

Tumor-Specific Mitochondrial DNA Variants Are Rarely Detected in Cell-Free DNA^{1,2}



M.J.A. Weerts^{*}, E.C. Timmermans[†],
A. van de Stolpe[†], R.H.A.M. Vossen[‡], S.Y. Anvar^{‡,§,¶},
J.A. Foekens^{*}, S. Sleijfer^{*} and J.W.M. Martens^{*}

^{*}Department of Medical Oncology and Cancer Genomics Netherlands, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, The Netherlands; [†]Philips Research Laboratories, High Tech Campus 11, Eindhoven, The Netherlands; [‡]Leiden Genome Technology Center (LGTC), Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands; [§]Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands; [¶]Department of Clinical Pharmacy and Toxicology, Leiden University Medical Center, Leiden, The Netherlands

Abstract

The use of blood-circulating cell-free DNA (cfDNA) as a “liquid biopsy” in oncology is being explored for its potential as a cancer biomarker. Mitochondria contain their own circular genomic entity (mitochondrial DNA, mtDNA), up to even thousands of copies per cell. The mutation rate of mtDNA is several orders of magnitude higher than that of the nuclear DNA. Tumor-specific variants have been identified in tumors along the entire mtDNA, and their number varies among and within tumors. The high mtDNA copy number per cell and the high mtDNA mutation rate make it worthwhile to explore the potential of tumor-specific cf-mtDNA variants as cancer marker in the blood of cancer patients. We used single-molecule real-time (SMRT) sequencing to profile the entire mtDNA of 19 tissue specimens (primary tumor and/or metastatic sites, and tumor-adjacent normal tissue) and 9 cfDNA samples, originating from 8 cancer patients (5 breast, 3 colon). For each patient, tumor-specific mtDNA variants were detected and traced in cfDNA by SMRT sequencing and/or digital PCR to explore their feasibility as cancer biomarker. As a reference, we measured other blood-circulating biomarkers for these patients, including driver mutations in nuclear-encoded cfDNA and cancer-antigen levels or circulating tumor cells. Four of the 24 (17%) tumor-specific mtDNA variants were detected in cfDNA, however at much lower allele frequencies compared to mutations in nuclear-encoded driver genes in the same samples. Also, extensive heterogeneity was observed among the heteroplasmic mtDNA variants present in an individual. We conclude that there is limited value in tracing tumor-specific mtDNA variants in blood-circulating cfDNA with the current methods available.

Neoplasia (2018) 20, 687–696

Introduction

Mitochondria are organelles within our cells responsible for a variety of functions, including energy production and initiating apoptosis. Their small circular genome (mitochondrial DNA, mtDNA) encodes for proteins essential in the oxidative phosphorylation system and the tRNA and rRNA molecules of the mitochondrial translation apparatus. Within a single cell, multiple copies of mtDNA exist (mtDNA content), but due to its small size, the mtDNA represents only a minor fraction of the total cellular DNA (<0.1%). In general, cells with high energy demand (e.g., muscle cells) have a higher mtDNA content than cells with lower energy demand (e.g., blood

Abbreviations: mtDNA, mitochondrial DNA; nDNA, nuclear DNA; cfDNA, cell-free DNA; SMRT, single-molecule real-time sequencing; dPCR, digital PCR

Address all correspondence to: M.J.A. Weerts, Department of Medical Oncology and Cancer Genomics Netherlands, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, The Netherlands. E-mail: m.weerts@erasmusmc.nl

¹Competing Interests: The authors declare no conflicts of interest.

²Funding: This work was supported by a grant from Philips Research (Eindhoven, The Netherlands).

Received 6 February 2018; Revised 7 May 2018; Accepted 8 May 2018

© 2018 . Published by Elsevier Inc. on behalf of Neoplasia Press, Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). 1476-5586

<https://doi.org/10.1016/j.neo.2018.05.003>

cells) [1]. In human cancer, changes in mtDNA content have been reported when tumor specimens are compared to their normal counterparts [2]. The polyploid nature of mtDNA invokes the concept of only a single (homoplasmy) or two or more mitochondrial genotypes (heteroplasmy) within a cell. It has been shown that heteroplasmy patterns within an individual can differ between tissues, even in an allele-specific manner [3–6]. Also within cancer, tumors harbor mtDNA that is genetically different to their normal counterparts, either at a homo- or heteroplasmic level, and their number and position vary among tumors [7–9]. Interestingly, since the mutation rate of mtDNA is several orders of magnitude higher than that of nuclear DNA (nDNA) [10], it is very informative to assess phylogenetic distance not only intra- and interspecies but also interindividual.

Within oncology, the use of blood-circulating cell-free DNA (cfDNA) as a “liquid biopsy” is being explored for its potential as a screening tool, to establish prognosis, or as a marker for response to treatment. The origin of cfDNA is mainly from apoptotic cells, hence its typical fragmentation pattern representing DNA cleavage between nucleosomes or chromatosomes (~146–166 base pairs and multiples thereof) [11]. The physical characteristics of cf-mtDNA have not been studied as extensively as its nuclear counterpart. Since mtDNA is packed into nucleoids [12], which are not fragmented during apoptosis [13], the fragmentation pattern as seen for nDNA does not apply to mtDNA. Indeed, the majority of the cf-mtDNA in human plasma appears associated with particles of at least 0.45 μm in diameter [14], and a fraction of it is severely fragmented down to at least 30 base pairs [15–18]. If not fragmented, the circular nature of mtDNA might render it less susceptible to enzymatic cleavage and thus more stable within the circulation. The total amount of cfDNA is often increased in cancer patients compared to healthy individuals, for both DNA from the nucleus as well as mtDNA [16,19–23]. The detection of tumor-specific cfDNA is aided by the aberrations present in the cancer's genome and thus by the detection of tumor-specific mutations within the cfDNA. A few studies have attempted to detect mtDNA variants in blood-derived cfDNA [4,24–29] or other bodily fluids [24,30–33]. However, in the studies on blood-derived cf-mtDNA, used methods were either not very sensitive (i.e., conventional Sanger sequencing), or the variants were not truly tumor-derived (i.e., already present in matched normal specimens). Also, quantitative variant allele frequencies were not reported in all assessed samples, making interpretation of these results difficult. Nevertheless, the combination of a high copy number per cell, a high mutation rate, and potentially high stability within the circulation makes it worthwhile to explore the potential of tumor-specific variants in cf-mtDNA as a cancer biomarker.

In this study, we used a targeted single-molecule real-time (SMRT) sequencing approach to profile the entire mtDNA of the primary tumor and/or metastatic sites, tumor-adjacent normal tissue, and cfDNA of eight cancer patients. We have recently shown that the SMRT sequencing approach is able to reliably detect unknown variants $\geq 1.0\%$ allele frequency and to trace known low-frequent variants down to at least 0.1% allele frequency [34]. In our cohort, we observed tumor-specific mtDNA variants for each patient and explored the feasibility to trace these tumor-specific variants in cf-mtDNA as a cancer biomarker.

Materials and Methods

Patient Selection and Sampling

We used material from our biobank at the department of Medical Oncology of the Erasmus MC Cancer Institute, Rotterdam, the

Netherlands. Patient selection was based on availability of a frozen blood derivative (plasma or serum) to obtain cfDNA and fresh-frozen resection material of tumor tissue (primary or metastasis). For all except one case, fresh-frozen material of normal tissue originating from the same resection material was available. Blood sampling was done in either serum separation tubes according to routine procedures in our hospital or in EDTA tubes followed by cell separation within 24 hours after blood draw (10 minutes at 800 $\times g$). Obtained serum or plasma samples had been stored at -80°C until use. After thawing, plasma samples underwent additional sedimentation at 1020 $\times g$ for 10 minutes at 4°C , of which the supernatant was used. Use of the patient material was approved by the medical ethics committee of the Erasmus MC (MEC 02.953 and MEC 06.089) and conducted in accordance to the Code of Conduct of the Federation of Medical Scientific Societies in the Netherlands.

