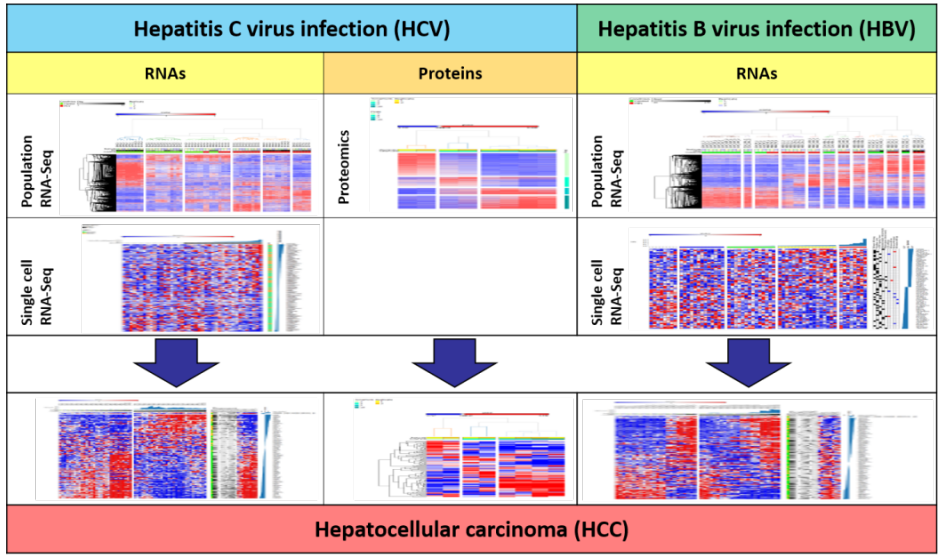
Integrative genomic framework for dissecting regulatory mechanisms underlying hepatocellular carcinoma

Nathalie Pochet, Olivier Gevaert, Thomas Baumert

Case study: Identify cell circuits driving virus-induced reprogramming of the host hepatocyte transcriptome and proteome to identify candidate drivers of HCC

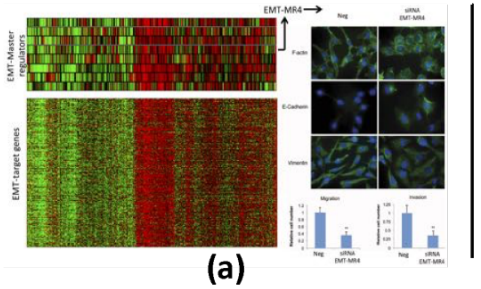
Data: Time course population & single-cell RNA & protein levels of genome-wide host responses to HCV & HBV



Inferring regulatory networks: Probabilistic graphical models are powerful tools to infer regulatory networks

AMARETTO algorithm & software tool

- Integrates genetic, epigenetic & transcriptomic or proteomic data
- Optimal prediction performance, data size & time complexity



(a) (b)

General limitations inherent to current formulations

1. Analysis is based on assumption that samples are independent
2. Analysis is driven by modeling one functional genomics level
3. Analysis is driven by modeling one biological system
4. Analysis is based on assumption that genes are independent

Can we learn better models via integration of host responses to HCV & HBV infection to improve our understanding of HCC?

Can we gain power by modeling known dependencies, multiple functional genomics levels and multiple biological systems?

Integrative genomic framework for dissecting regulatory mechanisms underlying hepatocellular carcinoma

Nathalie Pochet, Olivier Gevaert, Thomas Baumert

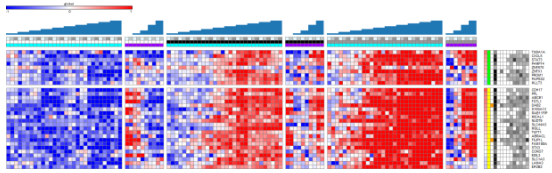
Reformulations to AMARETTO

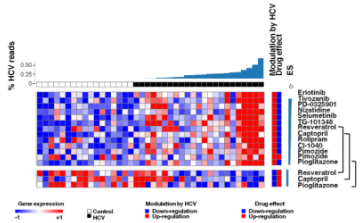
1. Analysis is based on assumption that samples are independent
➢ Model known dependencies between samples
e.g., time courses & single cells with viral load

2. Analysis is driven by modeling one functional genomics level
➢ Model regulatory interactions between RNAs & proteins

3. Analysis is driven by modeling one biological system
➢ Joint modeling of multiple viral infections & etiologies

4. Analysis is based on assumption that genes are independent
➢ Model known dependencies between genes
e.g., genetic & chemical perturbations





Goal 1: Develop integrative genomic framework to study virus-induced cancer

- Reformulate AMARETTO
- Infer regulatory networks
- Select candidate drivers
- Experimental validation of drivers
- Iterative optimization of the framework

Goal 2: Develop statistical benchmarking framework for generalizability in cancer

- Assess previous methods reformulations
- Optimize models for specific case studies

Dissemination

- Source code in R via GitHub
- Analytic modules in GenePattern & GenomeSpace