

Master's Degree Thesis for Sapienza University of Rome

Advisor: Prof. Manuela Petti

Co-Advisor: Prof. Lorenzo Farina

Faculty: Information Engineering, Computer Science and Statistics

Course: Data Science

Academic Year: 2022-2023

Author: Davide Mascolo, 2001991



SAPIENZA
UNIVERSITÀ DI ROMA

Integrating Genomic and Transcriptomic data of checkpoint blockade response in Non-Small Cell Lung Cancer through Multilayer Network Modelling

Introduction

- **Global Impact:** Lung cancer as leading cause of cancer-related deaths worldwide.
- **Personalized Treatment:** Shift from a one-size-fits-all approach to personalized treatments.
- **Targeted Therapy:** Biomarkers identification.
- **Multidata Approach:** Not all lung cancers are the same.

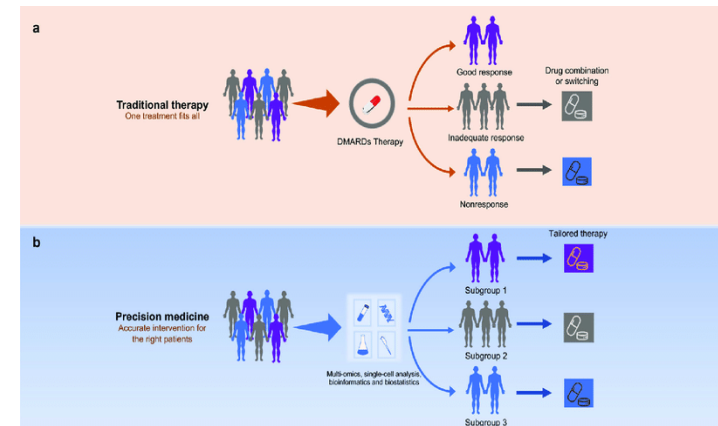


Fig. 1: Precision Medicine Approach

Case study

- **Inspiration:** Project initiated based on Arvind Ravi et co. *Genomic and transcriptomic analysis of checkpoint blockade response in advanced non-small cell lung cancer*, *Nature Genetics* – 06 Apr 2023
- **Data Types:** Genomic and Transcriptomic data.
- **Cohort Size:** 393 NSCLC patients treated with anti-PD-(L)1 immunotherapy.
- **Objective:** Multinetwork approach for patients stratification.

