Low-depth native-molecule whole-genome nanopore-sequencing of lung cancer cell-free DNA samples with *EGFR* and *TP53* co-mutations

David T. Mulder, Adrian Kense, Pedro Feijao, Fatih Karaoglanoglu, Sahand Khakabi, Andréanne Morin, Rosalía Aguirre-Hernández, Melissa K. McConechy

Imagia Canexia Health, Vancouver, BC, Canada;

Contact: dmulder@canexiahealth.com *http://www.canexiahealth.com*

TP53 and *EGFR* gene mutations are drivers of lung cancer

Genetic mutation of the tyrosine kinase *EGFR* and tumour suppressor *TP53* is common in advanced non-small cell lung cancer (NSCLC) (*The Cancer Genome Atlas Research Network, 2014, Nature*). Lung cancers with *EGFR* and *TP53* co-mutations are aggressive, are resistant to treatment, and have poor prognoses (*Liu et al, 2022, Front. Oncol.*).

Sequencing of cell-free DNA (cfDNA) from blood plasma provides a low intervention alternative to tissue biopsy (*Bonanno et al, 2022, Br. J. Cancer.*). Targeted sequencing of cancer cfDNA has emerged as a major diagnostic method but involves specialized lab work.

Low-depth native-molecule whole-genome nanopore-sequencing offers a rapid and inexpensive alternative for profiling of cancer cfDNA. We evaluated this strategy with ten normal and ten lung cancer cfDNA samples with *EGFR* and *TP53* co-mutations and assessed somatic copy number alterations (SCNAs), fragment lengths, and differential methylation.

METHODS AND MATERIALS

NSCLC plasma cfDNA samples were assayed using the *Follow It* liquid biopsy assay (*Imagia Canexia Health)* to identify samples with *EGFR* and *TP53* co-mutations.

Native-molecule WGS libraries were prepared from up to 100 ng of cfDNA from ten advanced NSCLC cfDNA samples with *EGFR* and *TP53* co-mutations, and ten normal cfDNA samples using the *SQK-LSK110* ligation kit (*Oxford Nanopore Technologies*) with a short-fragment optimized protocol. Libraries were sequenced using *MinION r9* flow cells on a *MinION Mk1b* or *GridION* instrument. Data was high-accuracy basecalled and aligned to the human genome (*Guppy* v6.2.1; dna_r9.4.1_450bps_modbases_5hmc_5mc_cg_hac.cfg; GRCh38).

SCNAs and tumour fraction were called (*IchorCNA* v0.2.0; *Adalsteinsson et al, 2017, Nat. Comm.*) with a custom pool of 10 normal cfDNA samples and a window size of 1Mb. Methylation calls were extracted for analysis (*modbam2bed v0.6.2*) and beta values were calculated for non-overlapping genomic windows of 100kb. Analyses of methylation profiles and cfDNA fragment profiles were performed using a custom analysis pipeline.

Low-depth native-molecule whole-genome nanopore-sequencing of cfDNA

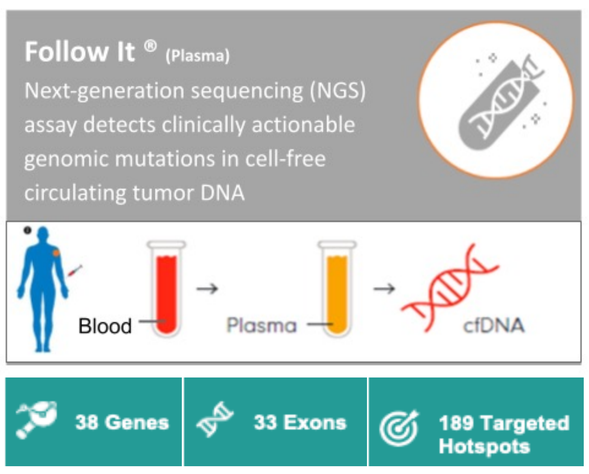
We assayed cfDNA lung cancer cfDNA samples with the *Follow It* liquid biopsy assay which interrogates 189 hotspots in 38 genes and identified ten samples with *EGFR* and *TP53* co-mutations (**Figure 1a**). The *Follow It* assay was analytically validated in a CLIA setting (99D2111438).

We performed low-depth native-molecule whole-genome nanopore sequencing on these ten cfDNA samples as well as ten normal cfDNA samples with one sample library sequenced per flow cell (**Figure 1b**). We obtained 22-80M reads with genome coverage depths of 1.5-4.0X, and an average Phred score of 11.0 (12.5 for aligned reads).