DNA Extraction

For fresh-frozen tissue specimens, a DNA extraction method that enriches for mtDNA was performed as described before [34]. Briefly, 20 cryosections of 30 μm (average input of 30 mg tissue, range of 16–59 mg) per specimen were lysed to solubilize cellular membrane and release all cellular compartments (10 minutes, 1 ml of 0.5 \times TBE containing 0.5% (v/v) Triton X-100). Cell nuclei were removed (10 minutes 1020 $\times g$), and DNA was extracted from the remaining supernatant using the QIAamp Circulating Nucleic Acid Kit (*Qiagen*) according to the suppliers' protocol. For four specimens, an additional sample was obtained by an independent DNA extraction (as described above) with subsequent enzymatic degradation for linear DNA as described before [34]. Briefly, those DNA extracts (max. 250 ng) were incubated with ATP-dependent exonuclease PlasmidSafe (*Epicenter*) (40 U, 3 hours at 37°C), heat-inactivated (30 minutes 70°C), and purified (ethanol precipitation, 70% ethanol). For some frozen tissue specimens, DNA extracts were already available in our biobank and had been obtained by either the PureLink Genomic DNA kit (*Invitrogen*) or the DNeasy Tissue Kit (*Qiagen*) as described by the supplier. For each tissue sample, 5- μm sections were obtained on microscopy slides and hematoxylin and eosin (HE) stained to estimate the percentage of tumor cells within the sections used for DNA extraction. For the blood derivatives, after thawing at 4°C , DNA was extracted using the QIAamp Circulating Nucleic Acid Kit (*Qiagen*) according to the suppliers' protocol. Serum input ranged from 450 to 500 μl , and plasma supernatant input was 1000 μl . Specifications for each sample are provided in Supplementary Table 1.

DNA Quantification and mtDNA Purity Assessment

All DNA extracts were quantified using the Qubit dsDNA HS assay kit (*Life Technologies*) according to the suppliers' protocol. Purity of mtDNA was measured in duplicate runs of a multiplex qPCR assay targeting a nuclear and a mitochondrial encoded gene to calculate the ratio of mtDNA molecules opposed to nDNA molecules by the relative quantitation method ($2^{-\Delta\Delta\text{Cq}}$) as described before [35].

SMRT Sequencing

SMRT sequencing was performed as described before [34]. Briefly, amplicons covering the complete mtDNA were generated by singleplex (tissue DNA) or multiplex (cfDNA) PCR with initial denaturation for 3 minutes at 98°C , 15 or 18 cycles of a three-step PCR with 10-second denaturation (98°C), 30-second annealing (67°C) and 90-second extension (72°C), followed by a final extension (72°C) for 5 minutes. DNA input was set to contain at least 100,000

Table 1. Patient Characteristics and Number of Detected Heteroplasmic mtDNA Variants for Each Specimen by SMRT Sequencing

Patient	Primary Tumor Type	Sex	Age at Diagnosis	Stage	Clinicopathological	Specimen	Heteroplasmic mtDNA Variants
P1	Breast	Female	75	T4 cN+ → ypN0 Mx	ER: positive PR: positive HER2: not done Primary tumor size: 2 cm	Primary tumor	4
						Normal mammary	3
						Serum 1	5
						Serum 2	not done
						Serum 3	not done
P2	Breast	Female	81	T3 pN2 M0	ER/PR/HER2: not done Primary tumor size: 8 cm	Serum 4	not done
						Serum 5	3
						Primary tumor	7
						Normal mammary	14
						Serum	4
P3	Breast	Female	57	T3 cN+ → ypN0 M0	ER: positive PR: positive HER2: balanced Primary tumor size: 8.5 cm	Primary tumor	3
						Normal mammary	4
						Serum	0
P4	Breast	Female	49	T3 pN1 M0	ER: positive PR: positive HER2: not done Primary tumor size: 5.2 cm	Primary tumor	3
						Normal mammary	4
						Serum	2
P5	Breast	Female	64	ND	ER/PR/HER2: unknown	Primary tumor	1
						Serum	7
P6	Colon	Male	68	T3 pN1 M1	Right hemicolon Dukes: D Differentiation: moderate Primary tumor size: 3.3 cm	Primary tumor	5
						Liver metastasis 1	4
						Liver metastasis 2	7
						Normal colon	6
						Normal liver	7
						Plasma	5
P7	Colon	Female	63	T4 pN1 M1	Sigmoid Dukes: D Differentiation: moderate Primary tumor size: 5.5 cm	Primary tumor	1
						Omental metastasis	4
						Normal colon	0
						Plasma	1
P8	Colon	Male	84	T3 and T3 pN0 and pN1 Mx	Sigmoid and distal colon Dukes: B and C Differentiation: moderate Primary tumor size: 6 and 5 cm	Liver metastasis	3
						Normal liver	5
						Plasma	9

but maximally 50,000,000 copies mtDNA/reaction based on the mtDNA content. Each 50- μ l reaction contained DNA and 1 U of Hot-Start Q5 High Fidelity DNA polymerase (*NEB*) in 1 \times Q5 reaction buffer, 200 μ M dNTPs, and 0.5 μ M of 5'-M13 tailed forward and reverse primer (Supplementary Table 2). Specificity of the generated products was confirmed using microchip electrophoresis (DNA-12000 reagent kit, *Shimadzu*). Amplicons were equimolar pooled per sample and purified using AMPure PB paramagnetic beads (*Pacific Biosciences*) with a 0.6 bead:sample ratio according to the SMRTbell Template Prep Kit protocol and eluted in 10 mM Tris-HCl pH 8.5. The 5'-M13 universal sequence tail of the primers allowed barcoding of each sample by performing five amplification cycles of the three-step PCR as described above but with an annealing temperature of 58°C. Specificity of the generated products was confirmed using microchip electrophoresis (BioAnalyzer High Sensitivity DNA kit, *Agilent*). A final mix of barcoded fragments was obtained by pooling of multiple samples and subsequently purified using AMPure PB paramagnetic beads with a 0.6 bead: sample ratio. Concentration of the final mix was determined using the Qubit dsDNA HS assay kit, and SMRTbell library was generated according to the Amplicon Template Preparation and Sequencing guide (*Pacific Biosciences*). Sequencing was performed on Pacific Biosciences RSII with P6-C4 sequencing chemistry and 360-minute movie-time. For the tissue samples, a total of 15 SMRT cells were used to reach a read depth estimated at 600 \times per sample, and for the cfDNA samples, a total of 28 SMRT cells were used to reach a read depth estimated at 3000 \times per sample. Specifications for each sample are provided in Supplementary Table 1.

Digital PCR

Digital PCR (dPCR) was performed on the Quantstudio 3D digital PCR system (*Thermo Fisher*) according to the supplier's protocol. Detection of KRAS p.G12D, KRAS p.G12 V, TP53 p.R248Q, TP53 p.R273H, and PIK3CA p.H1047R was done by validated Taqman SNP genotyping assays (*Thermo Fisher*). Detection of mtDNA variants 664G > A, 6255G > A, 1924 T > C, and 2305 T > C was done by custom assays (Supplementary Table 2) and carried out with an adaption to the DNA input due to high mtDNA copy number (set to contain at least 1E+ 3 but maximally 2E+ 4 copies mtDNA/reaction based on mtDNA content). Reactions contained DNA in 1 \times dPCR mastermix v2, 0.9 μ M of each primer and 0.2 μ M of each probe. After initial denaturation for 10 minutes at 96°C, the 40-cycle 2-step PCR was performed at 30-second denaturation (98°C) and 120-second annealing/extension (56°C) followed by a final 2-minute extension (56°C). To calculate allele frequency of the alternative variant, the threshold for signal dots was set to at least two dots per dye. For samples where the variant was not detected, the limit of detection was calculated based on the total number of positive wildtype dots and two mutant signal dots (e.g., 5000 wild-type dots would correspond to a detection limit of 2 / 5000 = 0.04%).