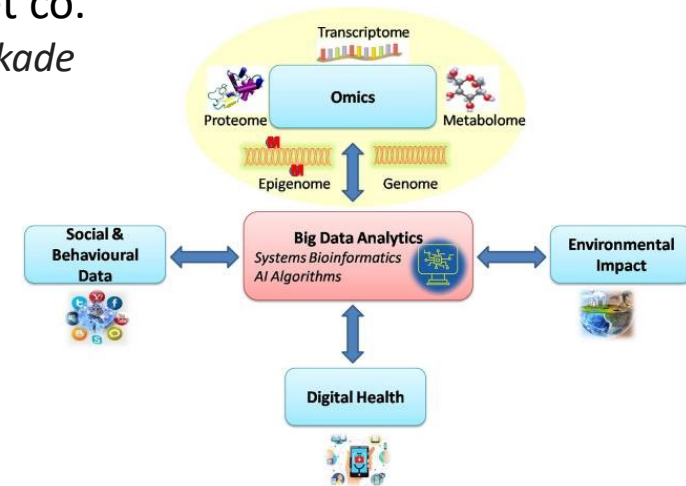


Fig. 2: Multidata Approach



Materials

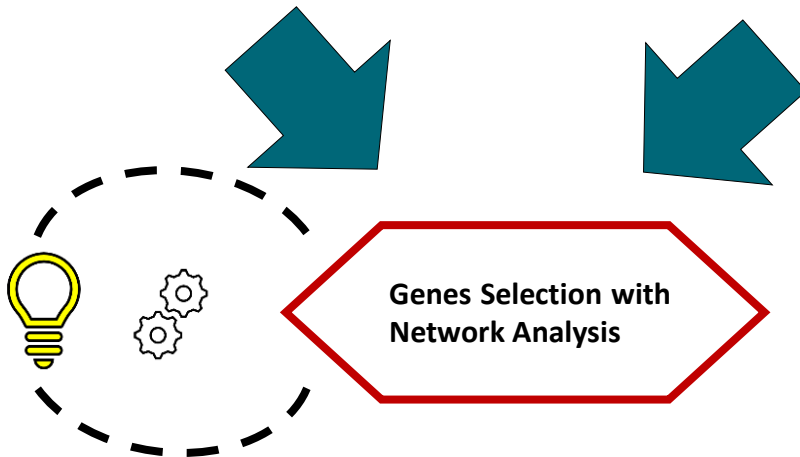
Data

Metadata

Genomic

Transcriptomic

Signatures



Curated

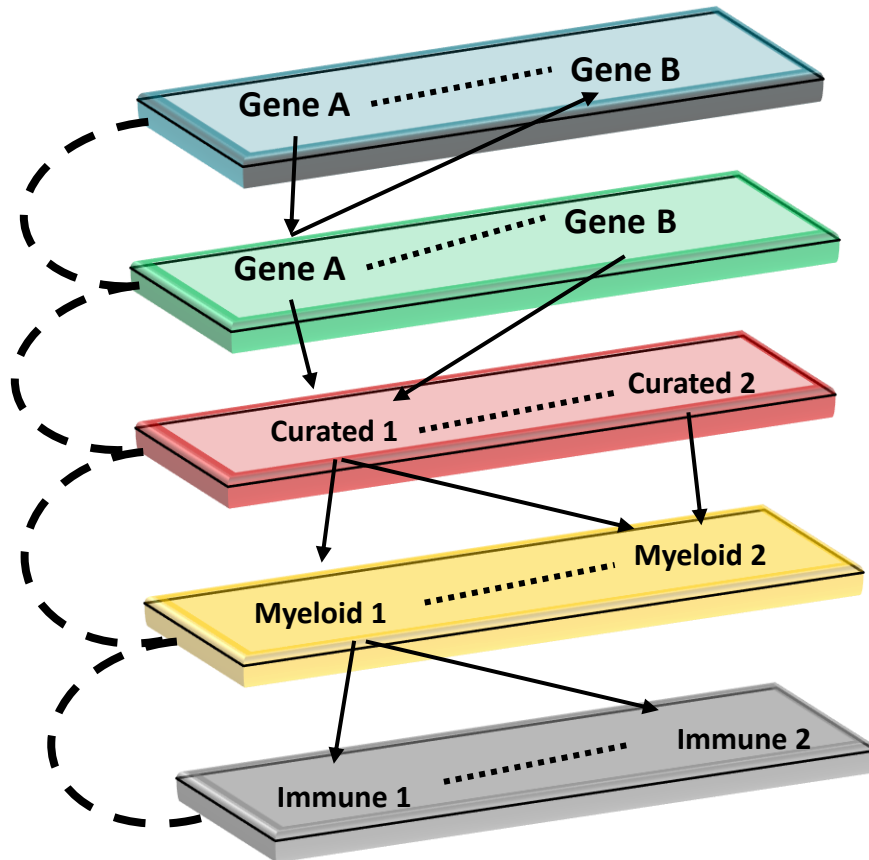
Myeloid

Immune



Materials

Multinetwork Approach - Data



Genomic Layer

Transcriptomic Layer

Curated Layer

Myeloid Layer

Immune Layer



Methods

DEGs & Network Analysis

- **Objective:** Identify DEGs w.r.t. *Resistance* condition.
- **Data:**
 - RNA Counts filtered by DEGs.
 - *Response*: 52 patients
 - *Resistance*: 84 patients
- **Differential Co-Expression Network:** Capture alterations in connections between nodes across different conditions.



Methods

Multinetwork Approach

- **Data:**
 - Genomic Data filtered by 10% differential co-expression hubs.
 - Transcriptomic Data filtered by 10% differential co-expression hubs.
 - Clinical Signatures.
- **Algorithm:** Similarity Network Fusion and Spectral Clustering.
- **Validation:**
 - Cox log rank test, with $\alpha = 0.05$.
 - Hypergeometric Test to enrich cluster significance.



