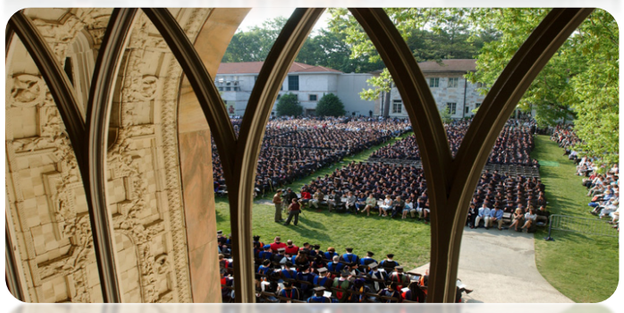
Disease diagnosis for AD using cell-free DNA

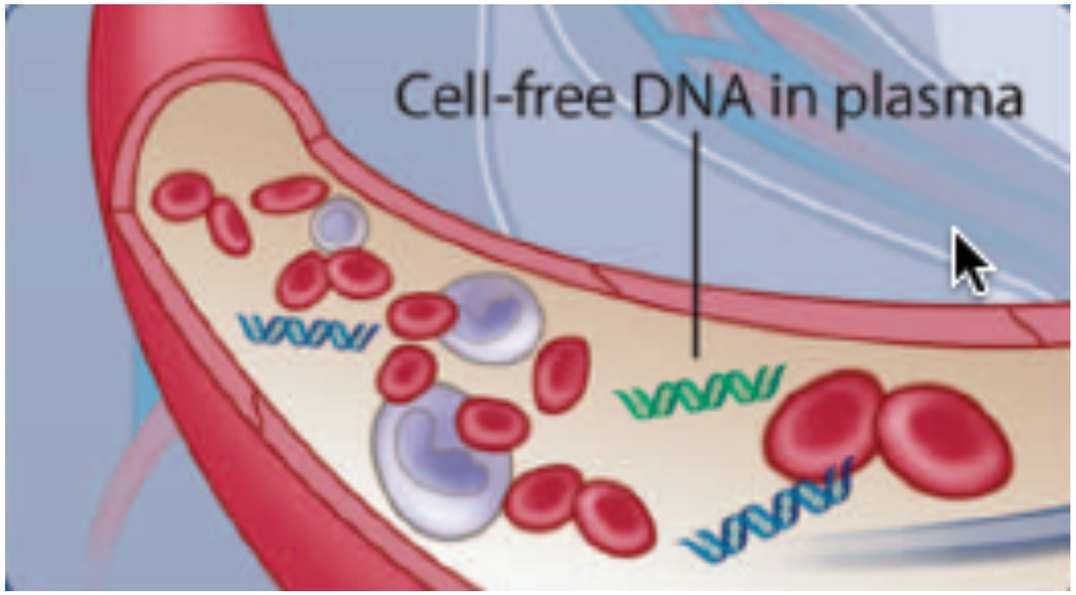


EMORY UNIVERSITY

Goizueta Alzheimer's Disease Research Center

www.alzheimers.emory.edu

Cell-free DNA (cfDNA)



