Detection of cell-free DNA fragmentation and copy number alterations in cerebrospinal fluid from glioma patients

Mouliere, Florent; Mair, Richard; Chandrananda, Dineika; Marass, Francesco; Smith, Christopher G.; Su, Jing; Morris, James; Watts, Colin; Brindle, Kevin M.; Rosenfeld, Nitzan

DOI: 10.15252/emmm.201809323

License: Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Mouliere, F, Mair, R, Chandrananda, D, Marass, F, Smith, CG, Su, J, Morris, J, Watts, C, Brindle, KM & Rosenfeld, N 2018, 'Detection of cell-free DNA fragmentation and copy number alterations in cerebrospinal fluid from glioma patients', EMBO Molecular Medicine, vol. 10, e9323. https://doi.org/10.15252/emmm.201809323

Link to publication on Research at Birmingham portal
Detection of cell-free DNA fragmentation and copy number alterations in cerebrospinal fluid from glioma patients

Florent Mouliere¹,²,³,†, Richard Mair¹,²,⁴, Dineika Chandrananda¹,²,‡, Francesco Marass¹,²,⁵,⁶, Christopher G Smith¹,², Jing Su¹,², James Morris¹,², Colin Watts⁷, Kevin M Brindle¹,²,⁸,⁹,‡ & Nitzan Rosenfeld¹,²,*,‡,‖

Abstract

Glioma is difficult to detect or characterize using current liquid biopsy approaches. Detection of cell-free tumor DNA (cfDNA) in cerebrospinal fluid (CSF) has been proposed as an alternative to detection in plasma. We used shallow whole-genome sequencing (sWGS, at a coverage of < 0.4×) of cell-free DNA from the CSF of 13 patients with primary glioma to determine somatic copy number alterations and DNA fragmentation patterns. This allowed us to determine the presence of cfDNA in CSF without any prior knowledge of point mutations present in the tumor. We also showed that the fragmentation pattern of cell-free DNA in CSF is different from that in plasma. This low-cost screening method provides information on the tumor genome and can be used to target those patients with high levels of cfDNA for further largescale sequencing, such as by whole-exome and whole-genome sequencing.

Keywords cell-free DNA; cerebrospinal fluid; fragmentation; glioma; shallow WGS

Subject Categories Biomarkers & Diagnostic Imaging; Cancer

DOI 10.15252/emmm.201809323 | Received 11 May 2018 | Revised 11 October 2018 | Accepted 15 October 2018 | Published online 6 November 2018

EMBO Mol Med (2018) 10: e9323

See also: AP Cheng et al (December 2018)

Introduction

The detection rate and concentration of cell-free tumor DNA (cfDNA) in the plasma of patients with glioma are extremely low (Bettegowda et al, 2014), making glioma refractory to current plasma-based liquid biopsy approaches. Recent studies using DNA sequencing of CSF samples showed that the fraction of mutant DNA in CSF was higher than in the plasma of glioma patients. However, these methods were based on identification of mutations in tumor biopsies and targeted sequencing (Wang et al, 2015) or large-scale, untargeted, and expensive analyses using whole-exome sequencing (WES; De Mattos-Arruda et al, 2015; Pentsova et al, 2016). Despite these advances, detection of cfDNA in CSF remains challenging as it is affected by both tumor heterogeneity and the location of the tumor (Wang et al, 2015). The lack of somatic alteration hotspots in gliomas, other than IDH mutations (Brennan et al, 2013), further limits the utility of targeted assays.

Recent studies have examined fragmentation patterns in total cell-free DNA (cfDNA) and specifically the tumor fraction of cfDNA (called cftDNA) in plasma and urine samples (Jiang & Lo, 2016). Distinct mechanisms of DNA release have been proposed, identifiable through the patterns of cfDNA fragmentation that they produce (Thierry et al, 2016; Wan et al, 2017). Apoptosis is associated with a pattern of chromatin cleavage that generates DNA fragments of around 167 bp (Lo et al, 2010) and multiples thereof. While fragments of DNA in the plasma generally display a mean fragment size of 167 bp, samples from individuals with a high tumor burden show a trend toward shorter sizes (Mouliere et al, 2011; Underhill et al, 2016). Non-tumor cfDNA is present
mainly as fragments of around 167 bp, while cftDNA fragments are around 145 bp (Jiang et al., 2015; Underhill et al., 2016; preprint: Mouliere et al., 2017). Furthermore, an oscillatory pattern has been observed in the distribution of fragment lengths below 150 bp with a periodicity of 10 bp, which likely corresponds to the length of DNA sections wrapped around the nucleosome (Fan et al., 2010; Lo et al., 2010). The size profile of tumor-derived DNA in CSF has not previously been characterized.

