

Integrated clonal analysis reveals circulating tumor DNA in urine and plasma of glioma patients.

Florent Mouliere^{1,2,3,*}, Katrin Heider^{1,2}, Christopher G. Smith^{1,2}, Jing Su^{1,2}, Mareike Thompson^{1,2}, James Morris^{1,2}, Jonathan C.M. Wan^{1,2}, Dineika Chandrananda^{1,2}, James Hadfield^{1,2,§}, Marta Grezlak^{1,2}, Irena Hudecova^{1,2}, Wendy Cooper^{1,2}, Davina Gale^{1,2}, Matt Eldridge^{1,2}, Colin Watts^{4,†}, Kevin Brindle^{1,2,*}, Nitzan Rosenfeld^{1,2,*}, Richard Mair^{1,2,4,*}

1. Cancer Research UK Cambridge Institute, University of Cambridge, CB2 0RE, Cambridge, UK.

2. Cancer Research UK Major Centre – Cambridge, Cancer Research UK Cambridge Institute, CB2 0RE, Cambridge, UK.

3. Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Pathology, Cancer Centre Amsterdam, 1081 HV, Amsterdam, The Netherlands.

4. Division of Neurosurgery, Department of Clinical Neurosciences, University of Cambridge, CB2 0QQ, Cambridge, UK.

* To whom correspondence should be addressed: f.mouliere@amsterdamumc.nl;
kmb1001@cam.ac.uk; nitzan.rosenfeld@cruk.cam.ac.uk; richard.mair@cruk.cam.ac.uk;

§ Current affiliation: Precision Medicine, R&D Oncology Unit, AstraZeneca, CB4 0WG, Cambridge, UK.

† Current affiliation: Institute of Cancer Genomics Science, University of Birmingham, B15 2TT, Birmingham, UK.

One Sentence Summary: Circulating tumor DNA can be detected in the majority of plasma and urine samples from primary brain tumor patients using sequencing guided by mutations detected in multi-region tumor biopsies.

Abstract: Glioma-derived cell-free tumor DNA is challenging to detect using standard liquid biopsy techniques as its levels in body fluids are very low, similar to those in patients with early stage carcinomas. By sequencing cell-free DNA across thousands of clonal and private mutations identified individually in each patient's tumor we detected tumor-derived DNA in plasma (10/12, 83%) and urine samples (8/11, 72%) from the majority (7/8, 87.5%) of glioma patients tested.

Introduction

Primary brain tumors, which are diagnosed in over 260,000 patients annually (1), have a poor prognosis and lack effective treatments. Better methods for early detection and identification of tumor recurrence may enable the development of novel treatment strategies including the use of immunotherapy (2, 3). The development of new treatments would also benefit from minimally invasive methods that characterise the glioma genome (4–6). DNA analysis in liquid biopsies has the potential to replace or supplement current ineffective imaging-based monitoring techniques and reduce the morbidity associated with repeated biopsy while providing the genomic information required by precision medicine (4, 7, 8). However, cell-free tumor DNA (ctDNA) is extremely challenging to detect in the plasma of patients with brain tumors as its fractional concentration (mutant allele fractions, MAF) are in the same range as those observed in plasma of patients with early stage carcinomas (9, 10). Reported detection rates for glioma are typically around 15%-30% (9), although higher rates of detection have been claimed (10). Tumor DNA has been detected in urine for some cancers, however this has been largely limited to urothelial cancers, or patients with advanced cancers having high tumor burden detected in plasma (11–13). Cerebrospinal fluid (CSF) has been proposed as an alternate medium for ctDNA analysis (14–18), however detection rates have been low as reported in a recent large study (42/85 patients detected, 49.4%) (19). In addition, CSF sampling via lumbar puncture is an invasive procedure, which limits its use for longitudinal sampling (20, 21).

Here we demonstrate that tumor DNA can be detected in the majority of plasma and urine samples from primary brain tumor patients using sequencing guided by mutations detected in multi-region tumor biopsies. When using this sensitive technique, we found no improvement in detection of tumor-derived DNA by analysing CSF samples. We also identified mutations in ctDNA obtained from plasma and urine during follow-up sampling post-surgery. We propose therefore that this method could also be used to track tumor recurrence.

Results

In this study, we analysed a total of 73 samples including tumor subparts (3 to 6 per patient), CSF, plasma, buffy coat and urine, from 8 patients: 7 patients with primary glioblastomas and 1 patient with anaplastic oligodendrogloma (**Table S1** and **Table S2**). DNA from 34 tumor subparts from 8 biopsies were analysed by whole exome sequencing (WES), resulting in detection of 8838 single nucleotide variants (SNV) (**Fig. 1A**) at an average of 1105 SNV per patient (range 435-1725) (**Table S3** and **Fig. S1**). Mutations were called from the different tumor subparts individually, and also collectively by merging the sequencing reads from these subparts in order to increase the sensitivity for detecting low abundance stem-mutations of the common clones. Hybrid-capture sequencing panels were designed to target the SNVs detected by WES in the tumor (driver and passenger mutations, a total of 8838 loci) and this list was supplemented by comprehensive coverage of the 52 most frequently mutated genes in glioma (5). Sequencing data were analyzed using the INVAR (INtegration of VAriants Reads) pipeline, which combines locus-based noise filtering, strand selection, and enrichment of mutant fragments using ctDNA biological characteristics (**Fig. S2**, **Fig. S3** and **Table S4**) (22). Fractions of tumor-derived DNA were estimated by integrated mutant allele fractions (IMAF) which indicate the fraction of reads covering target loci which carried the mutant alleles identified in the matched patient's tumor specimens.

ctDNA was detected in plasma from 7/8 patients prior to surgery, and in a total of 10/12 samples when including follow-up samples (**Fig. 1B**). ctDNA was detected in urine from 6/8 patients, in a total of 8/11 samples. Although more mutations were in general detected in CSF samples, when integrating data across all mutation loci there was no improvement in ctDNA detection rate by analysing CSF collected prior to surgery (6/8 patients detected). This was despite the mean ctDNA fraction being 243-fold lower in plasma (range 9-fold to 1343-fold) (**Fig. 1C**) and 389-fold lower in urine (range 50-fold to 834-fold) compared to CSF, with IMAF of 3.1×10^{-5} (plasma), 4.72×10^{-5} (urine), and 6.4×10^{-3} (CSF) (**Fig. 1B**).

