Pan-cancer blood plasma cell-free RNA profiles reveal cancer type and patient-specific changes

Annelien Morlion[1,2], Philippe Decruyenaere[1,2,3], Kathleen Schoofs[1,2,4], Jasper Anckaert[1,2], Justine Nuytens[1,2], Eveline Vanden Eynde[1,2], Kimberly Verniers[1,2], Celine Everaert[4], Fritz Offner[3], Jo Van Dorpe[5], Jo Vandesompele[1,2,*], Pieter Mestdagh[1,2,*]

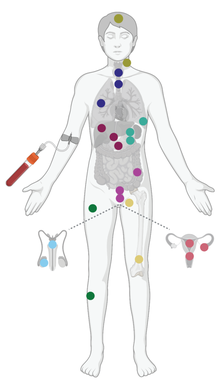
Annelien.Morlion@UGent.be

[1] Department of Biomolecular Medicine, Ghent University, Ghent, Belgium
[2] OncoRNALab, Cancer Research Institute Ghent (CRIG), Ghent, Belgium
[3] Department of Hematology, Ghent University Hospital, Ghent, Belgium
[4] TOBI Lab, Center for Medical Biotechnology, VIB-UGent, Zwijnaarde, Belgium
[5] Department of Pathology, Ghent University Hospital, Ghent, Belgium
* joined last authors

introduction

Liquid biopsy-based biomarkers, such as circulating nucleic acids in blood plasma, have emerged as promising tools for cancer diagnosis and monitoring. Contrary to tissue biopsies, liquid biopsies are minimally invasive and allow serial collection. Most studies to date focused on circulating tumor DNA (ctDNA). However, cell-free RNAs (cfRNA) may provide an additional cancer biomarker resource given their dynamic nature and tissue specificity. Even though the presence of cfRNA in liquid biopsies is firmly established, the utility for cancer detection and monitoring remains underexplored.

cancer signal in cell-free mRNA profiles



- brain tumor (GBM & ANA)
- head and neck cancer (HNSC)
- thyroid cancer (THCA)
- esophageal cancer (ESCA)
- lung cancer (LUAD & LUSC)
- breast cancer (BRCA)
- stomach cancer (STAD)
- kidney cancer (KIRC)
- liver cancer (LIHC)
- cholangiocarcinoma (CHOL)
- pancreas cancer (PAAD)
- colon cancer (COAD)
- bladder cancer (BLCA)
- rectal cancer (READ)
- lymphoma (DLBCL)
- leukemia (AML)
- sarcoma (SARC)
- melanoma (SKCM)
- prostate cancer (PRAD)
- testicular cancer (TGCT)
- ovarian cancer (OV)
- cervical cancer (CESC)
- uterine cancer (UCEC)

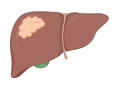
Blood plasma cell-free mRNA (capture sequencing) profiles from late stage to metastatic cancer patients and cancer-free controls in two independent cohorts: a pan-cancer cohort (n=208, 25 cancer types) and a three-cancer cohort (n=58; ovarian (OV), prostate (PRAD) and uterine cancer (UCEC))

Pan-cancer systemic signals

Distinct acute myeloid leukemia profiles

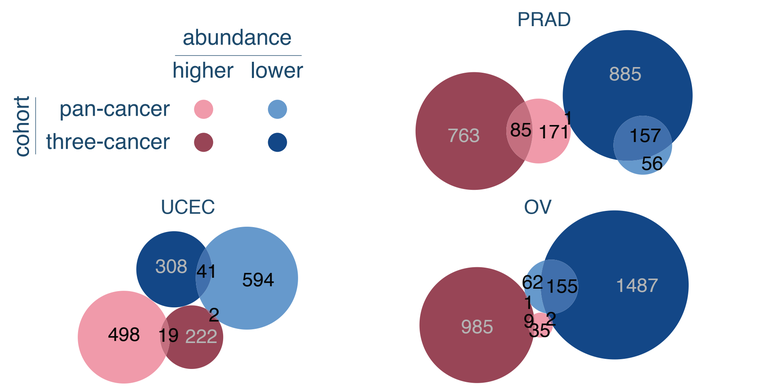
5310 differentially abundant mRNAs
PML::RARA fusion transcripts
Myeloid progenitor cell contribution
Lower platelet contribution