Results and Discussion

We used double-stranded library preparation and paired-end shallow whole-genome sequencing (sWGS) at low depth of coverage (<0.4×) to infer somatic copy number alterations (SCNAS) and DNA fragment size distributions in CSF samples from 13 glioma patients (Fig 1A and Appendix Fig S1). SCNAS were observed in 5/13 (39%) patients (Fig 1A and B). Despite the small cohort size and limited sensitivity of sWGS, which generally detects the presence of cftDNA only down to concentrations of ~5% of total cftDNA (Heitzer et al., 2013; Adalsteinsson et al., 2017), our rate of detection was similar to previous studies performed with more expensive genome-wide sequencing methods (De Mattos-Arruda et al., 2015; Pentsova et al., 2016). Targeted methods that track tumor mutations previously identified in each patient by DNA sequencing of tissue biopsies have a higher rate of cftDNA detection (Wang et al., 2015), but the requirement for invasive biopsies makes these methods impractical and expensive in a clinical setting. The levels of cftDNA in the CSF were associated with the presence of SCNAS, and the CSF samples with the highest cftDNA concentrations were those from patients with tumor SCNAS (Fig 1A), but not associated with tumor volume (Appendix Table S1). For one patient, tumor tissue DNA was available from multiple subregions within the tumor (Fig 1C and...
Appendix Fig S2). Shared alterations between the different tumor subparts were observed in the CSF sample, even if some private alterations were not detected. The SCNAs detected by sWGS in 28 genes commonly altered in glioblastoma showed good concordance between the CSF and tumor tissue DNA samples (Fig 2A). Interestingly, three genomic alterations, notably a loss of chromosome 10 and gain of chromosome 7, were missed in one of the four tumor biopsies (G1_T1) but were detected in CSF (Fig 2A). In addition, we compared the SCNAs spanning EGFR and PTEN between tumor samples and the corresponding CSF samples for the 12 other patients (Fig 2B). Detection and confirmation of the SCNAs in the CSF sample was influenced by the size of the tumor, the level of cfDNA, and the glioma grade.

Paired-end sWGS also allowed analysis of the size distribution of cfDNA fragments in the CSF samples (Fig 3A and B). For patient G1, we analyzed the fragmentation pattern of cfDNA in matched CSF, plasma, and urine samples (Fig 3A). The fragmentation pattern in CSF differed from that observed in the plasma and urine samples (Fig 3A). Eight CSF samples (62%) exhibited enrichment of DNA fragments at \(<145\) bp, and a low or absent peak at \(-167\) bp (Fig 3B and Appendix Fig S3). Analysis of the fragment size distribution showed that \(>50\)% of the fragments were smaller than 150 bp in the CSF samples (compared to \(<20\)% for the plasma samples; Fig 3C). SCNAs were observed primarily in patients with a DNA fragmentation pattern that had a peak at \(-145\) bp (Fig 3D). A peak at \(-133\) bp was also noted in one patient, which may represent a fragmentation pattern associated with a histone variant (Melters et al., 2015; Appendix Fig S2). The patients with detected cfDNA showed a shift toward shorter fragment sizes and also an apparent reduction in the amplitude of the 10 bp periodic peaks observed below 145 bp (Fig 3D and E). Patients with detected SCNAs also had a relative enrichment of fragments in the size range 220–320 bp (Fig 3E). In addition to the shift toward shorter cfDNA fragment sizes in CSF (Mouliere et al., 2011, 2014; Jiang et al., 2015; Underhill et al., 2016), we also observed a negative correlation between the amplitude of the 10 bp periodic peak patterns in fragment size distributions, calculated from the size profile of sWGS data, and the levels of SCNAs in these CSF samples (Fig 3F), (Pearson, \(-0.85, P = 0.0002\)). Thus, an overall decrease in the peak fragment size was associated with a reduction in the amplitude of the subnucleosomal peaks. The origin of the 10 bp oscillatory pattern is believed to be due to variable accessibility of the DNA due to its winding around the histone cores (Jiang & Lo, 2016). Alternative nuclease activities in cancer or alternate mechanisms of DNA release may produce an alteration in this fragmentation pattern. The use of this subnucleosomal fragmentation signature enabled us to identify those CSF samples in which there were higher concentrations of cfDNA, irrespective of the analysis of point mutations or SCNAs.

