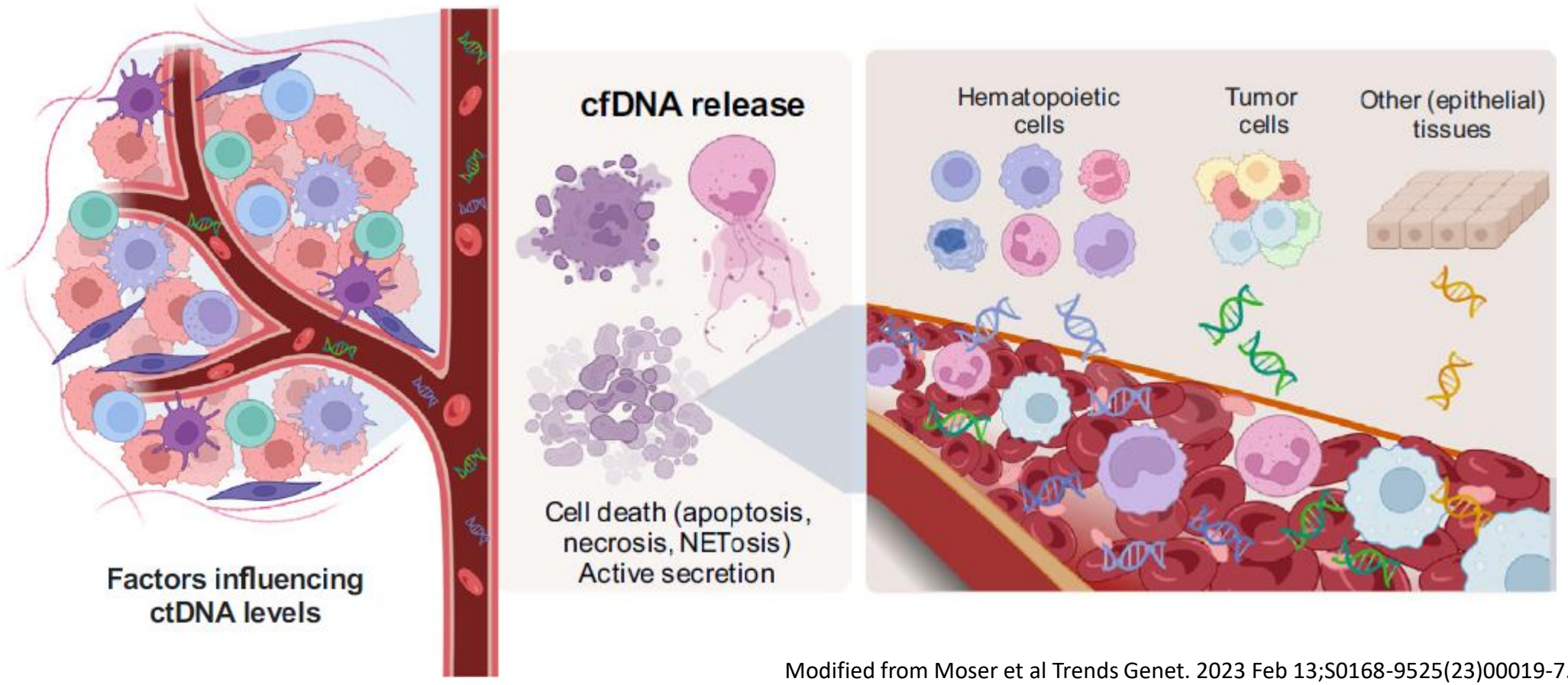
” Ultra-sensitiv cirkulerende tumor DNA detektion - muliggjort gennem anvendelse af maskinlæring til reduktion af DNA sekventeringsstøj ”

Professor Claus Lindbjerg Andersen
Department of Molecular Medicine
Director of the National Danish ctDNA Research Center
Aarhus University Hospital
Aarhus University
Aarhus, Denmark
Email: cla@clin.au.dk

Sources of cell free DNA



Modified from Moser et al Trends Genet. 2023 Feb 13;S0168-9525(23)00019-7.

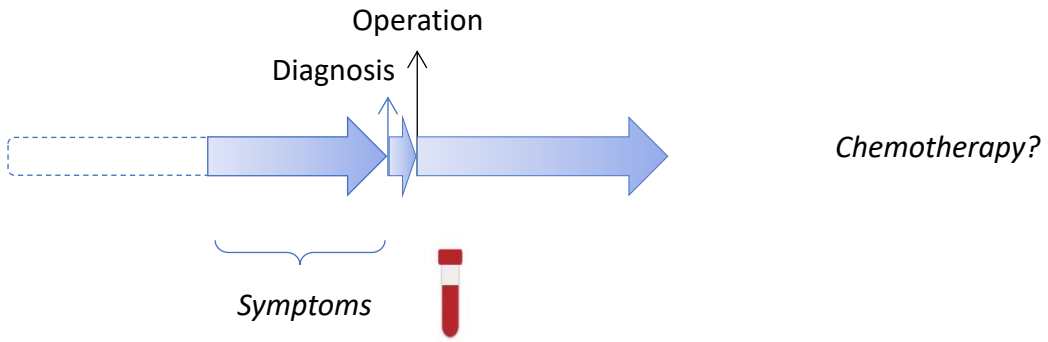
Annals of Oncology
Volume 35, Issue 2, February 2024, Pages 229-239

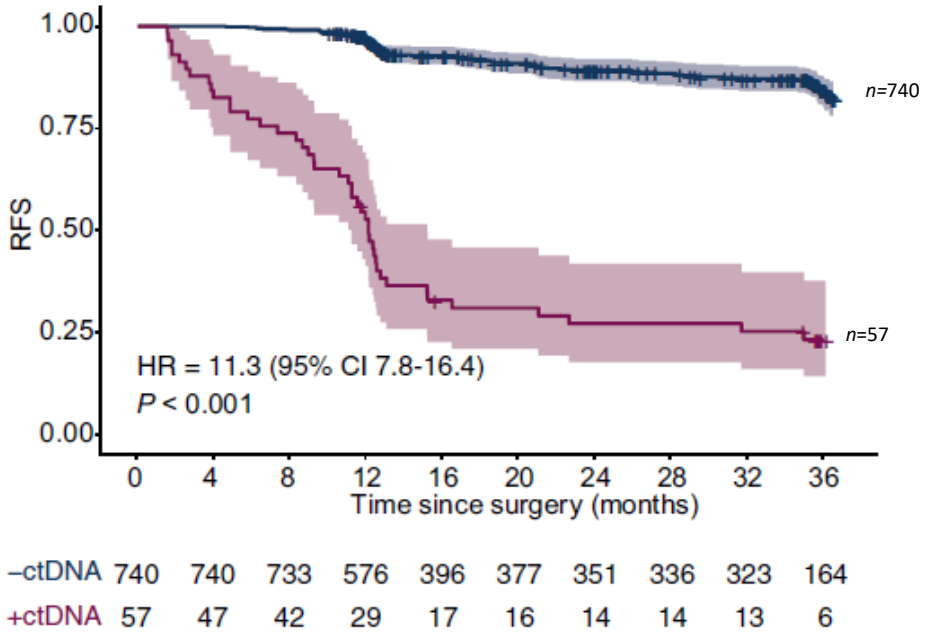
Original article

Unraveling the potential clinical utility of circulating tumor DNA detection in colorectal cancer—evaluation in a nationwide Danish cohort

T.V. Henriksen [1,2], C. Demuth [1,2], A. Frydendahl [1,2], J. Nors [1,2], M. Nesic [1,2], M.H. Rasmussen [1,2], T. Reinert [1,2], O.H. Larsen [1,2], C. Jaensch [3], U.S. Løve [4], P.V. Andersen [5], T. Kolbro [6], O. Thorlacius-Ussing [7], A. Monti [8], M. Gögenur [9], J. Kildsig [10], P. Bondeven [11], N.H. Schlesinger [12], L.H. Iversen [13], K.A. Gotschalck [14]...C.L. Andersen [1,2]

Clinical challenge:
Who is at risk of relapsing ?



