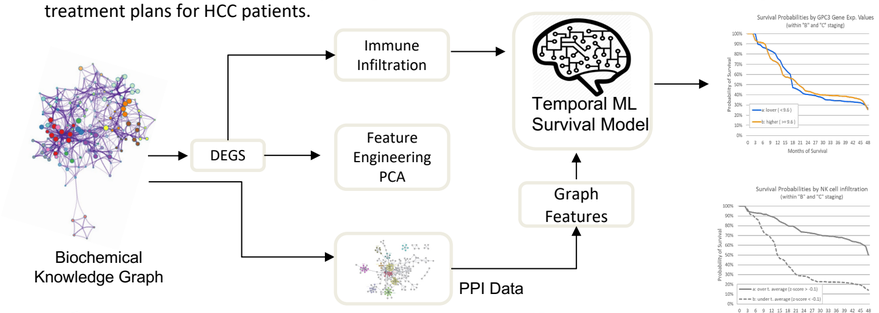
1243 Leveraging Graph AI to analyze the influence of GPC3 Gene expression and NK-Cell tumor infiltration on Hepatocellular Carcinoma Survival rates

Abhik Seal[1], Yaron Ilan[2], Sergii Zakharov[1], Olga Lyudovyk[3], Marton Szel[1], Gyorgy Lajtai[1], Armin Rath[4], Dan Chiche[4], Tony Arulanandam[5], Daniel Teper[4]

[1]Lynx analytics PTE ltd , [2]Department of medicine, Hadassah medical center, Jerusalem, Israel , [3]Computational biology and medicine program, Weill Cornell medicine, New York, NY, [4]Naya Biosciences Inc., Miami, FL, [5]Cytovia Therapeutics, Natick, MA

Introduction

- The study utilizes Graph AI and digital cytometry to analyze gene expression and immune cell infiltration in hepatocellular carcinoma (HCC), providing a comprehensive view of the tumor microenvironment and its impact on patient survival l[1].
- Glypican-3 (GPC3) is identified as a significant biomarker, with higher expression levels associated with shorter survival times. The study highlights the mitigating effect of natural killer (NK) cell infiltration on the adverse outcomes linked to high GPC3 expression, suggesting enhanced immune surveillance [2-3]. Ongoing studies are evaluating the effects of other immune cell types.
- The study demonstrates that models incorporating gene expression data and graph features outperform those based solely on demographic and laboratory results, with PCA-derived features showing clear differentiation between tumor and non-tumor samples.
- The findings underscore the potential of targeting GPC3 and enhancing NK cell activity as therapeutic strategies, with implications for improving the efficacy of immunotherapies and developing personalized treatment plans for HCC patients.



Methods

Using Graph Features to Model Survival

- For each gene in each sample, the difference in normalized expression between tumorous and non-tumorous states was calculated using the formula $D_{g,s} = E_{T,g,s} - E_{NT,g,s}$ This difference was then normalized using the standard deviation of expression levels in non-tumorous samples.
- We constructed a PPI based subgraph from Biochemical Graph, where nodes represent genes or proteins, and edges represent their interactions. The method combined gene expression data with network topology features (degree, PageRank, centrality) derived from protein-protein interaction networks. . Weighted averages of these features, based on correlation thresholds, were calculated to create comprehensive gene profiles The ECLAT algorithm [5]was used to identify frequent itemsets within gene expression data .
- We tested different feature sets in the context of predicting patient outcomes like demography, medical laboratory results, immune infiltration data , graph features and gene expression values
- **Survival Curves and log-rank tests**: survival curves were compared of different patient cohorts (e.g., split by gender, age, tumor size, staging, AFP-level, etc.) for identifying the important factors
  - ✓ **Using digital cytometry**: NK-cell infiltration is estimated by digital cytometry, and the survival probabilities were compared of patients with low/high NK-cell infiltration values
  - ✓ **Finding relevant gene expressions and conditions**: those conditions and gene expression values were collected which makes the NK-cell infiltration more significantly separate the survival curves
  - ✓ Testing with time-dependent ROC-AUC: for each month (between 6 and 48), predict whether a patient survives or dies, and calculate the models' ROC-AUC value on the test set of these binary classification models (0.5=random prediction, 1=perfect prediction).