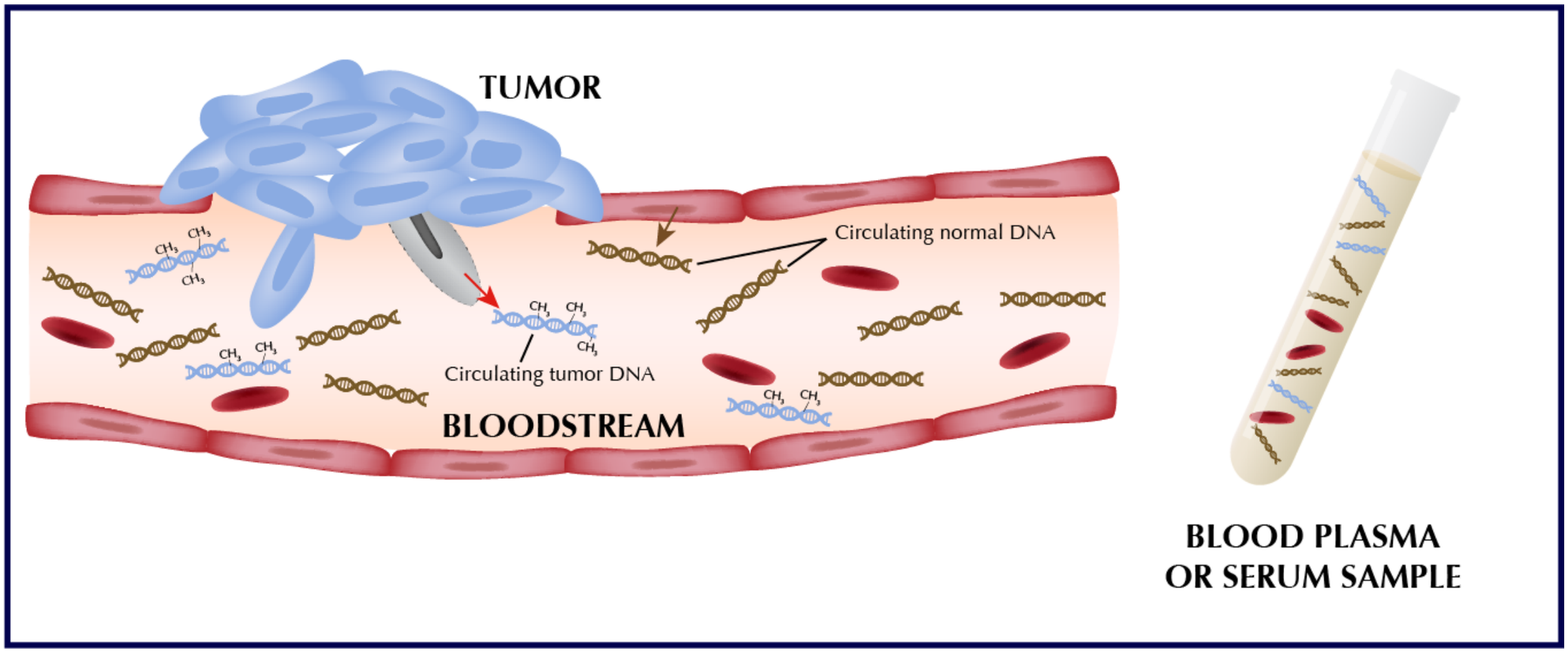
- Short DNA fragments in plasma that don’t belong to any cell.
- Released from blood cells/dead tissue cells/external cells.
- Examples of cfDNA:
  - In cancer patients, circulating tumor DNA (ctDNA).
  - In pregnant women, cell-free fetal DNA (cffDNA).

Motivation to study cfDNA

- Has great potentials to be diagnostic biomarker:
  - Non invasive.
  - Early detection.
  - Cheap and easy.

"liquid biopsy".

Cancer detection and assessment



cfDNA as diagnostic biomarker

- The essence is to diagnose based on “abnormal” cfDNA.
- Currently the information of “abnormality” is based on mutations.
  - cfDNA with unusually high number of mutations are deemed abnormal, and can be linked to disease.
- So far the application of cfDNA are mostly limited to “mutation-rich” diseases such as cancer, NIPT, etc.
- The principal cannot be directly applied for mutation-poor diseases, such as AD.

cfDNA epigenetics

- It is known that epigenetic signatures are highly tissue-specific.
- It is possible to explore the epigenetic information on cfDNA, and then link those to diseases.
- Existing works:
  - DNA methylation (Sun et al. 2015 *PNAS*).
  - Nucleosome position (Snyder et al. 2016 *Cell*).

cfDNA epigenetic biomarker

- cfDNA is a mixture of DNA from different tissues.
- AD leads to change of mixing proportions (greater proportion of cfDNA from brain in AD cases), thus the marginal epigenetic profiles.
- Use epigenetic profiles at selected genomic regions as predictors for disease.

Prediction method

- We investigated and compared methods to
  - Use cfDNA markers
  - Use estimated mixing proportions

Briefings in Bioinformatics, 2018, 1–13

doi: 10.1093/bib/bby029
Review Article

Disease prediction by cell-free DNA methylation

Hao Feng, Peng Jin and Hao Wu

Corresponding author: Hao Wu, Department of Biostatistics and Bioinformatics, Emory University Rollins School of Public Health, Atlanta, GA 30322, USA. Tel.: +1 404-727-8633. E-mail: hao.wu@emory.edu

cfDNA 5hmC data for AD

- Sample:
  - 10 healthy young people
  - 10 healthy old people
  - 10 AD patients
- Sample source: plasma
- Sequencing technology: hmC-Seal sequencing
- Measurement: genome-wide 5-hydroxymethylcytosine (5hmC) level

Data processing

- Genome was cut into bins with 5kb length each
- 5hmC level was calculated for all the bins for each sample
- Find the differentially hydroxymethylation regions (DhMR) between
  - 10 old samples vs 10 young samples (209 DhMRs)
  - 10 old samples vs 10 AD samples (296 DhMRs)
- Use DhMR as predictors for AD prediction
  - Marker selection is the key

Preliminary prediction results

	Pred AD	Pred normal
True AD	9	1
True normal	2	8

85% accuracy

Future plan

- Larger sample size.
- 5mC sequencing of cfDNA.
- New statistical method development:
  - Feature selection.
  - Sample deconvolution.
  - Integrated 5mC/5hmC/nucleosome positioning model.

Acknowledgement

- Emory ADRC
- Peng Jin (Emory Human Genetics)
- Harry Feng (Emory Biostatistics and Bioinformatics)