For 3 patients, samples of both urine and plasma were obtained 6 months following surgery in addition to the baseline samples collected immediately prior to surgery. Mutated (tumor-derived) DNA was detected 6 months post-surgery in both plasma and urine samples for 2/3 patients (**Fig. 1D**). Contrast agent-enhanced T_1 -weighted MRI demonstrated that the two positive patients had residual or recurrent disease, whereas the patient with no detectable ctDNA had no evidence of recurrence. These preliminary data suggest that tumor-guided sequencing of urine or plasma samples may also be used to detect recurrent disease.

Of the 52 genes frequently mutated in glioma, mutations were detected at least once for the majority of genes across all tumor subparts (**Fig. 2A**). By limiting analysis in body fluids to these 52 most commonly mutated genes, which thus does not require pre-analysis of the tumor samples, we were able to detect mutations in plasma samples from 4/8 of cases, which was not improved by CSF analysis (also 4/8). Mutations in these genes were detected in the urine of 1/8 cases. Amongst the large number of tumor-specific mutations detected, several clinically actionable mutations were detected (**Fig. 2A** and **Fig. 2B**). Notably, *EGFR* and *MUC16* mutations were identified in plasma samples, both being potentially targetable with precision therapy (**Fig. 2A** and **Table S4**).

Detection of ctDNA in bio-fluids is thought to be affected by intra-tumoral genomic heterogeneity (23). The mutant allele fraction (MAF) in tumor DNA correlated with that of matched mutations detected in CSF ($R=0.56$, $p<0.0001$), but not in plasma ($R=0.14$, $p>0.05$) or urine ($R=0.12$, $p>0.05$) (**Fig. 2C** and **Fig. S4**). We subsequently compared the frequency of variant detection in plasma, urine and CSF samples with the presence of the mutation in one or more tumor sub-parts (**Fig. 2D**). Mutations common to all sub-parts, or that were shared across several sub-parts, were more frequently detected in CSF when compared with mutations that were private to a single region (**Fig. 2D**). This resulted in a trend for increased representation of clonal mutations in CSF samples (**Fig. 2E** and **Fig. S5**). However, in plasma and urine samples the opposite was observed where mutations tended to be representative of clones identified within a single tumor region (**Fig. 2F** and **Figure S6**).

Discussion

We have shown that ctDNA is can be detected, with suitably sensitive methods, in the urine and plasma for the majority of patients with high grade glioma; and that analysis of CSF may not provide added benefit for detection of tumor-derived DNA when such methods are employed. ctDNA was detected in the urine (6/8) and plasma (7/8) of patients with

glioma. By tracking a large number of mutations, we demonstrate that the sensitivity of detection can be improved. The median ctDNA fractions in plasma and urine were very low (IMAF of 3.1×10^{-5} and 4.72×10^{-5} respectively).

Whilst methylation-based detection, cell-free DNA genome-wide fragmentation, tumor-derived mitochondrial DNA (mtDNA), exosomes, vesicles and tumor-educated platelets have all been proposed as alternate methods for plasma-based detection of glioma-derived mutant DNA, these alternate strategies provide limited information about the tumor genome (7, 24–28). Through tumor-guided sequencing we have identified mutations in the plasma and urine of GB patients, which has the potential to track tumor-specific actionable mutations that may be important for targeting therapies and for monitoring tumor recurrence.

In other cancers, ctDNA has been shown to be representative of treatment-driven clonal tumor evolution, with real time measurements of clonal dynamics reflecting clonal tumor hierarchy (23, 29). The use of multi-region tumor sampling allowed us to identify clonal, shared and private mutations within the primary cancer (30), and on the matched CSF, plasma and urine samples. We show that the representation of these mutations varies between the plasma and urine samples, which mostly represented private clones, and the CSF which mostly represented clonal or shared mutations.

Overall, we demonstrated that ctDNA can be detected in the plasma and urine of patients with GB, a cancer that has previously been beyond the reach of non-invasive liquid biopsy.

Materials and Methods:

Study design:

Patients were recruited at Addenbrooke's Hospital, Cambridge, UK as part of the BLiNG (Biopsy of Liquids in New Gliomas) study (REC reference number: 15/EE/0094) (Supplementary Table 1). Matched tumor tissue DNA, CSF, plasma, urine and buffy coat samples were collected for each patient. Written informed consent was obtained from the patients; the studies were conducted in accordance with the Declaration of Helsinki and were approved by an Institutional Review Board. Lumbar puncture was performed immediately prior to craniotomy for tumor debulking. After sterile field preparation, the thecal sac was cannulated between the L3 and L5 intervertebral spaces using a 0.61 mm gauge lumbar puncture needle, and 10 ml of CSF was removed. After collection, CSF and whole blood samples were immediately placed on ice and then rapidly transferred to a pre-chilled centrifuge for processing. Samples were centrifuged at 1500 g at 4C for 10 minutes. Supernatant was removed and further centrifuged at 20,000 g for 10 minutes, and aliquoted into 2 mL microtubes for storage at -80 °C (Sarstedt, Germany).

Sample preparation:

Tumor tissue DNA were extracted and isolated as described previously (15). Fluids were extracted using the QIAsymphony platform (Qiagen, Germany). Up to 10 mL of plasma, 10 mL of urine and 8 mL of CSF was used per sample. DNA from plasma, urine and

CSF samples was eluted in 90 μ L, and further concentrated down to 30 μ L using a Speed-Vac concentrator (Eppendorf, Germany).

