

ARTICLE OPEN

Characterisation of the changing genomic landscape of metastatic melanoma using cell free DNA

Anthony Cutts^{1,2}, Oliver Venn^{3,4}, Alexander Dilthey^{4,7}, Avinash Gupta^{5,8}, Dimitris Vavoulis^{1,2}, Helene Dreau^{1,2}, Mark Middleton^{2,5}, Gil McVean⁴, Jenny C. Taylor⁴ and Anna Schuh^{2,5,6}

Cancer is characterised by complex somatically acquired genetic aberrations that manifest as intra-tumour and inter-tumour genetic heterogeneity and can lead to treatment resistance. In this case study, we characterise the genome-wide somatic mutation dynamics in a metastatic melanoma patient during therapy using low-input (50 ng) PCR-free whole genome sequencing of cell-free DNA from pre-treatment and post-relapse blood samples. We identify de novo tumour-specific somatic mutations from cell-free DNA, while the sequence context of single nucleotide variants showed the characteristic UV-damage mutation signature of melanoma. To investigate the behaviour of individual somatic mutations during proto-oncogene B-Raf -targeted and immune checkpoint inhibition, amplicon-based deep sequencing was used to verify and track frequencies of 212 single nucleotide variants at 10 distinct time points over 13 months of treatment. Under checkpoint inhibition therapy, we observed an increase in mutant allele frequencies indicating progression on therapy 88 days before clinical determination of non-response positron emission tomography-computed tomography. We also revealed mutations from whole genome sequencing of cell-free DNA that were not present in the tissue biopsy, but that later contributed to relapse. Our findings have potential clinical applications where high quality tumour-tissue derived DNA is not available.

npj Genomic Medicine (2017)2:25; doi:10.1038/s41525-017-0030-7

INTRODUCTION

Whole exome sequencing (WES) and whole genome sequencing (WGS) of DNA derived from tumour biopsies has generated atlases of somatic mutations across cancer types, revealing deep genetic heterogeneity within and between tumours.^{1–3} These catalogues have improved the molecular classification of cancer types and are a resource for drug development.^{4, 5} However, for WES and WGS, high quality DNA from fresh tumour material is required, which has limited their widespread application in clinical practice.^{6–8} Establishing the acquired genetic profile of a cancer by using targeted sequencing of specific cancer genes is increasingly used to both help diagnosis and guide treatment decisions and commonly uses formalin fixed paraffin-embedded tumour material. However, tumour sequencing has its limitations: biopsies are invasive limiting treatment monitoring, obtaining tumour material can be difficult and provides only limited representations of the potential underlying heterogeneity through inherent sampling bias. In the era of molecularly targeted high-cost continuous maintenance therapy, monitoring of treatment response is essential to avoid continuing ineffective therapies, to prevent unnecessary side effects from inappropriate therapy, and to determine the benefit of new therapeutics as early as possible in the treatment course. Therefore, there is an urgent need for new approaches to comprehensively and non-invasively identify and

track changes in the mutational landscape across primary tumours and metastases.

Cell-free DNA (cfDNA) molecules are short nuclear DNA fragments present in several bodily fluids generated predominantly through cell apoptosis.⁹ The abundance of cfDNA in blood plasma is highly variable and is influenced by a number of covariates, including post blood draw contamination.¹⁰ The average length of cfDNA are 167 bases with a periodicity of 10 bases, with increasing evidence suggesting the majority of tumour fragments may be shorter than non-tumour derived cfDNA.^{11–13} In cancer patients, tumour-originated circulating tumour DNA (ctDNA) typically constitute a minority population of cfDNA in plasma.¹⁴ Observed correlations between ctDNA abundance and imaged tumour size,¹⁵ stage and type¹⁴ suggest a potential molecular representation of disease. To address the challenges of low ctDNA abundance high efficiency and high fidelity detection approaches are required for which various digital PCR and targeted next generation sequencing (NGS) approaches have been developed.^{15–18} Additionally, WES studies have been used to identify low level variants and monitor tumour evolution.^{19–22} So far, genome-wide sequencing efforts of cfDNA have focussed on detection of a limited set of tumour specific mutations in patients with advanced cancers.^{21, 23–27}

In the current study we tracked mutations from pre-treatment, post-treatment and relapse cfDNA samples from a patient with malignant melanoma across 10 different time points. We

¹Nuffield Division of Clinical Laboratory Sciences (NDCLS), Oxford Molecular Diagnostics Centre, University of Oxford, Oxford, UK; ²NIHR Oxford Biomedical Research Centre, Oxford, UK; ³Lighthouse Cancer Diagnostics Ltd, Oxford, UK; ⁴The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; ⁵University of Oxford Department of Oncology, Churchill Hospital, Oxford, UK and ⁶Department of Haematology, Oxford University Hospital Trust, Oxford, UK

Correspondence: Anna Schuh (Anna.Schuh@oncology.ox.ac.uk)

⁷Present address: NHGRI-NIH, 20892 Bethesda, MD, USA

⁸Present address: Department of Medical Oncology, The Christie NHS Foundation Trust, Manchester, UK
Anthony Cutts, Oliver Venn and Alexander Dilthey contributed equally to this work