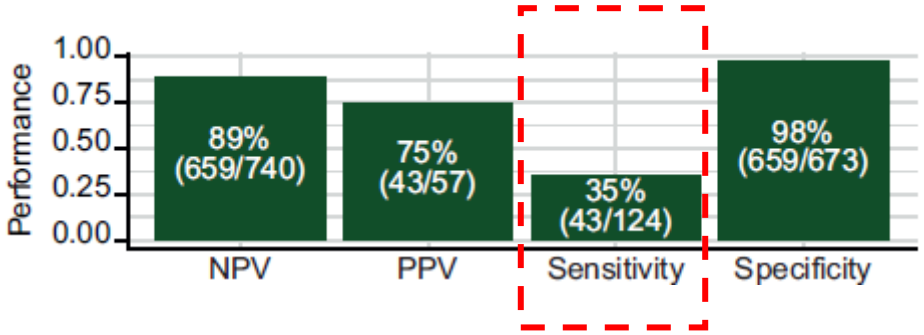


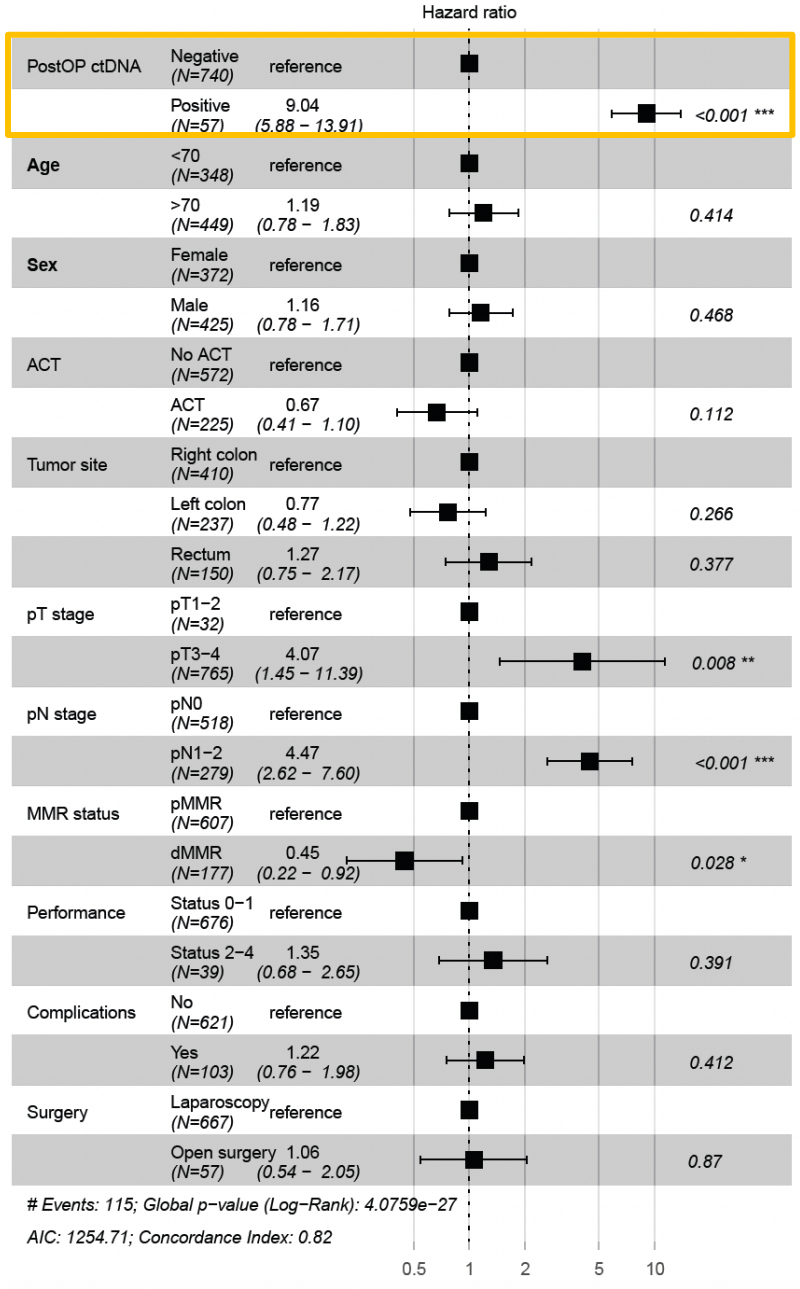
Annals of Oncology
Volume 35, Issue 2, February 2024, Pages 229-239

Original article

Unraveling the potential clinical utility of circulating tumor DNA detection in colorectal cancer—evaluation in a nationwide Danish cohort

T.V. Henriksen [1,2], C. Demuth [1,2], A. Frydendahl [1,2], J. Nors [1,2], M. Nesic [1,2], M.H. Rasmussen [1,2], T. Reinert [1,2], O.H. Larsen [1,2], C. Jaensch [3], U.S. Løve [4], P.V. Andersen [5], T. Kolbro [6], O. Thorlacius-Ussing [7], A. Monti [8], M. Gögenur [9], J. Kildsig [10], P. Bondeven [11], N.H. Schlesinger [12], L.H. Iversen [13], K.A. Gotschalck [14]...C.L. Andersen [1,2]





Whole genome sequencing based tumor-informed ctDNA detection

nature medicine

Explore content ▾ About the journal ▾ Publish with us ▾

nature › nature medicine › articles › article

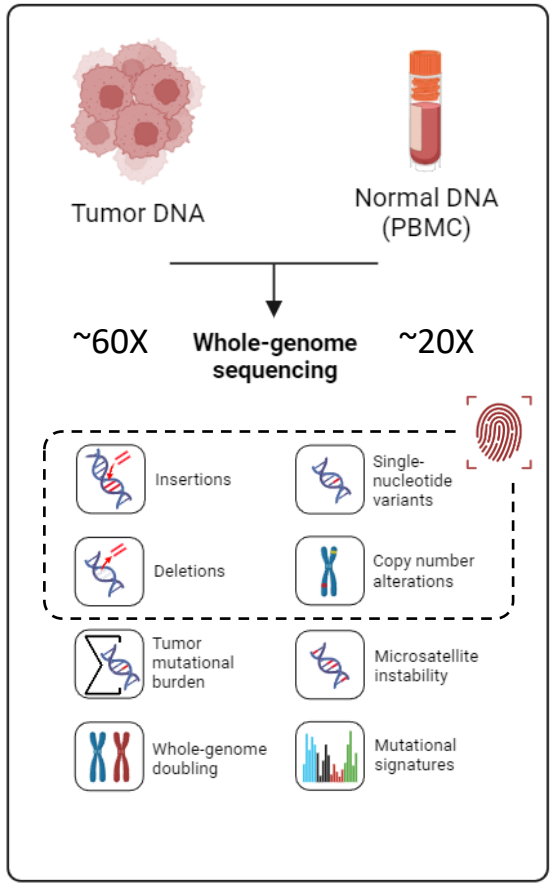
Article | Published: 14 June 2024

Ultrasensitive plasma-based monitoring of tumor burden using machine-learning-guided signal enrichment

