel	NCT00703625

* Final therapeutic association performed post-mortem

care of at least three patients at the time of this writing. At least eight patients died or experienced significant functional decline before their physicians received NY-GGC results.

Concordance of WGS/RNA-seq and panel-based diagnostic reports

Twenty of the 30 GBM patients had targeted panel-based sequencing performed (Fig. 5). Significantly, we identified targetable variants not identified by panels in 18 of 20 (90%) cases. There were a median of 4 (IQR = 2) WGS/RNA-seq unique treatment targets, 5 (IQR = 5.5) calls in common, and 0 (IQR = 0.25) panel-unique calls per case. The number of common calls varied by panel type; WGS/RNA-seq and Panel 1, which incorporates a matched normal, had 82% of calls in common. Thirteen percent of calls were in common with Panel 3, which does not include a matched normal; we identified germline variants falsely described as tumor variants by the panel in two of five individuals. Panel 2, which also does not sequence a matched normal, reported a total of 10 tumor variants across five reports (providing therapeutic implications for one) which WGS/RNA-seq identified as germline variants. In four of the 20 cases, the panel identified one unique variant that WGS/RNA-seq did not, largely due to low VAF.

In sum, 84 additional clinically actionable calls were made using WGS/RNA-seq that were not identified by panels, compared to four made by panels and missed by WGS/RNA-seq. Of all 200 potentially actionable variants identified, panels did not identify 39.5%, and WGS/RNA-seq did not identify 2.5%. Out of the 44 calls made by two panels without matched normals, 13 (30%) were germline.

Concordance with WGA

WGA achieved good agreement with NY-GGC when comparing the reported variants (mean sensitivity = 0.71,

SD = 0.18/PPV = 0.880, SD = 0.20 across patients, Fig. 6a). WGA and NY-GGC maintained different thresholds for identifying certain types of variants. For example, NY-GGC often identified genes with copy number gains of less than five and heterozygous losses as potentially targetable while WGA did not. NY-GGC normalized some CNV calls, for example to account for tumor purity. NY-GGC also manually reviewed known cancer genes to identify SNVs that were below threshold (15% VAF) and had at least 40 total read count. Moreover, at the time of initial analysis, WGA was just beginning to be trained to identify SVs as targetable, whereas NY-GGC comprehensively noted SVs and associated therapies.