Sequencing library preparation and WES for tissue DNA:

In order to identify patient specific somatic mutations, we first performed whole exome sequencing (WES) of all tumor tissue and germline buffy coat DNA samples. Fifty nanograms of DNA were fragmented to ~120bp by acoustic shearing (Covaris) according to the manufacturer's instructions. Libraries were prepared using the Thruplex DNA-Seq protocol (Rubicon Genomics) with 5x cycles of PCR. Libraries were quantified using quantitative PCR (KAPA library quantification, KAPA biosystems) and pooled for exome capture (TruSeq Exome Enrichment Kit, Illumina). Exome capture was performed with the addition of i5 and i7 specific blockers (IDT) during the hybridization steps to prevent adaptor 'daisy chaining'. Pools were concentrated using a SpeedVac vacuum concentrator (Eppendorf, Germany). After capture, 8x cycles of PCR were performed. Enriched libraries were quantified using quantitative PCR (KAPA library quantification, KAPA Biosystems), DNA fragment sizes were assessed by Bioanalyzer (2100 Bioanalyzer, Agilent Genomics) and captures were pooled in equimolar ratio for paired-end next generation sequencing on a HiSeq4000 (Illumina).

Sequencing reads were de-multiplexed, allowing zero mismatches in barcodes. The reference genome was the GRCh37/b37/hg19 human reference genome - 1000 Genomes GRCh37-derived reference genome, which includes chromosomal plus unlocalized and unplaced contigs, the rCRS mitochondrial sequence (AC:NC_012920), Human herpesvirus 4 type 1 (AC:NC_007605) and decoy sequence derived from HuRef, Human Bac and Fosmid clones and NA12878. The sequence data of the patient samples were aligned to the reference genome using BWA-MEM v0.7.15. Duplicate reads were marked using Picard v1.122 (<http://broadinstitute.github.io/picard>). Somatic SNV and indel mutations were called using GATK Mutect2 (Genome Analysis Toolkit), (<https://www.broadinstitute.org/gatk>) in tumor-normal pair mode using buffy coat as the normal.

MAFs for each single-base locus were calculated with MuTect2 for all bases with PHRED quality ≥ 30 . After MuTect2, we applied filtering parameters so that a mutation was called if no mutant reads for an allele were observed in germline DNA at a locus that was covered at least 10x, and if at least 4 reads supporting the mutant were found in the plasma data with at least 1 read on each strand (forward and reverse). Variants were annotated using Ensembl Variant Effect Predictor with details about consequence on protein coding, accession numbers for known variants and associated allele frequencies from the 1000 Genomes project.

Tumor-guided capture sequencing:

Hybrid-based capture for the different body fluids (CSF, plasma, urine) analysis was designed to cover the variants identified above for each patient using the SureDesign software (Agilent). In addition, 52 genes of interest for glioma were included in the tumor-guided sequencing panel based on the TCGA databases. Patients were separated into 2 panels covering all the mutations included for these patients (4 patients per panel). Patients GB1, GB2, GB9, GB16 were grouped in panel 1, and patients GB7, GB11, GB12, GB19 were grouped in panel 2. Panel 1 covers in total 526 kbp (5841 regions) and panel 2 covers 526 kbp as well (5701 regions). Custom panels ranged in size between 1.404-1.430 Mb with 120

bp RNA baits. Baits were designed with 5x tiling density, moderately stringent masking and balanced boosting. 99.7% of the targets had baits designed successfully.

Indexed sequencing libraries were prepared using the ThruPLEX tag-seq kits (Takara). Libraries were captured either in 1-plex for plasma and urine samples or 3-plex for CSF samples (to a total of 1000 ng capture input) using the Agilent SureSelectXT HS protocol, with the addition of i5 and i7 blocking oligos (IDT), as recommended by the manufacturer for compatibility with ThruPLEX libraries. Custom Agilent SureSelectXT HS baits were used. 13 cycles were used for amplification of the captured libraries. Post-capture libraries were purified with AMPure XT beads, then quantified using quantitative PCR (KAPA library quantification, KAPA Biosystems), and DNA fragment sizes controlled by Bioanalyzer (2100 Bioanalyzer, Agilent Genomics). Capture libraries were then pooled in equimolar ratios for paired-end next generation sequencing on a HiSeq4000 (Illumina).

Capture sequencing analysis and INVAR:

Sequencing reads were de-multiplexed, allowing zero mismatches in barcodes. Cutadapt v1.9.1 was used to remove known 5' and 3' adaptor sequences specified in a separate FASTA of adaptor sequences. Trimmed FASTQ files were aligned to the UCSC hg19 genome using BWA-mem v0.7.13 with a seed length of 19. Error-suppression was carried out on ThruPLEX Tag-seq library BAM files using CONNOR. The consensus frequency threshold -f was set as 0.9 (90%), and the minimum family size threshold -s was varied between 2 and 5 for characterization of error rates. For capture data, a minimum family size of 2 was used.

Patient-specific sequencing data consists of informative reads at multiple known patient-specific loci that were identified from tumor sequencing (see above). Because each panel comprised mutations from multiple patients, we can compare mutant allele fractions across loci as a means of error-suppression. Patients could control for another patient's panel as long as the tumor sequencing did not identify any overlapping mutations. The distribution of signal across loci potentially allows for the identification of noisy loci not consistent with the overall signal distribution. Loci that carried signal in more than 10% of control samples or a mean allele fraction >1% were blacklisted as noisy and removed from the analysis. Each locus was also annotated with trinucleotide error rate, the corresponding tumor allele fraction, fragment size and whether that locus passes an additional outlier suppression filter as identified by INVAR (INtegration of VAriant Reads), (version 0.7).

For each sample, the IMAF (Integrated Mutant Allelic Fraction) was determined across all loci passing pre-INVAR data processing filters with mutant allele fraction at that locus of <0.25; loci with signal >0.25 mutant allele fraction were not included in the calculation because (i) loci would not be expected to have such high mutant allele fractions in body fluids of glioma patients (unless they are mis-genotyped SNPs), and (ii) if the true IMAF of a sample is >0.25, when a large number of loci are tested, they will show a distribution of allele fractions such that detection is still supported by having many low allele fraction loci with signal. Based on the ctDNA level of the sample, the binomial probability of observing each individual locus given the IMAF of that sample was calculated. Loci with a Bonferroni corrected P-value <0.05 (corrected for the number of loci interrogated) were excluded in that sample, thereby suppressing outliers.

Supplementary Materials

Fig. S1. Quality control assessment of the WES data from the tumor tissue DNA.