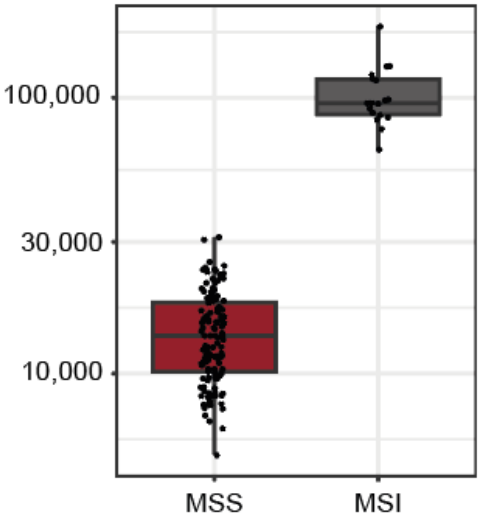
Adam J. Widman, Minita Shah, Amanda Frydendahl, Daniel Halmos, Cole C. Khamnei, Nadia Øgaard, Srinivas Rajagopalan, Anushri Arora, Aditya Deshpande, William F. Hooper, Jean Quentin, Jake Bass, Mingxuan Zhang, Theophile Langanay, Laura Andersen, Zoe Steinsnyder, Will Liao, Mads Heilskov Rasmussen, Tenna Vesterman Henriksen, Sarah Østrup Jensen, Jesper Nors, Christina Therkildsen, Jesus Imielinski, Michael F. Berger, Sam Aparicio, Nasser K. Altorki, Michael A. Postow, Nicolas Robine, Claus Lindbjerg Andersen & Dan A. Landau — Show fewer authors

Nature Medicine **30**, 1655–1666 (2024) | Cite this article

9036 Accesses | **4** Citations | **253** Altmetric | Metrics



Markers in the mutational fingerprint



Whole genome sequencing based tumor-informed ctDNA detection

nature medicine

Explore content ▾ About the journal ▾ Publish with us ▾

nature › nature medicine › articles › article

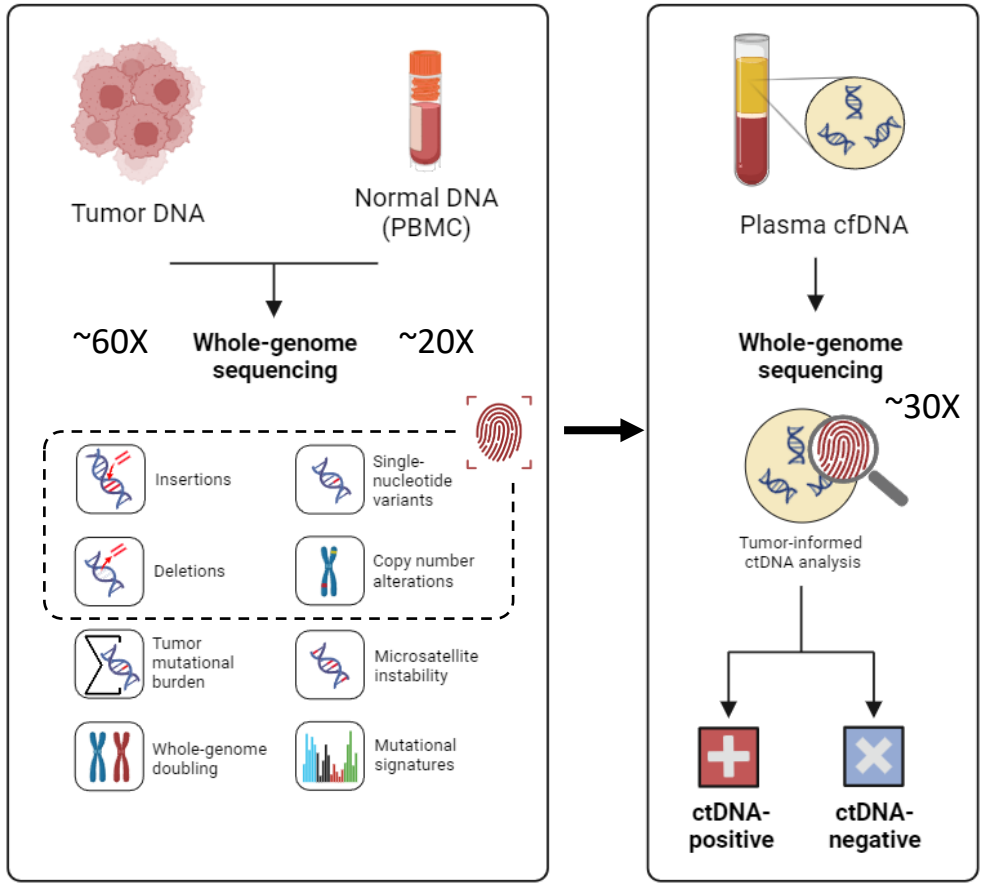
Article | Published: 14 June 2024

Ultrasensitive plasma-based monitoring of tumor burden using machine-learning-guided signal enrichment

Adam J. Widman, Minita Shah, Amanda Frydendahl, Daniel Halmos, Cole C. Khamnei, Nadia Øgaard, Srinivas Rajagopalan, Anushri Arora, Aditya Deshpande, William F. Hooper, Jean Quentin, Jake Bass, Mingxuan Zhang, Theophile Langanay, Laura Andersen, Zoe Steinsnyder, Will Liao, Mads Heilskov Rasmussen, Tenna Vesterman Henriksen, Sarah Østrup Jensen, Jesper Nors, Christina Therkildsen, Jesus Imielinski, Michael F. Berger, Sam Aparicio, Nasser K. Altorki, Michael A. Postow, Nicolas Robine, Claus Lindbjerg Andersen & Dan A. Landau — Show fewer authors

Nature Medicine **30**, 1655–1666 (2024) | Cite this article

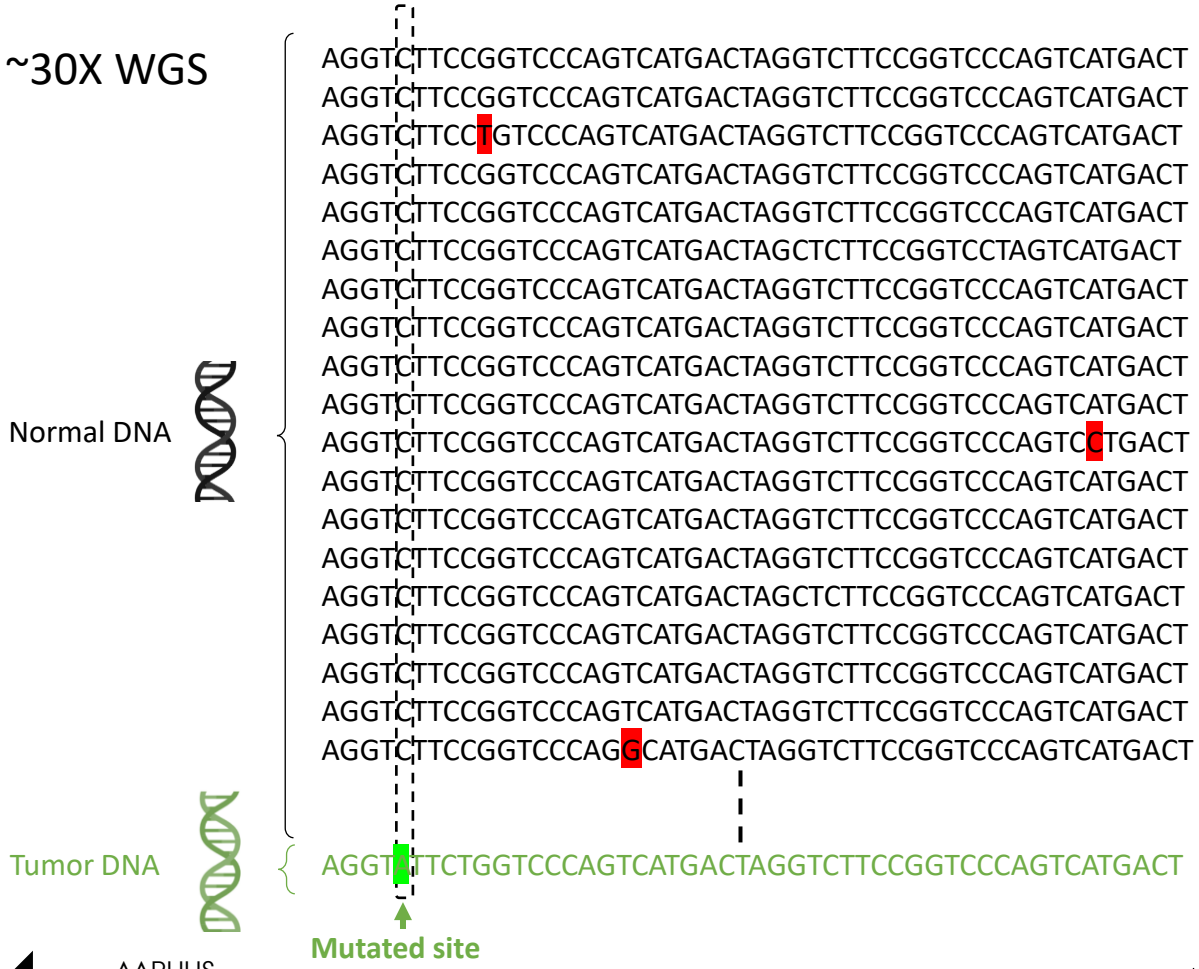
9036 Accesses | **4** Citations | **253** Altmetric | Metrics