Received: 10 February 2017 Revised: 17 July 2017 Accepted: 25 July 2017

Published online: 04 September 2017

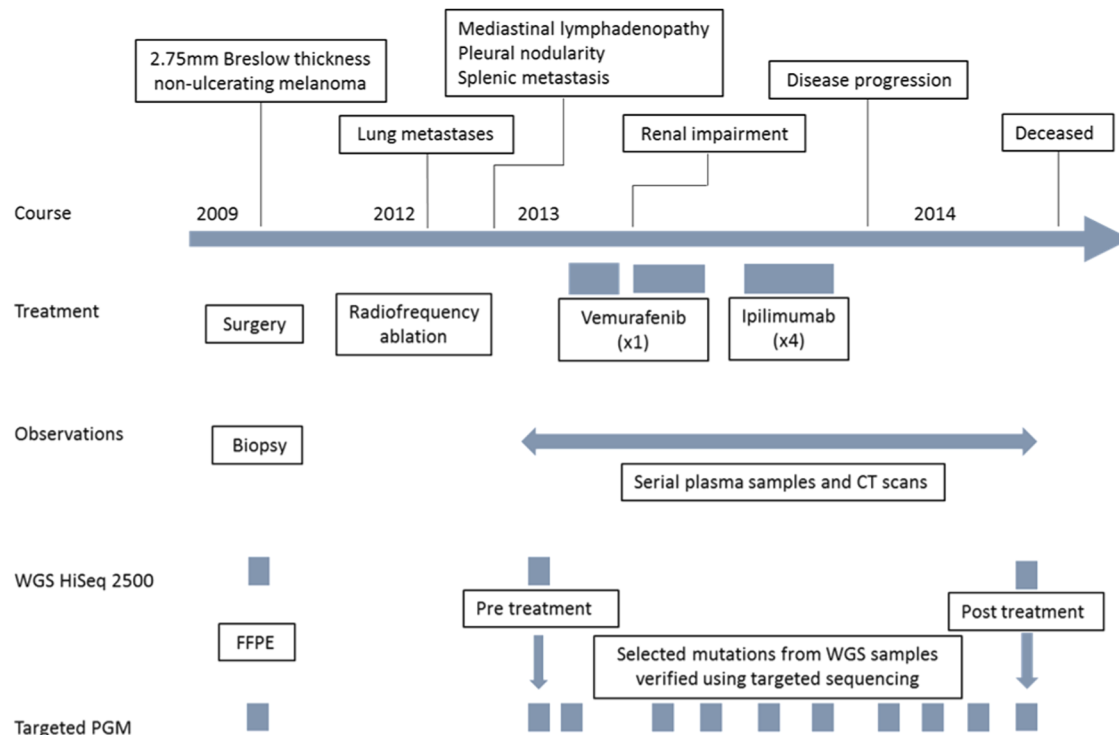


Fig. 1 Outline of study design. Timeline indicates, when samples were collected, the type of sequencing that was undertaken, treatment administered and key stages in disease development

developed a low-input PCR-free protocol and bioinformatics pipeline to facilitate WGS analysis of cfDNA. We provide proof-of-principle that this approach is feasible and reliably identifies a broad spectrum of different types of mutations including those in the non-coding regions of the genome. Finally, we identify mutation signatures associated with exposure to UV-light from cfDNA and identify clinically actionable mutations that can be used to direct treatment and molecular monitoring of treatment response.

RESULTS

WGS sequencing quality

To characterise the impact of therapy on tumour genetic diversity, we performed WGS using cfDNA from samples taken before vemurafenib and after ipilimumab treatment, a formalin-fixed paraffin-embedded (FFPE) tumour biopsy and genomic DNA from peripheral blood mononuclear cells (Fig. 1).

To address the low abundance of ctDNA in cfDNA we developed a low input (50 ng) DNA library preparation method. Post-sequencing metrics showed similar depth of read uniformity and depth of coverage distribution (Fig. 2a and b) compared to libraries prepared from genomic DNA. As expected, insert sizes of ctDNA were small (Fig. 2c). The pre-therapy and post-therapy cfDNA samples were sequenced to an average depth of 100× and 71× respectively. Ninety-eight percent of the genomes were covered to a depth greater than 30×. The normal sample, obtained from peripheral blood mononuclear cells, was sequenced to a mean depth of coverage of 34×, with 96% of the genome covered 30× or more. The archival tumour tissue obtained from FFPE material performed poorly compared to the normal and cfDNA samples however, with much greater variability observed in coverage across the genome.

Identification of somatic genomic alterations from cfDNA using WGS

WGS was applied to cfDNA extracted from the plasma sample taken at time point 1 and an optimised bioinformatics pipeline was developed and applied to detect somatic single nucleotide variants (SNVs). Briefly, to control sequencing artefacts, we developed a two stage somatic mutation calling algorithm. First, we identified candidate mutations using a full probabilistic model applied to sequencing alignments correcting for sequencing errors using empirical confusion matrices. Second, we applied de novo assembly to regions in which candidates were identified to guard against artefacts caused by genome divergence from the human reference genome sequence.

Time point 1

After filtering candidate somatic mutations to control false positive rate (SOM), we detected 91 non-synonymous exonic SNVs by WGS of the cfDNA that ranged from 5–16% variant allele frequency (VAF) (Supplementary Table 1). Using an orthogonal custom-made Ion Torrent targeted deep sequencing panel (TSP) at an average depth greater than 3000× to cover the positions of the 91 SNVs, we were able to validate 84 somatic SNVs in this cfDNA sample. Not unexpectedly, the FFPE sample was of insufficient quality to yield accurate WGS data.^{28, 29} We therefore performed TSP sequencing on the tumour sample. The vast majority of mutant alleles detected by WGS of cfDNA were present in the primary tumour DNA (88/91), with VAFs in the primary tumour varying from 3 to 96%. Notably three variants were validated in the cfDNA that were not detected in the primary tumour DNA using TSP at 3000× (Supplementary Table 2). Conversely, seven variants detected in the biopsy were not seen in the WGS of the pre-treatment cfDNA sample. These were all subclonal with VAFs ranging from 4 to 33%. Where these discrepancies were observed and variants were not detected, only the major alleles were present as opposed to variants being present but not being called.