Fig. S2. Quality control and background sequencing error rates of the two capture sequencing panels.

Fig. S3. Mutant and non-mutant DNA fragment size distribution determined from the capture sequencing data.

Fig. S4. The number and MAF of detected mutations increase with the number of tumor sub-parts in CSF.

Fig. S5. Matched mutations detected and intra-tumor heterogeneity for each patient depending on the type of samples.

Fig. S6. MAF of the detected mutations depending on the tumor subparts and bio-fluids.

Table S1. Characteristics of the patients included in the study.

Table S2. Characteristics of the samples included in the study.

Table S3. List of SNVs called by Mutect2 after WES of the tumor tissue DNA individual sub-parts.

Table S4. Patient-specific variants passing filters after tumor-guided sequencing of the CSF, plasma and urine samples.

References and Notes:

1. Wesseling, P., and Capper, D. (2018) WHO 2016 Classification of gliomas. *Neuropathol. Appl. Neurobiol.* **44**, 139–150
2. Stewart, C. M., Kothari, P. D., Mouliere, F., Mair, R., Somnay, S., Benayed, R., Zehir, A., Weigelt, B., Dawson, S. J., Arcila, M. E., Berger, M. F., and Tsui, D. W. (2018) The value of cell-free DNA for molecular pathology. *J. Pathol.* **244**, 616–627
3. Volckmar, A.-L., Sültmann, H., Riediger, A., Fioretos, T., Schirmacher, P., Endris, V., Stenzinger, A., and Dietz, S. (2018) A field guide for cancer diagnostics using cell-free DNA: From principles to practice and clinical applications. *Genes, Chromosom. Cancer.* **57**, 123–139
4. Westphal, M., and Lamszus, K. (2015) Circulating biomarkers for gliomas. *Nat. Rev. Neurol.* **11**, 556–566
5. Brennan, C. W., Verhaak, R. G. W., McKenna, A., Campos, B., Noushmehr, H., Salama, S. R., Zheng, S., Chakravarty, D., Sanborn, J. Z., Berman, S. H., Beroukhim, R., Bernard, B., Wu, C.-J., Genovese, G., Shmulevich, I., Barnholtz-Sloan, J., Zou, L., Vegesna, R., Shukla, S. A., Ciriello, G., Yung, W. K., Zhang, W., Sougnez, C., Mikkelsen, T., Aldape, K., Bigner, D. D., Van Meir, E. G., Prados, M., Sloan, A., Black, K. L., Eschbacher, J., Finocchiaro, G., Friedman, W., Andrews, D. W., Guha, A., Iacocca, M., O'Neill, B. P., Foltz, G., Myers, J., Weisenberger, D. J., Penny, R., Kucherlapati, R., Perou, C. M., Hayes, D. N., Gibbs, R., Marra, M., Mills, G. B., Lander, E., Spellman, P., Wilson, R., Sander, C., Weinstein, J., Meyerson, M., Gabriel, S., Laird, P. W., Haussler, D., Getz, G., Chin, L., and TCGA Research Network (2013) The somatic genomic landscape of glioblastoma. *Cell.* **155**, 462–77
6. McLendon, R., Friedman, A., Bigner, D., Van Meir, E. G., Brat, D. J., Mastrogianakis,

G. M., Olson, J. J., Mikkelsen, T., Lehman, N., Aldape, K., Yung, W. K. A., Bogler, O., Weinstein, J. N., VandenBerg, S., Berger, M., Prados, M., Muzny, D., Morgan, M., Scherer, S., Sabo, A., Nazareth, L., Lewis, L., Hall, O., Zhu, Y., Ren, Y., Alvi, O., Yao, J., Hawes, A., Jhangiani, S., Fowler, G., San Lucas, A., Kovar, C., Cree, A., Dinh, H., Santibanez, J., Joshi, V., Gonzalez-Garay, M. L., Miller, C. A., Milosavljevic, A., Donehower, L., Wheeler, D. A., Gibbs, R. A., Cibulskis, K., Sougnez, C., Fennell, T., Mahan, S., Wilkinson, J., Ziaugra, L., Onofrio, R., Bloom, T., Nicol, R., Ardlie, K., Baldwin, J., Gabriel, S., Lander, E. S., Ding, L., Fulton, R. S., McLellan, M. D., Wallis, J., Larson, D. E., Shi, X., Abbott, R., Fulton, L., Chen, K., Koboldt, D. C., Wendl, M. C., Meyer, R., Tang, Y., Lin, L., Osborne, J. R., Dunford-Shore, B. H., Miner, T. L., Delehaunty, K., Markovic, C., Swift, G., Courtney, W., Pohl, C., Abbott, S., Hawkins, A., Leong, S., Haipek, C., Schmidt, H., Wiechert, M., Vickery, T., Scott, S., Dooling, D. J., Chinwalla, A., Weinstock, G. M., Mardis, E. R., Wilson, R. K., Getz, G., Winckler, W., Verhaak, R. G. W., Lawrence, M. S., O'Kelly, M., Robinson, J., Alexe, G., Beroukhim, R., Carter, S., Chiang, D., Gould, J., Gupta, S., Korn, J., Mermel, C., Mesirov, J., Monti, S., Nguyen, H., Parkin, M., Reich, M., Stransky, N., Weir, B. A., Garraway, L., Golub, T., Meyerson, M., Chin, L., Protopopov, A., Zhang, J., Perna, I., Aronson, S., Sathiamoorthy, N., Ren, G., Yao, J., Wiedemeyer, W. R., Kim, H., Sek, W. K., Xiao, Y., Kohane, I. S., Seidman, J., Park, P. J., Kucherlapati, R., Laird, P. W., Cope, L., Herman, J. G., Weisenberger, D. J., Pan, F., Van Den Berg, D., Van Neste, L., Joo, M. Y., Schuebel, K. E., Baylin, S. B., Absher, D. M., Li, J. Z., Southwick, A., Brady, S., Aggarwal, A., Chung, T., Sherlock, G., Brooks, J. D., Myers, R. M., Spellman, P. T., Purdom, E., Jakkula, L. R., Lapuk, A. V., Marr, H., Dorton, S., Yoon, G. C., Han, J., Ray, A., Wang, V., Durinck, S., Robinson, M., Wang, N. J., Vranizan, K., Peng, V., Van Name, E., Fontenay, G. V., Ngai, J., Conboy, J. G., Parvin, B., Feiler, H. S., Speed, T. P., Gray, J. W., Brennan, C., Soccia, N. D., Olshen, A., Taylor, B. S., Lash, A., Schultz, N., Reva, B., Antipin, Y., Stukalov, A., Gross, B., Cerami, E., Wei, Q. W., Qin, L. X., Seshan, V. E., Villafania, L., Cavatore, M., Borsig, L., Viale, A., Gerald, W., Sander, C., Ladanyi, M., Perou, C. M., Hayes, D. N., Topal, M. D., Hoadley, K. A., Qi, Y., Balu, S., Shi, Y., Wu, J., Penny, R., Bittner, M., Shelton, T., Lenkiewicz, E., Morris, S., Beasley, D., Sanders, S., Kahn, A., Sfeir, R., Chen, J., Nassau, D., Feng, L., Hickey, E., Barker, A., Gerhard, D. S., Vockley, J., Compton, C., Vaught, J., Fielding, P., Ferguson, M. L., Schaefer, C., Zhang, J., Madhavan, S., Buetow, K. H., Collins, F., Good, P., Guyer, M., Ozenberger, B., Peterson, J., and Thomson, E. (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*. **455**, 1061–1068