There was also good agreement in the drugs recommended when the same variants were identified and when considering similarly scoped therapies (mean sensitivity = 0.74, SD = 0.34/PPV = 0.79, SD = 0.23 across patients, Fig. 6b). As an example of one discrepancy, in NYGC-GBM13, NY-GGC prioritized *MDM2* amplification as a potential targetable variant, while WGA selected *CDK4* amplification. WGA did not associate *MDM2* with an inhibitor because open trials were not for patients with GBM. NY-GGC prioritized *MDM2* over *CDK4* therapy based on in vitro studies that *CDK4* amplification alone is not sufficient for CDK4 inhibitor sensitivity in cell lines. [57] The CDK4 inhibitor, palbociclib, also had contradicting evidence regarding blood-brain barrier penetration, [58] and tumor board discussions of preliminary data of the GBM clinical trial of palbociclib, later terminated, revealed that the physicians were less likely to recommend this drug. WGA focused on reporting only molecularly targeted therapies; thus chemotherapies and immunotherapies were outside its scope. WGA also did not offer therapeutic options for VUS where NY-GGC did, based on literature suggesting potential oncogenicity and targetability. Finally, in creating drug databases, WGA used an automated, comprehensive, and
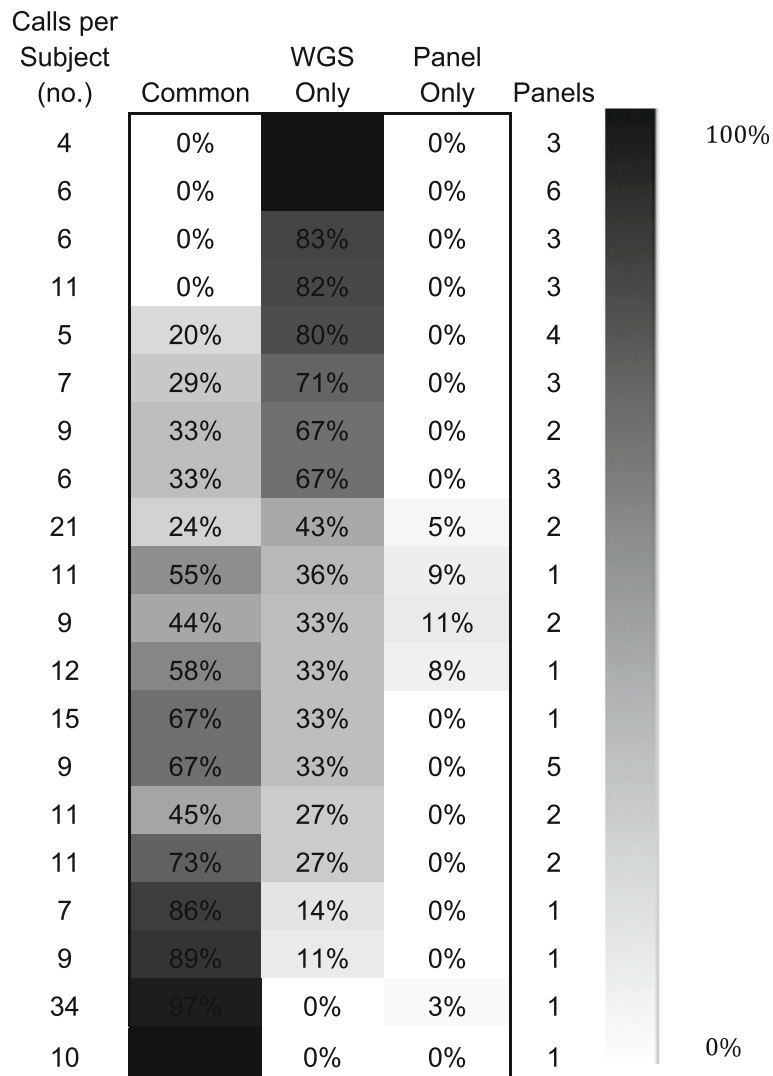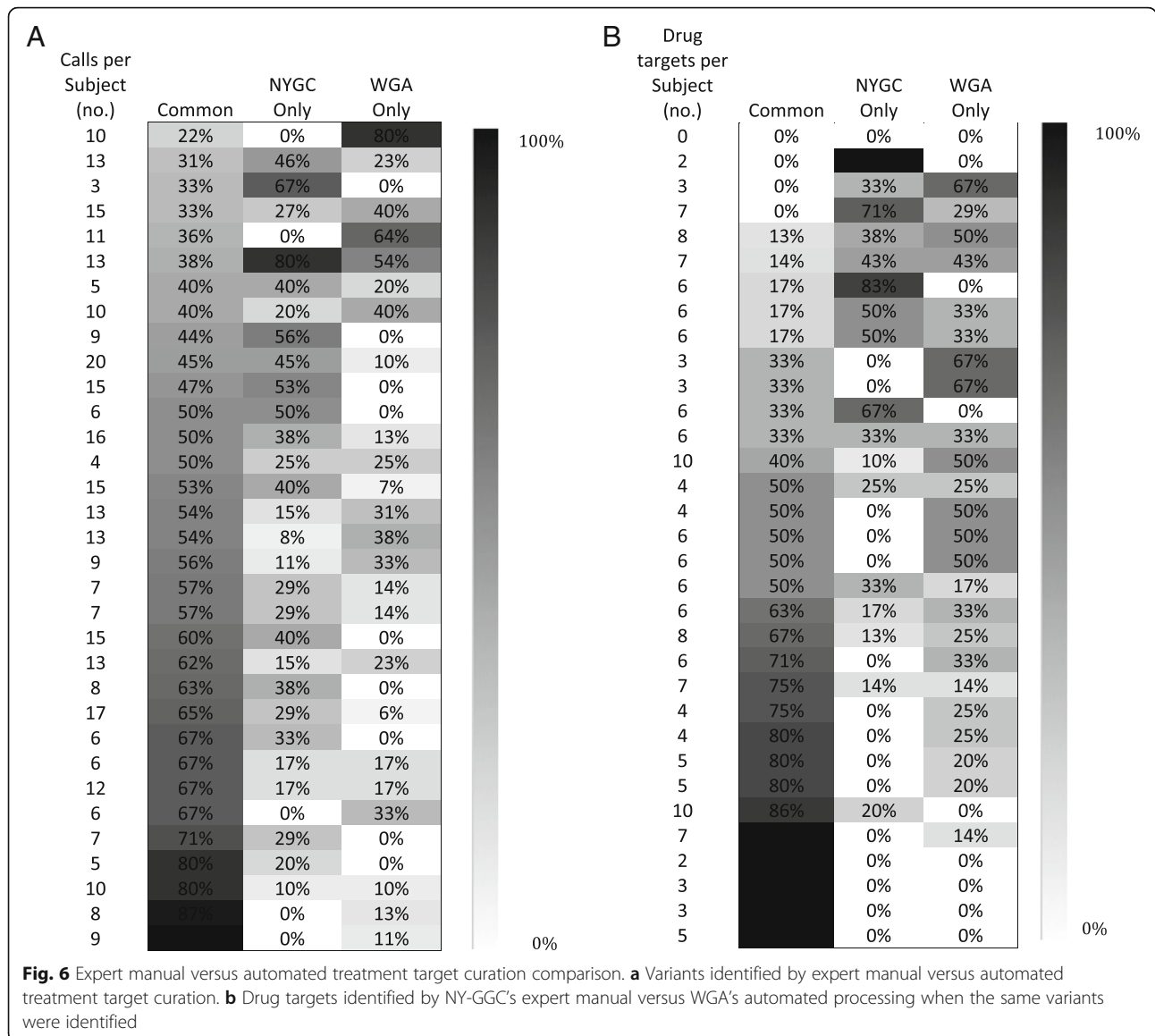
Calls per Subject (no.)	Common	WGS Only	Panel Only	Panels
4	0%	100%	0%	3
6	0%	100%	0%	6
6	0%	83%	0%	3
11	0%	82%	0%	3
5	20%	80%	0%	4
7	29%	71%	0%	3
9	33%	67%	0%	2
6	33%	67%	0%	3
21	24%	43%	5%	2
11	55%	36%	9%	1
9	44%	33%	11%	2
12	58%	33%	8%	1
15	67%	33%	0%	1
9	67%	33%	0%	5
11	45%	27%	0%	2
11	73%	27%	0%	2
7	86%	14%	0%	1
9	89%	11%	0%	1
34	97%	0%	3%	1
10	100%	0%	0%	1