Average
730 mill reads per sample

The prominent obstacle to WGS-based detection of ctDNA SNVs

Distinguishing True Tumor Mutations from **far more abundant** Sequencing Errors (Error rate = 1e-4)



With
- Tumor fractions << 1%
- ~30X sequencing depth

Each **mutated site** is at best supported by **a single fragment**.

Consequence:

Cannot rely on classical **consensus** **mutation calling** to identify ctDNA SNVs

Whole-genome sequencing for ctDNA detection

nature medicine

Explore content · About the journal · Publish with us

nature › nature medicine › articles › article

Article | Published: 14 June 2024

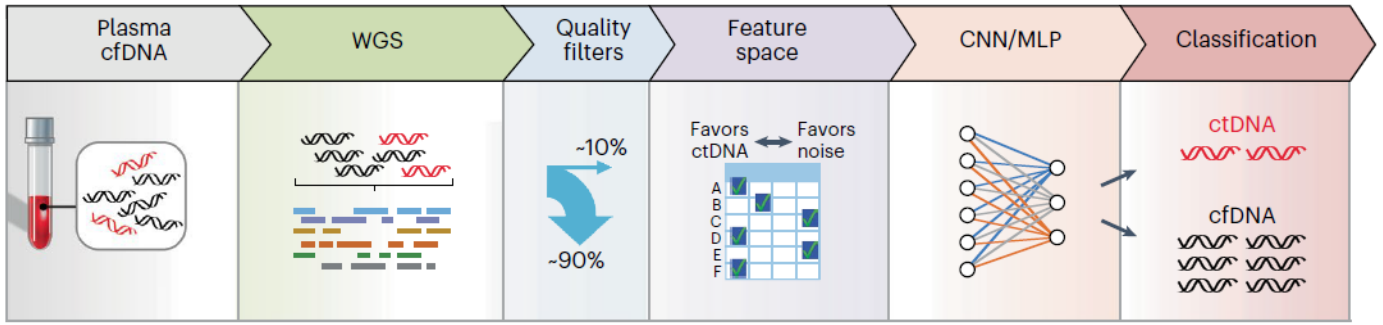
Ultrasensitive plasma-based monitoring of tumor burden using machine-learning-guided signal enrichment

Adam J. Widman, Minita Shah, Amanda Frydendahl, Daniel Halmos, Cole C. Khamnei, Nadia Øgaard, Srinivas Rajagopalan, Anushri Arora, Aditya Deshpande, William F. Hooper, Jean Quentin, Jake Bass, Mingxuan Zhang, Theophile Langanay, Laura Andersen, Zoe Steinsnyder, Will Liao, Mads Heilskov Rasmussen, Tenna Vesterman Henriksen, Sarah Østrup Jensen, Jesper Nors, Christina Therkildsen, Jesus Imielinski, Michael F. Berger, Sam Aparicio, Nasser K. Altorki, Michael A. Postow, Nicolas Robine, Claus Lindbjerg Andersen & Dan A. Landau — Show fewer authors

Nature Medicine **30**, 1655–1666 (2024) | Cite this article

9036 Accesses | **4** Citations | **253** Altmetric | Metrics

Fragment level classifier



Tumor informed ctDNA analysis

Tumor SNV detection rate:

$$\frac{\text{Number of SNVs detected in cfDNA}}{\text{Number of reads checked}}$$

Features include:

- Alignment quality
- Trinucleotide contexts
- Variant position in read
- Fragments length
- …and more…

Whole-genome sequencing for ctDNA detection

nature medicine

Explore content ∨ About the journal ∨ Publish with us ∨

nature › nature medicine › articles › article

Article | Published: 14 June 2024

Ultrasensitive plasma-based monitoring of tumor burden using machine-learning-guided signal enrichment

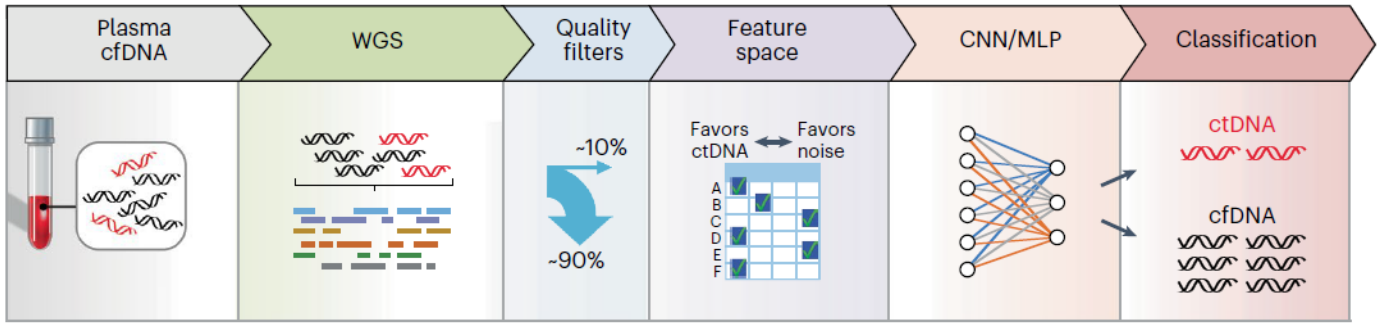
Adam J. Widman, Minita Shah, Amanda Frydendahl, Daniel Halmos, Cole C. Khamnei, Nadia Øgaard, Srinivas Rajagopalan, Anushri Arora, Aditya Deshpande, William F. Hooper, Jean Quentin, Jake Bass, Mingxuan Zhang, Theophile Langanay, Laura Andersen, Zoe Steinsnyder, Will Liao, Mads Heilskov Rasmussen, Tenna Vesterman Henriksen, Sarah Østrup Jensen, Jesper Nors, Christina Therkildsen, Jesus Imielinski, Michael F. Berger, Sam Aparicio, Nasser K. Altorki, Michael A. Postow, Nicolas Robine, Claus Lindbjerg Andersen & Dan A. Landau — Show fewer authors

Nature Medicine **30**, 1655–1666 (2024) | Cite this article

9036 Accesses | **4** Citations | **253** Altmetric | Metrics

Sample ctDNA calls:

Fragment level classifier



deviation (μ,σ)