Bioinformatics

RS bax.h5 files were converted to Sequel BAM files, of which circular consensus reads (CCS) were generated using the CCS2 algorithm, and attributed to each sample using the sample-specific barcode [36]. Next, a minimum quality threshold of 99% and at least five passes of the SMRTbell were applied to select for highly accurate

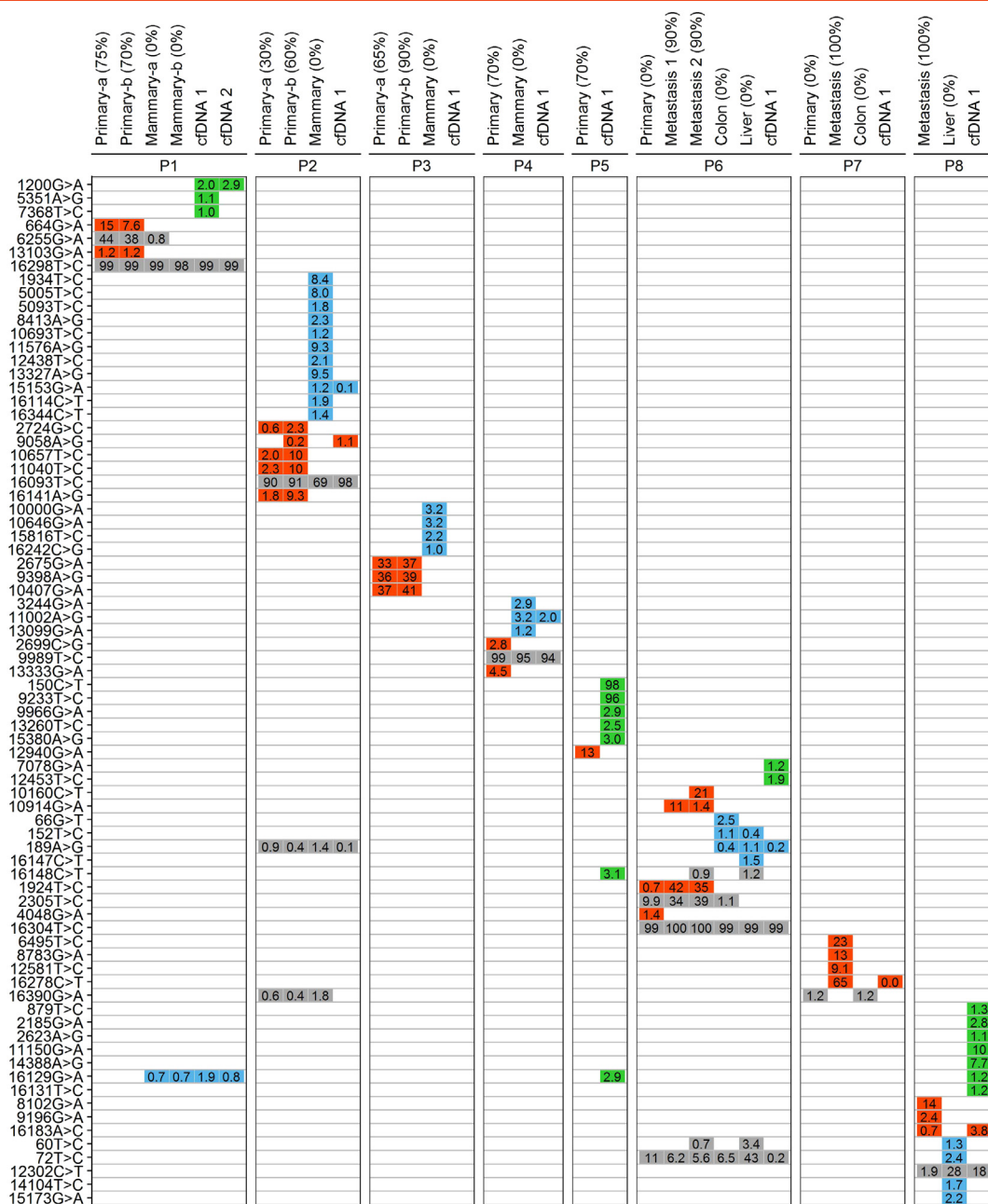


Figure 1. Heteroplasmic mtDNA variants in tumor and normal tissue. Heteroplasmic mtDNA variants (vertical) detected by SMRT sequencing of tumor and normal tissue (horizontal) of eight cancer patients (P1 to P8). Ubiquitous, normal-only, and tumor-only variants in, respectively, gray, blue, and red, cfDNA variants of unknown tissue origin in green. Within the squares, allele frequency (%) of the variant is indicated. The percentage of tumor cells in the analyzed sections based on morphological estimations in HE-stained slides between brackets behind tissues.

single-molecule reads. Selected CCS reads were trimmed (Cutadapt [37] for primers tails) and subsequently aligned against a reference sequence (BWA-MEM parameters -k17 -W40 -r10 -A1 -B1 -O1 -E1 -L0 [38]). As reference sequence, we used an extended version of rCRS to compensate for mapping bias due to circularity of the mitochondrial genome. Positions alternative to the reference sequence in pileup files (Bioconductor Rsamtools 1.26.2 pileup function with pileupParam min_base_quality = 30, min_mapq = 0, min_nucleotide_depth = 5, min_minor_allele_depth = 0, distinguish_strands = TRUE, distinguish_nucleotides = TRUE, ignore_query_Ns = TRUE, include_deletions = FALSE, include_insertions = FALSE) were converted back to

rCRS positions and used for analyses. Allele frequency was calculated based on the alternative variant (alternative reads / total reads). First, all homoplasmic and high heteroplasmic alternative variants (>50% allele frequency) were used for haplotyping (HaploGrep2 v2.1.0) to determine patient specificity for each sample (Supplementary Figure 1). Then, initial variant selection was performed on the pileup files using a threshold of 1% to 99% alternative allele frequency. Those variants were manually inspected in Integrative Genomics Viewer (IGV, Broad Institute) [39] to exclude mapping artifacts (as elaborated on in [34]). Of the remaining variants, their presence within the initial pileup file was determined in all examined samples of that patient to generate a final list

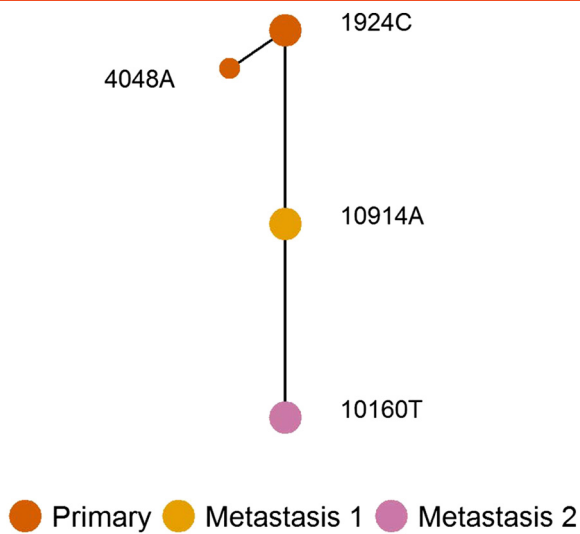


Figure 2. Schematic of phylogenetic relationship between the sequenced colorectal cancer specimens of P6.

of detected variants (Supplementary Table 3). Also, the detected heteroplasmic variants were used in a nucleotide BLAST against the human reference sequence (NCBI's nucleotide web blast, <https://blast.ncbi.nlm.nih.gov>) with the surrounding reference sequence (30 bases 5' and 30 bases 3') to uncover potential NUMT events, but none were recovered. For the samples in which variants were not called within the final list, limit of detection at that position was calculated based on the read depth at that position and an alternative variant read depth of 5 (e.g., a position with 5000 \times read depth would correspond to a detection limit of $5 / 5000 = 0.1\%$).

