Cell-free DNA analysis by **whole genome sequencing**

ctDNA detection

Outline

- Why is whole genome sequencing interesting ?
- 4 “proof of concept” studies

Successful detection of ctDNA

Depends on 2 consecutive sampling processes, each with its own statistical probability:

1) **Sampling probability:** the probability that the sample contains a tumor DNA fragment

2) **Detection probability:** the probability that the ctDNA detection approach can detect the marker fragment, given
   - The Tumor Fraction of the total cfDNA
   - The number of fragments analyzed
   - The number of markers analyzed
   - Technical variation in the used detection method

Successful detection of ctDNA

Depends on 2 consecutive sampling processes, each with its own statistical probability:

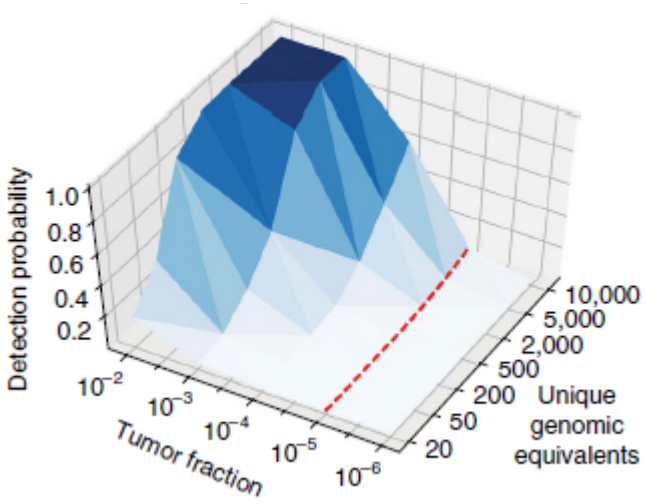
1) **Sampling probability:** the probability that the sample contains a tumor DNA fragment

2) **Detection probability:** the probability that the ctDNA detection approach can detect the marker fragment, given
   - The tumor fraction of the total cfDNA
   - The number of fragments analyzed
   - The number of markers analyzed
   - Technical variation in the used detection method

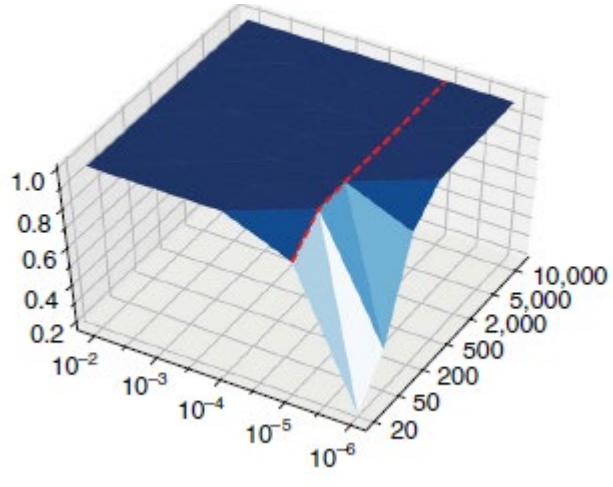
Factors affecting the sensitivity of ctDNA detection methods

- Tumor fraction of the TOTAL cell free DNA
- Number of genome equivalents examined (plasma volume)
- Number of markers

Single marker



10,000 markers



Modified from Nat Med. 2020 Jul;26(7):1114-1124.

Advantage of whole genome analysis

Targeted analysis

Tumor genome

Commonly ONE fragment per mutated tumor genome

Cell free DNA (Tumor genome)

Whole genome analysis

All mutated fragments per tumor genome (thousands)

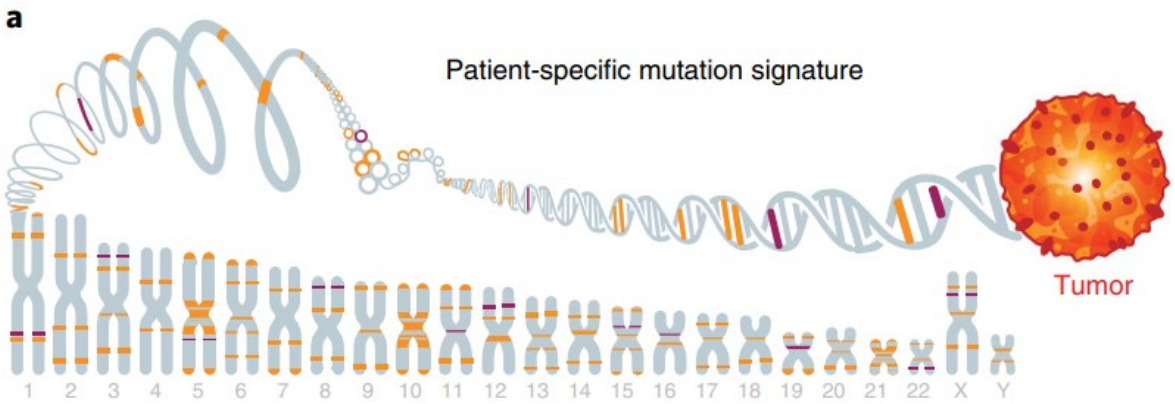
Potentially millions of informative fragments

Tumor informed analysis

ARTICLES
https://doi.org/10.1038/s41591-020-0915-3
nature medicine
Check for updates

Genome-wide cell-free DNA mutational integration enables ultra-sensitive cancer monitoring

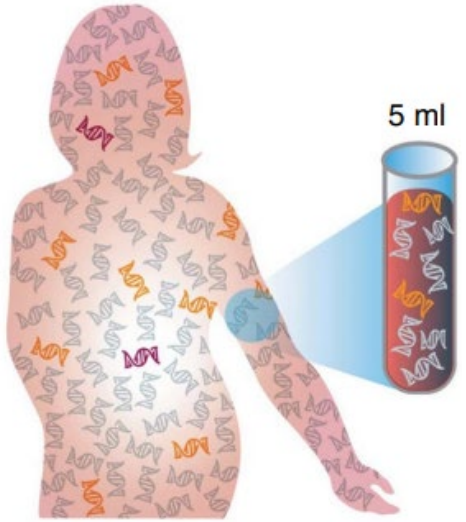
Asaf Zviran[1,2,7], Rafael C. Schulman[1,2,7], Minita Shah[1,7], Steven T. K. Hill[1,2], Sunil Deochand[1,2], Cole C. Khamnei[1,2], Dillon Maloney[1], Kristofer Patel[1,2], Will Liao[1], Adam J. Widman[1,2,3], Phillip Wong[3], Margaret K. Callahan[3], Gavin Ha[4], Sarah Reed[5], Denisse Rotem[5], Dennie Frederick[6], Tatyana Sharova[6], Benchun Miao[6], Tommy Kim[6], Greg Gydush[5], Justin Rhoades[5], Kevin Y. Huang[1,2], Nathaniel D. Omans[1,2], Patrick O. Bolan[2], Andrew H. Lipsky[2], Chelston Ang[1,2], Murtaza Malbari[2], Catherine F. Spinelli[2], Selena Kazancioglu[1], Alexi M. Runnels[1], Samantha Fennessey[1], Christian Stolte[1], Federico Gaiti[1,2], Giorgio G. Inghirami[2], Viktor Adalsteinsson[5], Brian Houck-Loomis[1], Jennifer Ishii[1], Jedd D. Wolchok[3], Genevieve Boland[6], Nicolas Robine[1], Nasser K. Altorki[2] and Dan A. Landau[1,2]✉

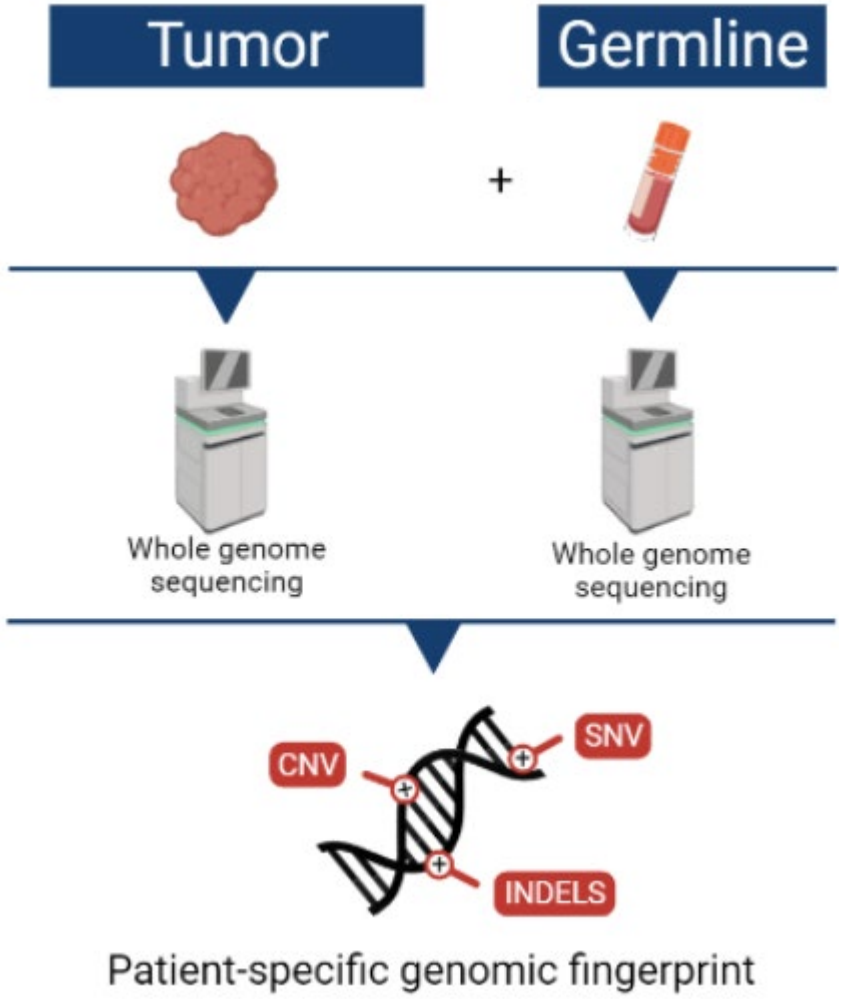


MRDetect

Approach:
Tumor informed whole genome sequencing (30x)

