Inference of Gene Expression using Fragmentation Patterns From Targeted High-Depth Sequencing of Cell-Free DNA

Abstract number 6258

Emily Leff*, Alexander Tseng*, Victoria Cheung, Ehsan Tabari, Kim Walter, Alex Lovejoy, Jimmy Lin
Freenome, 279 East Grand Avenue 5th Floor, South San Francisco, CA
** These authors contributed equally*

Freenome

BACKGROUND

- Cell-free DNA (cfDNA) contains epigenetic signatures of the cells from which it was produced. As a result, cfDNA enables gene activation inference and tissue-of-origin classification with potential applications for non-invasive cancer detection[1,2]
- Because of low depth of coverage of sites of interest, current whole genome sequencing (WGS) methods have not been able to infer expression of individual genes or limited gene sets. Additionally, low tumor cell contribution in cfDNA increases the difficulty of detecting tumor signal above healthy background
- We developed a method to measure transcriptional start site gene activation probabilities (TSS-GAP) using a targeted deep sequencing approach

OBJECTIVES

- Develop a method to determine sets of expressed genes specific to a cancer cell line and detectable above the background signal level observed in cfDNA derived from healthy donors
- Determine gene-level limits of detection (LoD) of our TSS-GAP method

METHODS

Samples

- Micrococcal nuclease (MNase)-digested DNA has been shown to produce nucleosome-bound fragments comparable to cfDNA[2,3,4]
- Healthy peripheral blood mononuclear cells (PBMC) (n=8) served as the healthy background signal while blends of MNase-digested DNA from a CRC epithelial cell line (LS180) and sorted immune cell populations (CD4 and CD14) at varying concentrations (n=18) served as contributing tumor and immune cell signal that can be found in cfDNA (**Figure 1**)

KEY FINDINGS AND CONCLUSIONS

- We developed a method to predict gene activation at transcriptional start sites within plasma cell-free DNA from fragment location and position (TSS-GAP)
- The method presented here can be used to assess the sensitivity of TSS- GAP at the gene level
- An LoD for TSS-GAP can be obtained on the individual gene level. LS180 signal can be detected above healthy PBMC signal at very low concentrations of LS180 - down to 0.1%
- Colorectal cancer epithelial lineage genes can be detected at very low levels (0.1%) by TSS-GAP

Figure 1. MNase-digested blends of 100% cell lines

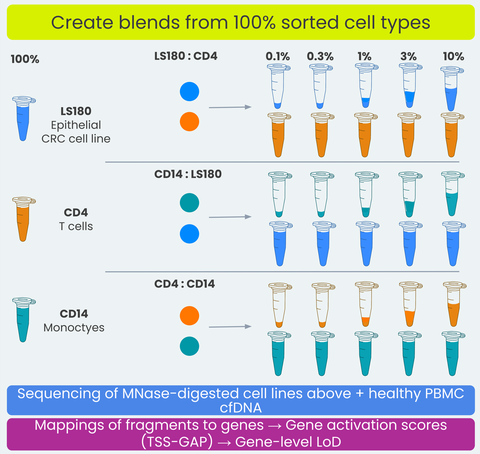
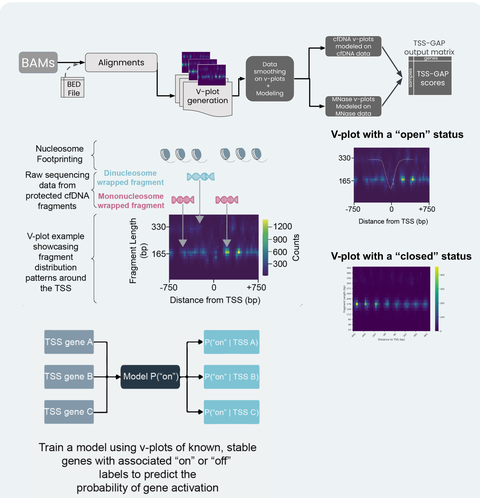
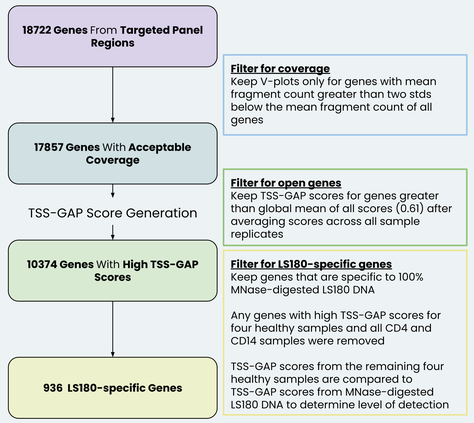


Figure 2. Predicting gene activation from plasma cell-free DNA from fragment length and fragment position