7. Kros, J. M., Mustafa, D. M., Dekker, L. J. M., Smitt, P. A. E. S., Luider, T. M., and Zheng, P. P. (2015) Circulating glioma biomarkers. *Neuro. Oncol.* **17**, 343–360
8. Wan, J. C. M., Massie, C., Garcia-Corbacho, J., Mouliere, F., Brenton, J. D., Caldas, C., Pacey, S., Baird, R., and Rosenfeld, N. (2017) Liquid biopsies come of age: Towards implementation of circulating tumour DNA. *Nat. Rev. Cancer.* **17**, 223–238
9. Bettegowda, C., Sausen, M., Leary, R. J., Kinde, I., Wang, Y., Agrawal, N., Bartlett, B. R., Wang, H., Luber, B., Alani, R. M., Antonarakis, E. S., Azad, N. S., Bardelli, A., Brem, H., Cameron, J. L., Lee, C. C., Fecher, L. A., Gallia, G. L., Gibbs, P., Le, D., Giuntoli, R. L., Goggins, M., Hogarty, M. D., Holdhoff, M., Hong, S. M., Jiao, Y., Juhl, H. H., Kim, J. J., Siravegna, G., Laheru, D. A., Lauricella, C., Lim, M., Lipson, E. J., Marie, S. K. N., Netto, G. J., Oliner, K. S., Olivi, A., Olsson, L., Riggins, G. J., Sartore-Bianchi, A., Schmidt, K., Shih, I. M., Oba-Shinjo, S. M., Siena, S.,

Theodorescu, D., Tie, J., Harkins, T. T., Veronese, S., Wang, T. L., Weingart, J. D., Wolfgang, C. L., Wood, L. D., Xing, D., Hruban, R. H., Wu, J., Allen, P. J., Schmidt, C. M., Choti, M. A., Velculescu, V. E., Kinzler, K. W., Vogelstein, B., Papadopoulos, N., and Diaz, L. A. (2014) Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl. Med.* **6**, 224ra24

10. Zill, O. A., Banks, K. C., Fairclough, S. R., Mortimer, S. A., Vowles, J. V., Mokhtari, R., Gandara, D. R., Mack, P. C., Odegaard, J. I., Nagy, R. J., Baca, A. M., Eltoukhy, H., Chudova, D. I., Lanman, R. B., and Talasaz, A. A. (2018) The landscape of actionable genomic alterations in cell-free circulating tumor DNA from 21,807 advanced cancer patients. *Clin. Cancer Res.* **24**, 3528–3538

11. Patel, K. M., Van Der Vos, K. E., Smith, C. G., Mouliere, F., Tsui, D., Morris, J., Chandrananda, D., Marass, F., Van Den Broek, D., Neal, D. E., Gnanapragasam, V. J., Forshaw, T., Van Rhijn, B. W., Massie, C. E., Rosenfeld, N., and Van Der Heijden, M. S. (2017) Association of Plasma and Urinary Mutant DNA with Clinical Outcomes in Muscle Invasive Bladder Cancer. *Sci. Rep.* **7**, 5554

12. Dudley, J. C., Schroers-Martin, J., Lazzareschi, D. V., Shi, W. Y., Chen, S. B., Esfahani, M. S., Trivedi, D., Chabon, J. J., Chaudhuri, A. A., Stehr, H., Liu, C. L., Lim, H., Costa, H. A., Nabet, B. Y., Liao, J. C., Alizadeh, A. A., and Diehn, M. (2019) Detection and surveillance of bladder cancer using urine tumor DNA. *Cancer Discov.* **9**, 500–509

13. Husain, H., Melnikova, V. O., Kosco, K., Woodward, B., More, S., Pingle, S. C., Weihe, E., Park, B. H., Tewari, M., Erlander, M. G., Cohen, E., Lippman, S. M., and Kurzrock, R. (2017) Monitoring Daily Dynamics of Early Tumor Response to Targeted Therapy by Detecting Circulating Tumor DNA in Urine. *Clin. Cancer Res.* **23**, 4716–4723

14. De Mattos-Arruda, L., Mayor, R., Ng, C. K. Y., Weigelt, B., Martínez-Ricarte, F., Torrejon, D., Oliveira, M., Arias, A., Raventos, C., Tang, J., Guerini-Rocco, E., Martínez-Saéz, E., Lois, S., Marín, O., De La Cruz, X., Piscuoglio, S., Towers, R., Vivancos, A., Peg, V., Cajal, S. R. y, Carles, J., Rodon, J., González-Cao, M., Tabernero, J., Felip, E., Sahuquillo, J., Berger, M. F., Cortes, J., Reis-Filho, J. S., and Seoane, J. (2015) Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. *Nat. Commun.* **6**, 8839