Limited type-specific signal in solid tumor profiles



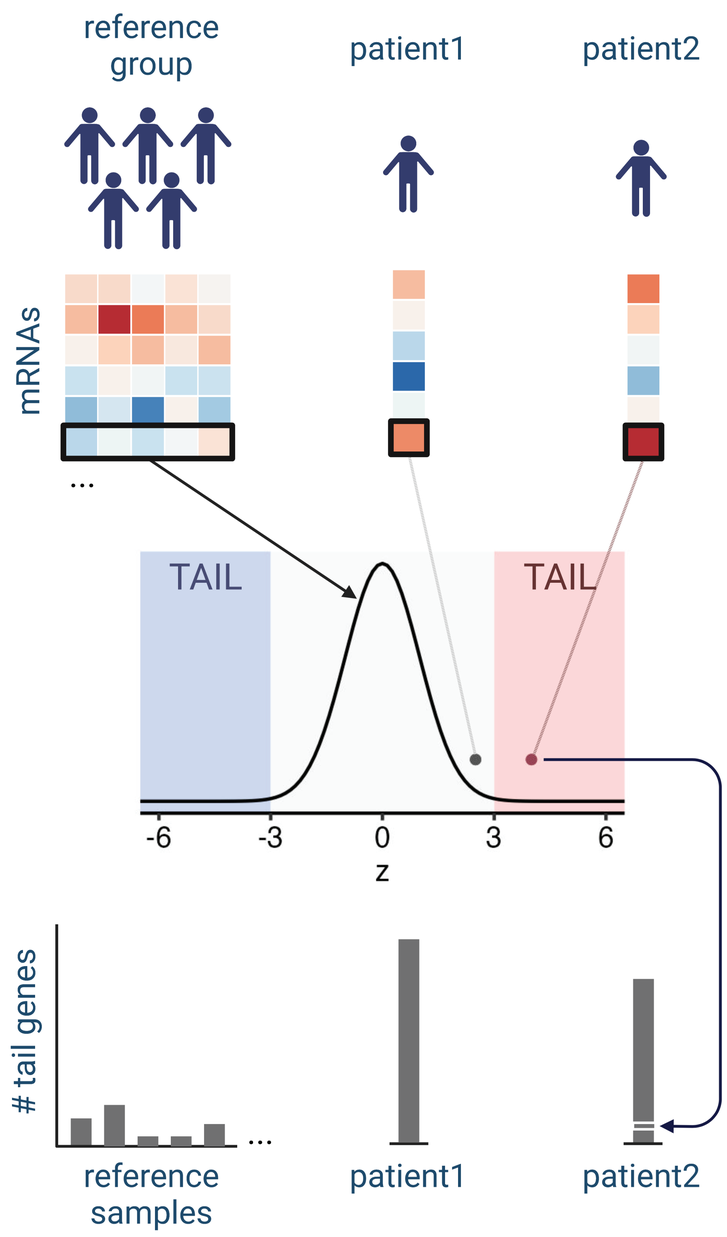
127 to 1154 differentially abundant mRNAs
Enrichment of liver-related gene sets in liver cancer patients

Heterogeneity among patients and cohorts



Significant though limited overlap of differentially abundant mRNAs between the same cancer type and controls in two cohorts (Jaccard indices < 15%).

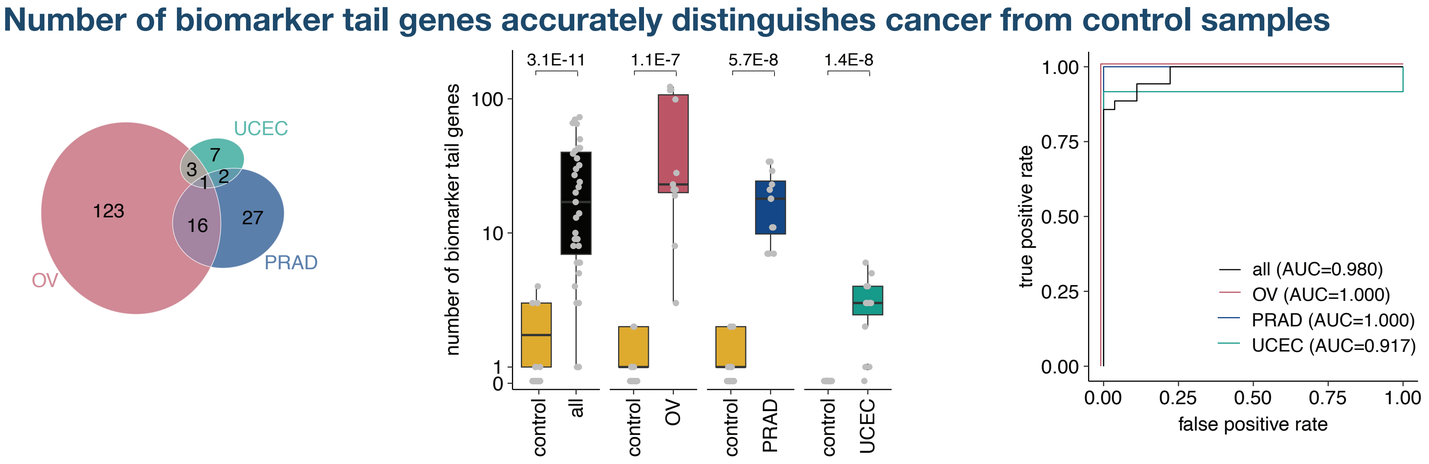
exploit heterogeneity for cancer classification



