Title: Computational genomics resource development for cancer diagnostics

Abstract: Cancer is one of the leading causes of death among adults. Most tumors initially respond to treatment before developing resistance and eventually grow back. Timely detection of changes in treatment response can help adjust clinical management plans and save lives. Tumor cells shed fragments of DNA, that circulate freely in blood, alongside fragments of DNA from other, normal cells. Molecular and fragmentation characteristics of cell-free DNA (cfDNA) from tumor and normal cells may differ, such that sensitive analysis of cfDNA can be used to get information about the tumor from a simple blood draw. Here, we propose to develop novel algorithm to (i) predict tumor burden and (ii) infer the site of tumor growth using machine learning approaches on cfDNA genomic data obtained from cancer patients. We have already obtained cfDNA data for >1000 cancer patients. The analytical techniques may involve (i) deconvolution using techniques such as non-negative matrix factorization, independent component analysis, singular value decomposition, and (ii) feature selection and classification using elastic net with support vector machine, random forest etc. The algorithms will be benchmarked against other tools and ground truth, before used to analyze cfDNA profiles of cancer patients over multiple time-points to non-invasively track the trajectory of cancer progression and predict recurrence. Accurate and early detection of cancer recurrence will benefit cancer diagnostics and has implications for treatment and advancement.

Reference:

1: De S. Signatures Beyond Oncogenic Mutations in Cell-Free DNA Sequencing for Non-Invasive, Early Detection of Cancer. Front Genet. 2021 Oct 14;12:759832. doi: 10.3389/fgene.2021.759832. PMID: 34721546; PMCID: PMC8551553.

2: Moser T, Kühberger S, Lazzeri I, Vlachos G, Heitzer E. Bridging biological cfDNA features and machine learning approaches. Trends Genet. 2023 Apr;39(4):285-307. doi: 10.1016/j.tig.2023.01.004. Epub 2023 Feb 13. PMID: 36792446.

Software references:

https://github.com/topics/cell-free-dna