Preliminary Analysis

DEGs Analysis

- **Key Points:**
 - *fold change* ≥ 0.5 and *p-value* < 0.05 .
 - *False Discovery Rate* correction
 - *DEGs* = 857
- **Relevant Genes:**
 - *SFTPC*
 - *PGC*
 - *CXC Chemokine Family*

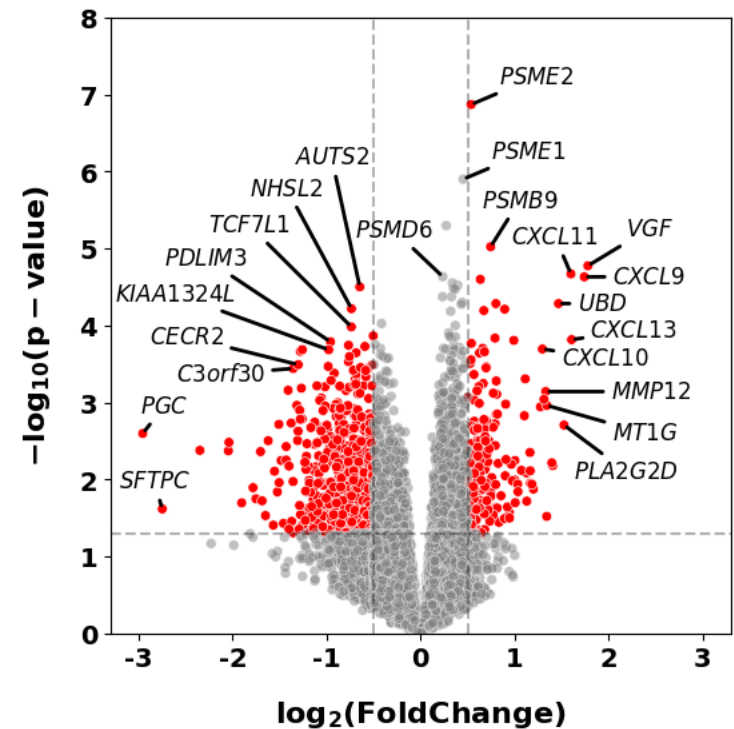


Fig. 3: DEGs with Volcano Plot



Results

Differential Co-Expression Network

- **Key Points:**
 - Quantile 90% (degree ≥ 46)
 - 10% most connected hubs ($N = 89$)

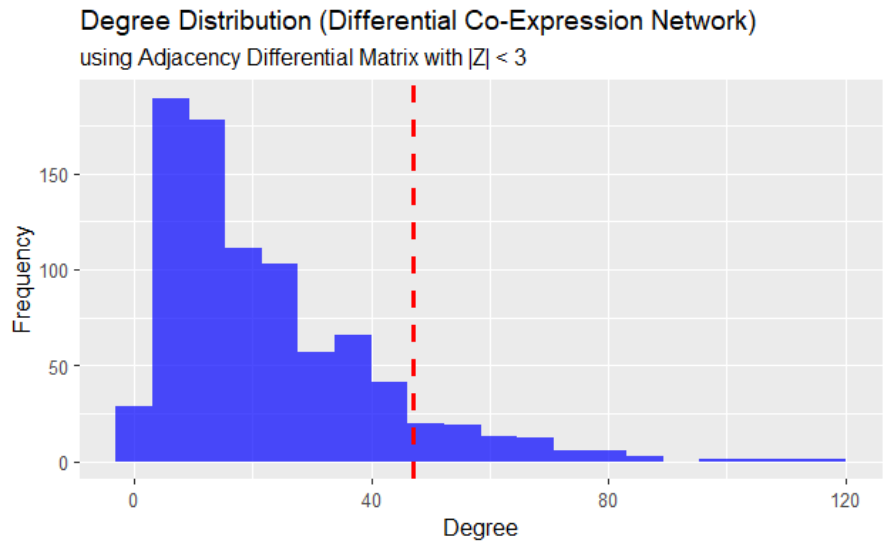


Fig. 6: Differential Co-Expression Network



Results

Multinetwork Analysis

- **Communities: 2**
 - *Resistance*
 - *Response*

| Community | Hypergeometric Test P-value |
|-----------|--------------------------------|
| 01 | 0.0014 |
| 02 | 0.0013 |

