Early cancer detection from liquid biopsy using cell-free RNA

Phil Sanders[1], Lucila González[1], Laura Sanchez[1], Beatriz Neto[1], Julien Lagarde[1], Marc Weber[1], Pablo Monteagudo[1], Cristina Tuñí[1], Giovanni Asole[1], Elena Cristina Rusu[1], Sol García[2], Pilar Simón-Extermera[2], Sonia Panadero[3], Maria Concepción Romero[3], Purificación Catalina[3], João Curado[1].

[1] Flomics Biotech, Roc Boronat 31, Barcelona 08005, Spain. [2] Health and Biomedicine, Leitat Technological Center, Barcelona Science Park, Barcelona 08028, Spain. [3] Andalusian Public Health System Biobank, Coordinating Node, Granada 18016, Spain

Introduction

Cell-free RNA (cfRNA) is a rich source of potential biomarkers in **liquid biopsies** as it is actively secreted into body fluids, including blood, by both **healthy and diseased cells**.

cfRNA analysis of a liquid biopsy therefore provides **insight into the health of an individual**.

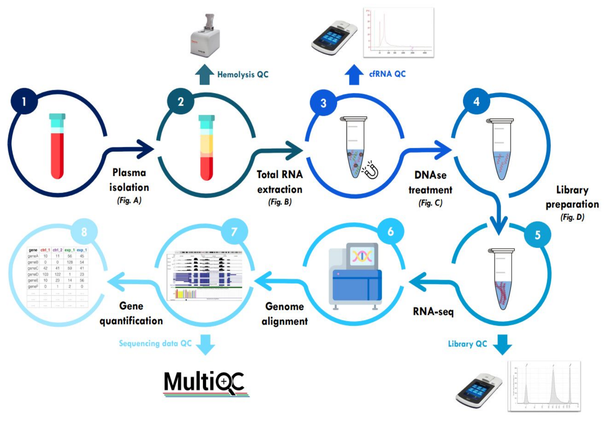
cfRNA is well suited to **early cancer detection,** and its advantages include:

- **Earliest detection:** Detectable RNA changes often occur before detectable DNA mutations, which is crucial for the detection of small and early cancers.
- **Tissue specificity:** RNA varies by tissue, facilitating cancer tissue of origin identification.
- **Personalized Medicine:** RNA provides functional information about the dysregulated signalling pathways in the cancer and guides treatment.
- **Dynamic molecule:** Recurring measurement of RNA levels allows the establishment of a “healthy baseline” which is key for Minimal Residual Disease (MRD) detection and monitoring.

At Flomics Biotech we have developed a high quality cfRNA-Seq platform that profiles human plasma cfRNA in a robust and reproducible manner. We are currently applying this platform in the ongoing LiquiDx pre-clinical study with the goal of developing a **cfRNA-based multi-cancer early detection test**.

Flomics cell-free RNA-Seq platform

In this ongoing study, we apply our cfRNA-Seq platform (Figure 1 - Left) to plasma samples from a cohort of over 1,000 donors (Figure 1 - Right).



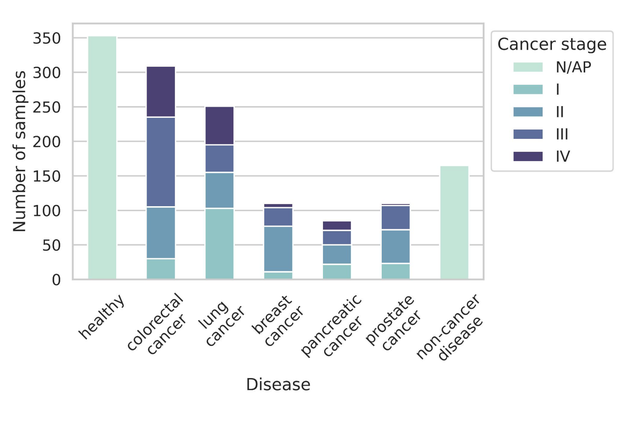
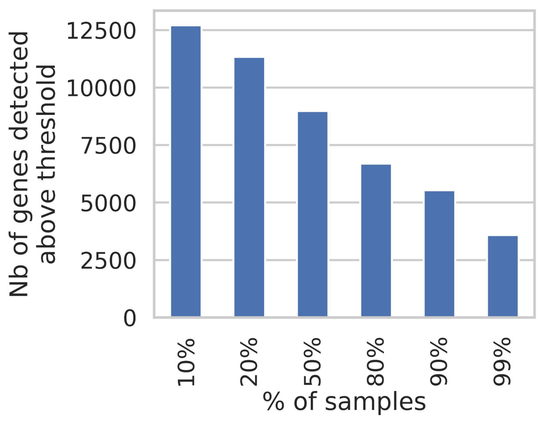
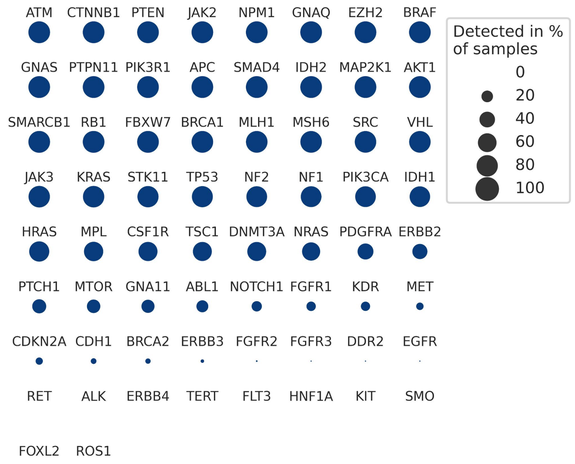