Tail gene identification workflow. A gene in a particular sample for which the mRNA abundance deviates more than 3 standard deviations from the reference group is a tail gene for that sample. Biomarker tail genes are a subset of tail genes specifically associated to cancer/control status in a sample cohort (Fisher's exact test). The number of (biomarker) tail genes per sample is used to determine cancer status.

patent application filed

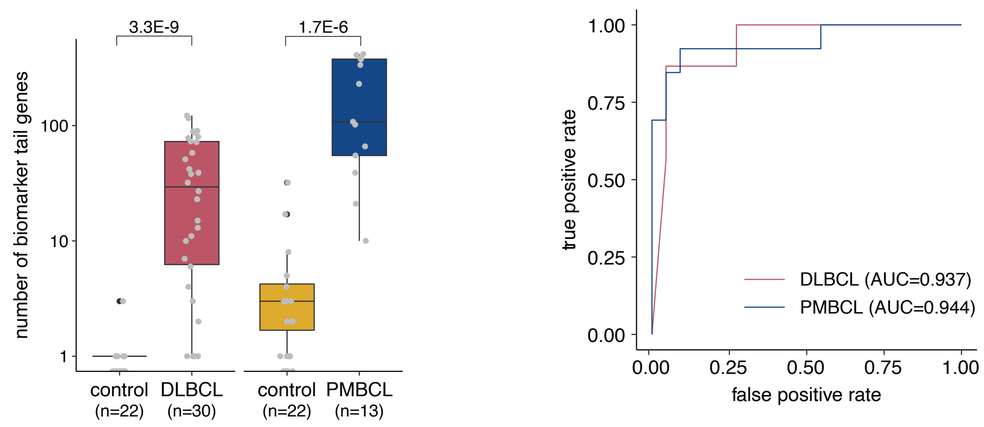
Number of biomarker tail genes accurately distinguishes cancer from control samples



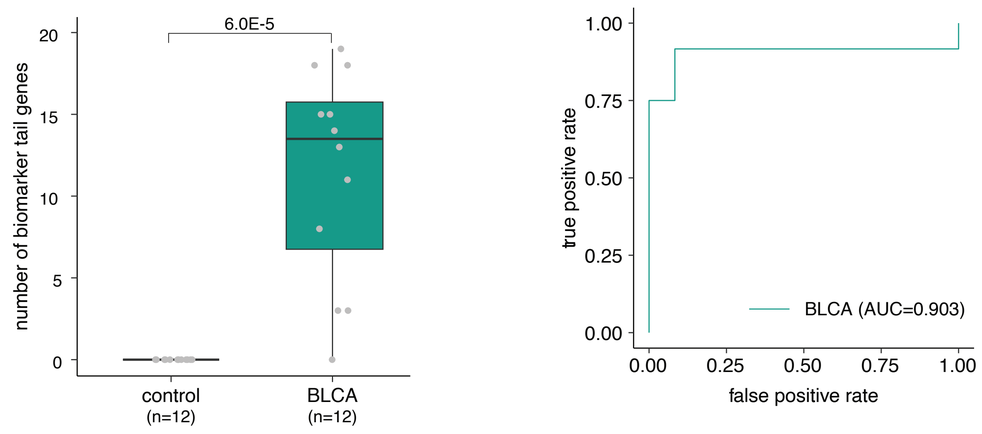
Higher number of biomarker tail genes in cancer patient plasma samples compared to controls. Left: Overlap of type-specific biomarker tail genes. Middle: Boxplot of number of biomarker tail genes per subset in control (n=27) and cancer samples (OV n=11, PRAD n=12, UCEC n=12, all n=35). 108 biomarker tail genes considered for all cancer versus control. Grey dots represent individual sample counts. Two-sided Wilcoxon rank-sum p-values indicated in plot. Right: ROC curve of leave-one-out cross-validated classifier for binary cancer/control classification based on the number of biomarker tail genes belonging to the respective biomarker tail gene set.

Concept validated in independent plasma and urine sample cohorts

plasma
total RNA sequencing
lymphoma patients (2 subtypes)
cancer-free controls
297 and 931 biomarker tail genes



urine
mRNA capture sequencing
12 bladder cancer patients
12 cancer-free controls
39 biomarker tail genes



Abbreviations
AUC: area under the ROC curve; cfRNA: cell-free RNA; ctDNA: circulating tumor DNA; DLBCL: diffuse large B-cell lymphoma; mRNA: messenger RNA; OV: ovarian cancer; PRAD: prostate cancer; UCEC: uterine cancer; PMBCL: primary mediastinal large B-cell lymphoma; ROC: receiver operating characteristic

Illustrations created with BioRender.com