Results

We sequenced the entire mtDNA of 19 tissue specimens and 9 cfDNA samples originating from 8 cancer patients, including at least 1 tumor and 1 cfDNA sample for each patient. For 4 tissue specimens (primary tumor of P1, P2 and P3, and normal tissue of P1), we performed independent resequencing of another (nearby) section of the specimen. For one patient (P1), a total of five consecutive cfDNA samples were profiled for mtDNA variants by dPCR. The study cohort consisted of two cancer types — breast and colon cancer — at variable disease stages. Patient characteristics are summarized in Table 1. After haplotyping the mtDNA of each specimen for each individual to make sure that samples were matched correctly (Supplementary Figure 1), the heteroplasmic mtDNA variants present in each sample were evaluated.

Heterogeneity in Heteroplasmic mtDNA Variants between Tumor and Normal Tissue

In the tumor and normal tissue specimens, the number of heteroplasmic mtDNA variants ranged from 0 up to 14 per specimen (Table 1), with allele frequencies between 0.2% and 99.4% (Figure 1) (Supplementary Table 3). Good concordance was observed in detected variants between different sections of a specimen (P1, P2, and P3; Figure 1), with 2 of the 18 variants missed due to coverage and thus limit of detection (6255G > A at 0.8% allele frequency in P1 normal mammary tissue and 9058A > G at 0.2% allele frequency in P2 primary tumor tissue Supplementary Table 3). Heterogeneity was observed between the tumor and normal tissues: majority of the variants (80%) were present either only within the tumor tissue

Table 2. Heteroplasmic mtDNA Variants Detected by dPCR in Two Patients

Patient	Tumor Type	Specimen	Variant 664G > A (Allele Frequency)	Variant 6255G > A (Allele Frequency)
P1	Breast	Primary tumor-a	12.9%	46.2%
		Primary tumor-b	5.4%	35.9%
		Normal mammary-a	0.01%	0.9%
		Normal mammary-b	0.05%	0.5%
		Serum 1	nd	nd
		Serum 2*	nd	0.03%
		Serum 3*	0.06%	0.3%
		Serum 4*	nd	nd
Serum 5	nd	nd		

Patient	Tumor Type	Specimen	Variant 1924 T > C (Allele Frequency)	Variant 2305 T > C (Allele Frequency)
P6	Colon	Primary tumor	0.9%	10.4%
		Liver metastasis 1	36.2%	40.1%
		Liver metastasis 2	46.2%	34.2%
		Normal colon	0.3%	1.1%
		Normal liver	0.08%	0.05%
		Plasma	nd	nd

Asterisks indicate the samples that had not been analyzed by SMRT sequencing. nd, not detected.

($n = 24$) or only within the normal tissue ($n = 27$) (Figure 1). Generally, tumor-only variants had higher allele frequency than normal-only variants [respectively, median (interquartile range) of 9.7% (23.7%) versus 1.9% (1.6%), Mann–Whitney U test $P < .001$]. Also, two of the ubiquitous variants ($n = 13$) showed heteroplasmic expansion between the normal and the tumor tissue (variants 6255G > A and 2305 T > C in, respectively, P1 and P6; Figure 1). Note that some of the ubiquitous variants at low heteroplasmy might not be present in the tumor cells but in normal cells residing at the tumor tissue (i.e., variants 189A > G and 16390G > A in P2, 60 T > C and 66G > T in P6, and 12302C > T in P8). A phylogenetic relationship based on the tumor-specific mtDNA variants was evident between the primary and the two metastatic tumor sites of P6 (Figure 2).

Due to the long read length of our sequencing approach — in the order of 2000 nucleotides — the detected variants could be grouped based on their presence on the same read or on separate reads, and thus, we could decipher if they originated from the same or another mtDNA molecule (phasing of variants). A total of 65 combinations of variants were close enough for phasing (Supplementary Table 4), of which 19 phased (partly) together and 46 were mutually exclusive. Interestingly, in tumor tissue, the heteroplasmic variants 10657 T > C and 11040 T > C in P2 and variants 9398A > G and 10407G > A in P3 phased together, but variants 1924 T > C and 2305 T > C in P6 were mutually exclusive.

Heteroplasmic mtDNA Variants within cfDNA

In the cfDNA samples, the number of heteroplasmic mtDNA variants (cf-mtDNA) ranged from 0 up to 9 per sample (Table 1), with detected allele frequencies between 0.04% and 99.4% (Figure 1) (Supplementary Table 3). Majority of the detected cf-mtDNA variants (59%) were not detected in the corresponding tissues we evaluated and thus are of unknown tissue origin ($n = 20$). Some of the variants were detected within only the normal tissue ($n = 5$) or both the tumor and normal tissue ($n = 8$), indicating that these are heteroplasmic patient-specific but not tumor-specific heteroplasmic mtDNA variants present as cf-mtDNA in the circulation. Of the 24

Table 3. Blood-Based Markers in Sera or Plasma of the Eight Evaluated Cancer Patients

Patient	Blood Draw	Tumor Sites <i>In Situ</i> at Blood Draw	Cancer Antigen		Circulating Tumor Cells	Tumor cf-nDNA [Allele Frequency]	Tumor cf-mtDNA [Allele Frequency]
			CA15.3	CA125			
P1 (breast)	1	Primary	20 kU/L	<i>not done</i>	<i>not done</i>	<i>not done</i>	0%
	2	Metastases (bone, lung, liver)	96 kU/L	<i>not done</i>	<i>not done</i>	<i>not done</i>	0.03%
	3	Metastases (bone, lung, liver)	883 kU/L	<i>not done</i>	<i>not done</i>	0% ^a	0.06%-0.3%
	4	Metastases (bone, lung, liver)	587 kU/L	<i>not done</i>	<i>not done</i>	<i>not done</i>	0%
	5	Metastases (bone, lung, liver)	129 kU/L	<i>not done</i>	<i>not done</i>	<i>not done</i>	0%
P2 (breast)	1	Metastases (bone, lung)	30 kU/L	<i>not done</i>	<i>not done</i>	0.06%	1.1% (?)
P3 (breast)	1	Metastases (colon, spleen, pancreas, omentum)	<i>not done</i>	97 kU/L	<i>not done</i>	0.3% ^a	0%
P4 (breast)	1	Primary and metastasis (lymph nodes)	<i>not done</i>	9 kU/L	<i>not done</i>	0%	0%
P5 (breast)	1	Metastasis (bone)	34 kU/L	<i>not done</i>	<i>not done</i>	47.5%	0%
P6 (colon)	1	Primary and metastases (lymph nodes, liver)	<i>not done</i>	<i>not done</i>	2 / 7.5 mL	2.4%	0%
P7 (colon)	1	Primary and metastases (lymph nodes, liver, small intestine, omentum)	<i>not done</i>	<i>not done</i>	0 / 7.5 mL	13.4%-18.5%	0.04%
P8 (colon)	1	Metastasis (liver)	<i>not done</i>	<i>not done</i>	35 / 7.5 mL	7.8%-15.0%	3.8% (?)

^a Note that for these samples, not the entire nDNA was evaluated, but only the subset of driver genes covered by the Oncomine Breast cfDNA assay.

tumor-specific and 2 tumor-expanded heteroplasmic mtDNA variants present in the tumor specimens, only 3 were detected by sequencing the cfDNA (Figure 1): in P2 variant, 9058A > G was present at 0.2% allele frequency in one of the two replicates of the primary tumor and 1.1% allele frequency in the cfDNA; in P7 variant, 16278C > T was present at 64.7% allele frequency in the liver metastasis and 0.04% allele frequency in the cfDNA; and in P8 variant, 16183A > C was present at 0.7% allele frequency in the liver metastasis and 3.8% allele frequency in the cfDNA (Supplementary Table 3). Note that in P2 and P8, the heteroplasmy level of the variant in the tumor tissue was very low. We confirmed by dPCR (orthogonal technique) the absence of one high-frequent tumor-specific or one high-frequent tumor-expanded variant in the sequenced cfDNA samples for both P1 and for P6 (Table 2). Note that the variant allele frequency of those variants in the tissue samples was comparable between SMRT sequencing and dPCR detection. For P1, we extended the number of cfDNA samples by three sera at different time points. The cancer antigen level in these three sera was extremely high (Figure 1, Table 3), indicative for a high tumor load at that point in time. In this patient, we detected by dPCR at low variant allele frequency the tumor-expanded cf-mtDNA variant prior to start of hormonal therapy (6255G > A 0.03% allele frequency) and both the tumor-expanded and tumor-specific cf-mtDNA variants prior to start of chemotherapy (6255G > A 0.3%, 664G > A 0.06% allele frequency) (Table 2, Figure 3).