Sample Z-score = $\frac{det\_rate-\mu}{\sigma}$

Whole-genome sequencing for ctDNA detection

nature medicine

Explore content · About the journal · Publish with us

nature › nature medicine › articles › article

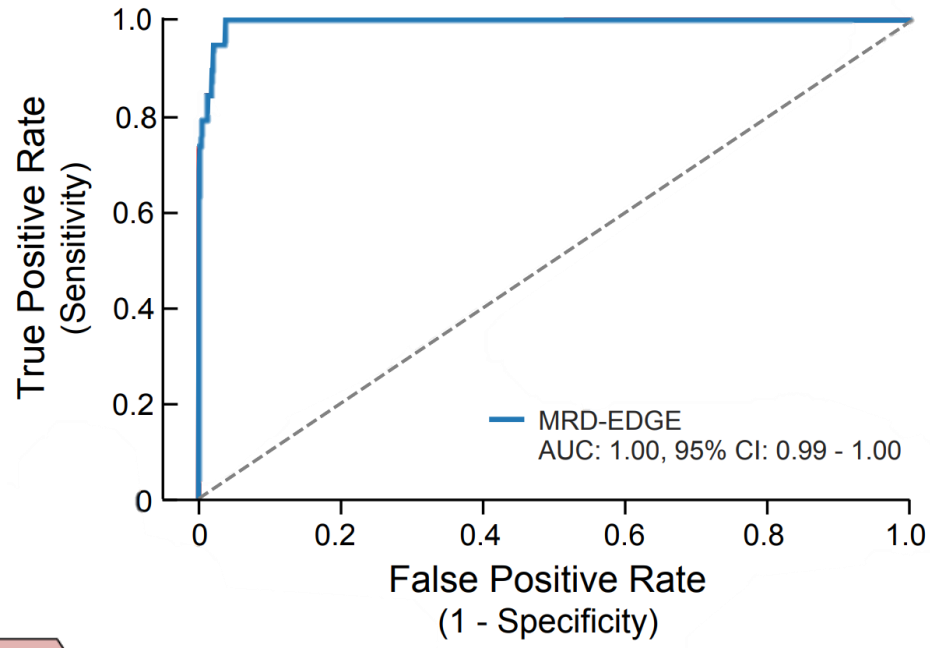
Article | Published: 14 June 2024

Ultrasensitive plasma-based monitoring of tumor burden using machine-learning-guided signal enrichment

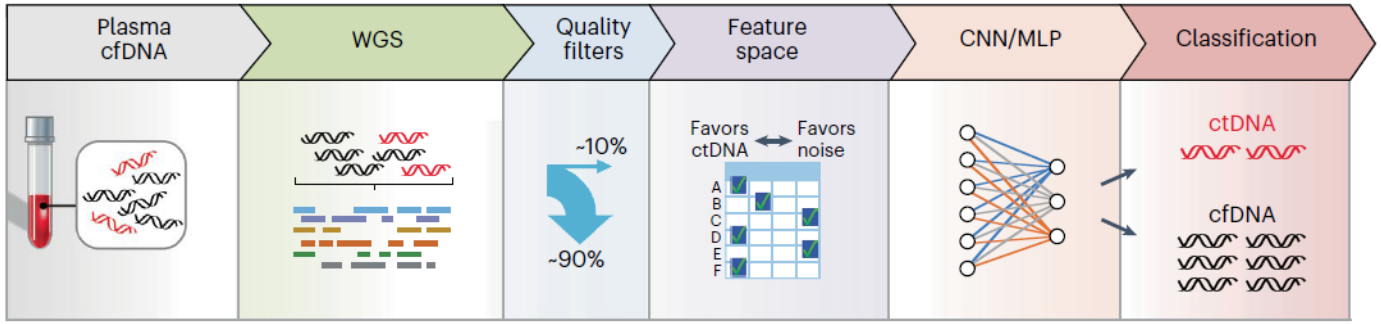
Adam J. Widman, Minita Shah, Amanda Frydendahl, Daniel Halmos, Cole C. Khamnei, Nadia Øgaard, Srinivas Rajagopalan, Anushri Arora, Aditya Deshpande, William F. Hooper, Jean Quentin, Jake Bass, Mingxuan Zhang, Theophile Langanay, Laura Andersen, Zoe Steinsnyder, Will Liao, Mads Heilskov Rasmussen, Tenna Vesterman Henriksen, Sarah Østrup Jensen, Jesper Nors, Christina Therkildsen, Jesus Imielinski, Michael F. Berger, Sam Aparicio, Nasser K. Altorki, Michael A. Postow, Nicolas Robine, Claus Lindbjerg Andersen & Dan A. Landau — Show fewer authors