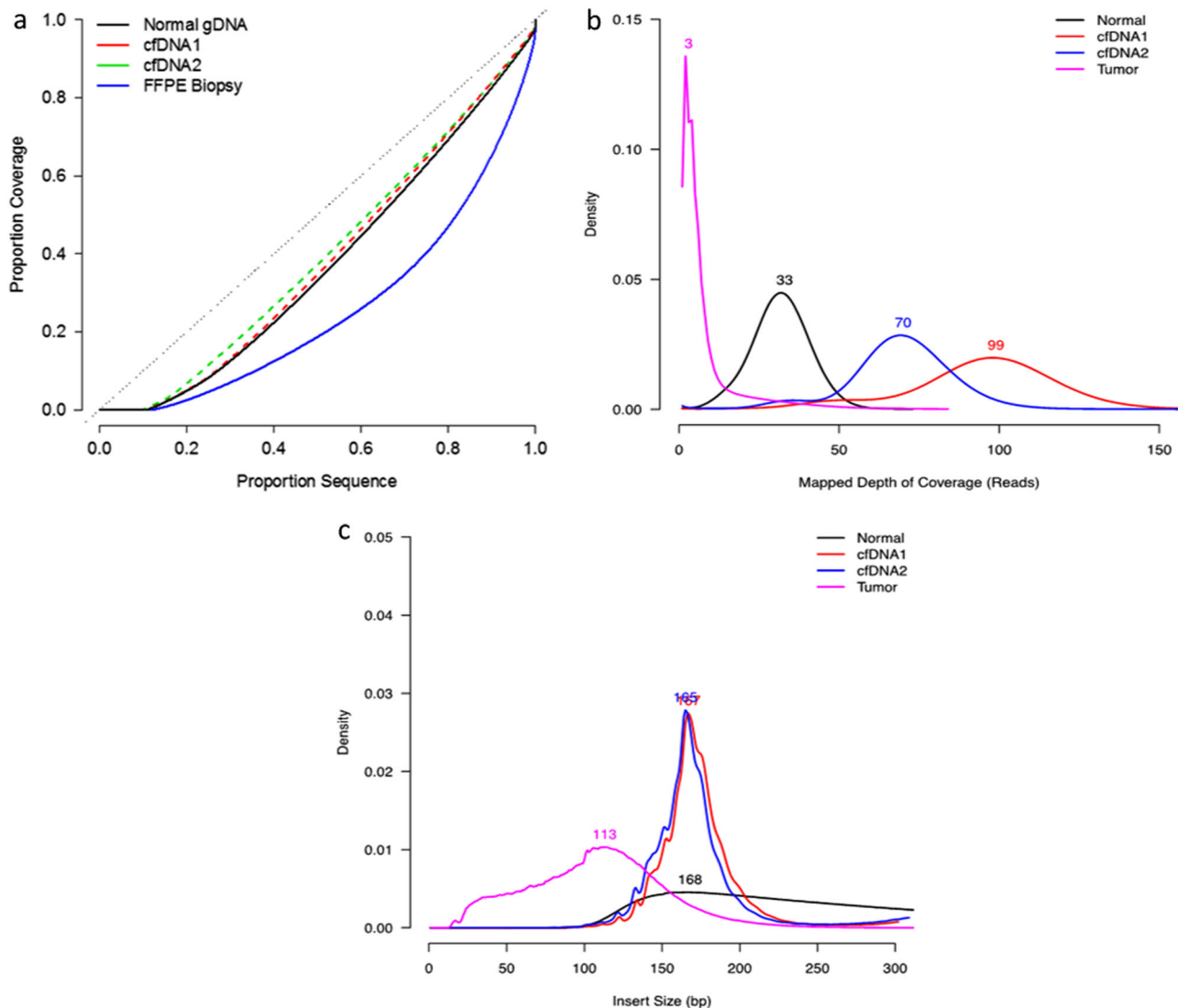


Fig. 2 Coverage uniformity of WGS libraries. **a** Represented is the cumulative proportion of sequencing coverage per cumulative proportion of sequence in whole genome sequencing across normal germline DNA (gDNA) from peripheral blood mononuclear cells, two cfDNA time points, and an archival FFPE tissue biopsy from a metastatic melanoma patient. If coverage was perfectly uniform across the genome coverage the relationship would be linear with gradient one. **b** Mapped depth of coverage distribution for WGS sequencing runs. The range of the coverage distribution is truncated at 150. Each trace is annotated with the mode of the distribution. **c** Insert size distribution of sequencing reads for the sequencing runs. The distribution is truncated at 300 base pairs. Each trace is annotated with the mode of the distribution

WGS of cfDNA allowed us to identify an activating mutation in proto-oncogene B-Raf (BRAF) (V600R) at 6% VAF, concordant with a VAF of 5% using TSP where the coverage was 131 \times and 7946 \times , respectively. No other exonic variants were detected that have previously been implicated in melanoma, although high-level mutations were detected in tumour suppressor genes (TRPS1 and CSMD1), pinpointing to their role as potential disease drivers in this patient. Importantly, using cfDNA WGS we were able to reveal non-coding mutations, including a pathogenic mutation in the TERT promoter that has been reported previously in melanoma.^{30, 31}

Time point 10

All of the exonic SNVs detected by WGS and verified using TSP1 in time point 1 were detected in time point 10, although in many instances there was considerable variation in the VAFs between the pre-treatment and post-treatment samples. Furthermore, all seven SNVs that were not detectable in cfDNA at time point 1 but were found in the biopsy were detectable in the cfDNA at time point 10.

Using WGS we also identified 121 exonic SNVs in the post-treatment sample that had not been detected in pre-treatment time point 1, with VAFs ranging between 5 and 23%. To investigate the properties of these variants, a second TSP panel was generated that targeted the 121 candidates. Ninety-nine percent (120) of the targeted candidate variants were confirmed in time point 10 cfDNA using the orthogonal TSP method and 93% (111) were also present in the tumour biopsy (Supplementary Table 3). One SNV (i.e., 1% of all validated cfDNA variants) in RP11-110A12.2.1 was found in the biopsy but not plasma time point 10, although it was detectable using TSP2 in earlier plasma time points. Ten SNVs (8%) were verified in the relapse cfDNA but were not present in the biopsy tumour sample (Supplementary Tables 3 and 4). Importantly, of these ten variants, five were detected in pre-treatment WGS from cfDNA. One of these variants occurred in INO80, a gene that is required for super-enhancer mediated oncogenic transcription. Furthermore it has been implicated in tumour growth in melanoma, where its protein expression is elevated.³²