Fig. 5 WGS vs focused NGS panel comparison. Number of calls made based on WGS versus calls made by focused NGS panel testing per patient. Blue indicates variant calls made uniquely by NY-GGC, green indicates variant calls made uniquely by panel testing, yellow indicates common calls made, purple indicates germline variants called as tumor variants by panel testing; *panel testing done on subsequent sample, **panel testing done on prior sample, ^partial report of panel testing available, Panel 1 = Memorial Sloan Kettering Cancer Center IMPACT Panel, Panel 2 = FoundationOne, Panel 3 = NYU Next Generation Tumor 50 Panel, Panel 4 = Weill Cornell Institute of Precision Medicine's whole exome sequencing assay, Panel 5 = University of California San Francisco's 500 Cancer Gene Panel, Panel 6 = Caris Molecular Intelligence's profile which includes next generation sequencing analysis of 44 genes as well as other assays such as immunohistochemistry

unbiased approach that included scanning PubMed abstracts with minimal manual filtering and incorporated publically available drug information resources, while NY-GGC's was created by manual review of the literature and refined through tumor boards and interdisciplinary discussions often including unpublished data. Therefore, NY-GGC's therapeutic associations adapted with knowledge and insight gained throughout the study.

Discussion

We undertook a comparative genomic-based study of GBM patients to investigate the utility of targeted panels versus WGS/RNA-seq and evaluated the feasibility and reliability of results obtained with manual versus automated curation by IBM's WGA. Together, WGS/RNA-seq was more sensitive than panels, detecting 39.5% more calls than panels, and 97.5% of the calls found by panels. Despite these findings, potentially clinically actionable calls, made in 100% of cases, altered treatment plans in only 10% of cases. This is likely a consequence of a number of factors, including availability of drug and concern of side effect profiles, lack of familiarity and slow acceptance of new technology, appropriate timing of sequencing within the patient's clinical course, lack of prior clinical

A

Calls per Subject (no.)	Common	NYGC Only	WGA Only
10	22%	0%	80%
13	31%	46%	23%
3	33%	67%	0%
15	33%	27%	40%
11	36%	0%	64%
13	38%	80%	54%
5	40%	40%	20%
10	40%	20%	40%
9	44%	56%	0%
20	45%	45%	10%
15	47%	53%	0%
6	50%	50%	0%
16	50%	38%	13%
4	50%	25%	25%
15	53%	40%	7%
13	54%	15%	31%
13	54%	8%	38%
9	56%	11%	33%
7	57%	29%	14%
7	57%	29%	14%
15	60%	40%	0%
13	62%	15%	23%
8	63%	38%	0%
17	65%	29%	6%
6	67%	33%	0%
6	67%	17%	17%
12	67%	17%	17%
6	67%	0%	33%
7	71%	29%	0%
5	80%	20%	0%
10	80%	10%	10%
8	87%	0%	13%
9		0%	11%

100% — 0% (scale)

B

Drug targets per Subject (no.)	Common	NYGC Only	WGA Only
0	0%	0%	0%
2	0%		0%
3	0%	33%	67%
7	0%	71%	29%
8	13%	38%	50%
7	14%	43%	43%
6	17%	83%	0%
6	17%	50%	33%
6	17%	50%	33%
3	33%	0%	67%
3	33%	0%	67%
6	33%	67%	0%
6	33%	33%	33%
10	40%	10%	50%
4	50%	25%	25%
4	50%	0%	50%
6	50%	0%	50%
6	50%	0%	50%
6	50%	33%	17%
6	63%	17%	33%
8	67%	13%	25%
6	71%	0%	33%
7	75%	14%	14%
4	75%	0%	25%
4	80%	0%	25%
5	80%	0%	20%
5	80%	0%	20%
10	86%	20%	0%
7		0%	14%
2		0%	0%
3		0%	0%
3		0%	0%
5		0%	0%

100% — 0% (scale)

Fig. 6 Expert manual versus automated treatment target curation comparison. **a** Variants identified by expert manual versus automated treatment target curation. **b** Drug targets identified by NY-GGC's expert manual versus WGA's automated processing when the same variants were identified

information to rigorously assess risk-vs-benefits concerns, and clinician perception of the usefulness and relevance of the sequencing data based on time elapsed between resection and sample submission; further research is warranted to identify the biggest barriers to implementation and ways to overcome them.

The majority of potentially therapeutic associations were identified in targetable genes (Tier 3) but with a variant unknown to be actionable. Such variants were prioritized when there was evidence that they would likely affect protein function, similarly to Tier 1–2 variants. These findings were typically complex and required manual curation of information from the literature, necessitating many person-hours per case. [18] The interpretation of splicing variants and structural variants resulting in splicing aberrations, only detected by

combining WGS [59] and RNA-Seq remains challenging, especially in the context of individual sample or small cohort analyses. There may also be additional connections between variants and pathways that we did not identify. Furthermore, the therapeutic associations made are based on what was available in the literature and in clinical trials at the time of analysis and interpretation of that individual sample and will evolve with new drug development.