15. Wang, Y., Springer, S., Zhang, M., McMahon, K. W., Kinde, I., Dobbyn, L., Ptak, J., Brem, H., Chaichana, K., Gallia, G. L., Gokaslan, Z. L., Groves, M. L., Jallo, G. I., Lim, M., Olivi, A., Quinones-Hinojosa, A., Rigamonti, D., Riggins, G. J., Sciubba, D. M., Weingart, J. D., Wolinsky, J. P., Ye, X., Oba-Shinjo, S. M., Marie, S. K. N., Holdhoff, M., Agrawal, N., Diaz, L. A., Papadopoulos, N., Kinzler, K. W., Vogelstein, B., and Bettegowda, C. (2015) Detection of tumor-derived DNA in cerebrospinal fluid of patients with primary tumors of the brain and spinal cord. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 9704–9709

16. Mouliere, F., Mair, R., Chandrananda, D., Marass, F., Smith, C. G., Su, J., Morris, J., Watts, C., Brindle, K. M., and Rosenfeld, N. (2018) Detection of cell-free DNA fragmentation and copy number alterations in cerebrospinal fluid from glioma patients. *EMBO Mol. Med.* **10**, e9323

17. Pentsova, E. I., Shah, R. H., Tang, J., Boire, A., You, D., Briggs, S., Omuro, A., Lin, X., Fleisher, M., Grommes, C., Panageas, K. S., Meng, F., Selcuklu, S. D., Ogilvie, S.,

Distefano, N., Shagabayeva, L., Rosenblum, M., DeAngelis, L. M., Viale, A., Mellinghoff, I. K., and Berger, M. F. (2016) Evaluating cancer of the central nervous system through next-generation sequencing of cerebrospinal fluid. *J. Clin. Oncol.* **34**, 2404–2415

18. Seoane, J., De Mattos-Arruda, L., Rhun, E. Le, Bardelli, A., and Weller, M. (2019) Cerebrospinal fluid cell-free tumour DNA as a liquid biopsy for primary brain tumours and central nervous system metastases. *Ann. Oncol.* **30**, 211–218

19. Miller, A. M., Shah, R. H., Pentsova, E. I., Pourmaleki, M., Briggs, S., Distefano, N., Zheng, Y., Skakodub, A., Mehta, S. A., Campos, C., Hsieh, W.-Y., Selcuklu, S. D., Ling, L., Meng, F., Jing, X., Samoila, A., Bale, T. A., Tsui, D. W. Y., Grommes, C., Viale, A., Souweidane, M. M., Tabar, V., Brennan, C. W., Reiner, A. S., Rosenblum, M., Panageas, K. S., DeAngelis, L. M., Young, R. J., Berger, M. F., and Mellinghoff, I. K. (2019) Tracking tumour evolution in glioma through liquid biopsies of cerebrospinal fluid. *Nature*. **565**, 654–658

20. Hasbun, R., Abrahams, J., Jekel, J., and Quagliarello, V. J. (2001) Computed Tomography of the Head before Lumbar Puncture in Adults with Suspected Meningitis. *N. Engl. J. Med.* **345**, 1727–1733

21. Engelborghs, S., Niemantsverdriet, E., Struyfs, H., Blennow, K., Brouns, R., Comabella, M., Dujmovic, I., van der Flier, W., Frölich, L., Galimberti, D., Gnanapavan, S., Hemmer, B., Hoff, E., Horts, J., Iacobaeus, E., Ingelsson, M., Jan de Jong, F., Jonsson, M., Khalil, M., Kuhle, J., Lleó, A., de Mendonça, A., Molinuevo, J. L., Nagels, G., Paquet, C., Parnetti, L., Roks, G., Rosa-Neto, P., Scheltens, P., Skårgård, C., Stomrud, E., Tumani, H., Visser, P. J., Wallin, A., Winblad, B., Zetterberg, H., Duits, F., and Teunissen, C. E. (2017) Consensus guidelines for lumbar puncture in patients with neurological diseases. *Alzheimer's Dement. Diagnosis, Assess. Dis. Monit.* **8**, 111–126

22. Wan, J. C. M., Heider, K., Gale, D., Murphy, S., Fisher, E., Mouliere, F., Ruiz-Valdepenas, A., Santonja, A., Morris, J., Chandrananda, D., Marshall, A., Gill, A. B., Chan, P. Y., Barker6, E., Young6, G., Cooper, W. N., Hudcová, I., Marass, F., Mair, R., Brindle, K. M., Stewart, G. D., Abraham, J., Caldas, C., Rassl, D. M., Rintoul, R. C., Bignell, G. R., Alifrangis, C., Middleton, M. R., Gallagher, F. A., Parkinson, C., Durrani, A., McDermott, U., Smith, C. G., Massie, C., Corrie, P. G., and Rosenfeld, N. (2019) ctDNA monitoring to parts per million using patient-specific sequencing and integration of variant reads. *BioRxiv*

23. Murtaza, M., Dawson, S. J., Pogrebniak, K., Rueda, O. M., Provenzano, E., Grant, J., Chin, S. F., Tsui, D. W. Y., Marass, F., Gale, D., Ali, H. R., Shah, P., Contente-Cuomo, T., Farahani, H., Shumansky, K., Kingsbury, Z., Humphray, S., Bentley, D., Shah, S. P., Wallis, M., Rosenfeld, N., and Caldas, C. (2015) Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. *Nat. Commun.* **6**, 8760

24. Moss, J., Magenheim, J., Neiman, D., Zemmour, H., Loyfer, N., Korach, A., Samet, Y., Maoz, M., Druid, H., Arner, P., Fu, K.-Y., Kiss, E., Spalding, K. L., Landesberg, G., Zick, A., Grinshpun, A., Shapiro, A. M. J., Grompe, M., Wittenberg, A. D., Glaser, B., Shemer, R., Kaplan, T., and Dor, Y. (2018) Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat. Commun.* **9**, 5068