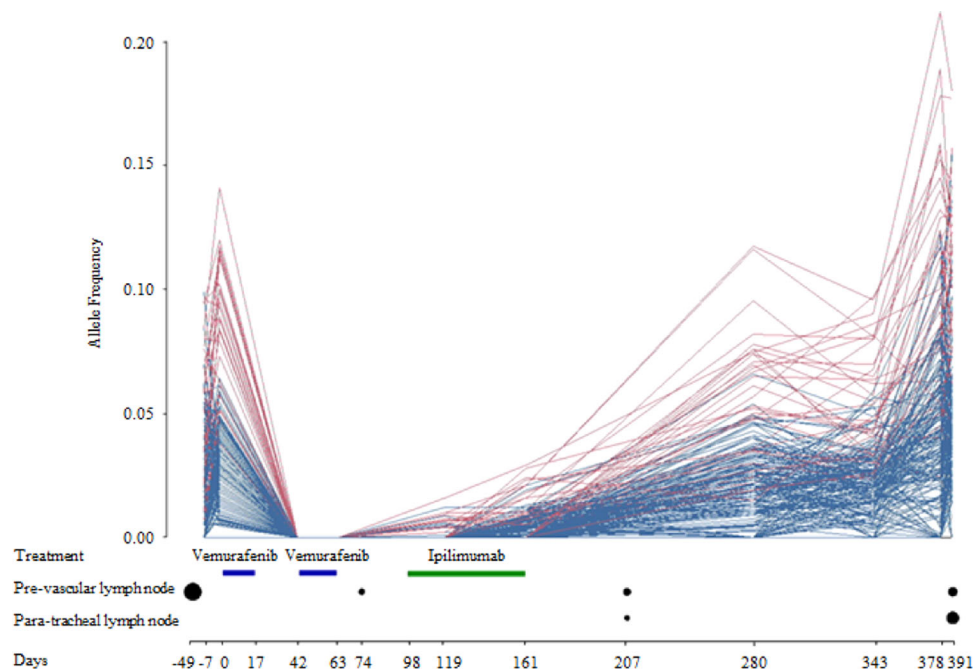


Fig. 3 Mutational profile of variants detected by targeted sequencing. VAFs of candidate mutations detected by WGS were assessed at 10 time points starting 7 days before treatment (of these, two time points were obtained before treatment and the remaining eight during therapy). This data was analysed in tandem with CT scan imaging of the patient taken over the same time period. Estimated volumes from CT imaging are represented as circles scaled by relative proportion in size. Two different major subclones defined by hierarchical agglomerative clustering are highlighted in *blue* and *red*

The pre-treatment time point 1 sample contained 96 of the 120 validated time point 10 SNVs. The majority of these were detectable at low levels, with VAFs typically below 5% and hence were likely below the limit of detection by WGS in time point 1. This suggests that sub-clones carrying these mutations were selected by the treatment. Aside from the exonic SNVs, the mutation in the promoter region of *TERT* was detectable at similar levels in time points 1 and 10 by WGS. Interestingly, due to the highly repetitive DNA sequence in this region and the short amplicons required in cfDNA panel design, it was not possible to design primers to verify this variant by targeted sequencing.

Correlation between VAF and imaging

In order to track quantitative changes in the mutation profile identified previously by WGS over time, we next used both the TSPs and compared this data with CT scans that were taken to monitor tumour size pre- and post-treatment on a three monthly basis (Fig. 3). In the two cfDNA samples sequenced prior to systemic treatment onset, the *BRAF* V600R variant, which is the only known driver mutation in melanoma detected in this patient, was consistently detected at similar levels of 5% VAF. Upon commencement with vemurafenib treatment, the *BRAF* V600R variant was no longer detectable by amplicon sequencing. CT imaging showed a decrease in tumour size in the same time period. Vemurafenib treatment was halted before disease progression because of toxicity and alternative therapy with ipilimumab was initiated. CT scans following completion of ipilimumab treatment showed an increase in tumour burden. This was reflected in *BRAF* V600R variant levels that increased from undetectable during vemurafenib treatment to 8% VAF at time point 10. A number of the exonic SNVs measured followed the *BRAF* V600R trajectory. Importantly, re-emergence of mutations (day 119) occurred nearly 3 months prior to clinical identification by imaging (day 207). Additionally hierarchical agglomerative clustering was used to group the mutations across all samples based on VAF values, with two major groups being

identified in the data (Fig. 3). Separately the similarity between cfDNA and tumour samples based on their VAFs was evaluated (Supplementary Figure 1), with the latter having much higher VAFs than the plasma samples either taken pre-treatment / post relapse or during the period the patient responded to treatment.

Mutational signature of melanoma

Finally, we evaluated the particular spectrum of mutations in this patient before and after treatment as large-scale analyses of cancer genomes have shown that distinct combinations of mutational signatures operate across cancer types.³³ In WGS time point 1 and time point 10 a total of 24,377 and 35,036 variants, respectively, were identified, which is consistent with the high somatic mutation burden previously described in melanoma.³⁴ Both WGS cfDNA samples also exhibit the classical mutation signature of C>T transition mutations based on the sequence context and notably, these mutations occur more frequently on the untranscribed strand than the transcribed strand.³³ This is due to the formation of pyrimidine dimers that arise due to ultraviolet light exposure and is referred to as ultraviolet B -signature (Fig. 4). Furthermore, though VAFs correlate with increasing tumour burden across all sequence contexts, we observe a significant ($p < 0.05$, BH corrected FET) decrease in the number of A-T>C-G mutations compatible with sub-clonal pruning during treatment.

DISCUSSION

The genomic classification of tumours and the resulting opportunity for molecularly informed treatment is driving much innovation in oncology. The utility of genomics will be determined by understanding the relationships between acquired genetic alterations, the type of treatment administered, and the dynamics of mutations during and after treatment. For solid tumours, progress is limited by the difficulty of collecting (a) repeat invasive biopsies during treatment and (b) obtaining high quality tumour material. CfDNA sequencing on the other hand enables