Nature Medicine **30**, 1655–1666 (2024) | Cite this article

9036 Accesses | **4** Citations | **253** Altmetric | Metrics



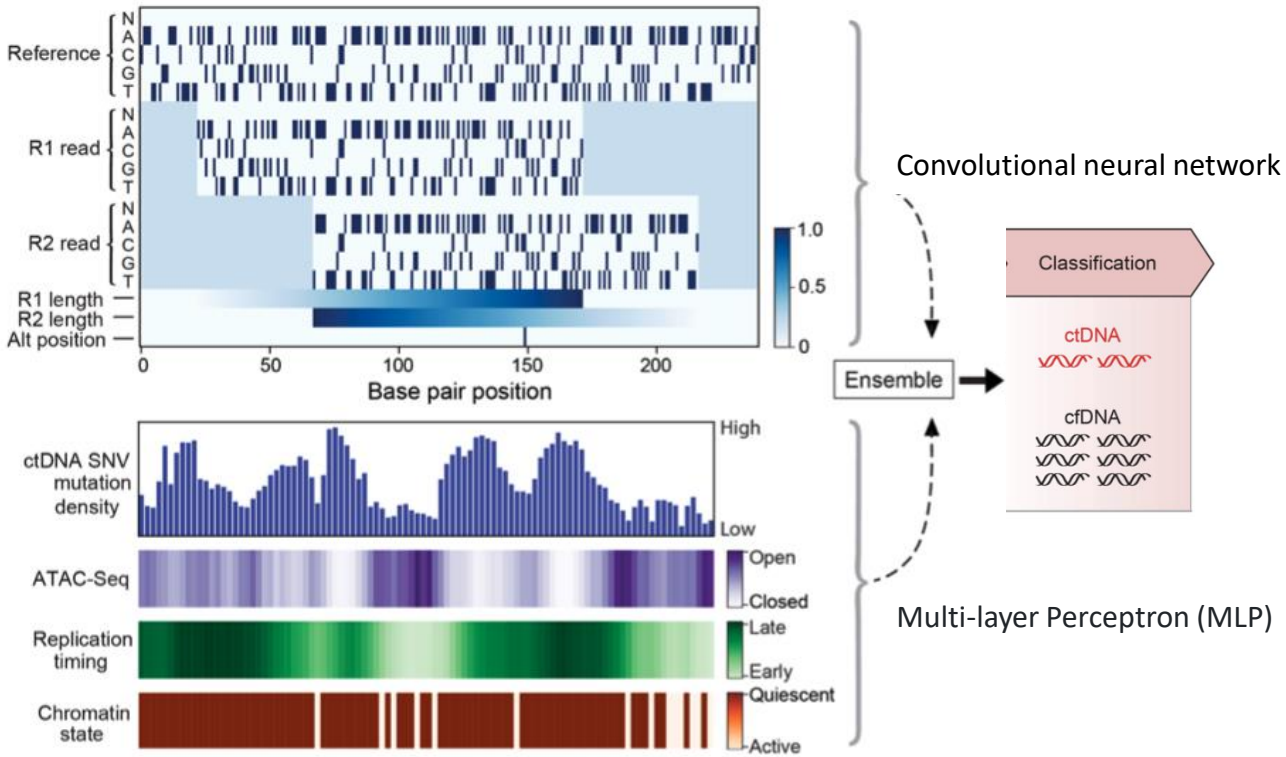
Fragment level classifier



Fragment level classifier

Error-suppression

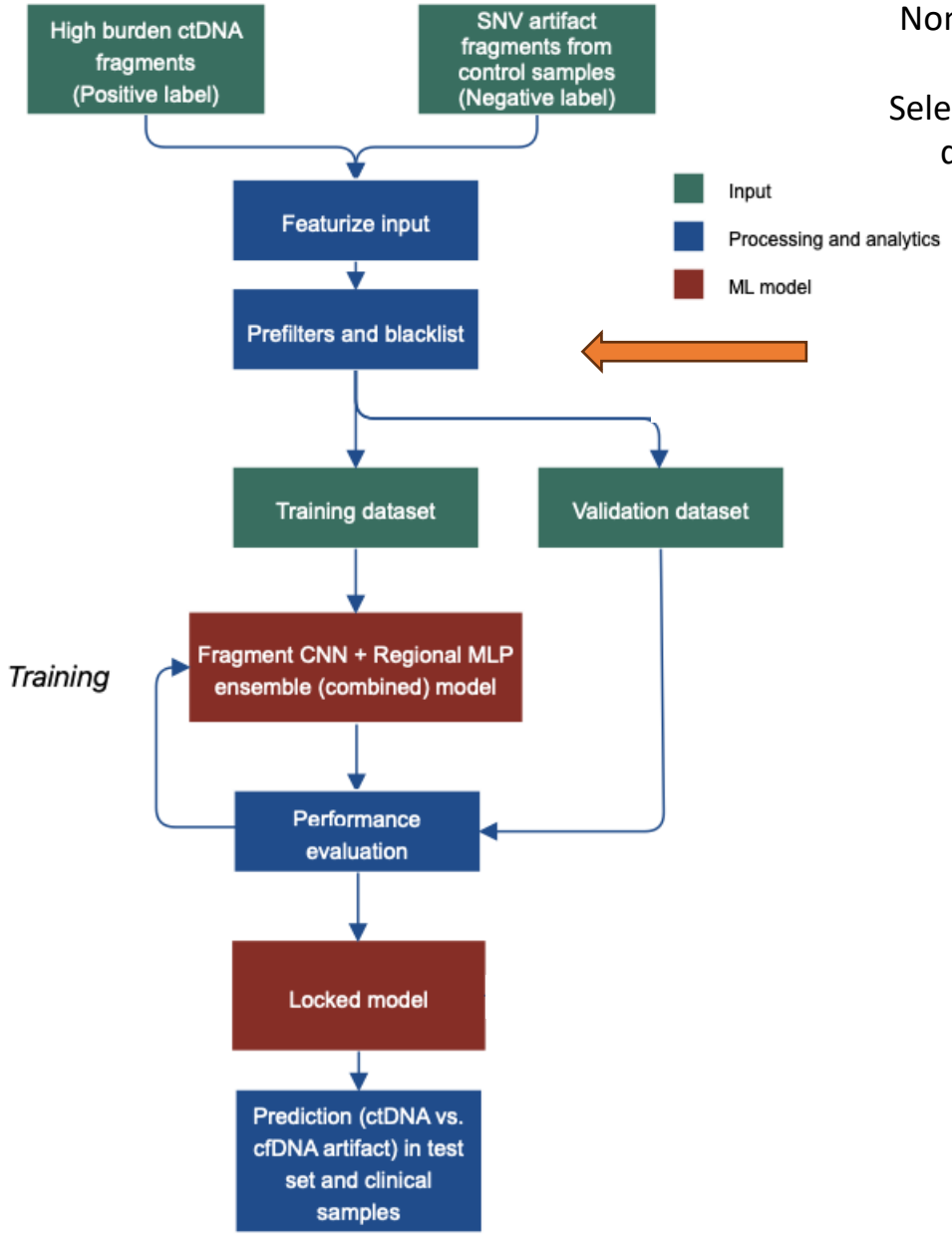
Signal to noise enrichment



MRD-EDGE SNV fragment-level model training

High tumor burden cfDNA samples (20X)

Select mutated fragments matching the tumor mutational compendia



Non-cancer control cfDNA samples (20X)

Select fragments with sequencing errors – defined as non-reference base calls

Training set:

- 5 high TF cfDNA samples
- 5 non-cancer control samples

>250.000 fragments, each of positive and negative labels

Validation set:

- 2 high TF cfDNA samples
- 2 non-cancer control samples

~100.000 fragments, each of positive and negative labels

TAKE HOME MESSAGE

The OBVIOUS errors are filtered.

- Residual germline SNPs
- recurrent plasma WGS artifact
- Variants with low
  - base quality scores
  - mapping quality scores

Removes the majority of cfDNA sequencing errors (mean 92.8%)