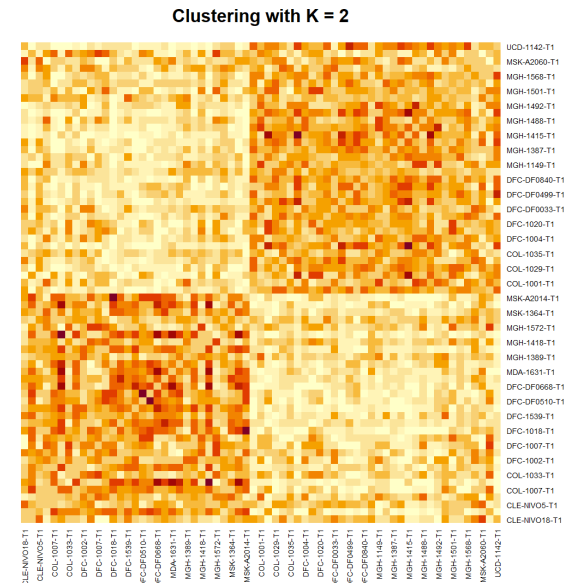


Fig. 7: Spectral Clustering with SNF



Results

Multinetwork Analysis

- **Communities: 2**
 - *Resistance*
 - *Response*

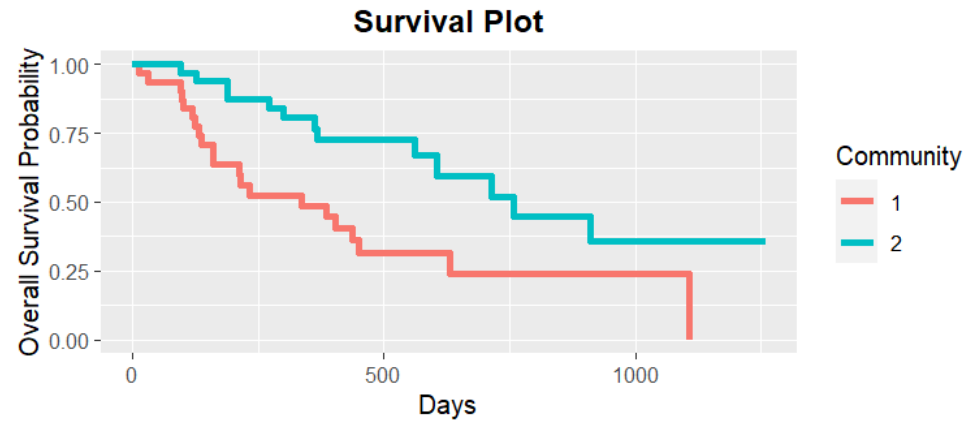


Fig. 8: Survival Analysis for SNF Clusters

| Community | Median Survival Time (Days) |
|-----------|-----------------------------|
| 01 | 214 |
| 02 | 470 |

| Cox Log Rank Test P-value |
|---------------------------|
| 0.004 |



Conclusions

- **Data Integration Significance:** Highlight the critical importance of integrating diverse data sources when studying complex and heterogeneous phenomena.
- **Power of Network Analysis:** The study demonstrates how network approaches can reveal hidden information within complex datasets.
- **Value of Preliminary Analyses:** Selection of genes to avoid the information redundancy expressed by the many genes.

Master's Degree Thesis for Sapienza University of Rome

Advisor: Prof. Manuela Petti

Co-Advisor: Dr. Lorenzo Farina

Faculty: Information Engineering, Computer Science and Statistics

Course: Data Science

Academic Year: 2022-2023

Author: Davide Mascolo, 2001991

Thank You For Your Attention



SAPIENZA
UNIVERSITÀ DI ROMA