Figure 1. Left: Summary of the Flomics cfRNA-Seq pipeline. Data generated by this pipeline is analysed using a combination of gold-standard bioinformatics tools and advanced machine learning methods to identify cancer type-specific biomarker signatures and develop predictive machine learning models for early cancer detection. QC = quality control. **Right:** Distribution of the samples in the study across different diseases. For each cancer type sample distribution across cancer stages is indicated. “non-cancer disease”: patients with non-cancer diseases of the same organs as the cancer types.

High-quality cancer biomarker detection

The Flomics cfRNA-Seq platform generates a high quality data set rich in potential biomarkers, with more than **6000 genes detected** in 80% of the samples (Figure 2 - top left). These genes are distributed across **a broad range of biotypes** (Figure 2 - bottom). Many **oncogenes and tumour suppressor genes are detected** (Figure 2 - top right), showing the potential of the Flomics cfRNA-Seq platform for the detection of cancer-related signals.





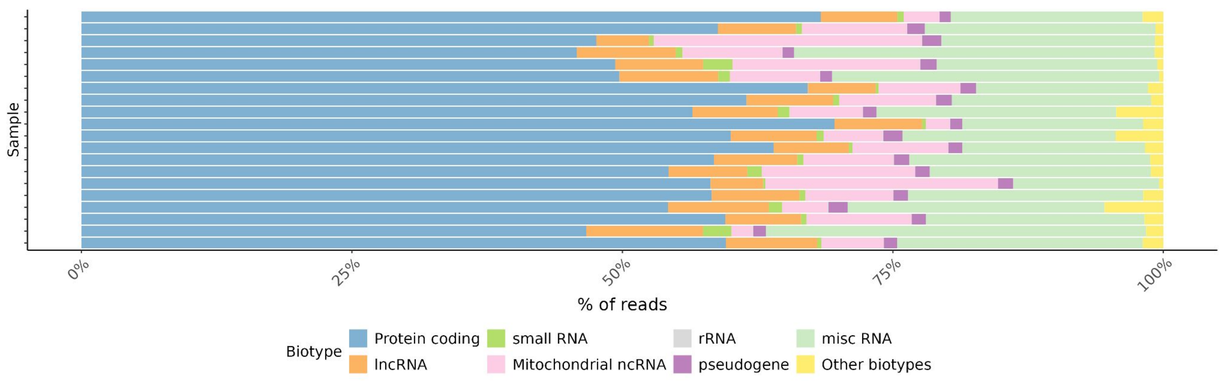
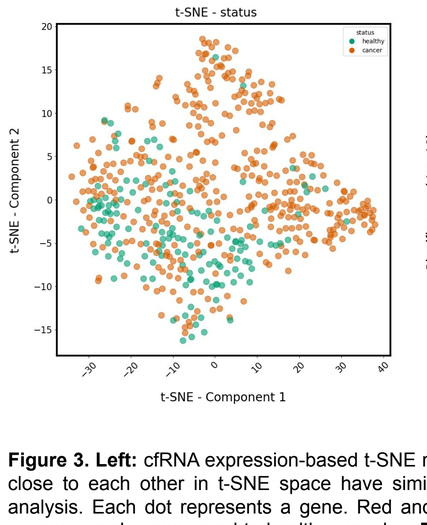
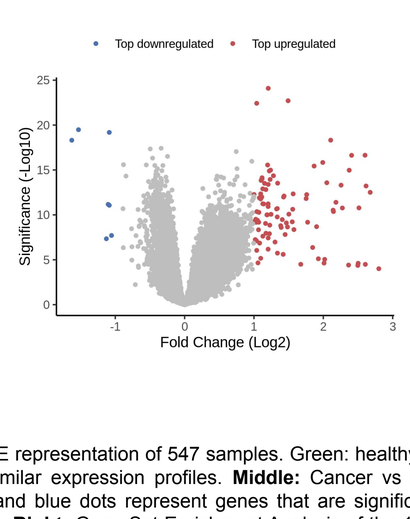