Cohort:
Healthy controls (n=38)
Cancer samples (n=60)
- LUAD; n=39, CRC; n=19 and melanoma; n=2

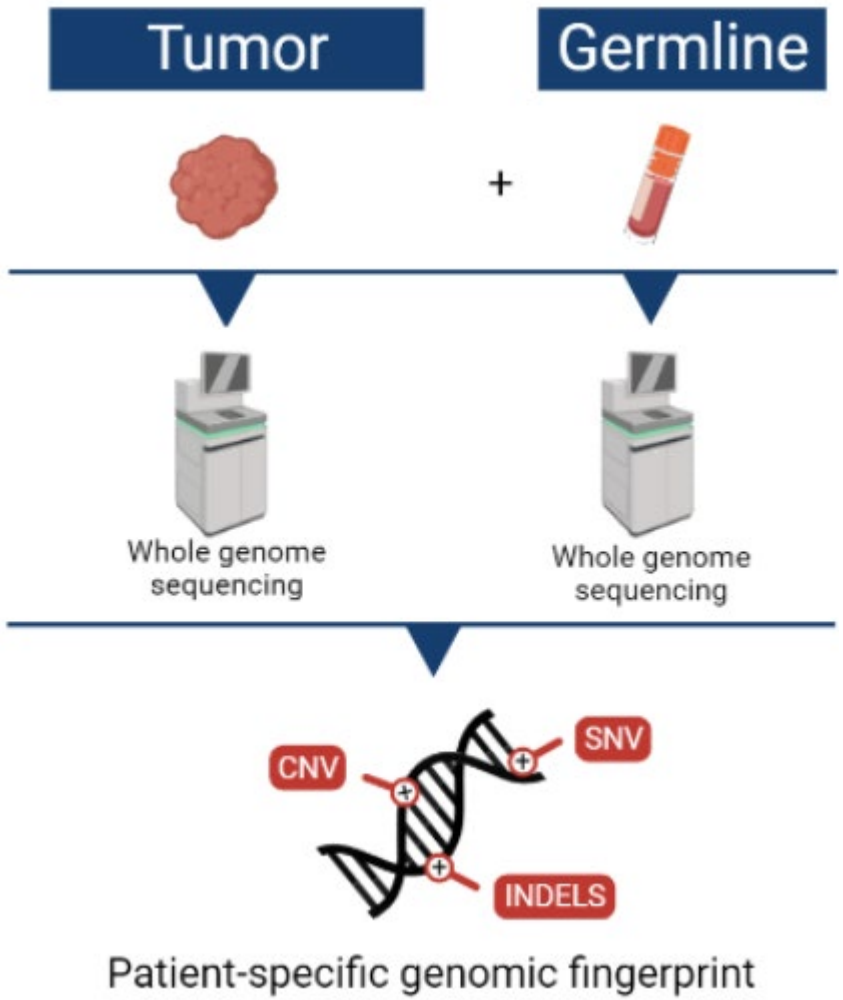


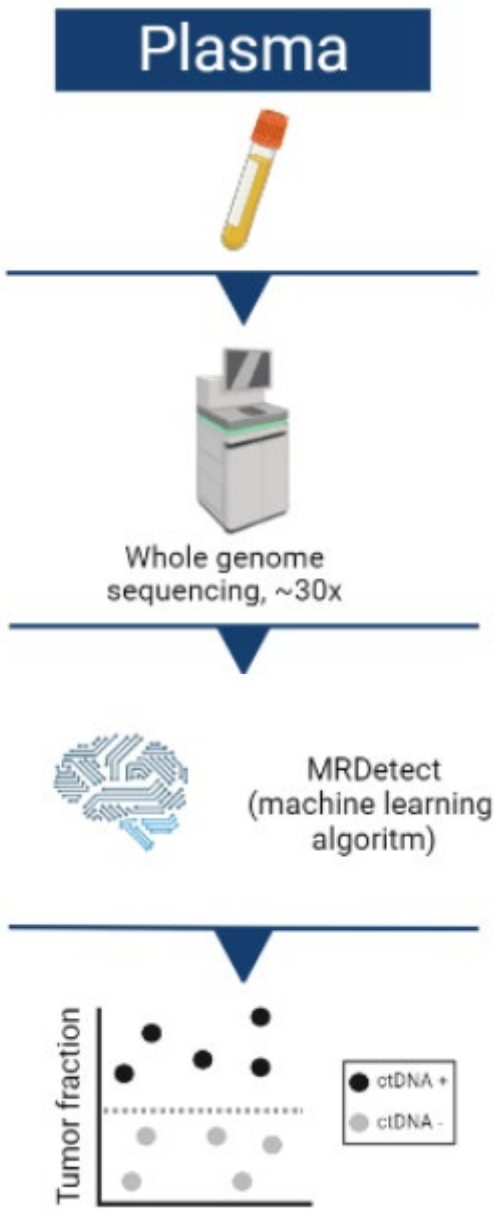


Patient-specific genomic fingerprint includes:

Single nucleotide vatiants (SNV), insertion/deletion (INDELS)

Copy number variations (CNV)

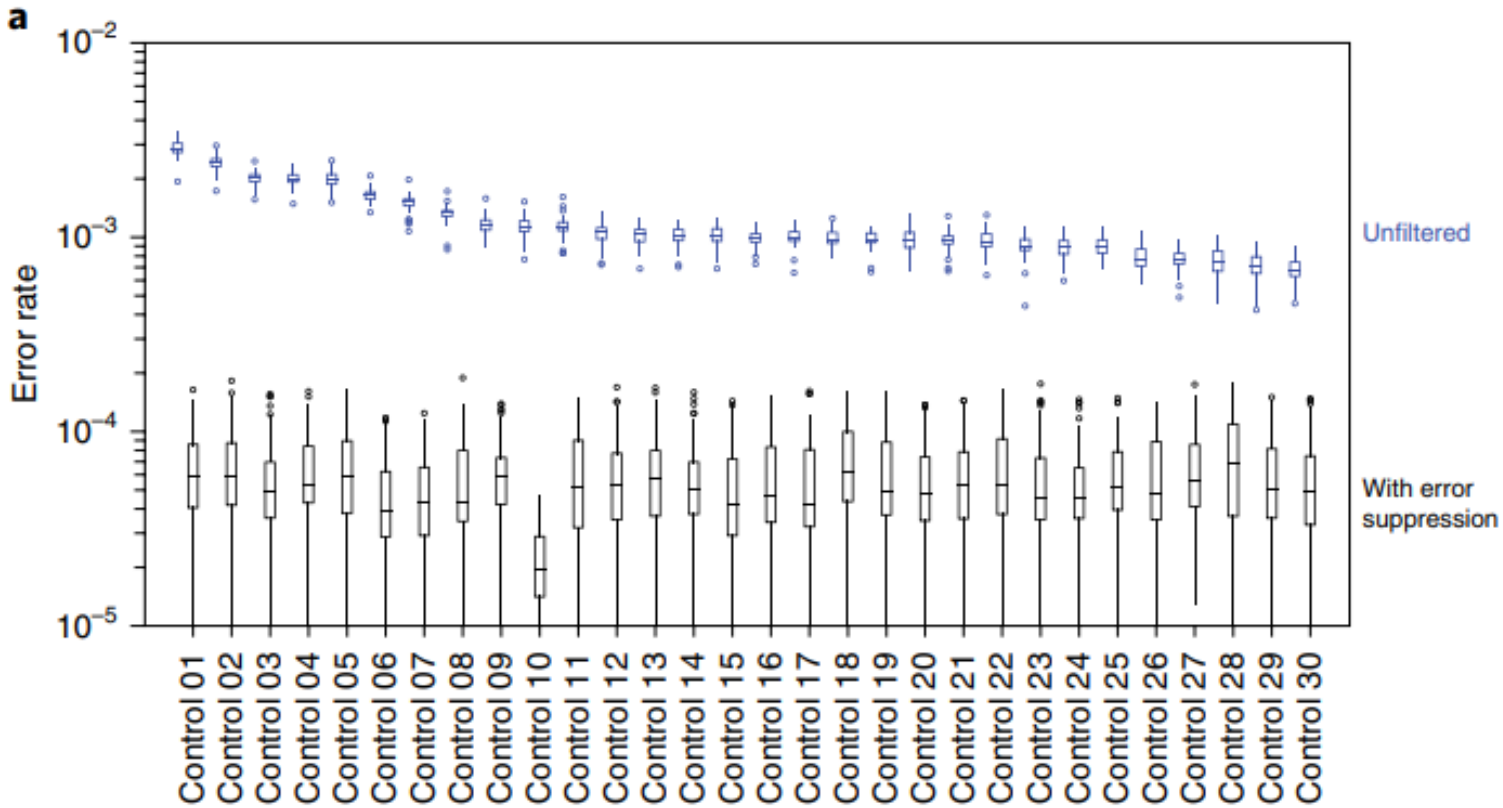




Note: with 30x WGS of plasma, there is not enough evidence to call mutations!