In summary, untargeted sWGS (<0.4× coverage) of CSF cfDNA from 13 patients with primary brain tumors identified SCNAs in 5/13 patients. In one patient where multiple tumor samples were analyzed, SCNAs detected in the CSF cfDNA recapitulated the overall alterations detected in tumor tissue DNA and further identified SCNAs absent in one tumor region. Analysis of cfDNA size profiles demonstrated relative enrichment in fragments of shorter length (145 bp) in the CSF of patients with primary brain tumors. This increase in the number of shorter DNA fragments correlated with the presence of tumor-derived SCNAs in the CSF. Finally, we observed a significant negative correlation between the amplitude of the periodic 10 bp peaks in fragment size distribution and the levels of SCNAs. This fragmentation signature provides an alternative way to detect the presence of cfDNA in CSF that requires no prior knowledge of point mutations or SCNAs within the tumor. The combination of SCNAs and fragmentation analysis of cfDNA by sWGS provides important genomic information at low cost and more rapidly when compared with WES or targeted sequencing guided by analysis of tumor biopsies (De Mattos-Arruda et al., 2015; Wang et al., 2015; Pentsova et al., 2016; Adalsteinsson et al., 2017). This approach could be used to identify samples for in-depth and more costly analysis by genome-wide or capture sequencing (Adalsteinsson et al., 2017).

Figure 2. Detection of SCNAs in CSF is influenced by tumor grade, cfDNA concentration, and contact between the tumor and CSF.

A Heat map summarizing the SCNAs detected by sWGS of 28 genes of interest in tumor biopsies and CSF from patient G1 (four tumor subparts and one CSF sample). Amplifications are shown in dark blue, deletions are in orange, and copy number neutral regions are in light gray.

B Heat map summarizing detection of EGFR and PTEN alterations in tumor tissue and in CSF samples. Shared detection in tissue and CSF is indicated in green, detection of the alteration only in tissue in orange, and non-detection in blue. The top bars indicate the cfDNA concentration (copies/ml); in a range of purples), the size of the tumors (in a range of brown), the type of glioma (in a range of blues), and whether the tumor was in direct contact with the CSF or not (based on MRI, green or red, respectively). Samples are ranked from the left to right by decreasing concentration of cfDNA (copies/ml).
Materials and Methods

Sample collection and DNA extraction

Thirteen glioma patients were recruited from Addenbrooke’s Hospital, Cambridge, UK, as part of the BLING study (REC—15/EE/0094). 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. All samples were immediately placed on ice and then rapidly transferred for a further 10 min, and aliquoted into 2-ml microtubes for storage at –80°C.

DNA was extracted from single aliquots (2 ml) of the pre-operative CSF samples using the QIAamp circulating nucleic acid kit (Qiagen). The elution volume was 50 μl. Internal oligonucleotide controls, based on sequence fragments from the *Xenopus tropicalis* genome, were used to estimate the efficiency of DNA extraction.

Shallow whole-genome sequencing

Indexed sequencing libraries were prepared using a commercially available kit (ThruPLEX Plasma-Seq, Rubicon Genomics). Libraries were pooled in equimolar amounts and sequenced to < 0.4× depth of coverage on a HiSeq 4000 (Illumina) generating 150-bp paired-end reads. Sequence data were analyzed using an in-house pipeline where paired-end sequence reads were aligned to the human reference genome (GRCh37) using BWA-MEM following the removal of contaminating adapter sequences. PCR and optical duplicates were marked using MarkDuplicates (Picard Tools) feature, and these were excluded from downstream analysis along with reads of low mapping quality and supplementary alignments. When necessary, reads were downsampled to 10 million in all samples for comparison purposes. The amplitude of the 10 bp periodic peaks was calculated from the sWGS following the manufacturer’s instructions. The elution volume was 50 μl. Internal oligonucleotide controls, based on sequence fragments from the *Xenopus tropicalis* genome, were used to estimate the efficiency of DNA extraction.

