Non-invasive multi-cancer detection targeting DNA hypomethylation of LINE-1 retrotransposons

Marine Gorse*[1], Charline Bianchi*[1], Victoria Dixon[1], Klaus Von Grafenstein[1], Chloé Azencott[2,3], and Charlottte Proudhon[1]

[1]Institut de recherche en santé, environnement et travail (IRSET) – Université de Rennes, École des Hautes Études en Santé Publique [EHESP], Institut National de la Santé et de la Recherche Médicale – 9 Av. du Professeur Léon Bernard, 35000 Rennes, France

[2]CBIO-Center for computational biology – MINES ParisTech, PSL Research University – 60 boulevard Saint-Michel 75006 Paris, France, France

[3]Cancer et génome: Bioinformatique, biostatistiques et épidémiologie d'un système complexe – Mines Paris - PSL (École nationale supérieure des mines de Paris), Institut Curie [Paris], Institut National de la Santé et de la Recherche Médicale – 26 rue d'Ulm - 75248 Paris cedex 05, France

Abstract

Circulating tumor DNA (ctDNA) released by tumor cells into the blood stream bears all the molecular alterations of the tumor of origin, enabling non-invasive molecular profiling and disease monitoring. However, targeting small fractions of tumor DNA when the tumor burden is low remains one of the biggest challenges of liquid biopsy. In addition, targeting mutations that may have less than one alteration copy per milliliter of plasma further complicates the chances of detection.

Therefore, we have developed a new highly sensitive strategy for detecting cancer-specific signatures in blood based on the methylation patterns of repeat sequences. We have centered our study on primate specific Long Interspersed Nuclear Element-1 family (L1PA and L1HS) which are known to be hypomethylated in multiple types of cancers. Firstly, we have designed a PCR-based bisulfite deep sequencing, targeting 35 CG sites. We then developed a prediction model integrating methylation haplotypes at the single molecule level. Resulting machine learning-based classifiers showed promising classification rates in 6 types of cancer to discriminate healthy from tumor plasmas.

We now want to further increase the detection sensitivity and the universality of our approach, by maximizing the number of CG sites targeted notably. To do so, we are transposing the method to a capture-based targeting of LINE-1 elements. Our preliminary results show that we are able to hit more LINE-1 copies with high specificity (89% on-target rate). Moreover, we obtain methylation information for healthy donors along the 95 L1HS CG sites. Based on these preliminary results, this approach could lead to the development of more efficient, non-invasive diagnostic tests that can be applied to multiple types of cancer.

Keywords: LINE, 1 elements, circulating DNA, DNA methylation, targeted sequencing, hybridization capture, multi, cancer detection, machine learning

---

*Speaker