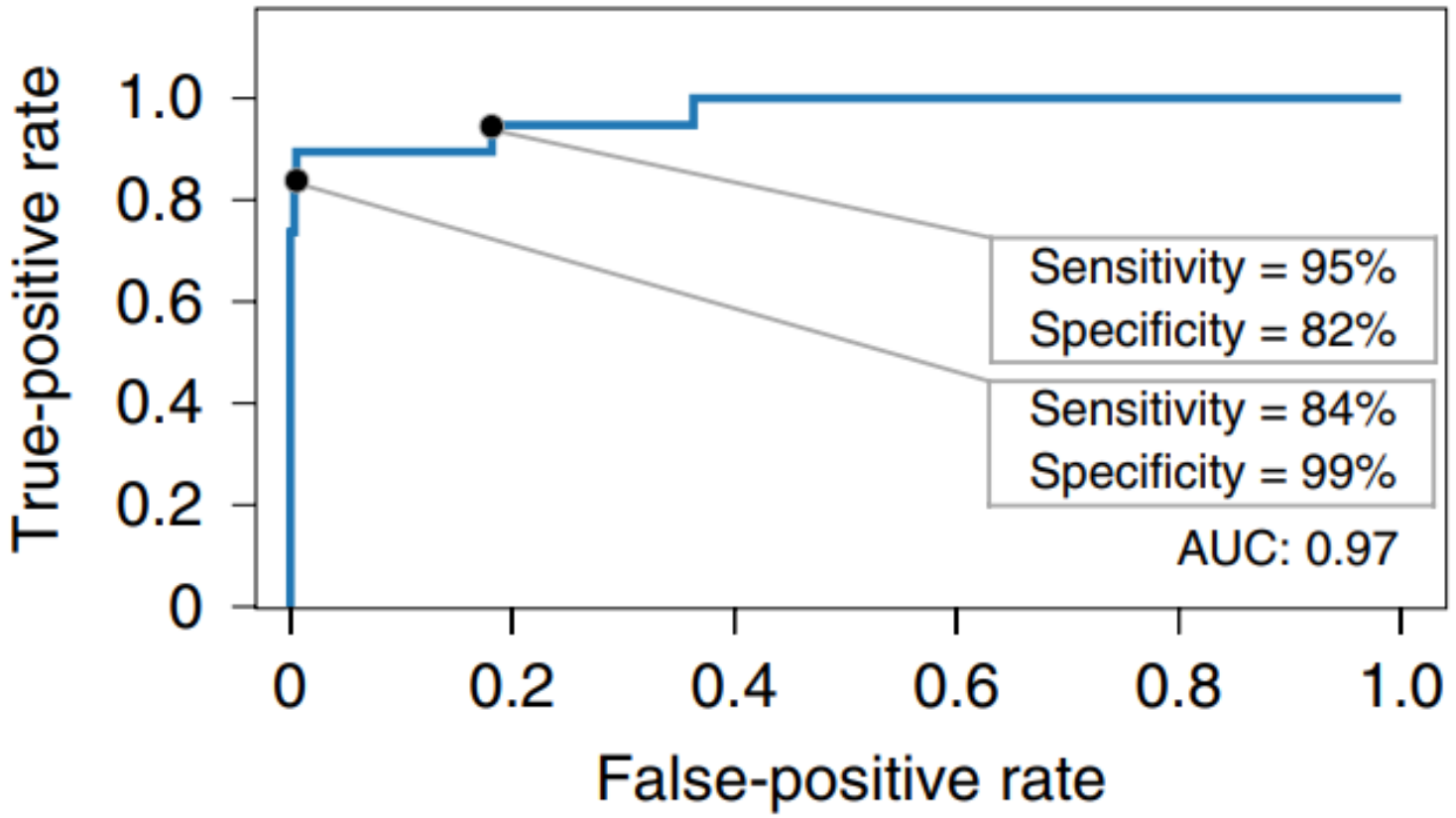
Instead we collect the cumulative signal of the fingerprint across the entire genome.

Error rate estimation in a cohort of test control plasma samples (n= 30) with and without error suppression



Error suppression and paired-end read concordance allow sequencing error reduction by a median of 21-fold
→ low error rate is essential for correct calling of the patient specific genomic fingerprint

MRDetect performance in colorectal cancer patients (n = 19)



For various clinical settings (e.g. treatment escalation/de-escalation), a different detection threshold may be relevant

Postive label: pre-operative plasma samples (n= 19)
Negative label: control plasma samples (n = 30) against all patient (n = 19) mutational compendia

Cold Spring Harbor Laboratory

bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

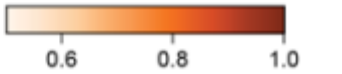
Machine learning guided signal enrichment for ultrasensitiv burden monitoring

Adam J. Widman, Minita Shah, Nadia Øgaard, Cole C. Khamnei, Amanda Fryder Anushri Arora, Mingxuan Zhang, Daniel Halmos, Jake Bass, Theophile Langanay, Sri Zoe Steinsnyder, Will Liao, Mads Heilskov Rasmussen, Sarah Østrup Jensen, Jesper Jesus Sotelo, Ryan Brand, Ronak H. Shah, Alexandre Pellan Cheng, Colleen Maher, Dennie T. Frederick, Murtaza S. Malbari, Melissa Marton, Dina Manaa, Lara Winterl Genevieve Boland, Jedd D. Wolchok, Ashish Saxena, Samra Turajlic, Marcin Imielins Nasser K. Altorki, Michael A. Postow, Nicolas Robine, Claus Lindbjerg Anderse

doi: https://doi.org/10.1101/2022.01.17.476508

MRD-Egde includes multiple features

		Used in MRDetect	Used in MRD-EDGE	LUAD	Colon	Melanoma
Mutational signature	Trinucleotide context	-	+	0.70	0.69	0.92
Regional context	ATAC-Seq accessibility	-	+	0.67	0.63	0.62
	PCAWG SNV density	-	+	0.67	0.60	0.69
	Replication timing	-	+	0.66	0.62	0.59
	RNA expression	-	+	0.63	0.60	0.64
	Plasma WGS error density	-	+	0.60	0.59	0.53
	Chromatin state	-	+	0.57	0.56	0.62
Quality metrics	Low quality bases on fragment	-	+	0.70	0.54	0.50
	Read edit distance	-	+	0.70	0.50	0.51
	Read alignment score	-	+	0.68	0.59	0.52
Fragment metrics	Variant position in read	+	+	0.61	0.59	0.50
	Fragment length	-	+	0.58	0.61	0.51



Single variable area under ROC

Cold Spring Harbor Laboratory

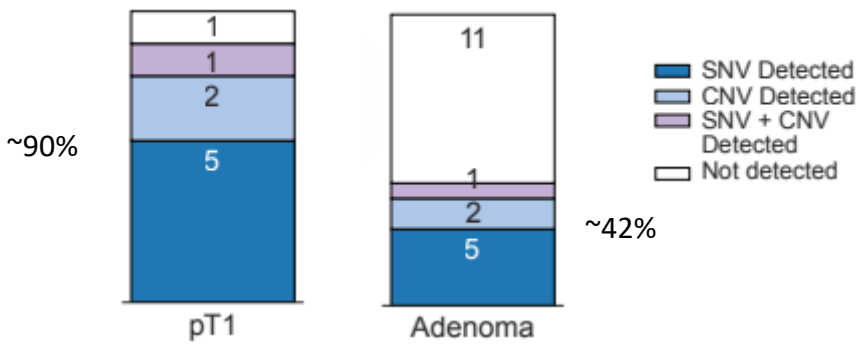
bioRxiv

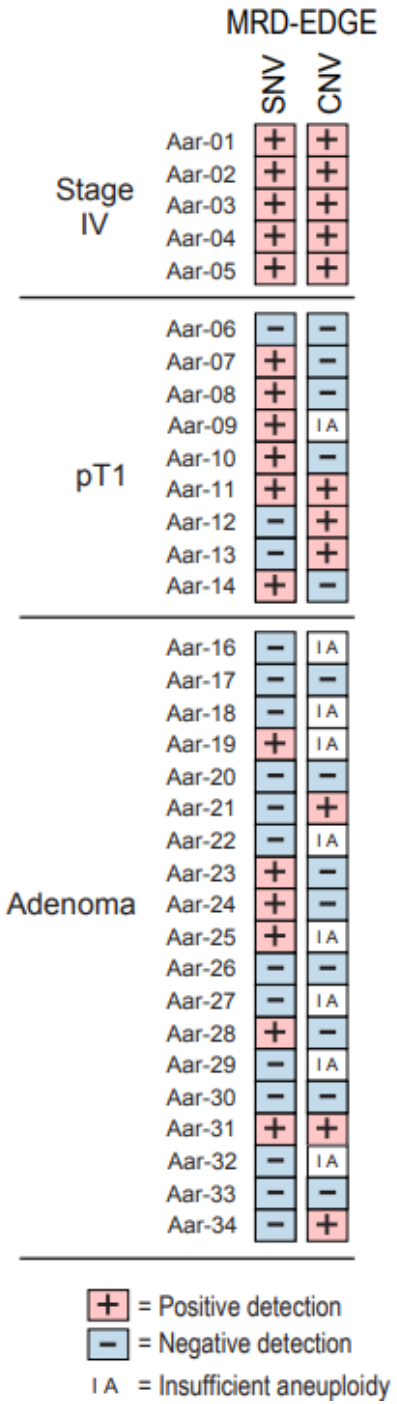
THE PREPRINT SERVER FOR BIOLOGY

Machine learning guided signal enrichment for ultrasensitive plasma tumor burden monitoring