Supplementary Methods

Sample collection and DNA extraction

Materials and Methods

Sample collection and DNA extraction

Thirteen glioma patients were recruited from Addenbrooke’s Hospital, Cambridge, UK, as part of the BLING study (REC—15/EE/0094). 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. All samples were immediately placed on ice and then rapidly transferred for a further 10 min, and aliquoted into 2-ml microtubes for storage at –80°C.

DNA was extracted from single aliquots (2 ml) of the pre-operative CSF samples using the QIAamp circulating nucleic acid kit (Qiagen) following the manufacturer’s instructions. The elution volume was 50 μl. Internal oligonucleotide controls, based on sequence fragments from the *Xenopus tropicalis* genome, were used to estimate the efficiency of DNA extraction.

Shallow whole-genome sequencing

Indexed sequencing libraries were prepared using a commercially available kit (ThruPLEX Plasma-Seq, Rubicon Genomics). Libraries were pooled in equimolar amounts and sequenced to < 0.4× depth of coverage on a HiSeq 4000 (Illumina) generating 150-bp paired-end reads. Sequence data were analyzed using an in-house pipeline where paired-end sequence reads were aligned to the human reference genome (GRCh37) using BWA-MEM following the removal of contaminating adapter sequences. PCR and optical duplicates were marked using MarkDuplicates (Picard Tools) feature, and these were excluded from downstream analysis along with reads of low mapping quality and supplementary alignments. When necessary, reads were downsampled to 10 million in all samples for comparison purposes. The amplitude of the 10 bp periodic peaks was calculated from the sWGS.
The paper explained

Problem
Detection of circulating cell-free tumor DNA (cfDNA) is difficult in the case of brain tumors, particularly in gliomas. Analysis of cfDNA in cerebrospinal fluid (CSF), rather than in blood plasma, has improved detection frequency. However, current detection methods are insensitive and require either prior knowledge of tumor mutations, determined by sequencing of tumor biopsies, or expensive whole-exome sequencing.

Results
We used untargeted, low-coverage whole-genome sequencing (< 0.4×) to detect somatic copy number alterations (SCNAs) in cell-free DNA in the CSF of glioma patients. Combining analyses of these SCNAs with DNA fragmentation patterns determined using paired-end sequencing allows detection of cfDNA in CSF using sWGS data.

Impact
The facilitated detection of cfDNA in the CSF of glioma patients can allow genomic analysis of their tumors using cost-effective sequencing methods.

Data availability
Sequencing data for this study are deposited in the EGA database, accession number EGAS00001003255 (https://www.ebi.ac.uk/ega/studies/EGAS00001003255). Other data associated with this study are present in the paper or Appendix.

Acknowledgements
We wish to thank for their help and support the Cancer Research UK Cambridge Institute core facilities, in particular the biological resource unit, bio-repository, bioinformatics, and genomics. We wish also to thank the Cambridge Molecular Diagnostic Laboratory, and in particular Dr. Mikel Velganan. 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.

Author contributions
Concept and design of the study: FMo and RM; methodology: FMo, RM, and CGS; investigation: FMo and RM; data analysis: FMo, RM, CGS, DC, FMa, KB, and NR; computational analysis: FMo, JM, CGS, FMa, JS, and DC; writing—original draft: F.M. and R.M.; writing—review and editing: F.M., R.M., D.C., CGS, FMa, JM, CW, KB, and NR; funding acquisition: CW, KB, and NR; supervision: FMo, KB, and NR.

Conflict of interest
N.R. is co-founder, shareholder, and officer/consultant of Inivata Ltd, a cancer genetics company that commercializes circulating DNA analysis. C.G.S. and F. Marass have consulted for Inivata Ltd. Inivata had no role in the conception, design, data collection, and analysis of the study. Other co-authors have no conflict of interests.

For more information
(i) https://www.cancerresearchuk.org/about-cancer/brain-tumours
(ii) https://www.thebraintumourcharity.org/
(iii) http://www.neurosurg.cam.ac.uk/