Data was generated for research use only from samples with ethics use approval and not for clinical use. Oxford Nanopore Technologies products are not intended for use for health assessment or to diagnose, treat, mitigate, cure, or prevent any disease or condition.

Figure 1. (A) *Follow It* assay description content. (B) Low-depth native-molecule whole-genome nanopore-sequencing summary table of 10 normal and 10 cancer cfDNA samples

A.



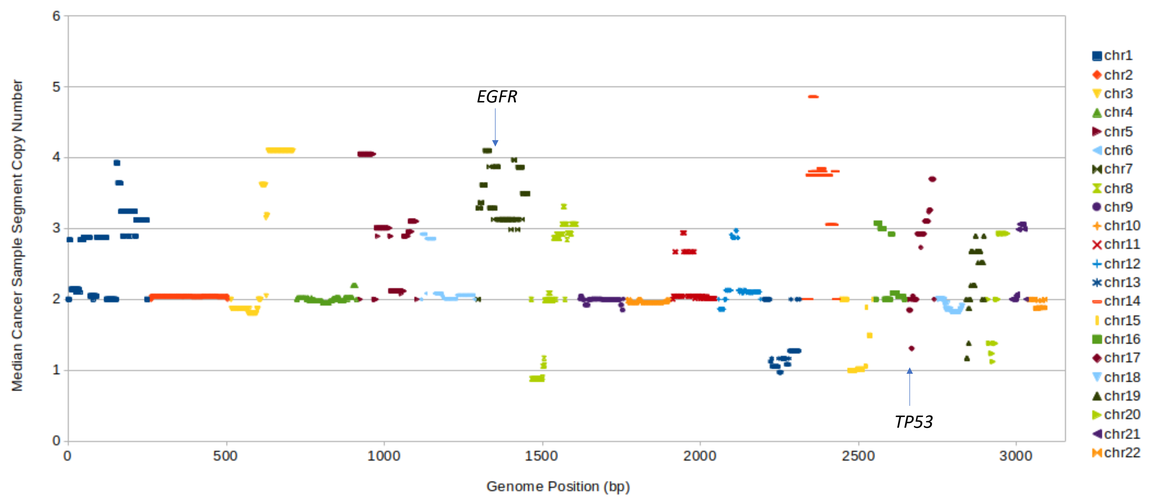
B.

Sample ID	Status	Platform	Reads (M)	Yield (Gb)	Depth (X)	N50 (bp)	Phred Score
N23	Normal	minION	50.0	11.9	2.8	209	13
N30	Normal	minION	46.0	10.0	2.3	204	12
N31	Normal	minION	80.0	17.9	4.0	206	13
N38	Normal	minION	45.8	12.0	2.6	212	12
N39	Normal	minION	64.2	16.1	3.5	210	13
N51	Normal	gridION	44.3	11.5	2.6	213	13
N52	Normal	gridION	35.9	9.1	1.9	210	12
N55	Normal	gridION	23.8	6.5	1.4	216	13
N56	Normal	gridION	22.8	5.8	1.2	212	12
N57	Normal	gridION	26.3	6.6	1.3	208	13
C41	Cancer	minION	38.4	8.4	1.8	199	12
C42	Cancer	gridION	36.1	8.8	1.9	205	13
C43	Cancer	gridION	27.5	6.7	1.5	207	13
C44	Cancer	gridION	30.4	7.2	1.6	204	12
C47	Cancer	gridION	40.4	8.4	1.7	197	12
C48	Cancer	gridION	32.4	7.1	1.5	200	12
C49	Cancer	gridION	38.1	9.7	2.2	207	13
C50	Cancer	gridION	39.3	9.0	2.1	204	13
C53	Cancer	gridION	28.0	7.6	1.7	213	13
C54	Cancer	gridION	28.0	6.7	1.5	207	13

Lung cancer cfDNA samples have extensive copy number alterations

We interrogated the ten lung cancer samples for large SCNAs using the ten normal cfDNA samples as a custom pool of normals. 9 of 10 cancer samples had alterations across multiple chromosomes with many arm-level SCNAs (**Figure 2**). One sample was below detectable level by *IchorCNA*. Observed global SCNA profile is consistent with the literature for lung carcinomas (*Staaf et al, 2013, Int. J. Canc.*).