25. Mouliere, F., Chandrananda, D., Piskorz, A. M., Moore, E. K., Morris, J., Ahlborn, L.

B., Mair, R., Goranova, T., Marass, F., Heider, K., Wan, J. C. M., Supernat, A., Hudcová, I., Gounaris, I., Ros, S., Jimenez-Linan, M., Garcia-Corbacho, J., Patel, K., Østrup, O., Murphy, S., Eldridge, M. D., Gale, D., Stewart, G. D., Burge, J., Cooper, W. N., Van Der Heijden, M. S., Massie, C. E., Watts, C., Corrie, P., Pacey, S., Brindle, K. M., Baird, R. D., Mau-Sørensen, M., Parkinson, C. A., Smith, C. G., Brenton, J. D., and Rosenfeld, N. (2018) Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci. Transl. Med.* **10**, eaat4921

26. Best, M. G., Sol, N., Kooi, I., Tannous, J., Westerman, B. A., Rustenburg, F., Schellen, P., Verschueren, H., Post, E., Koster, J., Ylstra, B., Ameziane, N., Dorsman, J., Smit, E. F., Verheul, H. M., Noske, D. P., Reijneveld, J. C., Nilsson, R. J. A., Tannous, B. A., Wesseling, P., and Wurdinger, T. (2015) RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. *Cancer Cell.* **28**, 666–676

27. Mair, R., Mouliere, F., Smith, C. G., Chandrananda, D., Gale, D., Marass, F., Tsui, D. W. Y., Massie, C. E., Wright, A. J., Watts, C., Rosenfeld, N., and Brindle, K. M. (2019) Measurement of plasma cell-free mitochondrial tumor DNA improves detection of glioblastoma in patient-derived orthotopic xenograft models. *Cancer Res.* **79**, 220–230

28. Perakis, S., and Speicher, M. R. (2017) Emerging concepts in liquid biopsies. *BMC Med.* **15**, 75

29. Gremel, G., Lee, R. J., Girotti, M. R., Mandal, A. K., Valpione, S., Garner, G., Ayub, M., Wood, S., Rothwell, D. G., Fusi, A., Wallace, A., Brady, G., Dive, C., Dhomen, N., Lorigan, P., and Marais, R. (2016) Distinct subclonal tumour responses to therapy revealed by circulating cell-free DNA. *Ann. Oncol.* **27**, 1959–1965

30. Sottoriva, A., Spiteri, I., Piccirillo, S. G. M., Touloumis, A., Collins, V. P., Marioni, J. C., Curtis, C., Watts, C., and Tavaré, S. (2013) Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 4009–14

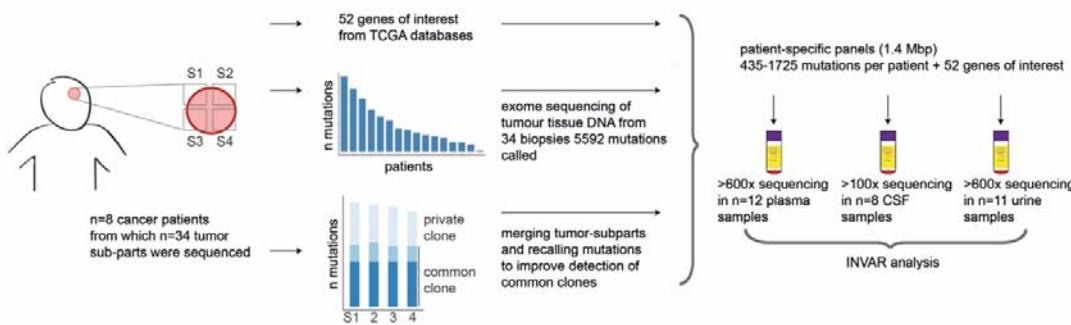
Acknowledgments: We wish to thank for their help and support the Cancer Research UK Cambridge Institute core facilities, in particular bio-repository, bioinformatics and genomics. We wish also to thank the Cambridge Molecular Diagnostic Laboratory. **Funding:** We would like also to acknowledge the support of The University of Cambridge, Cancer Research UK (grant numbers A11906, A20240, 17242, 16465). The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 337905. Colin Watts has received funding from the Brain Tumour Charity (grant 10/136). **Author contributions:** Concept and design of the study: F.M., N.R. and R.M.; Methodology: F.M., K.H., J.C.M.W., C.G.S. and R.M.; Investigation: F.M., K.H., M.T., C.G.S., D.G., I.H., W.C., M.G., J.C.M.W., I.H., W.C., D.G. and R.M.; Sample Collection: R.M. and C.W.; Data Analysis: F.M., K.H., N.R. and R.M.; Computational analysis: F.M., J.S., C.G.S., J.M., J.C.M.W., M.E., D.C. and K.H.; Writing – Original Draft: F.M. and R.M.; Writing – Review & Editing: F.M., R.M., K.H., I.H., C.G.S., C.W., K.B., and N.R.; Funding Acquisition: C.W., K.B., N.R., R.M.; Supervision: F.M., K.B., N.R. and R.M. **Competing interests:** N.R. and D.G are co-founder, shareholder and officer/consultant of Inivata Ltd, a cancer genomics company that commercialises circulating DNA analysis. C.G.S. has consulted for Inivata Ltd. Inivata and AstraZeneca had no role in the conception, design, data collection and analysis of the study.

N.R. is co-inventor of patent WO/2016/009224. Several of the authors may be listed as co-inventors on patent applications filed. Other co-authors have no conflict of interests. **Data and materials availability:** Liquid biopsy sequencing data will be made available at the European Genome-phenome archive. The INVAR pipeline will be made publicly accessible at <http://www.bitbucket.org/nrlab/invar>.