References
Adalsteinsson VA, Ha G, Freeman SS, Choudhury AD, Stover DG, Parsons HA, Cgdush G, Reed SC, Rotem D, Rhodes J et al (2017) Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nat Commun 8: 1324
Benjamin Y, Speed TP (2012) Summarizing and correcting the GC content bias in high-throughput sequencing. Nucleic Acids Res 40: e72
Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM et al (2014) Detection of circulating tumor DNA in early- and late-stage human malignancies. Sci Transl Med 6: 224ra24
Brennan CW, Verhaak RGW, McKenna A, Campos B, Noushmehr H, Salama SR, Zheng S, Chakravarty D, Sanborn JZ, Berman SH et al (2013) The somatic genomic landscape of glioblastoma. Cell 155: 462–477
Chandrananda D, Thorne NP, Ganesamoorthy D, Bruno DL, Benjamin Y, Speed TP, Slater HR, Bahlo M (2014) Investigating and correcting plasma DNA sequencing coverage bias to enhance aneuploidy discovery. PLoS One 9: e86993
De Mattos-Arruda L, Mayor R, Ng CKY, Weigelt B, Martinez-Ricarte F, Torrejon D, Oliveira M, Arias A, Raventos C, Tang J et al (2015) Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. Nat Commun 6: 8839
Fan HC, Blumenfeld YJ, Chitkara U, Hudgins L, Quake SR (2010) Analysis of the size distributions of fetal and maternal cell-free DNA by paired-end sequencing. Clin Chem 56: 1279–1286
Ha G, Roth A, Lai D, Bashashati A, Ding J, Goya R, Giuliany R, Rosner J, Oloumi A, Shumansky K et al (2012) Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. Genome Res 22: 1995–2007

Heitzer E, Ulz P, Belic J, Gutschi S, Quehenberger F, Fischедер K, Benezeder T, Auer M, Pischler C, Mannweiler S et al (2013) Tumor-associated copy number changes in the circulation of patients with prostate cancer identified through whole-genome sequencing. Genome Med 5: 30

jiang P, Chan CWM, Chan KCA, Cheng SH, Wong J, Wong VW-S, Wong GLH, Chan SL, Mok TSK, Chan HLY et al (2015) Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. Proc Natl Acad Sci USA 112: E1317 – E1325

jiang P, Lo YMD (2016) The long and short of circulating cell-free DNA and the Ins and outs of molecular diagnostics. Trends Genet 32: 360 – 371

Lo YMD, Chan KCA, Sun H, Chen EZ, Jiang P, Lun FMF, Zheng YW, Leung Ty, Lau TK, Cantor CR et al (2010) Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. Sci Transl Med 2: 61ra91

Melters D, Nye J, Zhao H, Dalal Y (2015) Chromatin dynamics in vivo: a game of musical chairs. Genes (Basel) 6: 751 – 776

Mouliere F, Robert B, Arnau Peyrotte E, Del Rio M, Ychou M, Molina F, Gongora C, Thierry AR (2011) High fragmentation characterizes tumour-derived circulating DNA. PloS One 6: e23418

Mouliere F, El Messaoudi S, Pang D, Dritschilo A, Thierry AR (2014) Multi-marker analysis of circulating cell-free DNA toward personalized medicine for colorectal cancer. Mol Oncol 8: 927 – 941

Mouliere F, Piskorz AM, Chandrananda D, Moore E, Morris J, Smith CG, Goranova T, Heider K, Mair R, Supernat A et al (2017) Selecting short DNA fragments in plasma improves detection of circulating tumour DNA. bioRxiv https://doi.org/10.1101/134437 [PREPRINT]

Olshen AB, Venkatraman ES, Lucito R, Wigler M (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 5: 557 – 572

Pentsova EI, Shah RH, Tang J, Boire A, You D, Briggs S, Omuro A, Lin X, Fleisher M, Grommes C et al (2016) Evaluating cancer of the central nervous system through next-generation sequencing of cerebrospinal fluid. J Clin Oncol 34: 2404 – 2415

Thierry AR, El Messaoudi S, Gahan PB, Anker P, Stroun M (2016) Origins, structures, and functions of circulating DNA in oncology. Cancer Metastasis Rev 35: 347 – 376

Underhill HR, Kitzman JO, Hellwig S, Welker NC, Daza R, Baker DN, Gligorich KM, Rostomily RC, Bronner MP, Shendure J (2016) Fragment length of circulating tumour DNA. PloS Genet 12: e1006162

Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, Pacey S, Baird R, Rosenfeld N (2017) Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nat Rev Cancer 17: 223 – 238

Wang Y, Springer S, Zhang M, McMahon KW, Kinde I, Dobbyn L, Ptak J, Brem H, Chaichana K, Gallia GL et al (2015) Detection of tumor-derived DNA in cerebrospinal fluid of patients with primary tumors of the brain and spinal cord. Proc Natl Acad Sci 112: 9704 – 9709

License: This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.