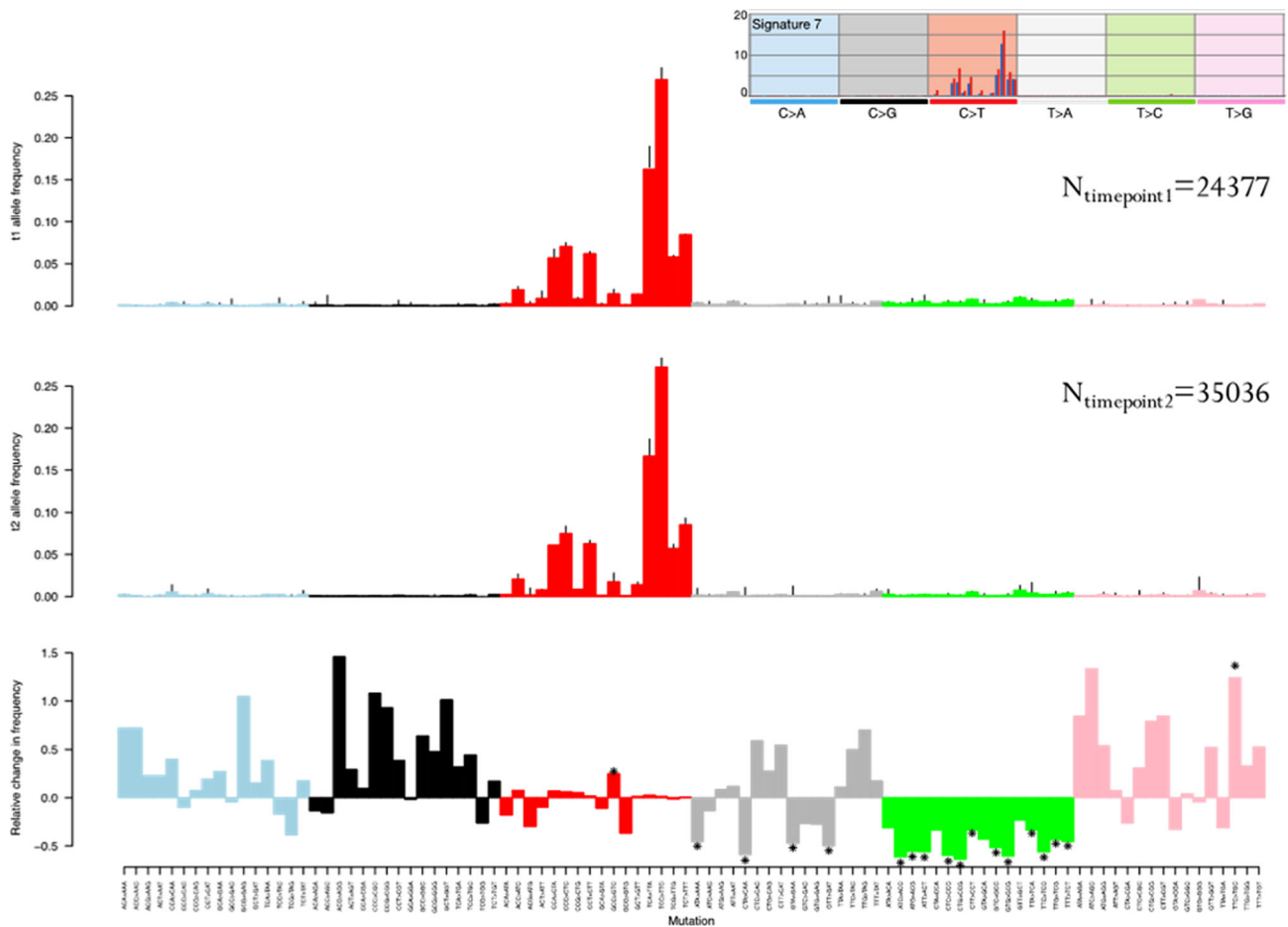


Fig. 4 Melanoma mutation signature from cfDNA WGS. The vertical axis in panels 1 and 2 indicate the percentage of each mutation type in WGS time point 1 and WGS time point 10, respectively, with the type of mutations shown on the horizontal axis. The third panel depicts the relative changes in mutation type between the two time points

standardised, minimally invasive serial monitoring of tumour variants without the need for fresh tumour biopsies.

In this study, we analysed the cfDNA genome using minimal input DNA in a PCR-free approach, thus demonstrating the feasibility of WGS on low amounts of input DNA without the requirement for PCR amplification. This compares favourably to current approaches as it avoids preferential GC bias, which can greatly affect outcome data³⁵ and gives greater confidence in accurate variant calling and quantification of variant allele frequencies. Moreover, detection of the TERT variant, which is one of the most common somatic variants in melanoma patients^{30, 31, 36} highlights one other key advantage of our PCR-free method compared to amplicon based approaches. Due to the highly fragmented nature of cfDNA, a prerequisite for primer design is that the amplified regions must be short, ideally ≤ 160 bp. This can make primer design extremely challenging especially for repetitive regions. Using targeted sequencing, it was not possible to design primers that specifically amplified the desired region of interest in the TERT promoter and this mutation would have been missed.

However, using a PCR-free approach also has potential disadvantages. Despite our optimisation of library preparation and bio-informatics pipelines, the DNA input requirements were still high at 50 ng. This means that our method on average requires at least 20–50 ml of peripheral blood.

An alternative to reduce the blood volume required would be to use reduced-error PCR by uniquely tagging individual cfDNA molecules. However not all PCR errors are removed via tagging,

particularly if they are introduced in the first round of PCR. Additionally as cfDNA is at such a low concentration, greatly reducing the input may increase sequencing biases.

A further consideration is that the majority of ctDNA is derived from cells undergoing apoptosis but not all malignant melanocytes will experience this process, meaning ctDNA analysis may not give complete representation of the tumour. To our knowledge though, no studies have been published detailing the average lifespan of malignant melanocytes. However previous studies have shown changes in ctDNA levels in melanoma patients within 3 weeks of treatment commencing.³⁷

The approach presented here may have utility in advanced cancers where median tumour DNA burden is high at 5–10 %^{17, 19, 38} and patients are often too frail to undergo tissue biopsy. This indication currently represents the major focus for drug development and drug approval from pharmaceutical industry and regulators, respectively. Compared to targeted deep sequencing, our approach is less sensitive. Increasing the depth of sequencing is technically feasible and would overcome this issue, but is currently still prohibitive due to cost. To detect mutation frequencies f below 1%, the depth of coverage X would need to be scaled as $X = \kappa / ((\alpha \times f))$ where α is the library efficiency and κ is the read count scaling efficiency which is a function of PCR duplication rate and variation in coverage.

There is increasing evidence that some cancers exhibit intratumoral and intermetastatic heterogeneity that arises due to clonal and subclonal evolution. Tumour mutation status can differ between metastatic sites in the same patient.¹ Several

studies have highlighted how NGS technologies can be used to identify and track large numbers of somatic aberrations in parallel, thus revealing vital information regarding tumour evolution.^{15, 24}

Here, we show that using WGS of cfDNA before therapy initiation we can confidently identify a subset of acquired mutations that were not present in the primary tumour biopsy using a deep sequencing approach, but that subsequently contributed to the relapse clone. The majority of putative mutations identified at relapse in WGS time point 2 that were not initially identified in WGS time point 1 were present in both biopsy and cfDNA at low level prior to treatment. We believe it is highly unlikely that the emergence of variants in WGS time point 2 that were not detectable in WGS time point 1 can be attributed to experimental variations, as the different samples were processed using exactly the same protocol. Also factors such as input cfDNA were kept the same to minimise any bias and post-sequencing metrics including depth of coverage were comparable. Separately, three relapse variants were not detectable in the biopsy but were present in WGS of pre-treatment cfDNA and verified by the orthogonal TSP approach, highlighting that the analysis of pre-treatment cfDNA by WGS reveals potential drivers of relapse early on in the disease course that would be missed by single biopsy sequencing due to spatial tumour heterogeneity. Furthermore, our data supports clonal selection as the major underlying mechanism for clonal evolution in this case and suggests that, contrary to chemotherapy, BRAF or checkpoint inhibitors are not mutagenic.