Adam J. Widman, Minita Shah, Nadia Øgaard, Cole C. Khamnei, Amanda Frydendahl, Aditya Deshpande, Anushri Arora, Mingxuan Zhang, Daniel Halmos, Jake Bass, Theophile Langanay, Srinivas Rajagopalan, Zoe Steinsnyder, Will Liao, Mads Heilskov Rasmussen, Sarah Østrup Jensen, Jesper Nors, Christina Therkildsen, Jesus Sotelo, Ryan Brand, Ronak H. Shah, Alexandre Pellan Cheng, Colleen Maher, Lavinia Spain, Kate Krause, Dennie T. Frederick, Murtaza S. Malbari, Melissa Marton, Dina Manaa, Lara Winterkorn, Margaret K. Callahan, Genevieve Boland, Jedd D. Wolchok, Ashish Saxena, Samra Turajlic, Marcin Imielinski, Michael F. Berger, Nasser K. Altorki, Michael A. Postow, Nicolas Robine, Claus Lindbjerg Andersen, Dan A. Landau

doi: https://doi.org/10.1101/2022.01.17.476508





Take home message
MRDetect

Advantages:

- Requires just 1 ml of plasma! – Just enough genome equivalents to reach 30x depth
- Same analysis for all cancers
- Whole genome sequencing is very simple
- Lab part easy to implement
- Can be run locally at any clinical sequencing facility

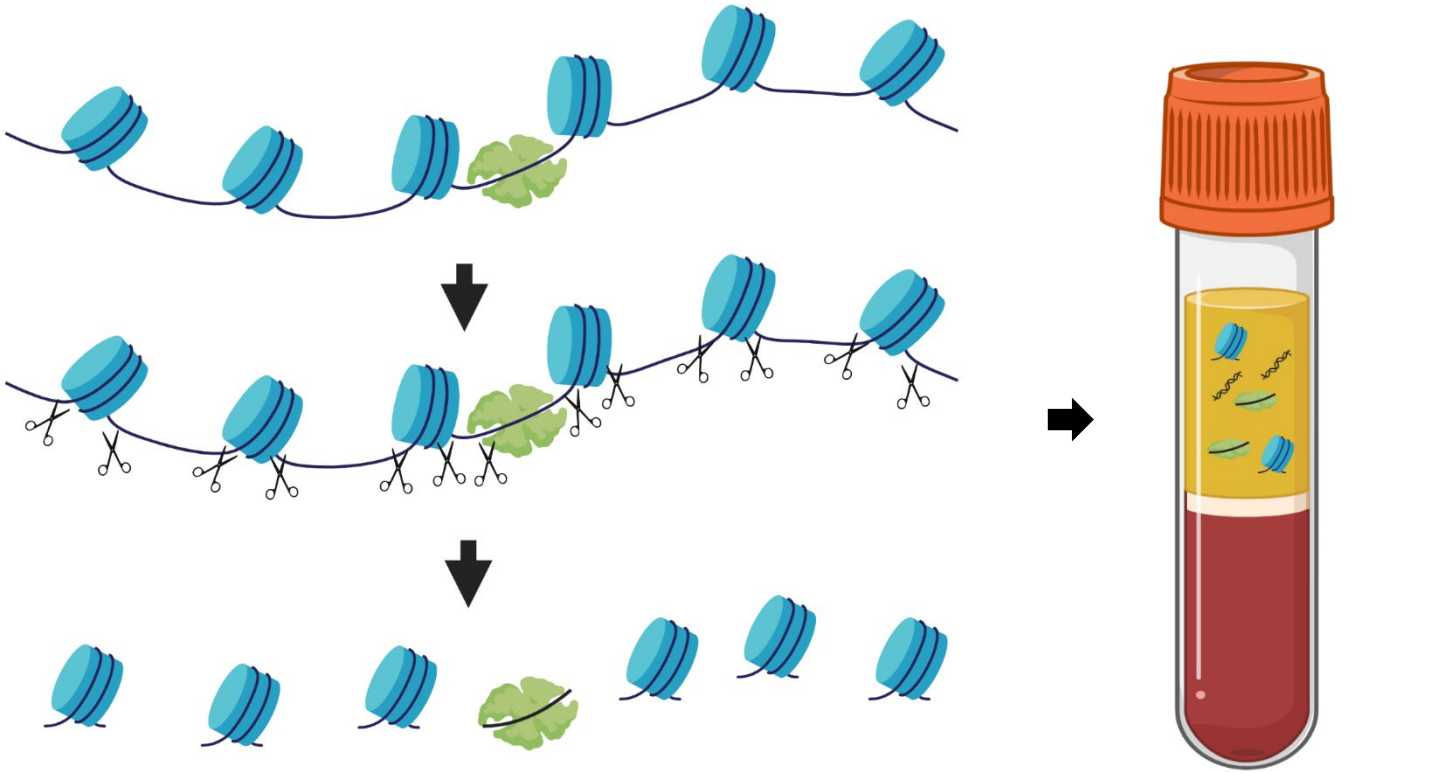
Disdvantages:

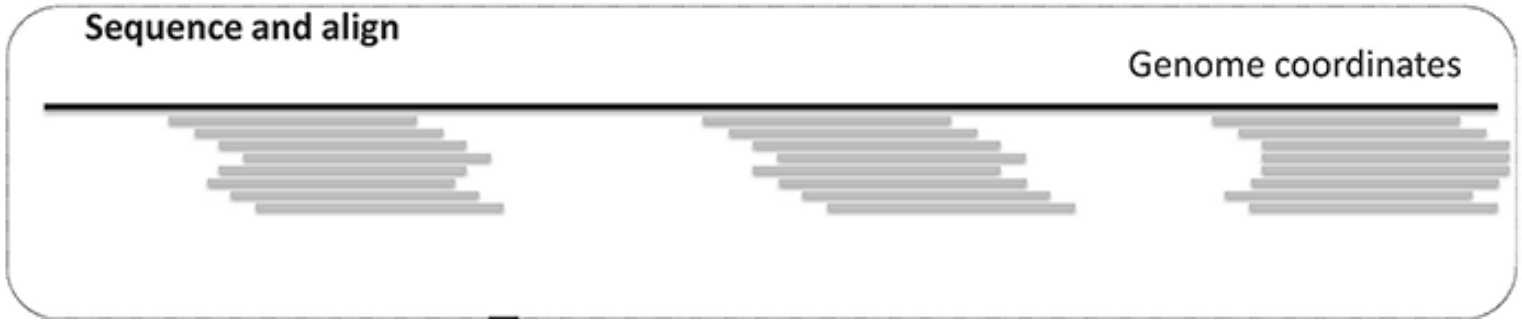
- Requires access to tumor tissue
- Cost of sequencing
- Lacks the ability to pinpoint specific mutations (targeted therapy etc.)

Non-tumor informed analysis

When DNA release into the circulation, it is fragmented and the unprotected parts are broken down

Fragmentation is non-random





Modified from Cell 164, 57–68, January 14, 2016

SCIENCE TRANSLATIONAL MEDICINE | RESEARCH ARTICLE

CANCER

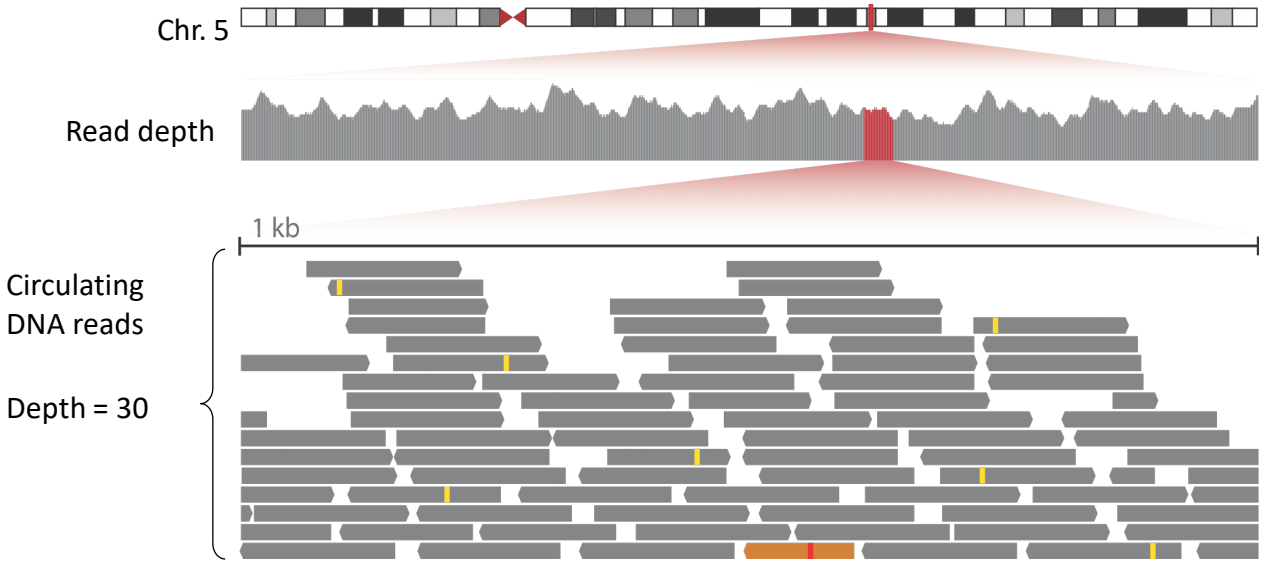
Enhanced detection of circulating tumor DNA by fragment size analysis