Thus, out of the 12 cfDNA samples, a total of 4 contained cf-mtDNA variants that were also present in the tumor tissue evaluated, as detected by either sequencing or dPCR. To put these results into perspective, we evaluated the levels of other blood-based cancer biomarkers in these samples. Tumor-specific mutations in nuclear-encoded driver genes were detected in the cfDNA (cf-nDNA) of three sera and three plasma samples (Supplementary Table 5), mostly at much higher allele frequencies than the detected tumor-specific cf-mtDNA variants (Table 3, Supplementary Figures 2 and 3). Also, the level of cancer antigen was increased in the blood of P1, P2, P3, and P5 (≥ 30 kU/L) (Table 3, Figure 3, Supplementary Figure 2), and circulating tumor cells were detected in the blood of P6 and P8 (Table 3, Supplementary Figure 3).

Discussion

In this work, we show that there is extensive heterogeneity in mtDNA variants between tumor and tumor-adjacent normal tissue and that

tumor-specific cf-mtDNA variants are hardly detectable in the circulation of cancer patients.

To detect mtDNA variants, we used a SMRT sequencing approach to evaluate the whole mitochondrial genome [34] and included dPCR as an orthogonal method to evaluate a subset of the detected variants. The limit of detection for variants by our sequencing approach is mainly dependent on the sequencing depth at a position, with variants called based on at least five highly accurate single-molecule reads containing the variant at that position. We have recently shown that the SMRT sequencing approach is able to trace known low-frequent variants $\geq 0.1\%$ allele frequency (sensitivity) and to reliably detect unknown variants $\geq 1.0\%$ allele frequency (specificity) [34]. Therefore, we start with calling variants with at least 1.0% allele frequency in the evaluated samples (detection of unknown variants) and subsequently evaluate the presence of those called variants in the complete dataset (tracing of known variants). For some variants, we obtained a sequencing depth that allowed for tracing down to 0.04% allele frequency (Supplementary Table 3). Note that false positives due to random PCR errors are unlikely to be detected: we used minimally 100,000 input molecules (Supplementary Table 1) corresponding to a 0.001% allele frequency of random PCR errors (1 / 100,000), a high fidelity polymerase (error rate of $\sim 10^{-7}$), and the number of PCR cycles was limited. To simultaneously evaluate the specificity of the method as well as the interference of nuclear insertions of mitochondrial origin (NUMTs), independent resequencing of four tissue specimens was performed after exonuclease treatment of the DNA to specifically degrade linear DNA and thus increase the circular mtDNA fraction. The latter is important since NUMTs can interfere with accurate variant detection due to their sequence similarity to mtDNA and thus complicate investigation of mitochondrial heteroplasmy. Good concordance was observed: independent resequencing revealed that of the 18 variants, 16 variants were confirmed. Two variants could not be confirmed: variant 6255G > A in P1 was detected first at 0.8% allele frequency but was below the limit of detection ($< 0.9\%$ allele frequency) in the resequenced specimen, and variant 9058A > G in P2 was not detected at first (limit of detection $< 0.2\%$ allele frequency) but was detected at 0.2% allele frequency in the resequenced specimen. These variants were called because they were present $\geq 1.0\%$ allele frequency in the cfDNA samples of the corresponding patients. Both variants are not at putative known NUMT positions as evaluated by nucleotide BLAST. From this, we concluded that the limiting factor in tracing variants by our approach is the sequencing depth, which influences the limit of detection.

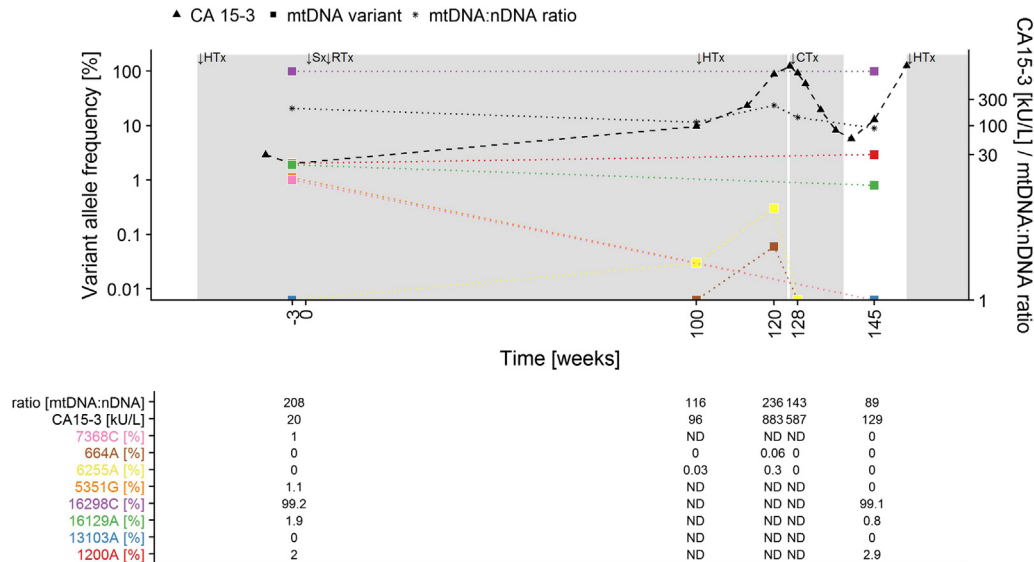


Figure 3. Timeline of patient 1. The allele frequency of detected mtDNA variants (squares), levels of CA 15-3 (triangles), and ratio of mtDNA:nDNA molecules (asterisks) on log-scale in the cfDNA samples (vertical) at five time points (horizontal). The table provides the specifics per variable (color coded). At the top of the graph, treatment is indicated (Sx, surgery; RTx, radiotherapy; HTx, hormone therapy; CTx, chemotherapy), where time = 0 corresponds to the surgery of the primary tumor. A gray background indicates that the patient was receiving systemic therapy. Note that allele frequency in the first and last sera was evaluated by SMRT sequencing, whereas in the second, third, and fourth sera, it was by dPCR. ND, not done.