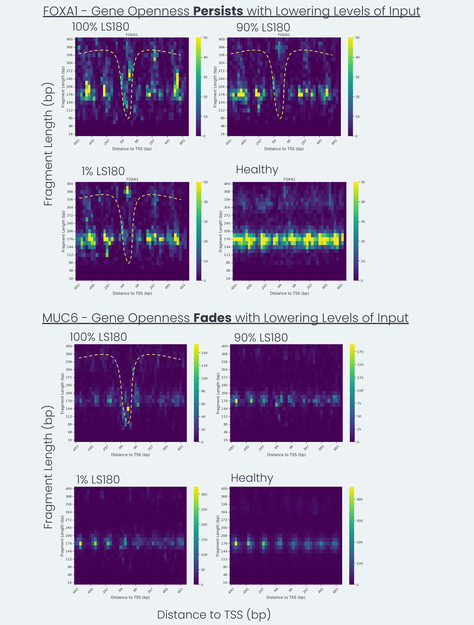
- This method of quantifying gene activation was applied to 18722 transcriptional start sites and **enables the generation of single holistic gene activation profiles per sample (Figures 2 & 3)**

Figure 3. Extracting LS180-specific expressed genes enables the measurement of TSS-GAP's sensitivity



RESULTS

Figure 4. TSS-GAP detects and predicts expression of different genes at different levels of sensitivity



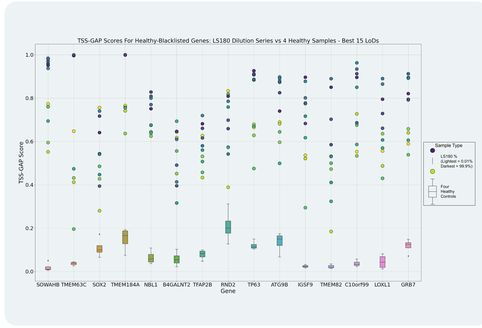
- At 0.1% LS180, expression signal in FOXA1 is detected while expression signal in MUC6 is not from their respective TSS-GAP scores (FOXA1 score: 0.73, MUC6 score: 0.02). This can be traced back to a difference in fragment length distributions at the TSS from the genes' associated V-plots (**Figure 4**)
- **This shows that not all LS180-associated genes are detectable at very low concentrations**

Table 1. Gene activation profiles are detectable even when low numbers of cells exhibit the signal

Concentration of LS180	0.10%	0.30%	1%	3%	10%	All other concentrations	**Total genes:**
# genes detectable at each concentration of LS180:	326	385	416	435	474	462	**936**

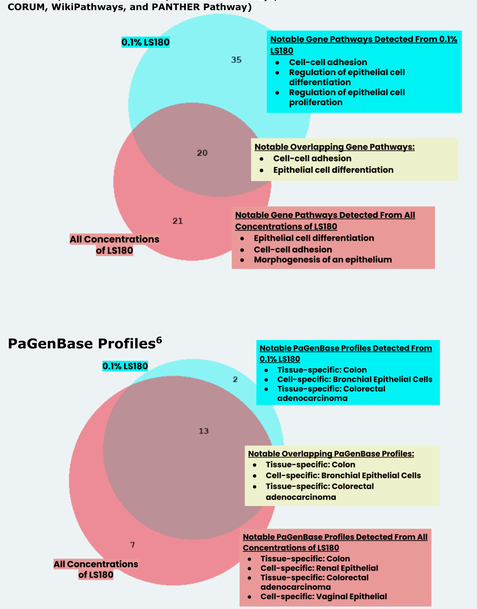
- Numbers of genes are cumulative across concentrations of LS180. Detectability is determined if the minimum TSS-GAP score for a gene at a concentration of LS180 is greater than the mean TSS-GAP score for the same gene among four healthy PBMCs (**Figure 3**)
- **TSS-GAP can infer expression of LS180-associated genes at high sensitivity**
- Genes with the 15 best LoDs show TSS-GAP can distinguish specific signal from one cell type from another at significantly low dilutions. Several genes are expressed in colorectal tissue and epithelial cells (TP63, SOX2; **Figure 5**)

Figure 5. Gene activation profiles reflect and track with cell signal even at low levels of contribution



- While TSS-GAP is able to detect epithelial CRC-associated pathways and gene expression profiles at all concentrations of LS180 (epithelial cell differentiation, colon tissue/epithelial cell signatures), it is still able to detect the same signal at significantly lower concentrations of LS180 (cell-cell adhesion, regulation of epithelial cell proliferation, colorectal adenocarcinoma tissue signatures and epithelial cell signatures at 0.1%) (**Figure 6**)

Figure 6. Colorectal cancer epithelial signatures can be detected at very low levels by TSS-GAP



References

1. Ulz et al. *Nature Genetics*, 2016.
2. Snyder et al. *Cell*, 2016.
3. Ivanov et al. *BMC Genomics*, 2015.
4. Zhang et al. *Clinical Chemistry*, 2017.
5. Zhou et al. *Nature Communications*, 2019
6. Pan et al. *PLOS ONE*, 2013

Acknowledgements

This study was sponsored by Freenome Holdings, Inc.

Disclosures

All authors are employees of Freenome Holdings, Inc.

Presented at the American Association for Cancer Research Annual Meeting, San Diego, CA. April 5-10, 2024.