Hence a critical component of improving scalability of data interpretation will be with automation. Here we explored this approach through a comparative analysis of manual expert with automated curation. The time required for WGA to match VCF calls with drug options improved to within 6 min over the course of the study, while manual expert curation was an average of 1.9

Frank *et al. BMC Medical Genomics*   (2019) 12:56

Page 13 of 16

months. While WGA requires further development, for example, to consider SVs as targets routinely or to expand the drug database beyond molecularly targeted therapies, the potential usefulness of this timeframe in the clinical setting is clear.

The per-patient cost of targeted panels at cancer centers are generally about $1000 and FoundationOne is currently ~$5800 [60]. By comparison, the all-in cost of clinical WGS/RNAseq and analysis by all platforms in the current study was ~$10,000. While financial cost-benefit analysis may currently favor panels or WES, our results suggest an imminent future in which the technical advantages and breadth of WGS/RNA-seq will increasingly provide improved cost-benefits by returning the most comprehensive analysis of tumor mutations. As costs and efficiencies improve, it may become reasonable to consider how to routinely apply WGS/RNA-seq to benefit cancer patients.

In addition to lowered cost and timely interpretation of sequencing data, other strategies towards improved implementation of WGS/RNA-seq include submission of samples as soon as possible after resection so that results are available for consideration at the time treatment decisions are made. This requires both clinician education about the utility of such an assay as well as the availability of clinically approved tests. NYGC has recently obtained such regulatory approval from the New York State Department of Health. One of the biggest challenges in doing so in WGS was in demonstrating reproducibility at far less depth (80X/40X for tumor normal pair) than depths of 100-200X for exomes and 500X for panels. Specifically, true negatives remain elusive even with very high depths because oncology variant callers look for variants and not for correct base calls, and remain a limit of this assay. It will be interesting to see how this influences adoption of WGS/RNA-seq for somatic variants in the coming months.

Limitations of this study include the inclusion of multiple sample preservation methods. This may have affected the variant calls, although previous studies have shown that 98% of actionable calls made in fresh frozen samples can also be made in FFPE samples. [61] Although actionable calls were made for all samples, some calls may have been missed in those samples that had tumor purity of less than 20%. [62] Another limitation is the small sample size. This sample may not have been representative of other patients with GBM. Furthermore, only 20 of the 30 patients enrolled had prior targeted panel with which the WGS analysis could be compared. To address this, we are conducting a follow-up study of 200 patients.

Conclusion

In sum, when we compared manual and automated searches for therapeutic options, WGA offered a broader array of options with a much faster processing time while maintaining significant sensitivity. Meanwhile, NY-GGC's manual curation associated clinical significance to a broader array of variants, suggested combination therapies, and incorporated feedback regarding individual patient clinical data as well as physician concerns discussed at tumor board meetings. Taken together, this study points to the potential of WGS/RNA-seq, combined with automated curation, to maximize therapeutic options which can be used in clinical decision making for the benefit of cancer patients.

Availability of data and materials
The datasets generated and/or analysed during the current study are available from the corresponding author until it becomes available from a public repository.

Author's contributions
MOF, AKR, and RBD contributed to the conception and design of the work. MOF, VF, DKhaira, AC, DKim, MFLE, CE, JB, AD, DGP, JGG, CB, JF, ABL, PC, CG, MD, ED, AO, and EP contributed to the acquisition of data. MOF, TK, KH, NR, FU, AKE, BJC, KA, MS, HG, VF, ED, SR, AF, VV, EAB, JLMV, CR, DKhaira, AC, DKim, MFLE, CE, PA, CS, JB, AD, DGP, JGG, CB, JF, ABL, PC, CG, MD, ED, AO, EP, DEO, SJH, JBP, VVM, VJ, MCZ, JK, LP, KOW, AKR and RBD contributed to the analysis and interpretation of the data and revision of manuscript. MOF, TK, KH, NR, JLMV, KOW, and RBD drafted the manuscript. All authors read and approved the manuscript.

Ethics approval and consent to participate
All participants provided written informed consent to participate. Protocols were approved by local or central Institutional Review Boards at: Rockefeller University, Biomedical Research Alliance of New York (on behalf of Northwell Health), Memorial Sloan Kettering Cancer Center, New York University School of Medicine, and Weill Cornell Medicine.

Consent for publication
Not applicable.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]New York Genome Center, 101 Avenue of the Americas, New York, NY 10013, USA. [2]Laboratory of Molecular Neuro-Oncology, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA. [3]IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA. [4]Northwell Health, Lenox Hill Hospital, 100 E. 77th Street, New York, NY 10075, USA. [5]Northwell Health, The Brain Tumor Center, 450 Lakeville Road, Lake Success, Lakeville, NY 11042, USA. [6]New York University, School of Medicine, 550 First Avenue, New York, NY 10016, USA. [7]Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. [8]Columbia University Medical Center, 710 West 168th Street, New York, NY 10032, USA. [9]Hospital for Special Surgery, 535 E. 70th Street, New York, NY 10021, USA. [10]IBM Watson Health, NW Broken Sound Bkwy, Boca Raton, FL 33487, USA. [11]Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA. [12]Present address: Google, 76 9th Avenue, New York, NY 10011, USA. [13]Present address: Rockefeller University, 1230 York Avenue, New York, NY 10065, USA. [14]Present address: 23&Me, 899 W Evelyn Ave, Mountain View, CA 94041, USA. [15]Present address: Max Planck Institute of Immunobiology and Epigenetics, Stübeweg 51 D-79108, Freiburg, Germany. [16]Present address: The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037, USA. [17]Present address: The Tisch Cancer Institute, 1470 Madison Avenue, New York, NY 10029, USA. [18]Present address: Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. [19]Present address: Harvard Medical School, 10 Shattuck Street, Boston, MA 02115, USA. [20]Present address: Yale School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA.