Importantly, thanks to the large number of somatic mutations detected by our genome-wide approach, we are able to detect the UV mutation signature in cfDNA. Future studies will be required to evaluate the clinical utility of measuring and tracking this and other signatures such as the smoking signature of lung cancer tumours in plasma samples from high-risk patients.³⁹

Finally, using the cfDNA WGS results to inform the design of a targeted deep sequencing panel allowed us to track mutations while the patient was going through treatment with targeted therapy and immunotherapy. Interestingly, between days 280 and 343 the allele frequencies of one of the major subclones generally decreases, whilst this effect was not observed in the other major subclone. We believe it is unlikely that the decrease in VAFs is due to a systematic error as all samples for targeted sequencing were processed and analysed in exactly the same manner. Therefore it is feasible that the growth rate of the subclones differed, with one slightly shrinking over this time period. Tracking mutations allowed us to detect potential loss of therapy response nearly 3 months prior to a CT scan demonstrated disease progression on completion of ipilimumab treatment. A repeat biopsy was not taken at this time point to compare with the cfDNA analysis as treatment would not have been altered according to current clinical guidelines and it would have been difficult to obtain due to the distribution of metastatic disease. Although the clinical implications of early detection of molecular relapse must be further evaluated in clinical trials, we can speculate that ipilimumab might have been stopped earlier as cfDNA showed an increase in the frequency of plasma mutations. The cost savings associated with early discontinuation of therapy far outweigh the cost of WGS sequencing in this situation. Conversely, if the patient had responded well to therapy, use of plasma DNA monitoring may have abrogated the need for regular CT scans with associated radiation exposure. Ipilimumab monotherapy has now mostly been replaced with PD-1 inhibitor treatment, either in combination with ipilimumab (if patients are fit enough), or alone (for less fit patients).⁴⁰ The role plasma DNA monitoring may play in this setting again requires further evaluation, although with potential for greater utility as currently PD-1 inhibitor treatment generally continues every 2 or 3 weeks until radiological disease progression, as assessed by 3 monthly CT scans.

In conclusion, we show that WGS can be carried out on cfDNA using a PCR free protocol with low input DNA and that

bioinformatics tools can be developed to accurately detect variants with high sensitivity. We reliably detect clinically relevant coding and non-coding mutations and were able to identify UV exposure as the disease causing mechanism from plasma. Combining our approach with targeted deep sequencing enabled us to describe the effects of targeted treatment with vemurafenib and ipilimumab on mutations detected by WGS from cfDNA. We establish that the re-emergence of mutations correlates with clinical disease progression. Finally, we propose that, provided costs of sequencing reduces, our approach could be used to systematically evaluate the clinical utility of cfDNA for guiding clinical decision making in clinical trials.

METHODS

Ethical approval

The National Health Service (NHS) Health Research Authority South Central—Oxford C Research Ethics Committee approved this study and all research was performed in accordance with relevant regulations and guidelines. Written informed consent was obtained from the patient according to current Oxford Radcliffe Biobank (ORB) guidelines (Oxford C Research Ethics Committee No: 09/H0606/5+5).

Study design

This case study was of an 84-year-old male patient with malignant melanoma. The patient's clinical treatment had five phases: detection of primary tumour, detection of metastases, molecularly targeted treatment, immunotherapy, and progression.

Detection of primary tumour. In October 2009 the patient presented with a 2.75 mm Breslow thickness non-ulcerating melanoma on his left ear. The melanoma was excised and preserved using formalin-fixation and paraffin-embedding (SOM).

Detection of metastases. In December 2012 lung metastases were detected, which were initially treated with radiofrequency ablation. Systemic treatment was held off while the patient was asymptomatic. Four months after ablation, further mediastinal lymphadenopathy plus pleural and splenic metastases were detected by CT scan.

Molecularly targeted treatment. Following detection of further metastases the patient's archival primary tumour biopsy tissue was screened for targetable mutations using a limited in-house amplicon sequencing panel targeting known cancer driver mutations. Amplicon sequencing detected a rare (c. 5%⁴¹) BRAF V600R mutation in the melanoma, predicted to be sensitive to the BRAF inhibitor vemurafenib.⁴² In May 2013 the patient started treatment with vemurafenib 960 mg twice daily continuous dosing, with radiological response. 17 days after initiation, the patient was admitted with vomiting and worsening renal function. Treatment was suspended for 25 days and then recommenced at 480 mg twice daily 2 weeks on, 2 weeks off. Due to recurrent renal impairment vemurafenib treatment was discontinued permanently on 8th July 2013. Despite intermittent treatment, imaging indicated partial response with reduction in a prevascular lymph node from 43 to 16 mm.

Immunotherapy. After a break of 39 days, second-line treatment with the CTLA-4 inhibitor ipilimumab was commenced on 15 August 2013, with completion of four treatment cycles by 17th October 2013.

Progression. Radiological imaging indicated disease burden increase 4 weeks after ipilimumab treatment was completed. Further disease progression was confirmed on CT scans in December 2013 and June 2014, with the patient succumbing to disease in the Autumn of 2014.

Clinical sample collection

Between May 2013 and June 2014 serial blood samples were collected in ethylenediamine tetraacetic acid tubes before, during and after treatment with vemurafenib and ipilimumab. 20–50 ml were collected across 10 time points and processed within 4 h of blood draw. Blood was centrifuged for 10 min at 3000 rpm to separate the plasma. The plasma was subsequently transferred to fresh microfuge tubes and centrifuged for 10 min at 7000 r.p.m. The supernatant was stored in aliquots at –80 °C until DNA extraction.