AARHUS UNIVERSITY
DEPARTMENT OF CLINICAL MEDICINE

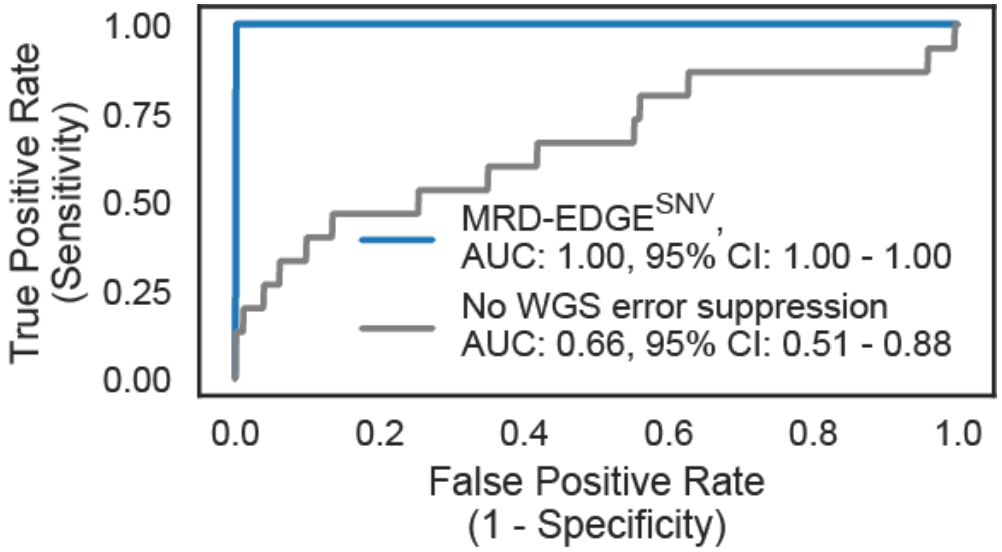
AARHUS UNIVERSITY HOSPITAL

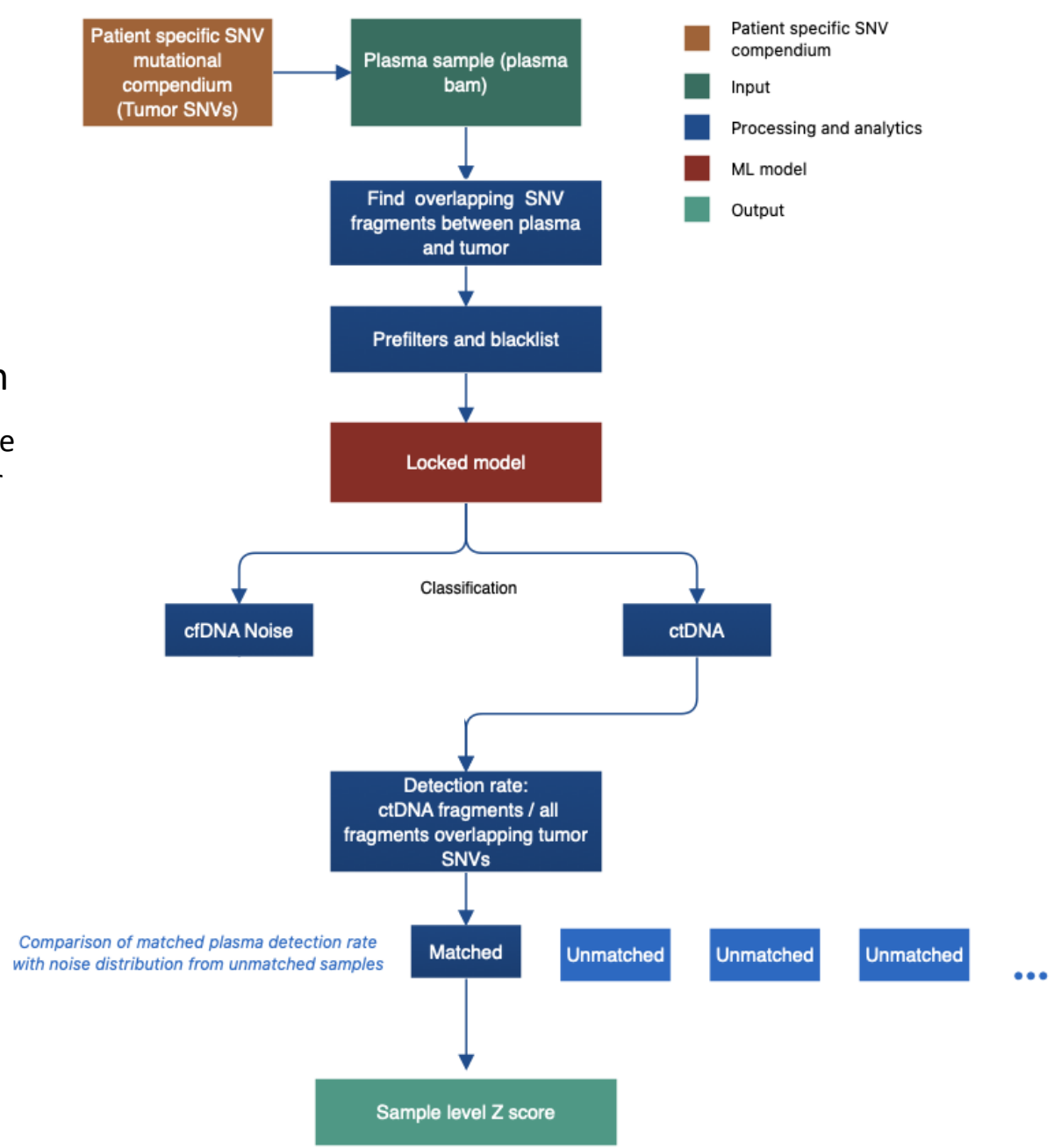
Sample classification using MRD-EDGE^{SNV}

Validation: Effect of MRD-EDGE^{SNV} error suppression

- 15 preOP CRC cfDNA samples
- 40 non-cancer control samples

} Independent of the fragment classifier training samples





Robustness assessment

NovaSeq platform
Data from Aarhus
Pos label: n=15
Neg label
Healthy

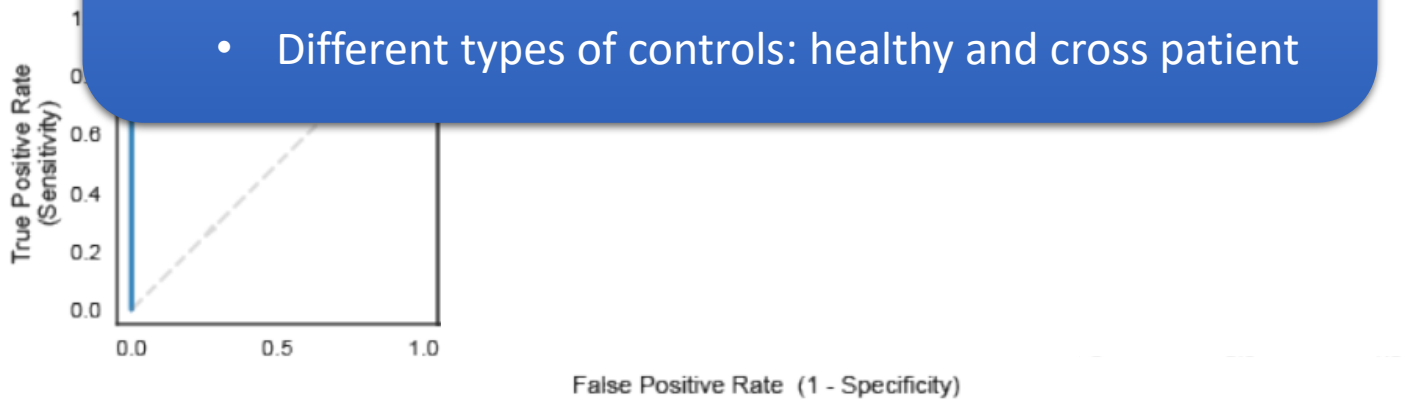
NovaSeq platform
Data from **New York**

HiSeq platform
Data from New York

NovaSeq platform
Data from Aarhus
Pos label: n=15
=210 (14x15)
t controls

Robust to

- Data generation at different sites
- Data generation using different sequencing platforms
- Different types of controls: healthy and cross patient



Supplementary Fig. 13: MRD-EDGE^{SNV} Z scores compared to 4 non-cancer control plasma cohorts

Post Doc Amanda Frydendahl et al
Accepted for publication in Molecular Oncology

How sensitive is WGS based ctDNA detection?

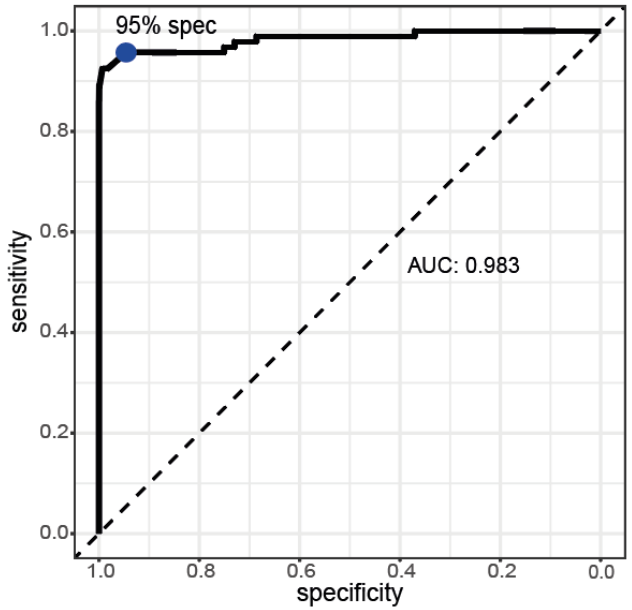
How does it compare to single marker digital PCR?

ctDNA detection in stage III CRC

Cohort A) Cancers

Stage III CRC (*n*=93)

Non-cancer controls (*n*=40)