Florent Mouliere[1,2*†], Dineika Chandrananda[1,2*], Anna M. Piskorz[1,2*], Elizabeth K. Moore[1,2,3*], James Morris[1,2], Lise Barlebo Ahlborn[4,5], Richard Mair[1,2,6], Teodora Goranova[1,2], Francesco Marass[1,2,7,8], Katrin Heider[1,2], Jonathan C. M. Wan[1,2], Anna Supernat[1,2,9], Irena Hudecova[1,2], Ioannis Gounaris[1,2,3], Susana Ros[1,2], Mercedes Jimenez-Linan[2,3], Javier Garcia-Corbacho[10], Keval Patel[1,2], Olga Østrup[5], Suzanne Murphy[1,2], Matthew D. Eldridge[1,2], Davina Gale[1,2], Grant D. Stewart[2,3,11], Johanna Burge[2,11], Wendy N. Cooper[1,2], Michiel S. van der Heijden[12,13], Charles E. Massie[1,2,14], Colin Watts[15], Pippa Corrie[3], Simon Pacey[3,14], Kevin M. Brindle[1,2,16], Richard D. Baird[17], Morten Mau-Sørensen[4], Christine A. Parkinson[1,2,3,18,19], Christopher G. Smith[1,2], James D. Brenton[1,2,3,18,19‡§], Nitzan Rosenfeld[1,2‡§]

Mouliere *et al.*, *Sci. Transl. Med.* **10**, eaat4921 (2018) 7 November 2018

30x coverage (depth) in MRDetect

Circulating free DNA data



<u>Approach:</u>
Shallow whole genome sequencing (1x)

<u>Cohort:</u>
Healthy controls (n=65)
Cancer samples, multiple cancer types (n=284)

SCIENCE TRANSLATIONAL MEDICINE | RESEARCH ARTICLE

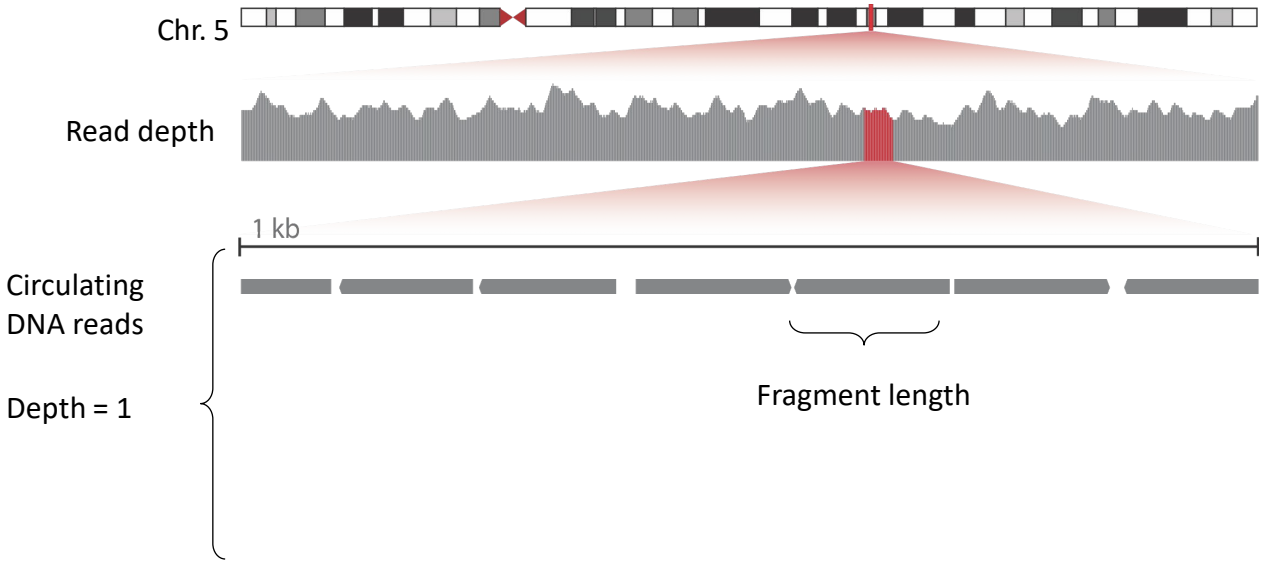
CANCER

Enhanced detection of circulating tumor DNA by fragment size analysis

Florent Mouliere[1,2*†], Dineika Chandrananda[1,2*], Anna M. Piskorz[1,2*], Elizabeth K. Moore[1,2,3*], James Morris[1,2], Lise Barlebo Ahlborn[4,5], Richard Mair[1,2,6], Teodora Goranova[1,2], Francesco Marass[1,2,7,8], Katrin Heider[1,2], Jonathan C. M. Wan[1,2], Anna Supernat[1,2,9], Irena Hudecova[1,2], Ioannis Gounaris[1,2,3], Susana Ros[1,2], Mercedes Jimenez-Linan[2,3], Javier Garcia-Corbacho[10], Keval Patel[1,2], Olga Østrup[5], Suzanne Murphy[1,2], Matthew D. Eldridge[1,2], Davina Gale[1,2], Grant D. Stewart[2,3,11], Johanna Burge[2,11], Wendy N. Cooper[1,2], Michiel S. van der Heijden[12,13], Charles E. Massie[1,2,14], Colin Watts[15], Pippa Corrie[3], Simon Pacey[3,14], Kevin M. Brindle[1,2,16], Richard D. Baird[17], Morten Mau-Sørensen[4], Christine A. Parkinson[1,2,3,18,19], Christopher G. Smith[1,2], James D. Brenton[1,2,3,18,19‡§], Nitzan Rosenfeld[1,2‡§]

Mouliere *et al.*, *Sci. Transl. Med.* **10**, eaat4921 (2018) 7 November 2018

Circulating free DNA data



Approach:

Shallow whole genome sequencing (1x)

Cohort:

Healthy controls (n=65)

Cancer samples (n=284)

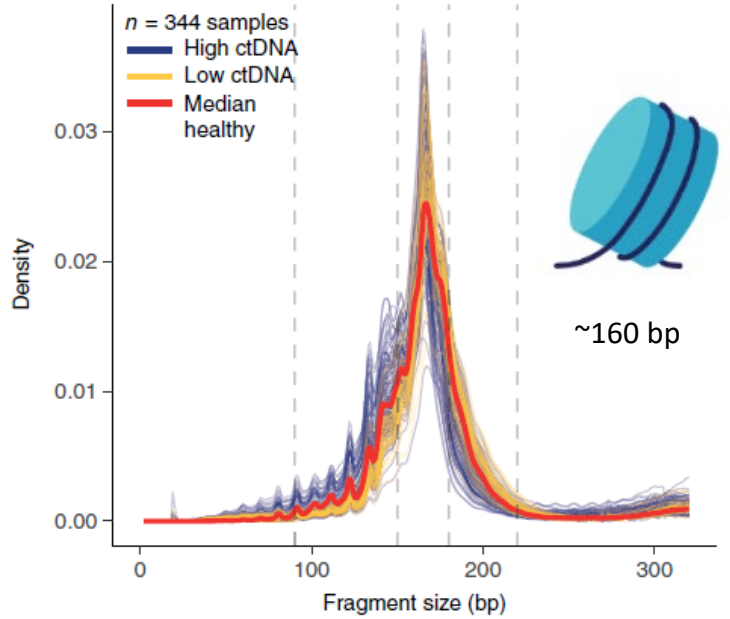
SCIENCE TRANSLATIONAL MEDICINE | RESEARCH ARTICLE

CANCER

Enhanced detection of circulating tumor DNA by fragment size analysis

Florent Mouliere[1,2*†], Dineika Chandrananda[1,2*], Anna M. Piskorz[1,2*], Elizabeth K. Moore[1,2,3*], James Morris[1,2], Lise Barlebo Ahlborn[4,5], Richard Mair[1,2,6], Teodora Goranova[1,2], Francesco Marass[1,2,7,8], Katrin Heider[1,2], Jonathan C. M. Wan[1,2], Anna Supernat[1,2,9], Irena Hudecova[1,2], Ioannis Gounaris[1,2,3], Susana Ros[1,2], Mercedes Jimenez-Linan[2,3], Javier Garcia-Corbacho[10], Keval Patel[1,2], Olga Østrup[5], Suzanne Murphy[1,2], Matthew D. Eldridge[1,2], Davina Gale[1,2], Grant D. Stewart[2,3,11], Johanna Burge[2,11], Wendy N. Cooper[1,2], Michiel S. van der Heijden[12,13], Charles E. Massie[1,2,14], Colin Watts[15], Pippa Corrie[3], Simon Pacey[3,14], Kevin M. Brindle[1,2,16], Richard D. Baird[17], Morten Mau-Sørensen[4], Christine A. Parkinson[1,2,3,18,19], Christopher G. Smith[1,2], James D. Brenton[1,2,3,18,19‡§], Nitzan Rosenfeld[1,2‡§]

Mouliere *et al.*, *Sci. Transl. Med.* **10**, eaat4921 (2018) 7 November 2018



Approach:

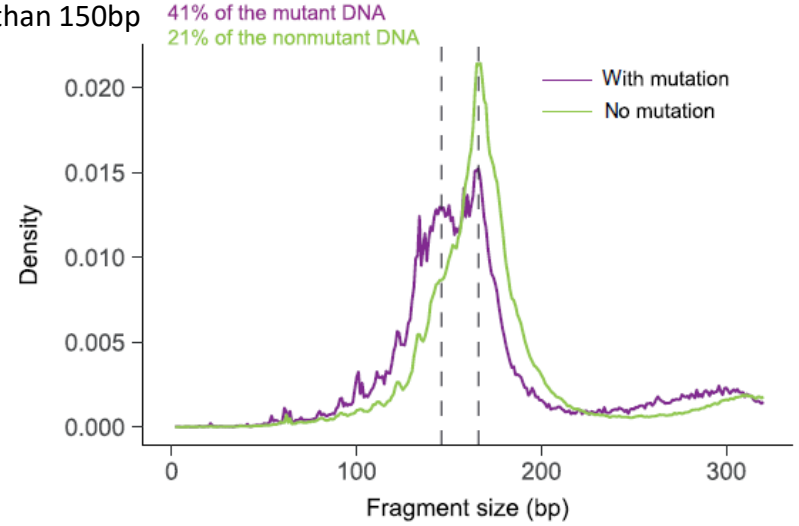
Shallow whole genome sequencing (1x)

Cohort:

Healthy controls (n=65)

Cancer samples (n=284)

Fraction of fragments shorter than 150bp

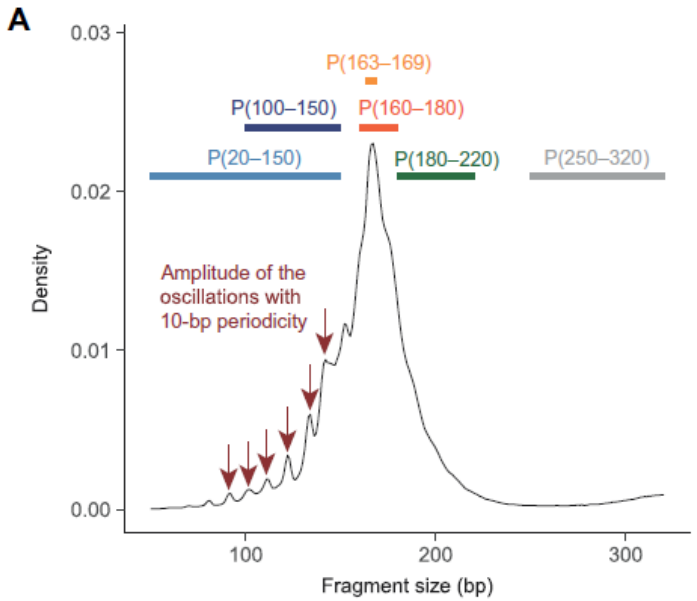


ctDNA Feature selection:

Fragment lengths

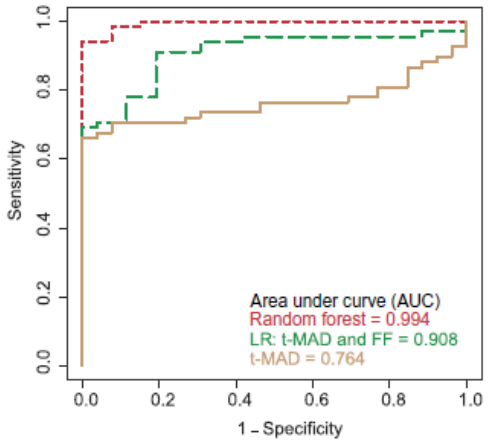
10 bp Osciliation

CNA (coverage skewness)

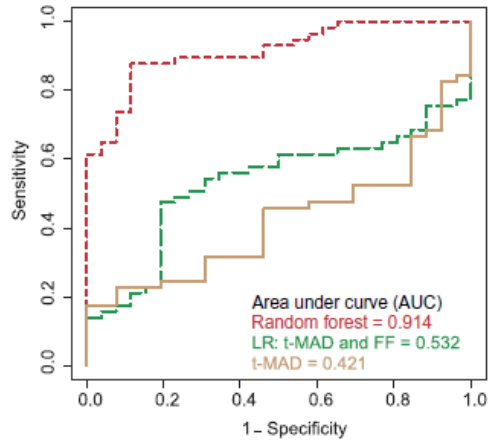


Independent validation (case/control)

Cancer types with
HIGH ctDNA levels



Cancer types with
LOW ctDNA levels

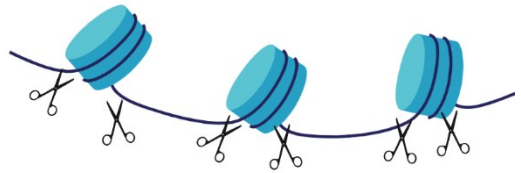


Preferred end coordinates and somatic variants as signatures of circulating tumor DNA associated with hepatocellular carcinoma

Peiyong Jiang[a,b,1], Kun Sun[a,b,1], Yu K. Tong[a,b], Suk Hang Cheng[a,b], Timothy H. T. Cheng[a,b], Macy M. S. Heung[a,b], John Wong[c], Vincent W. S. Wong[d,e], Henry L. Y. Chan[d,e], K. C. Allen Chan[a,b,f], Y. M. Dennis Lo[a,b,f,2], and Rossa W. K. Chiu[a,b,2]

[a]Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Shatin, New Terri

Proc Natl Acad Sci. 2018 Nov 13;115(46):E10925-E10933



Approaches:

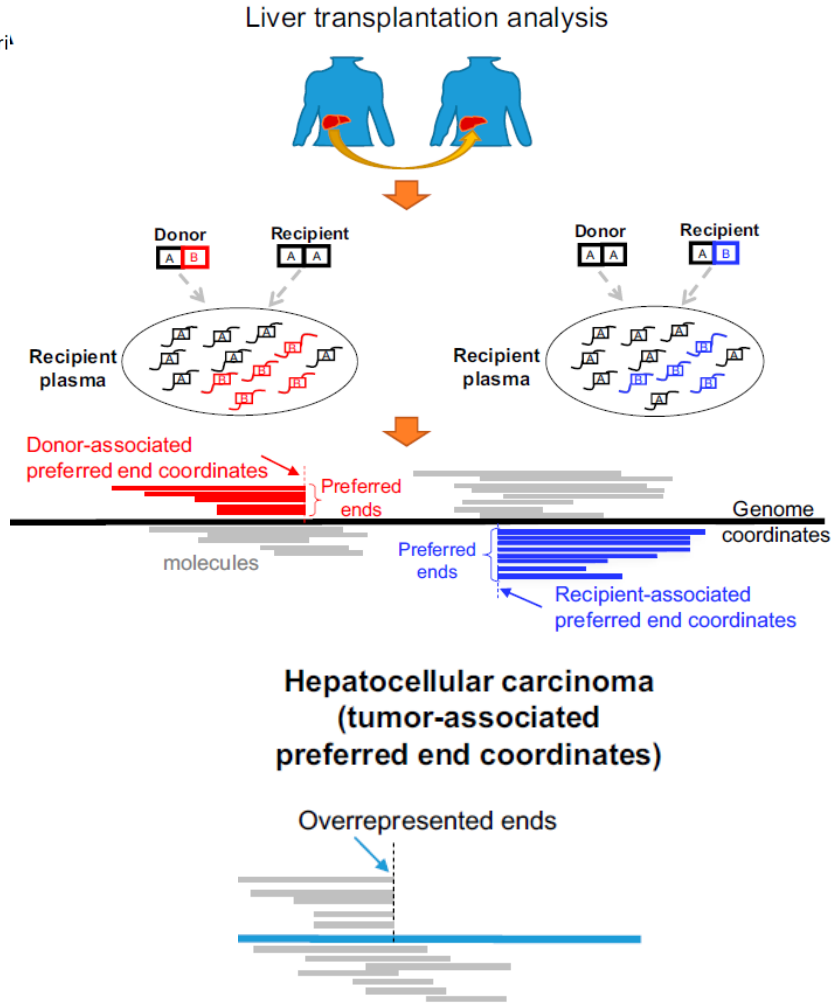
Deep whole genome sequencing (220x):

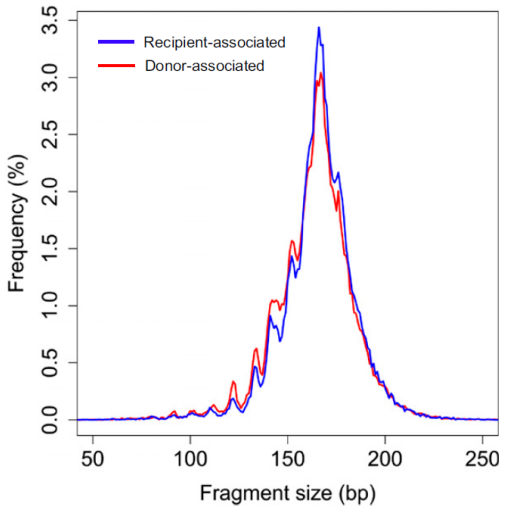
1 liver transplant patient

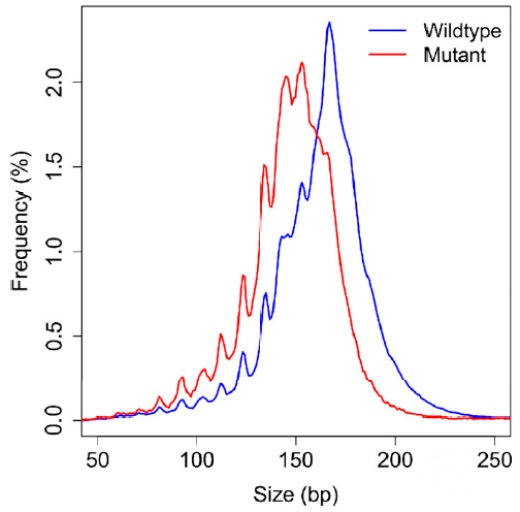
1 Hepatocellular carcinoma (HCC) patient