DNA extraction

cfDNA was extracted using the QIAamp circulating nucleic acid kit (Qiagen), according to the manufacturer's instructions. Input plasma volumes ranged from 5 ml to 20 ml per time point, in 5 ml batches. DNA was first eluted into 30 µl of buffer AVE and subsequently a second elution of 30 µl to maximise yield. Samples eluted into separate tubes from the same time point were pooled and stored at −20 °C. Germline DNA (gDNA) was extracted from the normal peripheral blood leucocytes using the QIAamp DNA mini kit (Qiagen), in accordance with the manufacturer's instructions. Matched tumour DNA (tDNA) was extracted from FFPE archival tumour material. Tumour area was marked out on ten 5 µm thick unstained sections and macro-dissection carried out to enrich for tumour material. TDNA was extracted using the QIAamp FFPE DNA extraction kit (Qiagen) and eluted into 20 µl buffer EB. All cfDNA and tDNA samples were quantified on a Qubit fluorometer (Invitrogen) using the high sensitivity dsDNA kit and gDNA was quantified using the broad range dsDNA kit.

Whole genome sequencing

Libraries from gDNA, cfDNA and tumour specimens were prepared using the TruSeq Nano LT DNA Sample Preparation Kit (Illumina Inc) 350 bp insert protocol according to the manufacturer's instructions, with some exceptions. As cfDNA is already highly fragmented, no further fragmentation of the samples was carried out. For each sample 50 ng of input DNA was used and the samples did not undergo any PCR amplification steps. Additionally the clean-up steps to remove small and large fragments were omitted. Final libraries were eluted in a volume of 22 µl resuspension buffer. Fragment distribution pre- and post-library preparation was determined by running 1 µl aliquots on a bioanalyzer 2100 (Agilent Technologies). Libraries were diluted 1:500 and quantified in triplicate by qPCR using the Kapa quantification kit (Kapa Biosystems) on an ABI 7500 (Life Technologies). Based on these quantifications, libraries were diluted to 2 nM and the final input concentration was 10 pM. DNA libraries were clustered onto Flow Cells using the c-Bot v3 kit, and sequencing with paired end reads was carried out on a HiSeq 2500 using v3 chemistry in High Throughput mode.

Targeted resequencing

Custom panels were designed to confirm the presence of SNVs identified by WGS in the cfDNA samples. Candidate chromosome positions based on the hg 19 human genome reference sequence were uploaded to Ion Ampliseq™ Designer version 3.0.1 and the DNA type selected 'FFPE', with all target amplicons being ≤175 bp. Sequencing was carried out on the Ion Torrent Personal Genome Machine (PGM, Life Technologies). Briefly, a multiplex PCR was performed using the Ion Ampliseq 2.0 kit, where 5 ng input DNA was used and amplified using 22 PCR cycles. After PCR, primer sequences were partially digested with FuPa and samples were indexed using Ion Xpress adaptors according to the manufacturer's instructions (Life Technologies). Libraries were purified using the Axyprep Mag PCR clean up kit (Axygen Biosciences) and then diluted 1:200 prior to quantification by qPCR on an ABI 7500 using the Ion Quantitation Kit (Life Technologies). Libraries were diluted to a 20 pM concentration and were pooled. Template positive ISPs were generated containing clonally amplified ISPs using the OneTouch 2 200 template kit v2 (Life Technologies). Samples were enriched on the Ion OneTouch ES and subsequently sequenced using 316 or 318 chips on the PGM using the Ion Sequencing 200 bp chemistry (Life Technologies).

Bioinformatics analysis

From the WGS data tumour-specific variants were identified through a combination of alignment-based and de novo assembly-based methods. Candidate variant positions in the cfDNA alignment were defined as positions with at least three high-quality non-genomic alleles and positions with unexpectedly high or low total coverage were ignored. For each candidate position, a range of frequentist and Bayesian statistics were computed, measuring the strength of evidence for the presence of non-genomic alleles. All statistics incorporate empirical confusion matrices for all binary alignment map (BAM) files involved. For each candidate position, a variant background haplotype was heuristically constructed, and the presence of this haplotype was verified in the cfDNA BAM and its absence in the genomic BAM using kmer-based statistics.

Sequence data from the PGM was analysed using the combined output of Strelka v1.0.15 and MuTect2. Target regions were defined using the

relevant bed files for the two custom panels. The positions of all candidate somatic variants were investigated in Integrated Genomics Viewer using BAM files aligned to the hg19 human genome reference sequence. For further information on bioinformatics analysis: see [Supplementary Methods](#).

Code availability

The code is currently not readily available as the authors are working on a separate manuscript for the software package.

Data availability

Sequence data has been deposited at the Sequence Read Archive (SRA) database (<https://www.ncbi.nlm.nih.gov/sra/>), under accession number SRP111727.

ACKNOWLEDGEMENTS

We thank the patients for their consent for samples to be used in this study and the Oxford Wellcome Trust Centre for Human Genetics for their assistance with sequencing. This research was funded and supported by the Wellcome Trust and Department of Health as part of the Health Innovation Challenge Fund scheme and also by the NIHR Biomedical Research Centre, Oxford. The views expressed are those of the authors and not necessarily those of the Department of Health or Wellcome Trust. The work was supported also by funding from the Wellcome Trust Core Award Grant Number 090532/Z/09/Z.

AUTHOR CONTRIBUTIONS

A.C., O.V., A.D., A.G., M.M., J.T. and A.S., were responsible for study design; A.C., O.V., A.D., A.G., D.V., H.D., M.M. and A.S. generated and analysed the data; A.C., O.V., A.D., A.G., M.M., G.M., D.V., J.T. and A.S., contributed to manuscript preparation; A.C., O.V. and A.D., contributed equally and are co-authors of this work.

ADDITIONAL INFORMATION

Supplementary Information accompanies the paper on the *npj Genomic Medicine* website (doi:[10.1038/s41525-017-0030-7](https://doi.org/10.1038/s41525-017-0030-7)).

Competing interests: A.S., A.D., O.V. are shareholders of Lighthouse and GRAIL. O.V. is an employee of GRAIL.