Figure 2. Manhattan SCNA plot of median of 1Mbp segments for 9 lung cancer samples



Lung cancer samples have copy gain or copy loss genotypes

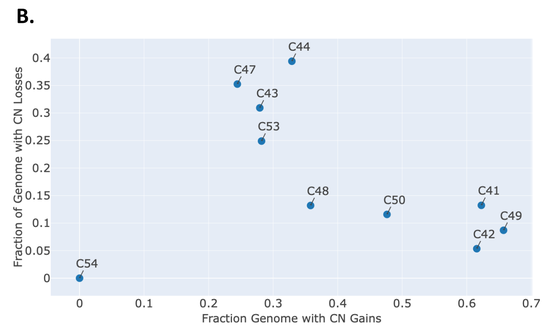
In 9 cancer samples, tumor content was calculated by as 11-50% with fraction genome copy altered of 49-76% (**Figure 3a**). Four samples had relatively more losses while three samples had relatively more gains (**Figure 3b**).

Figure 3. Cancer cfDNA sample summaries (a) Variant allele frequencies (VAF), copy number, and genome fraction alteration. (b) Plot of total fraction genome alteration.

A.

Sample ID	EGFR VAF % (Follow It)	TP53 VAF % (Follow It)	Tumor Content (IchorCNA)	% Genome Altered (IchorCNA)	EGFR SCNA (IchorCNA)	TP53 SCNA (IchorCNA)
C41	43%	41%	40%	75%	4.0	1.0
C42	62%	78%	50%	67%	4.6	2.6
C43	72%	20%	12%	59%	4.1	2.1
C44	86%	70%	47%	72%	2.9	0.5
C47	14%	9%	11%	60%	3.9	1.3
C48	16%	27%	28%	49%	2.9	1.1
C49	30%	74%	29%	74%	3.1	2.8
C50	44%	38%	33%	59%	4.0	2.9
C53	24%	37%	36%	53%	3.3	1.8
C54	14%	37%	0%	0	2.0	2.0

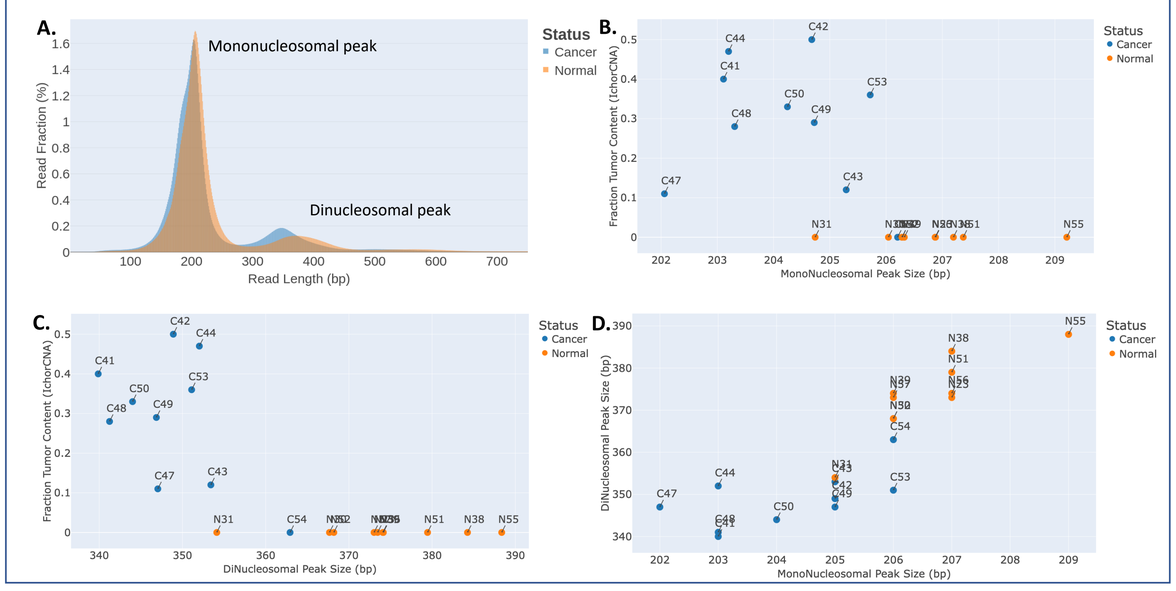
B.



Cancer cfDNA samples have smaller nucleosomal peak fragment lengths

Circulating tumor DNA (ctDNA) fragments are shorter than normal cfDNA fragments (*Underhill et al, 2016, PLoS Genet.; Mouliere et al, 2018, Sci. Transl. Med.*). When all reads were combined for normal or cancer cfDNA samples, cancer samples had smaller mononucleosomal and dinucleosomal peak fragment lengths (**Figure 4a**). Mononucleosomal profiles distinguished 5 cancer and 5 normal samples (**Figure 4b**) while dinucleosomal profiles distinguished 9 and 9 respectively (**Figure 4c**) showing promise for nucleosomal peak fragment lengths in classification of high tumor burden cfDNA samples (**Figure 4d**).