Shallow whole genome sequencing (1x):

- 32 healthy subjects
- 67 chronic Chronic Hepatitis B virus (HBV) carriers without cirrhosis
- 36 patients with HBV-related liver cirrhosis
- 90 patients with Hepatocellular carcinoma



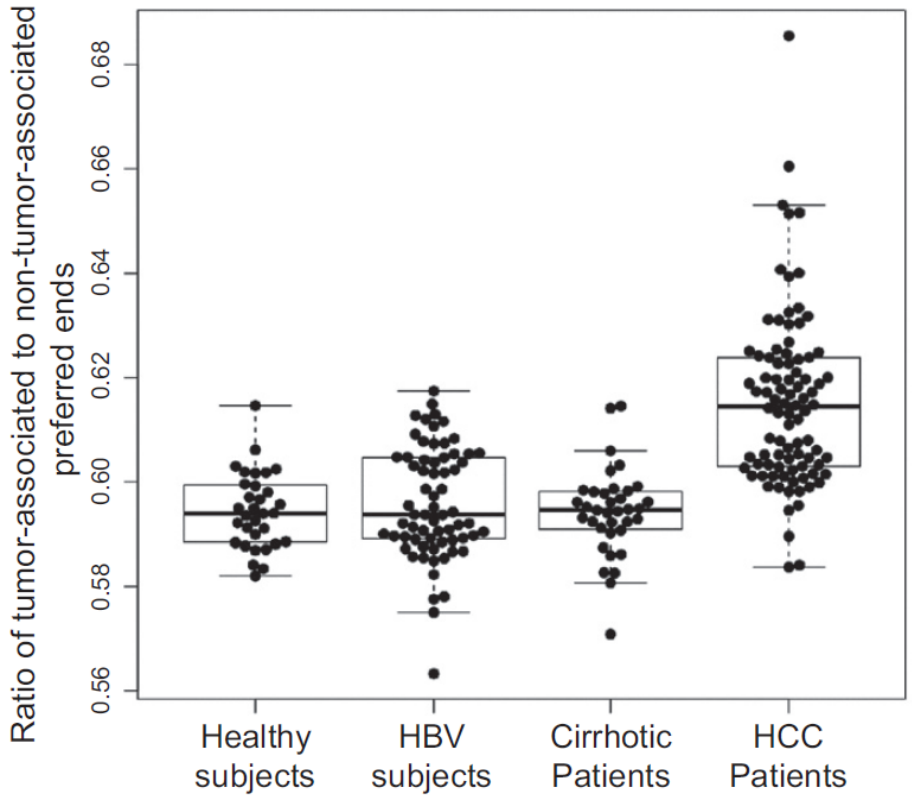


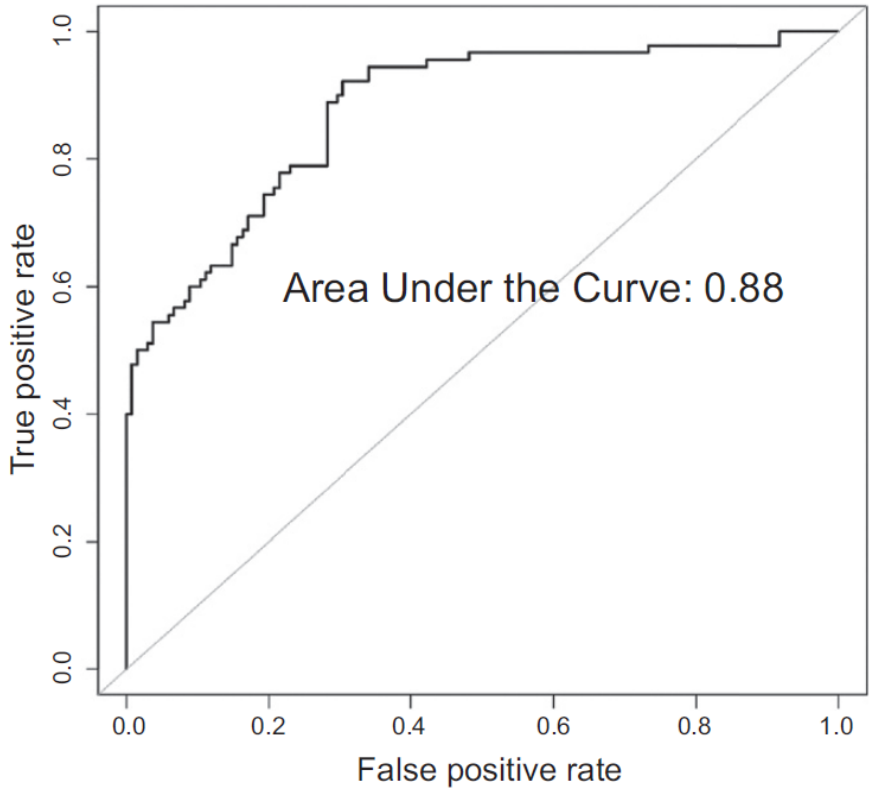


Preferred end coordinates and somatic variants as signatures of circulating tumor DNA associated with hepatocellular carcinoma

Peiyong Jiang[a,b,1], Kun Sun[a,b,1], Yu K. Tong[a,b], Suk Hang Cheng[a,b], Timothy H. T. Cheng[a,b], Macy M. S. Heung[a,b], John Wong[c], Vincent W. S. Wong[d,e], Henry L. Y. Chan[d,e], K. C. Allen Chan[a,b,f], Y. M. Dennis Lo[a,b,f,2], and Rossa W. K. Chiu[a,b,2]

[a]Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR, China; [b]Department of Chemical





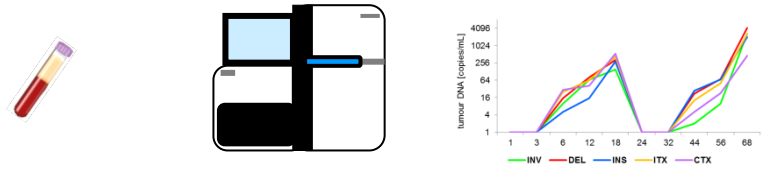
DELFI: DNA evaluation of fragments for early interception

LETTER 20 JUNE 2019 | VOL 570 | NATURE | 385
https://doi.org/10.1038/s41586-019-1272-6

Genome-wide cell-free DNA fragmentation in patients with cancer

Stephen Cristiano[1,2,15], Alessandro Leal[1,15], Jillian Phallen[1,15], Jacob Fiksel[1,2,15], Vilmos Adleff[1], Daniel C. Bruhm[1], Sarah Østrup Jensen[3], Jamie E. Medina[1], Carolyn Hruban[1], James R. White[1], Doreen N. Palsgrove[1], Noushin Niknafs[1], Valsamo Anagnostou[1], Patrick Forde[1], Jarushka Naidoo[1], Kristen Marrone[1], Julie Brahmer[1], Brian D. Woodward[4], Hatim Husain[4], Karlijn L. van Rooijen[5], Mai-Britt Worm Ørntoft[3], Anders Husted Madsen[6], Cornelis J. H. van de Velde[7], Marcel Verheij[8], Annemieke Cats[9], Cornelis J. A. Punt[10], Geraldine R. Vink[5], Nicole C. T. van Grieken[11], Miriam Koopman[5], Remond J. A. Fijneman[12], Julia S. Johansen[13], Hans Jørgen Nielsen[14], Gerrit A. Meijer[12], Claus Lindbjerg Andersen[3], Robert B. Scharpf[1,2]* & Victor E. Velculescu[1]*

Approach: Low pass WGS of plasma cfDNA (1x)



Plasma | **Shallow Whole genome seq** | **Plasma ctDNA profiles**

Observation: Cancer-derived cfDNA fragment lengths are more variable than non-cancer cfDNA fragments

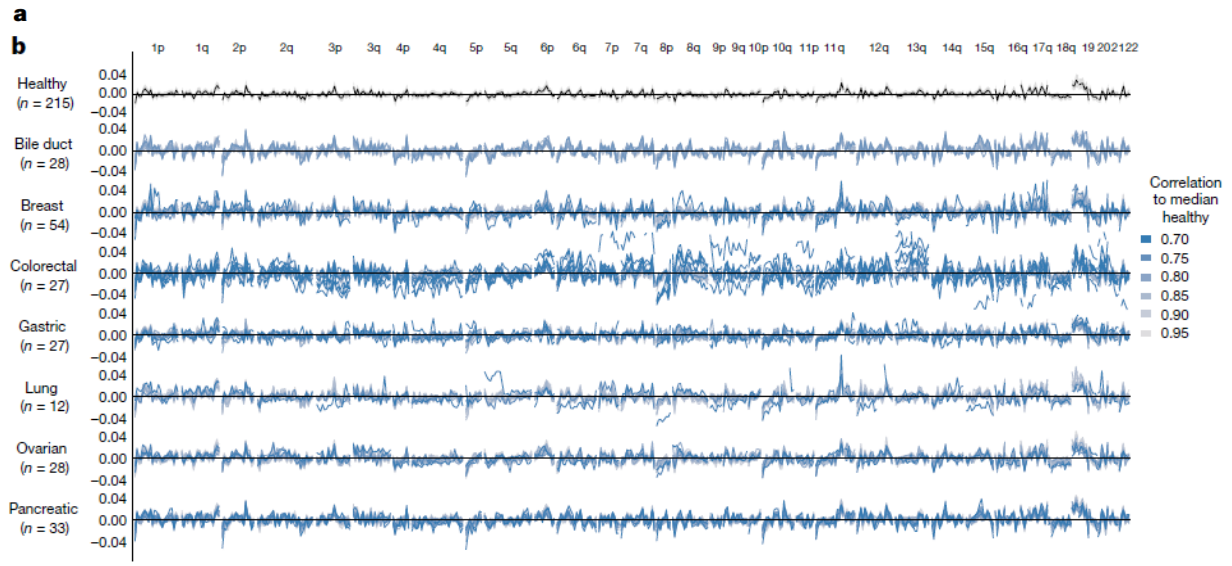
Hypothesis:

cfDNA fragmentation can serve as a biomarker for cancer detection

cfDNA fragmentation was measured as the coverage ratio:

$$\frac{\text{Short fragments (100-150 bp)}}{\text{Long fragments (151-220 bp)}}$$