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

REFERENCES

- Gerlinger, M. et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
- Aparicio, S. & Caldas, C. The implications of clonal genome evolution for cancer medicine. *N. Engl. J. Med.* **368**, 842–851 (2013).
- Rybicki, B. & Yun, K. Addressing intra-tumoral heterogeneity and therapy resistance. *Oncotarget* doi:[10.18632/oncotarget.11875](https://doi.org/10.18632/oncotarget.11875) (2016).
- Ahmed, A. A. & Abedalthagafi, M. Cancer diagnostics: the journey from histomorphology to molecular profiling. *Oncotarget* doi:[10.18632/oncotarget.11061](https://doi.org/10.18632/oncotarget.11061) (2016).
- Luke, J. J., Flaherty K. T., Ribas, A. & Long, G. V. Targeted agents and immunotherapies: optimizing outcomes in melanoma. *Nat. Rev. Clin. Oncol.* doi:[10.1038/nrclinonc.2017.43](https://doi.org/10.1038/nrclinonc.2017.43) (2017).
- Carrick, D. M. et al. Robustness of Next Generation Sequencing on Older Formalin-Fixed Paraffin-Embedded Tissue. *PLoS ONE* **10**, e0127353 (2015).
- Oh, E. et al. Comparison of accuracy of whole-exome sequencing with formalin-fixed paraffin-embedded and fresh frozen tissue samples. *PLoS ONE* **10**, e0144162 (2015).
- Horak, P., Frohling, S. & Glimm, H. Integrating next-generation sequencing into clinical oncology: strategies, promises and pitfalls. *ESMO Open* **1**, e000094 (2017).
- Jahr, S. et al. DNA fragments in the blood plasma of cancer patients: quantifications and evidence for their origin from apoptotic and necrotic cells. *Cancer Res.* **61**, 1659–1665 (2001).
- Schwarzenbach, H., Stoecklacher, J., Pantel, K. & Goekkurt, E. Detection and monitoring of cell-free DNA in blood of patients with colorectal cancer. *Ann. N.Y. Acad. Sci.* **1137**, 190–196 (2008).
- Underhill, H. R. et al. Fragment length of circulating tumor DNA. *PLoS Genet.* **12**, e1006162 (2016).

12. Mouliere, F. et al. High fragmentation characterizes tumour-derived circulating DNA. *PLoS ONE* **6**, e23418 (2011).
13. Jiang, P. et al. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc. Natl Acad. Sci. USA* **112**, E1317–E1325 (2015).
14. Bettegowda, C. et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci. Transl. Med.* **6**, 224ra224 (2014).
15. Newman, A. M. et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.* **20**, 548–554 (2014).
16. Tsao, S. C. et al. Monitoring response to therapy in melanoma by quantifying circulating tumour DNA with droplet digital PCR for BRAF and NRAS mutations. *Sci. Rep.* **5**, 11198 (2015).
17. Forshew, T. et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci. Transl. Med.* **4**, 136ra168 (2012).
18. Kinde, I., Wu, J., Papadopoulos, N., Kinzler, K. W. & Vogelstein, B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc. Natl Acad. Sci. USA* **108**, 9530–9535 (2011).
19. Murtaza, M. et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* **497**, 108–112 (2013).
20. Murtaza, M. et al. Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. *Nat. Commun.* **6**, 8760 (2015).
21. Ulz, P. et al. Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat. Genet.* doi:10.1038/ng.3648 (2016).
22. Gremel, G. et al. Distinct subclonal tumour responses to therapy revealed by circulating cell-free DNA. *Ann. Oncol.* doi:10.1093/annonc/mdw278 (2016).
23. Chan, K. C. et al. Cancer genome scanning in plasma: detection of tumor-associated copy number aberrations, single-nucleotide variants, and tumoral heterogeneity by massively parallel sequencing. *Clin. Chem.* **59**, 211–224 (2013).
24. Leary, R. J. et al. Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. *Sci. Transl. Med.* **4**, 162ra154 (2012).
25. Heitzer, E. et al. Tumor-associated copy number changes in the circulation of patients with prostate cancer identified through whole-genome sequencing. *Genome Med.* **5**, 30 (2013).
26. Belic, J. et al. Rapid identification of plasma DNA samples with increased ctDNA levels by a modified FAST-SeqS approach. *Clin. Chem.* **61**, 838–849 (2015).
27. Chan, K. C. et al. Noninvasive detection of cancer-associated genome-wide hypomethylation and copy number aberrations by plasma DNA bisulfite sequencing. *Proc. Natl Acad. Sci. USA* **110**, 18761–18768 (2013).
28. Williams, C. et al. A high frequency of sequence alterations is due to formalin fixation of archival specimens. *Am. J. Pathol.* **155**, 1467–1471 (1999).
29. Chang, H. T. et al. Enhancing mucosal immunity in mice by recombinant adenovirus expressing major epitopes of porcine circovirus-2 capsid protein delivered with cytosine-phosphate-guanosine oligodeoxynucleotides. *J. Vet. Sci.* **15**, 399–407 (2014).
30. Remke, M. et al. TERT promoter mutations are highly recurrent in SHH subgroup medulloblastoma. *Acta Neuropathol.* **126**, 917–929 (2013).
31. Horn, S. et al. TERT promoter mutations in familial and sporadic melanoma. *Science* **339**, 959–961 (2013).
32. Zhou, B. et al. INO80 governs superenhancer-mediated oncogenic transcription and tumor growth in melanoma. *Genes Dev.* **30**, 1440–1453 (2016).
33. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
34. Lawrence, M. S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
35. Benjamini, Y. & Speed, T. P. Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res.* **40**, e72 (2012).
36. Griewank, K. G. et al. TERT promoter mutation status as an independent prognostic factor in cutaneous melanoma. *J. Natl Cancer Inst.* **106**, doi:10.1093/jnci/dju246 (2014).
37. Girotti, M. R. et al. Application of sequencing, liquid biopsies, and patient-derived xenografts for personalized medicine in Melanoma. *Cancer Discov.* **6**, 286–299 (2016).
38. Diehl, F. et al. Detection and quantification of mutations in the plasma of patients with colorectal tumors. *Proc. Natl Acad. Sci. USA* **102**, 16368–16373 (2005).
39. Jamal-Hanjani, M. et al. Tracking the evolution of non-small-cell lung cancer. *N. Engl. J. Med.*, doi:10.1056/NEJMoa1616288 (2017).
40. Marquez-Rodas, I. et al. Immune checkpoint inhibitors: therapeutic advances in melanoma. *Ann. Transl. Med.* **3**, 267 (2015).
41. Bamford, S. et al. The COSMIC (Catalogue of somatic mutations in cancer) database and website. *Br. J. Cancer* **91**, 355–358 (2004).
42. McArthur, G. A. et al. Safety and efficacy of vemurafenib in BRAF(V600E) and BRAF(V600K) mutation-positive melanoma (BRIM-3): extended follow-up of a phase 3, randomised, open-label study. *Lancet Oncol.* **15**, 323–332 (2014).



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017