The extensive heterogeneity we observe between tumor and tumor-adjacent normal tissue is in line with the observation by others that heteroplasmy patterns can differ in an allele-specific manner between tissues within an individual [3–6]. This is also evident from the high fraction of cf-mtDNA variants we detect that are of unknown tissue origin. Since the number of cell generations of epithelial tumor cells greatly exceeds that of nontumor epithelial cells, it is likely that the intraindividual genetic drift observed between tissues within an individual likely also applies to tumor cells and their founder cells. Similarly, our observation that allele frequencies of tumor-specific variants are much higher than those observed in normal tissue corresponds to the hypothesis that more cell generations equal more opportunity for either loss or expansion of a heteroplasmic mtDNA variant [40]. In line with this, only a low fraction of variants phased together in our work (19 of the 65 variant combinations, 30%), indicative that extensive heterogeneity is present among heteroplasmic variants within an individual. Remarkably, P2 shows an exceptional high number of heteroplasmic variants detected in only the normal specimen ($n = 11$), and the tumor of this patient also contained the highest number of heteroplasmic variants ($n = 5$). Note that the primary tumor specimens of P6 and P8 did not contain tumor cells on the morphological level as evaluated by HE-stained slides (Figure 1, Supplementary Table 1), but on the molecular level, the primary tumor specimen of P6 did contain tumor cells as evaluated by the presence of *KRAS* mutated nDNA (Supplementary Table 5). Uncertainty in estimating tumor cell percentage in HE slides has been pointed out in literature [41], and fresh-frozen tissue sections are of lower morphological quality compared to formalin-fixed, paraffin-embedded tissue sections, hampering tumor cell percentage estimation. Notably, whereas in P6 also the normal colon and normal liver appeared nontumorous by morphological evaluation, mutated *KRAS* was present at the molecular level (Supplementary Table 5), indicating tumor cells are present in

these specimens as well. Thus, some of the variants defined as ubiquitous might actually be tumor-specific in P6 — especially variant 2305 T > C. Additionally, it is interesting to see in this case that even variants at high allele frequency can be present on different mtDNA molecules (1924 T > C and 2305 T > C in P6, Supplementary Table 4).

With regard to the number of tumor-specific mtDNA variants per tumor, we observe a higher number compared to other studies. Specifically, other studies using large sample sizes showed that 75% of the breast cancer patients and 60% of the colon cancer patients harbor at least one mtDNA variant in their primary tumor based on massive parallel sequencing studies [7–9], whereas we observe in 100% (8/8) of the patients at least one tumor-specific mtDNA variant. Those studies applied a threshold on variant allele frequency between $\geq 3\%$ and $\geq 15\%$ allele frequency. We used an initial threshold at $\geq 1.0\%$ and no threshold on allele frequency when the variant was called within another sample of the same patient. Based on our previous work, it is unlikely that this is due to false-positive calls since those appeared with $\leq 1.0\%$ variant allele frequency [34]. The higher number may thus be due to the higher sensitivity of our SMRT sequencing approach or statistical coincidence due to the relatively small series we analyzed. Noteworthy is that whereas other studies have reported on (near-)homoplasmic tumor-specific mtDNA variants [7–9], we only observed heteroplasmic tumor-specific mtDNA variants with $< 50\%$ allele frequency. This is likely due to the nontumor cells present in our specimens (i.e., infiltrating immune cells or tissue of origin): it would not be possible to reach 100% allele frequency for a tumor-specific mtDNA variant since nontumor mtDNA will be present as well.

Importantly, despite the presence of tumor-specific mtDNA variants for each tumor tissue analyzed, we were unable to detect the majority of these variants in blood-circulating cfDNA (83% not detected). The few cf-mtDNA variants we detected were present at

extremely low heteroplasmy levels (P1 and P7) or questionable in their true tumor-specific nature given the low heteroplasmy levels in the tumor tissue (P2 and P8). Specifically, in P2 variant, 9058A > G was detected at 1.1% allele frequency in the cfDNA, whereas it was present at 0.2% allele frequency in one of the two replicates of the primary tumor. In P2, the primary tumor also contained tumor-specific variants between 2% and 10% allele frequency (2724G > C, 10657 T > C, 11040 T > C and 16141A > G, Figure 1), but those were not detected as cfDNA in the circulation (Figure 1) (Supplementary Table 3). This would mean that — since the primary tumor was not *in situ* when blood was drawn — the metastases in this patient originated from a clone that contained only 9058A > G but not the other variants present in the primary tumor. Also, in P8 variant, 16183A > C present at 0.7% allele frequency in the liver metastasis was detected at 3.8% allele frequency in the cfDNA. In this patient, the liver metastasis was still *in situ* when blood was drawn, but the tumor-specific variants present at 2% and 14% allele frequency (8102G > A and 9196G > A, Figure 1) are not detected as cfDNA (Figure 1) (Supplementary Table 3). When comparing the detection of tumor-specific cf-mtDNA with other blood-based markers — including CTCs, cancer antigens, and mutations in cf-nDNA — the latter outperform cf-mtDNA since these could be detected in nearly all blood samples (Table 3).

Our hypothesis was that mutated cf-mtDNA would be easily detectable in the circulation due to high stability (if circular) and high copy number per cell and thus would require less sensitive methods to detect them. However, the fact that nuclear-encoded mutations are detected at much higher allele frequencies in the cfDNA could indicate either that more mtDNA from nontumor cells is released in the circulation or that the cf-mtDNA has a higher turnover than its nuclear counterpart, rendering it undetectable by the techniques we applied. Because the mtDNA content is variable per cell, it could be that tissues (nontumor) with higher mtDNA content than the tumor are shedding DNA into the circulatory system. This could in part be the tumor-adjacent normal tissue: three-quarters of the breast tumors harbor a reduction in tumor mtDNA content compared to adjacent normal mammary tissue [2,42–47], whereas approximately half of the colorectal tumors have a reduction in tumor mtDNA content when compared to adjacent normal colon or rectum tissue [2,48–52]. Another likely source of nontumor cf-mtDNA in blood are thrombocytes, which do not contain a nucleus but do contain mitochondria. Before being frozen, our plasma samples were obtained via a centrifugation force that should be enough to remove 90% of the thrombocytes [53]. After thawing, an additional centrifugation step was applied to obtain thrombocyte-poor plasma, but it could be that part of the thrombocytes had been damaged by freezing-thawing and thus released their mtDNA already into the plasma. Especially when comparing the mtDNA content of serum and plasma (order of hundred versus order of thousands, Supplementary Table 1), it seems evident that the capture of thrombocytes in the fibrin clot of serum results in less release of their mtDNA in the blood deriviate. Another probability is that cf-mtDNA is so severely fragmented that it is not detectable by our applied methods, which require DNA of at least 108 base pairs in length for dPCR and of at least 1700 base pairs in length for SMRT sequencing (see amplicon sizes in Supplementary Table 2). A study on the physical characteristics of plasma-derived cf-mtDNA indicates that the majority is associated with particles between 5 μm and 0.45 μm in size since filtering the plasma reduces the amount of cf-mtDNA whereas cf-nDNA is retained (note that the

diameter and length of mitochondria range from, respectively, 0.5 to 1 μm and 5 to 10 μm) [14]. Treatment of serum-derived cfDNA with an exonuclease that digests linear DNA but leaves circular DNA intact resulted in undetectable levels of nDNA, whereas mtDNA concentrations were reduced 5 \times to 10 \times (unpublished data), indicative that at least a fraction of the cf-mtDNA is in its circular form within serum. Other studies, based on whole genome sequencing of plasma-derived cfDNA without prior fragmentation of the DNA (thus, only short fragments smaller than ~600 base pairs are efficiently sequenced), indicate that part of cf-mtDNA is severely fragmented [15–18]. In our own hands, such a whole genome sequencing approach applied to plasma-derived cfDNA of cancer patients resulted in median 0.85 \times coverage (range 0.35–1.73 \times) of the mitochondrial contig, whereas the nuclear DNA was covered by ~1 \times (unpublished data). Taking into account the mtDNA content of cells (in the order of thousands in our plasma samples, Supplementary Table 2), this indicates that the largest proportion of cf-mtDNA is not sequenced by such a whole genome sequencing approach. This is in line with observations in hepatocellular cancer patients [16] and lung transplant recipients [17], where the fractional concentration of sequenced plasma cf-mtDNA was lower than expected. In the study on lung transplant recipients [17] and in another study on sepsis patients [18], the use of sequencing protocols that also include smaller DNA fragments (40–100 base pairs) increased the fraction of cf-mtDNA reads by 8- to 15-fold but still cannot fully explain the low abundance of mtDNA in those experiments. Thus, it seems that a large proportion of the mtDNA is not sequenced in these studies, likely due to the fact that intact mtDNA is not efficiently sequenced during such approaches. It must be kept in mind that those results apply to plasma and not necessarily to serum. Another possibility why cf-mtDNA was rarely detectable might not be related to physical characteristics but due to genetic drift: the tumor-specific variants present in the primary tumor are not present in the (micro-) metastases anymore. Especially when the elapsed time is large (e.g., >40 weeks in P2, > 240 weeks P3 and > 400 weeks P5), it might be possible that the genetic makeup of the mtDNA in the tumors has changed, similar to the heterogeneity we observed between the tumor and tumor-adjacent normal tissue. Given the heterogeneity in mtDNA variants between tissues within an individual [3–6], it is not possible to evaluate all cf-mtDNA variants present in the circulation of a patient as tumor-specific ones, as illustrated by the number of cf-mtDNA variants from tumor-adjacent and of unknown origin in our study. Noteworthy, it could be that the total number of heteroplasmic cf-mtDNA variants present in the circulation is increased for (advanced) cancer patients, but this potentially is also the case for a patient affected with other morbidities (e.g., liver cirrhosis resulting in liver-specific cf-mtDNA variants or colitis resulting in colon-specific cf-mtDNA variants).