Research subjects: Controls: 215
Cancers: 208 (7 different cancer types)
1 mL of plasma



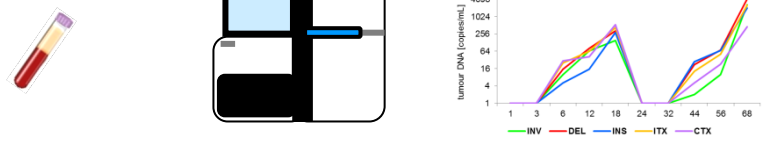
DELFI: DNA evaluation of fragments for early interception

LETTER 20 JUNE 2019 | VOL 570 | NATURE | 385
https://doi.org/10.1038/s41586-019-1272-6

Genome-wide cell-free DNA fragmentation in patients with cancer

Stephen Cristiano[1,2,15], Alessandro Leal[1,15], Jillian Phallen[1,15], Jacob Fiksel[1,2,15], Vilmos Adleff[1], Daniel C. Bruhm[1], Sarah Østrup Jensen[3], Jamie E. Medina[1], Carolyn Hruban[1], James R. White[1], Doreen N. Palsgrove[1], Noushin Niknafs[1], Valsamo Anagnostou[1], Patrick Forde[1], Jarushka Naidoo[1], Kristen Marrone[1], Julie Brahmer[1], Brian D. Woodward[4], Hatim Husain[4], Karlijn L. van Rooijen[5], Mai-Britt Worm Ørntoft[3], Anders Husted Madsen[6], Cornelis J. H. van de Velde[7], Marcel Verheij[8], Annemieke Cats[9], Cornelis J. A. Punt[10], Geraldine R. Vink[5], Nicole C. T. van Grieken[11], Miriam Koopman[5], Remond J. A. Fijneman[12], Julia S. Johansen[13], Hans Jørgen Nielsen[14], Gerrit A. Meijer[12], Claus Lindbjerg Andersen[3], Robert B. Scharpf[1,2]* & Victor E. Velculescu[1]*

Approach: Low pass WGS of plasma cfDNA



Plasma | **Shallow Whole genome seq** | **Plasma ctDNA profiles**

Observation: Cancer-derived cfDNA fragment lengths are more variable than non-cancer cfDNA fragments

Hypothesis:

cfDNA fragmentation can serve as a biomarker for cancer detection

Results:

Machine learning classifier:

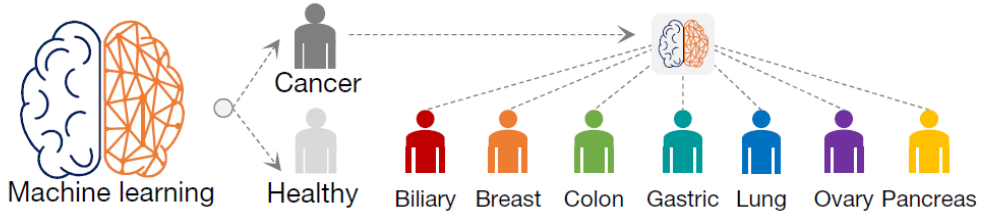
- **Fragmentation pattern**
- Copy number changes
- Mitochrondrial copy number changes

Sensitivity for CRC: 81%

stage I 73%
stage II 78%

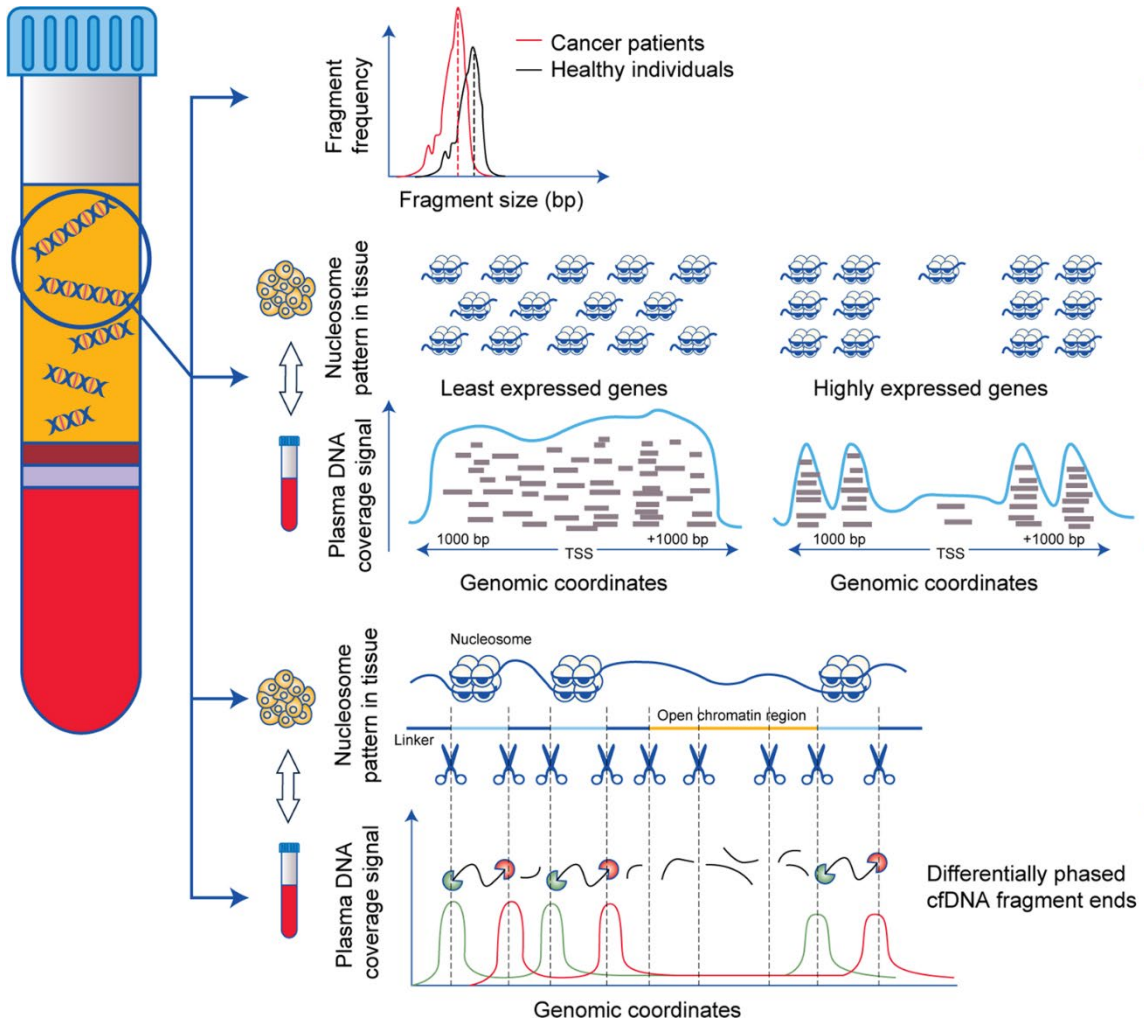
at 95% specificity

Tumor tissue of origin



75% accuracy at assigning the two most likely tissues of origin

Mutational signatures (personal/general) → Detection of cancer
Monitoring tumor burden



→ Early detection of cancer

→ Enhanced sensitivity for mutation detection

→ Inferring cancer driver genes expression

→ Inferring cfDNA tissue of origin

Take home message

cfDNA Fragment length and fragment pattern strategies

Advantages:

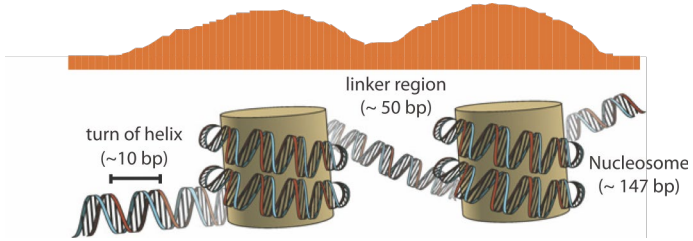
- Many markers
- Utilize the whole genome
- Speed
- Cost
- Indicates tissue of origin (chromatin organization)
- Generalizability (same test can be applied to “all” cancers)

Disadvantages:

- Specificity ??
- New territory – robust callers integrating features are being developed

Tumor fragments

Normal fragments



Group work

- List the differences between targeted and whole genome sequencing you can remember

- Which potential clinical applications do you see for whole genome sequencing?