References
1. Griffith M, Miller CA, Griffith OL, Krysiak K, Skidmore ZL, Ramu A, et al. Optimizing cancer genome sequencing and analysis. Cell Syst 2015;1(3):210–223. doi: https://doi.org/10.1016/j.cels.2015.08.015. PubMed PMID: 26645048; PubMed Central PMCID: PMCPMC4669575.
2. Hyman DM, Solit DB, Arcila ME, Cheng DT, Sabbatini P, Baselga J, et al. Precision medicine at memorial Sloan Kettering Cancer center: clinical next-generation sequencing enabling next-generation targeted therapy trials. Drug Discov Today 2015;20(12):1422–1428. doi: https://doi.org/10.1016/j.drudis.2015.08.005. PubMed PMID: 26320725; PubMed Central PMCID: PMCPMC4940024.
3. Consortium APG. AACR project GENIE: powering precision medicine through an international consortium. Cancer Discov 2017. doi: https://doi.org/10.1158/2159-8290.CD-17-0151. PubMed PMID: 28572459.
4. Schmidt KT, Chau CH, Price DK, Figg WD. Precision oncology medicine: the clinical relevance of patient-specific biomarkers used to optimize Cancer treatment. J Clin Pharmacol 2016;56(12):1484–1499. doi: https://doi.org/10.1002/jcph.765. PubMed PMID: 27197880; PubMed Central PMCID: PMCPMC5112148.
5. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 2017;23(6):703–713. doi: https://doi.org/10.1038/nm.4333. PubMed PMID: 28481359; PubMed Central PMCID: PMCPMC5461196.
6. Turner TN, Hormozdiari F, Duyzend MH, McClymont SA, Hook PW, Iossifov I, et al. Genome sequencing of autism-affected families reveals disruption of putative noncoding regulatory DNA. Am J Hum Genet 2016;98(1):58–74. doi: https://doi.org/10.1016/j.ajhg.2015.11.023. PubMed PMID: 26749308; PubMed Central PMCID: PMCPMC4716689.
7. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proc Natl Acad Sci U S A 2015;112(17):5473–5478. doi: https://doi.org/10.1073/pnas.1418631112. PubMed PMID: 25827230; PubMed Central PMCID: PMCPMC4418901.
8. Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Sci Transl Med 2017;9(386). doi: https://doi.org/10.1126/scitranslmed.aal5209. PubMed PMID: 28424332.
9. Honeyman JN, Simon EP, Robine N, Chiaroni-Clarke R, Darcy DG, Lim, II, et al. Detection of a recurrent DNAJB1-PRKACA chimeric transcript in fibrolamellar hepatocellular carcinoma. Science. 2014;343(6174):1010–1014. doi: https://doi.org/10.1126/science.1249484. PubMed PMID: 24578576; PubMed Central PMCID: PMCPMC4286414.
10. Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A, et al. Memorial Sloan Kettering-integrated mutation profiling of actionable Cancer targets (MSK-IMPACT): A Hybridization Capture-Based Next-Generation Sequencing Clinical Assay for Solid Tumor Molecular Oncology J Mol Diagn 2015;17(3):251–264. doi: https://doi.org/10.1016/j.jmoldx.2014.12.006. PubMed PMID: 25801821.
11. Jones S, Anagnostou V, Lytle K, Parpart-Li S, Nesselbush M, Riley DR, et al. Personalized genomic analyses for cancer mutation discovery and interpretation. Sci Transl Med 2015;7(283):283ra53. doi: https://doi.org/10.1126/scitranslmed.aaa7161. PubMed PMID: 25877891; PubMed Central PMCID: PMCPMC4442685.
12. Prados MD, Byron SA, Tran NL, Phillips JJ, Molinaro AM, Ligon KL, et al. Toward precision medicine in glioblastoma: the promise and the challenges. Neuro-Oncology 2015;17(8):1051–1063. doi: https://doi.org/10.1093/neuonc/nov031. PubMed PMID: 25934816; PubMed Central PMCID: PMCPMC4490873.
13. Xiu J, Piccioni D, Juarez T, Pingle SC, Hu J, Rudnick J, et al. Multi-platform molecular profiling of a large cohort of glioblastomas reveals potential therapeutic strategies. Oncotarget. 2016. doi: https://doi.org/10.18632/oncotarget.7722. PubMed PMID: 26933808.
14. Johnson DR, O'Neill BP. Glioblastoma survival in the United States before and during the temozolomide era. J Neuro-Oncol 2012;107(2):359–364. doi: https://doi.org/10.1007/s11060-011-0749-4. PubMed PMID: 22045118.
15. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 2010;17(1):98–110. doi: https://doi.org/10.1016/j.ccr.2009.12.020. PubMed PMID: 20129251; PubMed Central PMCID: PMCPMC2818769.
16. Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. Cell. 2016;164(3):550–563. doi: https://doi.org/10.1016/j.cell.2015.12.028. PubMed PMID: 26824661; PubMed Central PMCID: PMCPMC4754110.
17. Wang Q, Hu B, Hu X, Kim H, Squatrito M, Scarpace L, et al. Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment. Cancer Cell 2017;32(1):42–56 e6. doi: https://doi.org/10.1016/j.ccell.2017.06.003. PubMed PMID: 28697342; PubMed Central PMCID: PMCPMC5599156.
18. Wrzeszczynski KO, Frank MO, Koyama T, Rhrissorrakrai K, Robine N, Utro F, et al. Comparing sequencing assays and human-machine analyses in actionable genomics for glioblastoma. Neurol Genet 2017;3(4):e164. doi: https://doi.org/10.1212/NXG.0000000000000164. PubMed PMID: 28740869; PubMed Central PMCID: PMCPMC5506390.
19. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol 2006;7:3. doi: https://doi.org/10.1186/1471-2199-7-3. PubMed PMID: 16448564; PubMed Central PMCID: PMCPMC1413964.

20. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008; 456(7221):470–476. doi: https://doi.org/10.1038/nature07509. PubMed PMID: 18978772; PubMed Central PMCID: PMCPMC2593745.

21. Nicorici D, Satalan M, Edgren H, Kangaspeska S, Murumagi A, Kallioniemi O, et al. FusionCatcher - a tool for finding somatic fusion genes in paired-end RNA-sequencing data. bioRxiv. 2014. https://doi.org/10.1101/011650.

22. Xi R, Lee S, Xia Y, Kim TM, Park PJ. Copy number analysis of whole-genome data using BIC-seq2 and its application to detection of cancer susceptibility variants. Nucleic Acids Res 2016;44(13):6274–6286. doi: https://doi.org/10.1093/nar/gkw491. PubMed PMID: 27260798.

23. Rhrissorrakrai KK, T.; Parida, L. Watson for genomics: moving personalized medicine forward. Trends in Cancer 2016;2(8):392–395.

24. Patel NM, Michelini VV, Snell JM, Balu S, Hoyle AP, Parker JS, et al. Enhancing Next-Generation Sequencing-Guided Cancer Care Through Cognitive Computing. Oncologist. 2018;23(2):179–85. Epub 2017/11/22. doi: https://doi.org/10.1634/theoncologist.2017-0170. PubMed PMID: 29158372; PubMed Central PMCID: PMCPMC5813753.

25. Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics. 2012;28(18):i333-i3i9. doi: https://doi.org/10.1093/bioinformatics/bts378. PubMed PMID: 22962449; PubMed Central PMCID: PMCPMC3436805.

26. Johnson BE, Mazor T, Hong C, Barnes M, Aihara K, McLean CY, et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. Science. 2014;343(6167):189–193. doi: https://doi.org/10.1126/science.1239947. PubMed PMID: 24336570; PubMed Central PMCID: PMCPMC3998672.

27. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. Cancer Cell 2015;27(1):15–26. doi: https://doi.org/10.1016/j.ccell.2014.12.001. PubMed PMID: 25584892.

28. Smith JS, Tachibana I, Passe SM, Huntley BK, Borell TJ, Iturria N, et al. PTEN mutation, EGFR amplification, and outcome in patients with anaplastic astrocytoma and glioblastoma multiforme. J Natl Cancer Inst 2001;93(16): 1246–1256. PubMed PMID: 11504770.

29. Wang SI, Puc J, Li J, Bruce JN, Cairns P, Sidransky D, et al. Somatic mutations of PTEN in glioblastoma multiforme. Cancer Res 1997;57(19):4183–4186. PubMed PMID: 9331071.

30. Carico C, Nuno M, Mukherjee D, Elramsisy A, Dantis J, Hu J, et al. Loss of PTEN is not associated with poor survival in newly diagnosed glioblastoma patients of the temozolomide era. PLoS One 2012;7(3):e33684. doi: https://doi.org/10.1371/journal.pone.0033684. PubMed PMID: 22479427; PubMed Central PMCID: PMCPMC3315579.

31. Vinagre J, Almeida A, Populo H, Batista R, Lyra J, Pinto V, et al. Frequency of TERT promoter mutations in human cancers. Nat Commun 2013;4:2185. doi: https://doi.org/10.1038/ncomms3185. PubMed PMID: 23887589.

32. Arita H, Yamasaki K, Matsushita Y, Nakamura T, Shimokawa A, Takami H, et al. A combination of TERT promoter mutation and MGMT methylation status predicts clinically relevant subgroups of newly diagnosed glioblastomas. Acta Neuropathol Commun 2016;4(1):79. doi: https://doi.org/10.1186/s40478-016-0351-2. PubMed PMID: 27503138; PubMed Central PMCID: PMCPMC4977715.