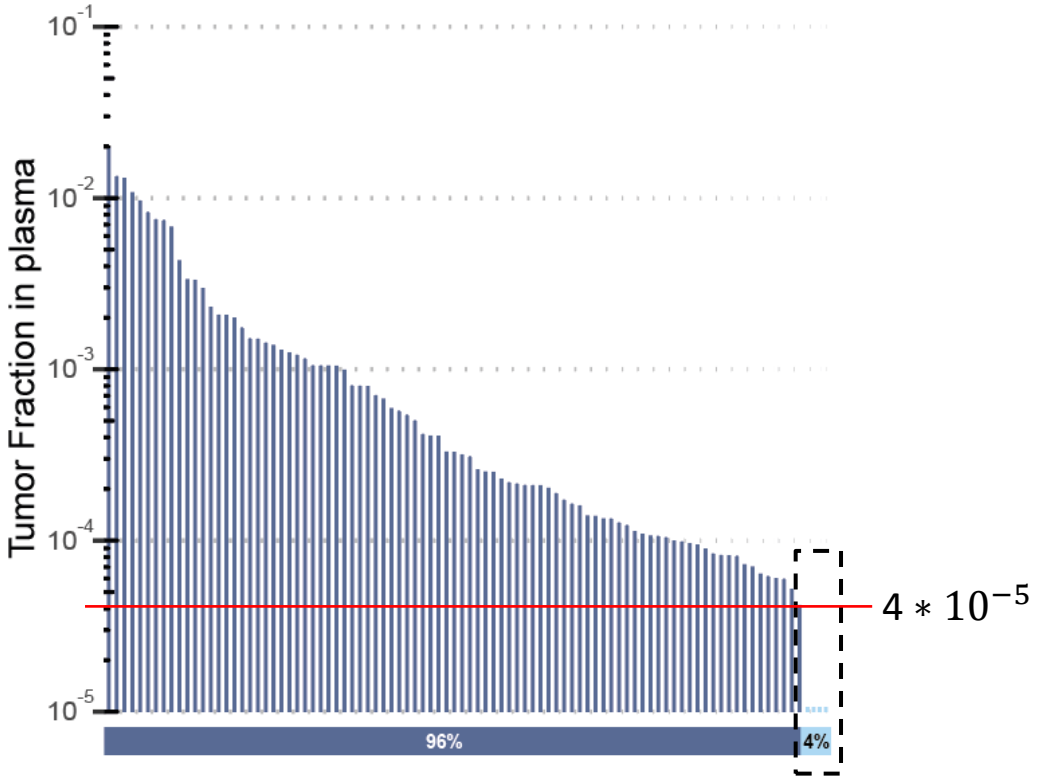
Positive labels: **N=93**

Pre-operative plasma from stage III CRC

Negative labels: **N= 3720**

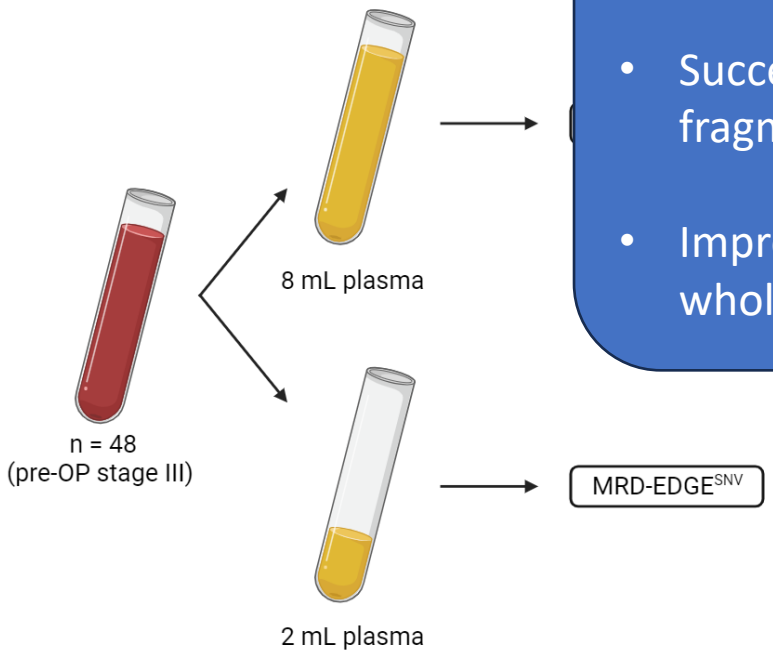
Patient mutational compendiums (n=93) applied to plasma from healthy controls (n=40)

Tumor fraction in pre-operative plasma (Stage III CRC)



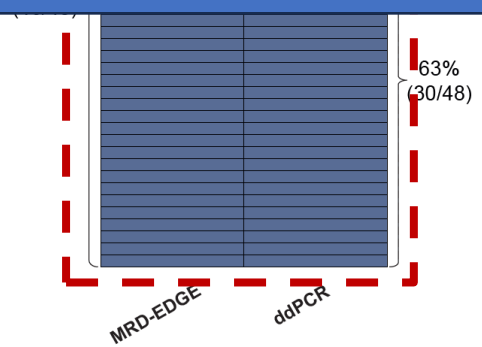
Superior sensitivity of MRD-EDGE^{SNV}

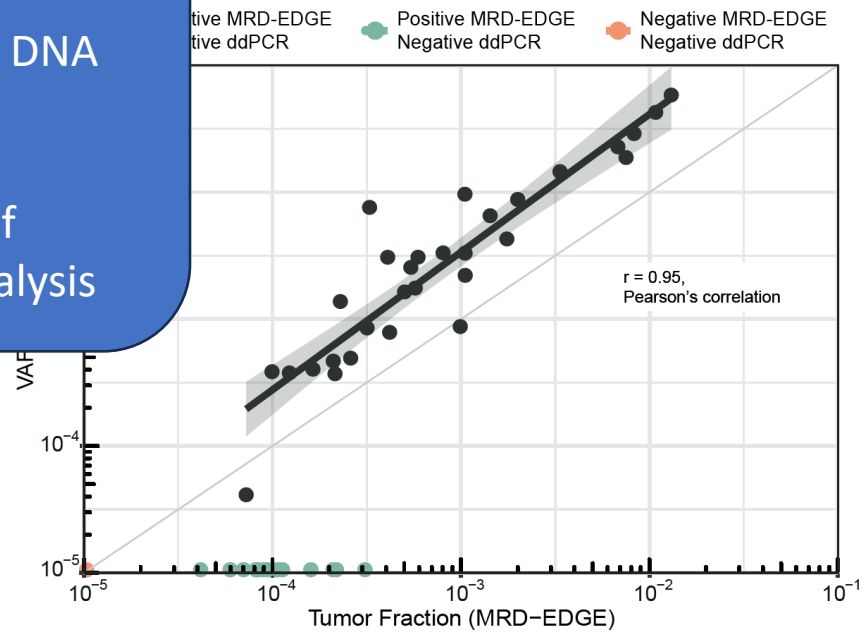
Comparison of MRD-EDGE^{SNV} to digital PCR



Conclusion:

- Successfully built a robust machine learning DNA fragment classifier
- Improved analytical and clinical sensitivity of whole genome sequencing based ctDNA analysis





Take home messages

Features

- Watch out for bias – be careful if data was not generated for purpose

Training

- The more data the more robust the models
- Remove the obvious, machine learning models may home in on the easy & forget the difficult

Validation:

- Analytical and clinical performance assessment
- Carefully assess robustness and generalizability
- Watch out for bias and confounding

AI/machine-learning predictive tools:
Process from development to clinical implementation

Well-establish process for development of biomarkers for guiding clinical decision making – **same should apply for AI/ML tools**

1. Discovery (screen for features correlated with clinical outcome)
2. Technology, method and test development – analytical validation
3. "Proof of concept" studies
4. Retrospective validation and optimization
   a) Performance assessment
   b) Method optimization (technical – analytical)

MRD-EDGE^{SNV}

5. Prospective validation – observational study
   a) Performance assessment in "real world" cohorts (explore generalizability and real-world performance)
   b) Method and test optimizations (technical – improved analytical performance)
   c) Product – a "locked" test
6. Prospective clinical utility assessment – intervention study
   a) clinical utility assessment – optimally through comparison to current clinical practice – randomized trial design
7. Regulatory approval and commercialization
8. Clinical implementation (incorporation in clinical guidelines, physician/patient education, approval by payers, and clinical authorities)

Acknowledgements

Department of Molecular Medicine, AUH

Amanda Frydendahl
Tenna V Henriksen
Nadia Øgaard
Mads H Rasmussen
Christina Demuth
Jesper Nors
Thomas Reinert
And many others

Weill Cornell Medicine and New York Genome Center

Adam Widman and Daniel Landau

C2i genomics, NY

Boris Oklander, Danielle Afterman, Santiago Gonzalez, Maja Kuzman, *Asaf Zviran*

Thank you to all the patients for participating!