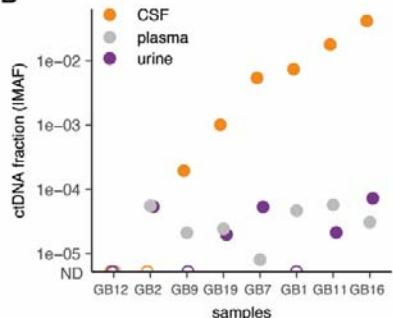
Figures:

Fig. 1:

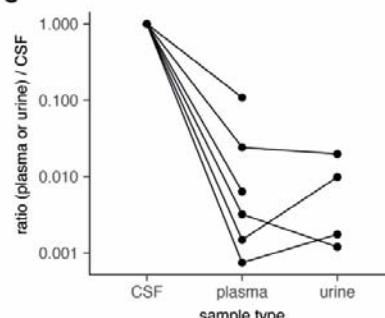
A



B



C



D

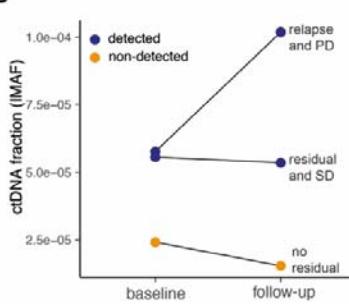


Fig. 1: Integrating tumor clonality analysis improves the detection of ctDNA in CSF and plasma from gliomas patients. **A:** Schematic of ctDNA detection in matched CSF, plasma and urine samples from glioma patients using INVAR. Depth of sequencing indicated is the mean across the samples analyzed. **B:** Estimated ctDNA fractions for the matched plasma (7/8 detected), CSF (6/8 detected) and urine samples (5/8 detected) collected immediately pre-surgery from 7 patients with primary glioblastoma and one patient (GB12) with anaplastic oligodendrogloma. The ctDNA fraction is expressed as IMAF (Integrated Mutant Allele Fraction). Detected cases are indicated by full circle and non-detected cases as an open circle. ND: non-detected. **C:** Estimated relative ratio of tumor DNA fractions (IMAF) in plasma and urine compared to the matched CSF samples. One patient (GB12) had tumor DNA detected in plasma and urine but not in CSF and is not shown. **D:** Detection of residual and progressive disease in plasma samples taken within 6 months of treatment. Follow-up plasma samples were available for 3 patients (GB2, GB11 and GB12). Blue data points indicate the IMAF of samples with detected ctDNA GB11 with progressive disease, PD, and GB2 with stable disease, SD), whilst orange data points indicate the IMAF of samples (from GB12 with no residual disease observed) that were not confirmed as being detected (below the detection limit).

the limit of calling). Data are annotated with clinical information (residual disease, relapse, PD and SD) determined by imaging.

Fig. 2:

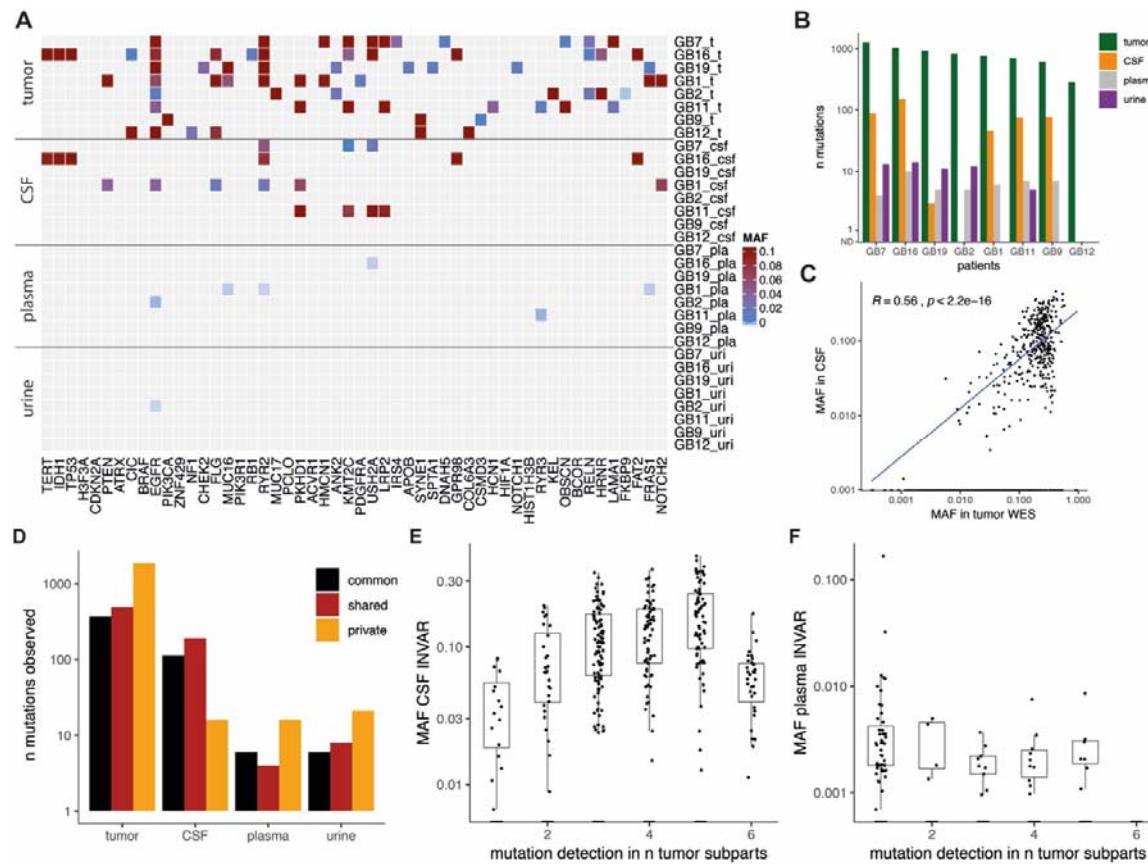


Fig. 2: Detection of ctDNA in plasma, urine and CSF is affected by intra-tumor heterogeneity in gliomas. **A:** Mutant allele fraction (MAF) in tumor samples and body fluids for the detected mutations in the 52 most frequently mutated genes in gliomas based on the TCGA databases. **B:** Summary of the number of mutations detected in DNA from tumor tissue, CSF, plasma and urine samples in the baseline pre-surgery samples from the 8 patients. **C:** Comparison of the MAF of detected mutation in tumor tissue DNA by WES, with the MAF of the mutation detected in CSF by targeted sequencing. **D:** Mutations are classified depending on their respective clones (in black commonly detected in all tumor subparts, in red shared across some tumor sub-parts, and in orange private mutations). **E, F:** Mutations common and shared across multiples tumor tissue subparts are more frequently detected and with a higher MAF than private mutations in CSF samples (E), but not in plasma samples (F).