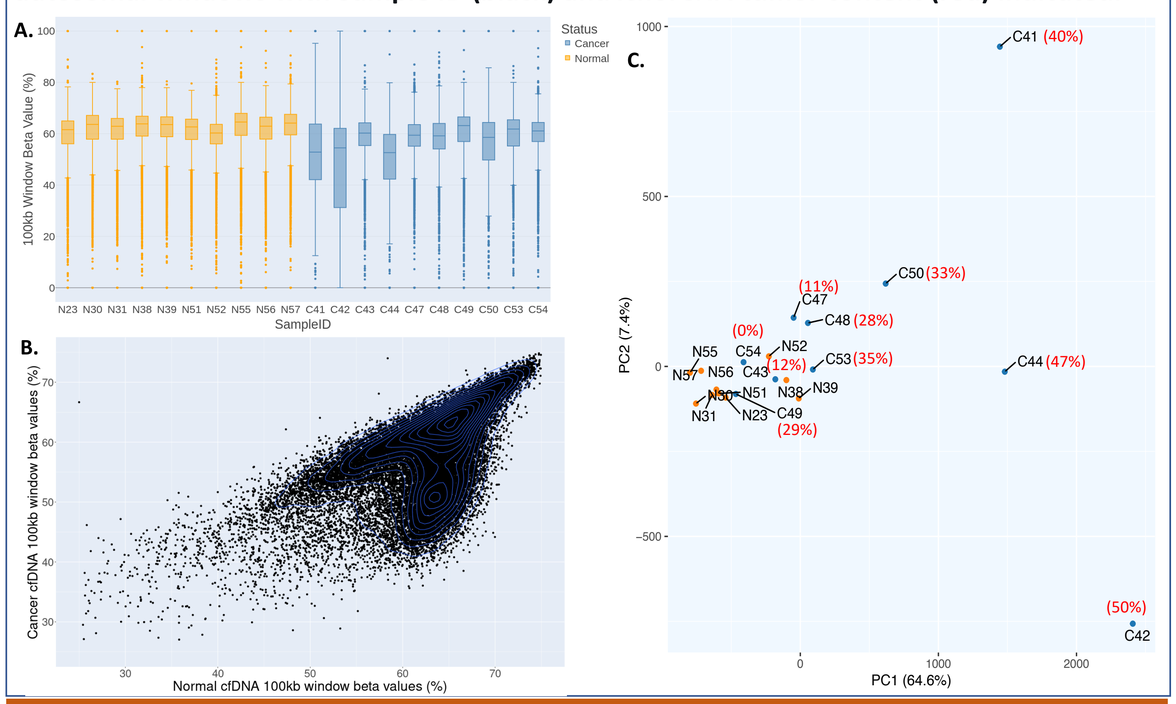
Figure 4. (a) Combined normal and cancer cfDNA nucleosomal peak fragment lengths. Comparison of tumor content against (b) mononucleosomal and (c) dinucleosomal peak fragment lengths. (d) Separation of samples based on nucleosomal peak fragment lengths.



High burden cfDNA samples are differentially methylated

Methylation beta values from low-depth WGS can be investigated by calculating beta values for 100kb genomic windows (*Zhou et al, 2018, Nat. Genet.*). High tumor content cancer samples had a wider distribution of 5mC beta values than normal samples (**Figure 5a**). When all reads were combined for samples, relative hypomethylation of the cancer samples was observed (**Figure 5b**). Of the 30,894 100kb windows, there were 962 windows with non-overlapping beta value distributions between normal and cancer samples. A principal component analysis (PCA) of 28,760 autosomal windows showed separation of high tumor content cancer samples (>40%) from normal and lower burden cancer samples (**Figure 5c**).

Figure 5. (a) Combined 100kb window beta value distributions for all cfDNA samples. (b) Comparison of cancer and normal samples in aggregate. (c) PCA analysis of 28,760 autosomal windows with Sample ID (black) and *IchorCNA* tumor content (red) indicated.



CONCLUSIONS

High tumor burden, advanced NSCLC cell-free DNA samples can be easily and rapidly differentiated from normal samples using low-depth whole-genome native-molecule nanopore-sequencing and analysis of copy number, fragment lengths, and methylation.

Classification and stratification of low tumor burden cancer cfDNA samples will require integration of multiple molecular features and development of advanced analytical methods.