Figure 2. Top left: Number of genes detected >1 TPM in the indicated percentage of samples. **Top right:** For a panel of **65 oncogenes and tumour suppressor genes**, the percentage of samples (size of the disk) in which the indicated gene is detected with >1 TPM. **Bottom:** Gene biotype distribution in 20 randomly selected samples (y axis) from the LiquiDx project. The x axis represents the percentage of total genes in a sample mapping to each biotype. The distribution of reads across biotypes observed here is representative of that in the full data set. The “small RNA” biotype includes all small RNAs such as miRNA, snRNA, and snoRNAs.

Identifying a cancer biomarker signature

Healthy and cancer samples have distinct cfRNA expression profiles that can be used to separate them (Figure 3 - Left). Differential gene expression analysis identifies 98 significantly upregulated genes and 7 significantly downregulated genes in the cancer sample group to give **a 105 gene cancer biomarker signature** (Figure 3 - Middle). Gene set enrichment analysis reveals the enrichment or under-representation of genes associated with various **cellular processes and signalling pathways implicated in cancer in the biomarker signature** (Figure 3 - Right). This supports the validity of our approach in using cfRNA to identify cancer biomarker signatures.





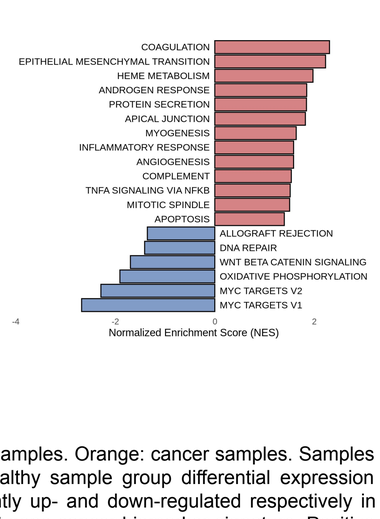


Figure 3. Left: cfRNA expression-based t-SNE representation of 547 samples. Green: healthy samples. Orange: cancer samples. Samples close to each other in t-SNE space have similar expression profiles. **Middle:** Cancer vs healthy sample group differential expression analysis. Each dot represents a gene. Red and blue dots represent genes that are significantly up- and down-regulated respectively in cancer samples compared to healthy samples. **Right:** Gene Set Enrichment Analysis of the 105 gene cancer biomarker signature. Positive and negative NES reflect enrichment and under-representation respectively of genes associated with the indicated cellular processes and signalling pathways.

Multi-cancer early detection

We leverage differences in cfRNA expression between cancer patients and healthy individuals to build a **machine learning model** to detect cancer. Training of the model is ongoing, however using currently available data (Figure 4 - Left) we have developed a model that **predicts cancer with promising results** (Figure 4 - Middle left). For the prediction of specific cancer types, **the model performs particularly well for lung and prostate cancer** (Figure 4 - Middle right). The model demonstrates a similar **high performance level for the prediction of all cancer stages**, (Figure 4 - Right).