33. Nguyen HN, Lie A, Li T, Chowdhury R, Liu F, Ozer B, et al. Human TERT promoter mutation enables survival advantage from MGMT promoter methylation in IDH1 wild-type primary glioblastoma treated by standard chemoradiotherapy. Neuro-Oncology 2017;19(3):394–404. doi: https://doi.org/10.1093/neuonc/now189. PubMed PMID: 27571882; PubMed Central PMCID: PMCPMC5464302.

34. Batista R, Cruvinel-Carloni A, Vinagre J, Peixoto J, Catarino TA, Campanella NC, et al. The prognostic impact of TERT promoter mutations in glioblastomas is modified by the rs2853669 single nucleotide polymorphism. Int J Cancer 2016;139(2):414–423. doi: https://doi.org/10.1002/ijc.30057. PubMed PMID: 26914704.

35. Arimoto K, Burkart C, Yan M, Ran D, Weng S, Zhang DE. Plakophilin-2 promotes tumor development by enhancing ligand-dependent and -independent epidermal growth factor receptor dimerization and activation. Mol Cell Biol 2014;34(20):3843–3854. doi: https://doi.org/10.1128/MCB.00758-14. PubMed PMID: 25113560; PubMed Central PMCID: PMCPMC4187709.

36. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 2015; 372(26):2509–2520. doi: https://doi.org/10.1056/NEJMoa1500596. PubMed PMID: 26028255; PubMed Central PMCID: PMCPMC4481136.

37. Ciunci CA, Perini RF, Avadhani AN, Kang HC, Sun W, Redlinger M, et al. Phase 1 and pharmacodynamic trial of everolimus in combination with cetuximab in patients with advanced cancer. Cancer. 2014;120(1):77–85. doi: https://doi.org/10.1002/cncr.28294. PubMed PMID: 24108668.

38. Kanteti R, Riehm JJ, Dhanasingh I, Lennon FE, Mirzapoiazova T, Mambetsariev B, et al. PI3 kinase pathway and MET inhibition is efficacious in malignant pleural mesothelioma. Sci Rep 2016;6:32992. doi: https://doi.org/10.1038/srep32992. PubMed PMID: 27623107; PubMed Central PMCID: PMCPMC5021085.

39. Maira S-M, Pecchi S, Huang A, Burger M, Knapp M, Sterker D, et al. Identification and characterization of NVP-BKM120, an orally available pan-class I PI3-kinase inhibitor. Mol Cancer Ther. 2012;11(2):317–28. https://doi.org/10.1158/1535-7163.mct-11-0474.

40. Chalhoub N, Baker SJ. PTEN and the PI3-kinase pathway in cancer. Annu Rev Pathol 2009;4:127–150. doi: https://doi.org/10.1146/annurev.pathol.4.110807.092311. PubMed PMID: 18767981; PubMed Central PMCID: PMCPMC2710138.

41. Deuker MM, Marsh Durban V, Phillips WA, McMahon M. PI3'-kinase inhibition forestalls the onset of MEK1/2 inhibitor resistance in BRAF-mutated melanoma. Cancer Discov. 2015;5(2):143–153. doi: https://doi.org/10.1158/2159-8290.CD-14-0856. PubMed PMID: 25472943; PubMed Central PMCID: PMCPMC4320669.

42. Sun C, Wang L, Huang S, Heynen GJ, Prahallad A, Robert C, et al. Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. Nature. 2014;508(7494):118–122. doi: https://doi.org/10.1038/nature13121. PubMed PMID: 24670642.

43. Straussman R, Morikawa T, Shee K, Barzily-Rokni M, Qian ZR, Du J, et al. Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. Nature. 2012;487(7408):500–504. doi: https://doi.org/10.1038/nature11183. PubMed PMID: 22763439; PubMed Central PMCID: PMCPMC3711467.

44. Pietrantonio F, Oddo D, Gloghini A, Valtorta E, Berenato R, Barault L, et al. MET-Driven Resistance to Dual EGFR and BRAF Blockade May Be Overcome by Switching from EGFR to MET Inhibition in BRAF-Mutated Colorectal Cancer. Cancer Discov. 2016;6(9):963–71. Epub 2016/06/22. doi: https://doi.org/10.1158/2159-8290.CD-16-0297. PubMed PMID: 27325282.

45. Rudin CM. Vismodegib. Clin Cancer Res 2012;18(12):3218–3222. doi: https://doi.org/10.1158/1078-0432.CCR-12-0568. PubMed PMID: 22679179; PubMed Central PMCID: PMCPMC3715061.

46. Zheng X, Zeng W, Gai X, Xu Q, Li C, Liang Z, et al. Role of the hedgehog pathway in hepatocellular carcinoma (review). Oncol Rep 2013;30(5):2020–2026. doi: https://doi.org/10.3892/or.2013.2690. PubMed PMID: 23970376.

47. Sloan AE, Nock CJ, Ye X, Kerstetter A, Supko J, Lamborn K, et al. Targeting glioma-initiating cells in GBM: ABTC-0904, a randomized phase 0/II study targeting the sonic hedgehog-signaling pathway. J Clin Oncol. 2014;32(15_suppl):2026. https://doi.org/10.1200/jco.2014.32.15_suppl.2026.