Conclusions

Our results demonstrate that extensive mtDNA heterogeneity is evident within an individual. We conclude that there is limited value in tracing tumor-specific cf-mtDNA variants as a blood-circulating biomarker with the current methods available.

Authors' Contributions

M. W., A. S., J. F., S. S., and J. M. conceived and designed the study. M. W., E. T., R. V., and S. A. designed experiments. J. F. and S. S. provided patient specimens. M. W. processed patient specimens and

carried out experiments. R. V. and S. A. led the SMRT sequencing. M. W. and S. A. performed data analyses. M. W., S. S., and J. M. prepared the manuscript, which was reviewed by all authors.

Acknowledgements

The authors are especially grateful to A. M. A. C. Trapman and A. M. Sieuwerts for their contribution in collecting and/or processing patient samples at the Department of Medical Oncology of the Erasmus MC Cancer Institute, to A. M. Timmermans and K. Biermann for their contribution in pathological assessment of tissue slides at the Department of Medical Oncology and Department of Pathology of the Erasmus MC Cancer Institute, to M. P. H. M. Jansen and R. van Marion for their contribution in Ion semiconductor sequencing at the Department of Pathology of the Erasmus MC Cancer Institute, and to H. Feitsma for constructive feedback on the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neo.2018.05.003>.

References

- Wachsmuth M, Hubner A, Li M, Madea B, and Stoneking M (2016). Age-related and heteroplasmy-related variation in human mtDNA copy number. *PLoS Genet* **12**(3):e1005939.
- Reznik E, Miller ML, Senbabaoglu Y, Riaz N, Sarungbam J, Tickoo SK, Al-Ahmadie HA, Lee W, Seshan VE, and Hakimi AA, et al (2016). Mitochondrial DNA copy number variation across human cancers. *Elife* **5**:e10769.
- Samuels DC, Li C, Li B, Song Z, Torstenson E, Boyd Clay H, Rokas A, Thornton-Wells TA, Moore JH, and Hughes TM, et al (2013). Recurrent tissue-specific mtDNA mutations are common in humans. *PLoS Genet* **9**(11):e1003929.
- He Y, Wu J, Dressman DC, Iacobuzio-Donahue C, Markowitz SD, Velculescu VE, Diaz Jr LA, Kinzler KW, Vogelstein B, and Papadopoulos N (2010). Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature* **464**(7288), 610–614.
- Li M, Schroder R, Ni S, Madea B, and Stoneking M (2015). Extensive tissue-related and allele-related mtDNA heteroplasmy suggests positive selection for somatic mutations. *Proc Natl Acad Sci U S A* **112**(8), 2491–2496.
- Calloway CD, Reynolds RL, Herrin Jr GL, and Anderson WW (2000). The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age. *Am J Hum Genet* **66**(4), 1384–1397.
- Ju YS, Alexandrov LB, Gerstung M, Martincorena I, Nik-Zainal S, Ramakrishna M, Davies HR, Papaemmanuil E, Gundem G, and Shlien A, et al (2014). Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer. *Elife* **3**:e02935.
- Larman TC, DePalma SR, Hadjipanayis AG, N. Cancer Genome Atlas Research Group, Zhang J, Gabriel SB, Chin L, Seidman CE, and Kucherlapati R, et al (2012). Spectrum of somatic mitochondrial mutations in five cancers. *Proc Natl Acad Sci U S A* **109**(35), 14087–14091.
- Stewart JB, Alaei-Mahabadi B, Sabarinathan R, Samuelsson T, Gorodkin J, Gustafsson CM, and Larsson E (2015). Simultaneous DNA and RNA mapping of somatic mitochondrial mutations across diverse human cancers. *PLoS Genet* **11**(6):e1005333.
- Brown WM, George Jr M, and Wilson AC (1979). Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci U S A* **76**(4), 1967–1971.
- Diaz Jr LA and Bardelli A (2014). Liquid biopsies: genotyping circulating tumor DNA. *J Clin Oncol* **32**(6), 579–586.
- Gilkerson R, Bravo L, Garcia I, Gaytan N, Herrera A, Maldonado A, and Quintanilla B (2013). The mitochondrial nucleoid: integrating mitochondrial DNA into cellular homeostasis. *Cold Spring Harb Perspect Biol* **5**(5):a011080.
- Murgia M, Pizzo P, Sandona D, Zanovello P, Rizzuto R, and Di Virgilio F (1992). Mitochondrial DNA is not fragmented during apoptosis. *J Biol Chem* **267**(16), 10939–10941.
- Chiu RW, Chan LY, Lam NY, Tsui NB, Ng EK, Rainer TH, and Lo YM (2003). Quantitative analysis of circulating mitochondrial DNA in plasma. *Clin Chem* **49**(5), 719–726.
- Chandrananda D, Thorne NP, and Bahlo M (2015). High-resolution characterization of sequence signatures due to non-random cleavage of cell-free DNA. *BMC Med Genomics* **8**, 29.
- Jiang P, Chan CW, Chan KC, Cheng SH, Wong J, Wong VW, Wong GL, Chan SL, Mok TS, and Chan HL, et al (2015). Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc Natl Acad Sci U S A* **112**(11), E1317–1325.
- Burnham P, Kim MS, Agbor-Enoh S, Luikart H, Valentine HA, Khush KK, and De Vlaminck I (2016). Single-stranded DNA library preparation uncovers the origin and diversity of ultrashort cell-free DNA in plasma. *Sci Rep* **6**, 27859.
- Zhang R, Nakahira K, Guo X, Choi AM, and Gu Z (2016). Very short mitochondrial DNA fragments and heteroplasmy in human plasma. *Sci Rep* **6**, 36097.
- Zachariah RR, Schmid S, Buerki N, Radpour R, Holzgreve W, and Zhong X (2008). Levels of circulating cell-free nuclear and mitochondrial DNA in benign and malignant ovarian tumors. *Obstet Gynecol* **112**(4), 843–850.
- Cheng C, Omura-Minamisawa M, Kang Y, Hara T, Koike I, and Inoue T (2009). Quantification of circulating cell-free DNA in the plasma of cancer patients during radiation therapy. *Cancer Sci* **100**(2), 303–309.
- Mead R, Duku M, Bhandari P, and Cree IA (2011). Circulating tumour markers can define patients with normal colons, benign polyps, and cancers. *Br J Cancer* **105**(2), 239–245.
- Mehra N, Penning M, Maas J, van Daal N, Giles RH, and Voest EE (2007). Circulating mitochondrial nucleic acids have prognostic value for survival in patients with advanced prostate cancer. *Clin Cancer Res* **13**(2 Pt 1), 421–426.
- Mahmoud EH, Fawzy A, Ahmad OK, and Ali AM (2015). Plasma circulating cell-free nuclear and mitochondrial DNA as potential biomarkers in the peripheral blood of breast cancer patients. *Asian Pac J Cancer Prev* **16**(18), 8299–8305.
- Jeronimo C, Nomoto S, Caballero OL, Usadel H, Henrique R, Varzim G, Oliveira J, Lopes C, Fliss MS, and Sidransky D (2001). Mitochondrial mutations in early stage prostate cancer and bodily fluids. *Oncogene* **20**(37), 5195–5198.
- Uzawa K, Baba T, Uchida F, Yamatoji M, Kasamatsu A, Sakamoto Y, Ogawara K, Shiiba M, Bukawa H, and Tanzawa H (2012). Circulating tumor-derived mutant mitochondrial DNA: a predictive biomarker of clinical prognosis in human squamous cell carcinoma. *Oncotarget* **3**(7), 670–677.
- Takeuchi H, Fujimoto A, and Hoon DS (2004). Detection of mitochondrial DNA alterations in plasma of malignant melanoma patients. *Ann N Y Acad Sci* **1022**, 50–54.
- Hibi K, Nakayama H, Yamazaki T, Takase T, Taguchi M, Kasai Y, Ito K, Akiyama S, and Nakao A (2001). Detection of mitochondrial DNA alterations in primary tumors and corresponding serum of colorectal cancer patients. *Int J Cancer* **94**(3), 429–431.
- Losanoff JE, Zhu W, Qin W, Mannello F, and Sauter ER (2008). Can mitochondrial DNA mutations in circulating white blood cells and serum be used to detect breast cancer? *Breast* **17**(5), 540–542.
- Okochi O, Hibi K, Uemura T, Inoue S, Takeda S, Kaneko T, and Nakao A (2002). Detection of mitochondrial DNA alterations in the serum of hepatocellular carcinoma patients. *Clin Cancer Res* **8**(9), 2875–2878.
- Fliss MS, Usadel H, Caballero OL, Wu L, Buta MR, Eleff SM, Jen J, and Sidransky D (2000). Facile detection of mitochondrial DNA mutations in tumors and bodily fluids. *Science* **287**(5460), 2017–2019.
- Zhu W, Qin W, Bradley P, Wessel A, Puckett CL, and Sauter ER (2005). Mitochondrial DNA mutations in breast cancer tissue and in matched nipple aspirate fluid. *Carcinogenesis* **26**(1), 145–152.
- Wong LJ, Lueth M, Li XN, Lau CC, and Vogel H (2003). Detection of mitochondrial DNA mutations in the tumor and cerebrospinal fluid of medulloblastoma patients. *Cancer Res* **63**(14), 3866–3871.
- Duberow DP, Brait M, Hoque MO, Theodorescu D, Sidransky D, Dasgupta S, and Mathies RA (2016). High-performance detection of somatic D-loop mutation in urothelial cell carcinoma patients by polymorphism ratio sequencing. *J Mol Med* **94**(9), 1015–1024.
- Weerts MJA, Timmermans EC, Vossen RHAM, van Strijp D, Van den Hout-van Vroonhoven MCGN, van Ijcken WFJ, van der Zaag PJ, Anvar SY, Sleijfer S, and Martens JWM (2018). Sensitive detection of mitochondrial DNA variants for analysis of mitochondrial DNA-enriched extracts from frozen tumor tissue. *Sci Rep* **8**, 2261.