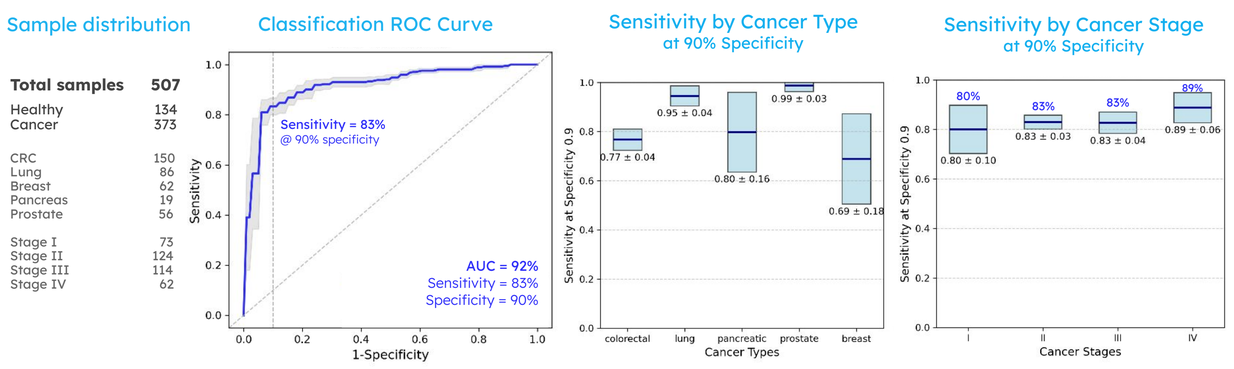
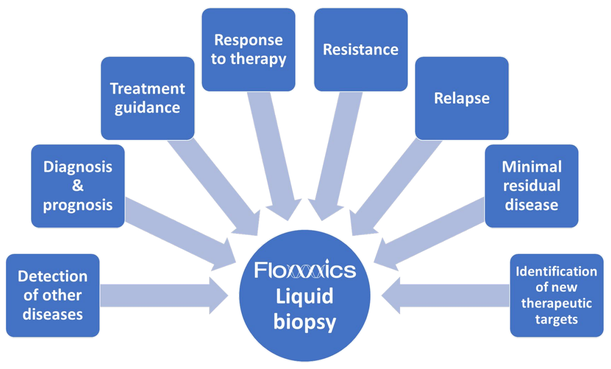


Figure 4. Left: Distribution of the samples used in this analysis across cancer type and stage. **Middle left:** Performance of our preliminary cancer vs healthy machine learning (ML) classifier, using the Extreme Gradient Boosting method. ROC curves (Sensitivity v.s. 1-Specificity) and Areas Under the Curve (AUCs) for the 4-folds cross-validation. **Middle right:** Performance of the ML classifier in predicting individual cancers at 90% specificity. **Right:** Performance of the ML classifier in predicting different stages of cancer at 90% specificity.

Applications beyond early cancer detection

The Flomics cfRNA-seq platform has many applications aside from early cancer detection , such as guiding cancer **treatment selection** and assessing **response to cancer therapies** such as immunotherapy (Figure 5 - Left). **Beyond cancer**, the platform can be applied to many other diseases and research areas such as infectious disease detection (Figure 5 - Right).



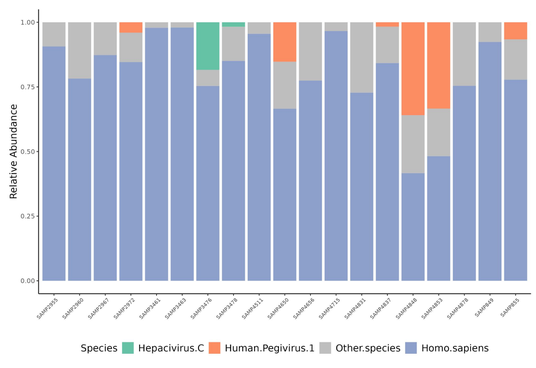


Figure 5. Left: Applications of Flomics’ cfRNA-based liquid biopsy technology beyond early cancer detection. **Right:** Detection of viral infections in LiquiDx samples. X-axis: Samples with detectable virus infection and randomly selected plasma samples without virus infection for comparison. Y-axis: Relative abundance of reads in each sample mapping to Homo sapiens (blue), Hepacivirus C (green), Human Pegivirus 1 (orange), and all other non-human species combined (grey). The Hepacivirus C-positive samples were confirmed as coming from Hepatitis C-positive donors via clinical data.

Conclusions

- Flomics has developed a world class cfRNA-Seq platform that can pioneer the use of cfRNA in clinical practice.
- We have developed a highly promising machine learning-based cancer detection model that maintains a high performance level across all cancer stages (including stage I), illustrating the benefits of using cfRNA for early cancer detection.
- With a diverse range of applications, the Flomics cfRNA-Seq pipeline can have a significant impact on global health.

Working with you!

Flomics is always looking for:

- Collaborations
- Opportunities to provide our technology and knowhow as a service
- Clinical expertise including sample provision

Contact us at info@flomics.com

Acknowledgements

The ‘LiquiDx’ project (reference CPP2021-008897) is financed by the Spanish Ministry of Science, Innovation and Universities (MCIN/AEI/10.13039/501100011033) and by the European Union “NextGenerationEU/PRTR”.