48. Tanaka S, Nakada M, Nobusawa S, Suzuki SO, Sabit H, Miyashita K, et al. Epithelioid glioblastoma arising from pleomorphic xanthoastrocytoma with the BRAF V600E mutation. Brain Tumor Pathol 2014;31(3):172–176. doi: https://doi.org/10.1007/s10014-014-0192-2. PubMed PMID: 24894018.

49. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013; 500(7463):415–421. doi: https://doi.org/10.1038/nature12477. PubMed PMID: 23945592; PubMed Central PMCID: PMCPMC3776390.

50. Hunter C, Smith R, Cahill DP, Stephens P, Stevens C, Teague J, et al. A hypermutation phenotype and somatic MSH6 mutations in recurrent human malignant gliomas after alkylator chemotherapy. Cancer Res 2006; 66(8):3987–3991. doi: https://doi.org/10.1158/0008-5472.CAN-06-0127. PubMed PMID: 16618716.

51. Wang J, Cazzato E, Ladewig E, Frattini V, Rosenbloom DI, Zairis S, et al. Clonal evolution of glioblastoma under therapy. Nat Genet 2016;48(7):768–776. doi: https://doi.org/10.1038/ng.3590. PubMed PMID: 27270107.

52. Shlien A, Campbell BB, de Borja R, Alexandrov LB, Merico D, Wedge D, et al. Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. Nat Genet 2015; 47(3):257–262. doi: https://doi.org/10.1038/ng.3202. PubMed PMID: 25642631.

53. Sampson JH, Vlahovic G, Sahebjam S, Omuro AMP, Baehring JM, Hafler DA, et al. Preliminary safety and activity of nivolumab and its combination with ipilimumab in recurrent glioblastoma (GBM): CHECKMATE-143. J Clin Oncol. 2015;33(15_suppl):3010. https://doi.org/10.1200/jco.2015.33.15_suppl.3010.

54. Vlahovic G, Fecci PE, Reardon D, Sampson JH. Programmed death ligand 1 (PD-L1) as an immunotherapy target in patients with glioblastoma. Neuro-Oncology 2015;17(8):1043–1045. doi: https://doi.org/10.1093/neuonc/nov071. PubMed PMID: 25964311; PubMed Central PMCID: PMCPMC4490880.

55. Yip S, Miao J, Cahill DP, Iafrate AJ, Aldape K, Nutt CL, et al. MSH6 mutations arise in glioblastomas during temozolomide therapy and mediate temozolomide resistance. Clin Cancer Res 2009;15(14):4622–4629. doi: https://doi.org/10.1158/1078-0432.CCR-08-3012. PubMed PMID: 19584161; PubMed Central PMCID: PMCPMC2737355.

56. Cahill DP, Levine KK, Betensky RA, Codd PJ, Romany CA, Reavie LB, et al. Loss of the mismatch repair protein MSH6 in human glioblastomas is associated with tumor progression during temozolomide treatment. Clin Cancer Res 2007;13(7):2038–2045. doi: https://doi.org/10.1158/1078-0432.CCR-06-2149. PubMed PMID: 17404084; PubMed Central PMCID: PMCPMC2873832.

57. Wiedemeyer WR, Dunn IF, Quayle SN, Zhang J, Chheda MG, Dunn GP, et al. Pattern of retinoblastoma pathway inactivation dictates response to CDK4/6 inhibition in GBM. Proc Natl Acad Sci U S A 2010;107(25):11501–11506. doi: https://doi.org/10.1073/pnas.1001613107. PubMed PMID: 20534551; PubMed Central PMCID: PMCPMC2895056.

58. Schroder LB, McDonald KL. CDK4/6 inhibitor PD0332991 in glioblastoma treatment: does it have a future? Front Oncol 2015;5:259. doi: https://doi.org/10.3389/fonc.2015.00259. PubMed PMID: 26649278; PubMed Central PMCID: PMCPMC4663246.

59. Wrzeszczynski KO, Felice V, Shah M, Rahman S, Emde AK, Jobanputra V, et al. Whole genome sequencing-based discovery of structural variants in glioblastoma. Methods Mol Biol 2018;1741:1–29. Epub 2018/02/03. doi: https://doi.org/10.1007/978-1-4939-7659-1_1. PubMed PMID: 29392687.

60. Chakradhar S. Tumor sequencing takes off, but insurance reimbursement lags. Nat Med 2014;20(11):1220–1221. doi: https://doi.org/10.1038/nm1114-1220. PubMed PMID: 25375911.

61. Robbe P, Popitsch N, Knight SJL, Antoniou P, Becq J, He M, et al. Clinical whole-genome sequencing from routine formalin-fixed, paraffin-embedded specimens: pilot study for the 100,000 genomes project. Genet Med 2018;20(10):1196–1205. Epub 2018/02/02. doi: https://doi.org/10.1038/gim.2017.241. PubMed PMID: 29388947.

62. Wrzeszczynski KO, Felice V, Abhyankar A, Kozon L, Geiger H, Manaa D, et al. Analytical validation of clinical whole-genome and transcriptome sequencing of patient-derived tumors for reporting targetable variants in Cancer. J Mol Diagn 2018;20(6):822–835. Epub 2018/08/24. doi: https://doi.org/10.1016/j.jmoldx.2018.06.007. PubMed PMID: 30138725; PubMed Central PMCID: PMCPMC6198246.