- [35] Weerts MJA, Sieuwerts AM, Smid M, Look MP, Foekens JA, Sleijfer S, and Martens JWM (2016). Mitochondrial DNA content in breast cancer: Impact on in vitro and in vivo phenotype and patient prognosis. *Oncotarget* **7**(20), 29166–29176.
- [36] Anvar SY, van der Gaag KJ, van der Heijden JW, Veltrop MH, Vossen RH, de Leeuw RH, Breukel C, Buermans HP, Verbeek JS, and de Knijff P, et al (2014). TSSV: a tool for characterization of complex allelic variants in pure and mixed genomes. *Bioinformatics* **30**(12), 1651–1659.
- [37] Martin M (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal* **17**(1), 10–12.
- [38] Li H and Durbin R (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**(5), 589–595.
- [39] Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, and Mesirov JP (2011). Integrative genomics viewer. *Nat Biotechnol* **29**(1), 24–26.
- [40] Coller HA, Khrapko K, Bodyak ND, Nekhaeva E, Herrero-Jimenez P, and Thilly WG (2001). High frequency of homoplasmic mitochondrial DNA mutations in human tumors can be explained without selection. *Nat Genet* **28**(2), 147–150.
- [41] Smits AJ, Kummer JA, de Bruin PC, Bol M, van den Tweel JG, Seldenrijk KA, Willems SM, Offerhaus GJ, de Weger RA, and van Diest PJ, et al (2014). The estimation of tumor cell percentage for molecular testing by pathologists is not accurate. *Mod Pathol* **27**(2), 168–174.
- [42] Mambo E, Chatterjee A, Xing M, Tallini G, Haugen BR, Yeung SC, Sukumar S, and Sidransky D (2005). Tumor-specific changes in mtDNA content in human cancer. *Int J Cancer* **116**(6), 920–924.
- [43] Yu M, Zhou Y, Shi Y, Ning L, Yang Y, Wei X, Zhang N, Hao X, and Niu R (2007). Reduced mitochondrial DNA copy number is correlated with tumor progression and prognosis in Chinese breast cancer patients. *IUBMB Life* **59**(7), 450–457.
- [44] Tseng LM, Yin PH, Chi CW, Hsu CY, Wu CW, Lee LM, Wei YH, and Lee HC (2006). Mitochondrial DNA mutations and mitochondrial DNA depletion in breast cancer. *Genes Chromosomes Cancer* **45**(7), 629–638.
- [45] Fan AX, Radpour R, Haghghi MM, Kohler C, Xia P, Hahn S, Holzgreve W, and Zhong XY (2009). Mitochondrial DNA content in paired normal and cancerous breast tissue samples from patients with breast cancer. *J Cancer Res Clin Oncol* **135**(8), 983–989.
- [46] Barekati Z, Radpour R, Kohler C, Zhang B, Toniolo P, Lenner P, Lv Q, Zheng H, and Zhong XY (2010). Methylation profile of TP53 regulatory pathway and mtDNA alterations in breast cancer patients lacking TP53 mutations. *Hum Mol Genet* **19**(15), 2936–2946.
- [47] McMahon S and LaFramboise T (2014). Mutational patterns in the breast cancer mitochondrial genome, with clinical correlates. *Carcinogenesis* **35**(5), 1046–1054.
- [48] Mohideen AM, Dicks E, Parfrey P, Green R, and Savas S (2015). Mitochondrial DNA polymorphisms, its copy number change and outcome in colorectal cancer. *BMC Res Notes* **8**, 272.
- [49] Feng S, Xiong L, Ji Z, Cheng W, and Yang H (2011). Correlation between increased copy number of mitochondrial DNA and clinicopathological stage in colorectal cancer. *Oncol Lett* **2**(5), 899–903.
- [50] Chen T, He J, Shen L, Fang H, Nie H, Jin T, Wei X, Xin Y, Jiang Y, and Li H, et al (2011). The mitochondrial DNA 4,977-bp deletion and its implication in copy number alteration in colorectal cancer. *BMC Med Genet* **12**, 8.
- [51] Gao J, Wen S, Zhou H, and Feng S (2015). De-methylation of displacement loop of mitochondrial DNA is associated with increased mitochondrial copy number and nicotinamide adenine dinucleotide subunit 2 expression in colorectal cancer. *Mol Med Rep* **12**(5), 7033–7038.
- [52] Cui H, Huang P, Wang Z, Zhang Y, Zhang Z, Xu W, Wang X, Han Y, and Guo X (2013). Association of decreased mitochondrial DNA content with the progression of colorectal cancer. *BMC Cancer* **13**, 110.
- [53] Perez AG, Lana JF, Rodrigues AA, Luzo AC, Belangero WD, and Santana MH (2014). Relevant aspects of centrifugation step in the preparation of platelet-rich plasma. *ISRN Hematol* **2014**, 176060.