Results

- The goal of the analysis is to check the different states (gene expression values, viral status, etc.) of B&C stage patients and identify subgroups where the NK-cell treatment could be more effective (Fig 2)
- Patients with Barcelona Clinic Liver Cancer (BCLC) stage B and C hepatocellular carcinoma (HCC) staging have 2.9-times higher 36-month survival probability if their **NK-cell infiltration is above average** (48 out of 228 patients had BCLC stage B or C) (Fig. 3A and 4A) This difference is even higher (4.6x) for the 29 B&C stage patients with **NK-cell infiltration and high GPC3** gene expression (above 9.6) vs B&C stage patients with low GPC3 (Fig. 3B and 4A).
- Over time, the performance of all models generally declines, as seen by the downward trends of the curves. The model with "Fuzzy PCA and Demography" (Purple line) tend to perform better over time compared to the other models, suggesting that the inclusion of fuzzy logic-derived features potentially improves the predictive accuracy of survival models. Fig 4B
- There's a general trend of increasing ROC-AUC scores over the first 10 to 20 months, after which the performance stabilizes. This could suggest that the models are more accurate at predicting short- to medium-term survival than long-term outcomes.
- Among B&C stage patients, for those with **high in-tumor GPC3 gene expression** values, the NK-cell density makes a bigger difference in survival → high in-tumor GPC3 values could be used to select patients for clinical trial enrollment [6] .

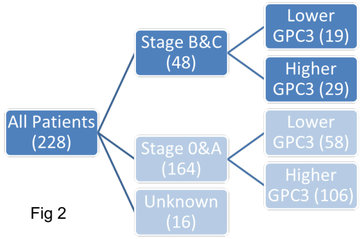
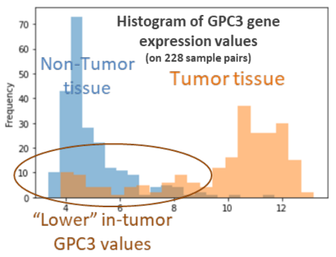


Fig 2



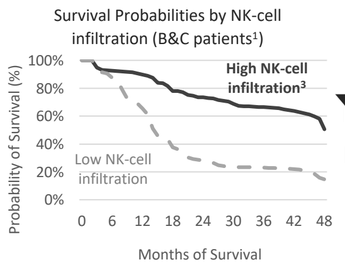


Fig 3 A

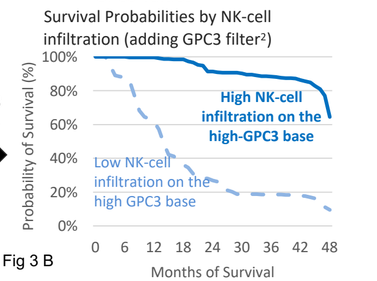


Fig 3 B

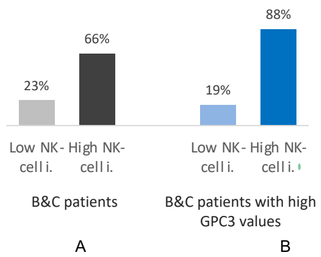


Fig 4A

36-month Survival Probabilities by NK-cell infiltration

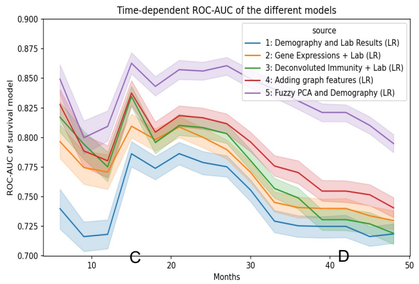


Fig 4B

Plot shows performance of Time dependent ROC-AUC curves using 5 different features with Logistic Regression.

- Survival probability is lower with **high Alpha-Fetoprotein (AFP)** level (over 300ng/ml) (Fig. 5A)
- For patients with **high AFP**, estimated NK-cell infiltration strongly differentiates patients with high/low survival probabilities (Fig. 5D)
- The models using gene expression data predict the survival chances significantly better than models relying only on only demography and medical laboratory and it improved more when data from immune deconvolution and graph where incorporated. **Table 1**

Data	ROC (Demography +Lab)	ROC (Gene Expressions + Lab)	ROC (Immune deconv + Lab)	ROC(Immune deconv + graphs
train	0.7705	0.8257	0.838	0.8418
test	0.750	0.7744	0.775	0.7783

Table 1

Survival Probabilities by AFP level (above/below 300mg/ml)

Survival Probabilities by NK-cell infiltration (all patients)

Survival Probabilities by NK-cell filtration (patients with low AFP)

Survival Probabilities by NK-cell infiltration (patients with high AFP)

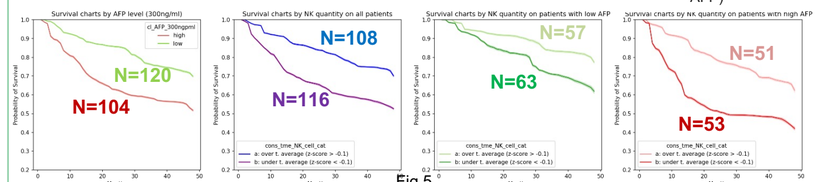


Fig 5

Figure 5 a) survival probabilities based on AFP levels using a threshold of 300 ng/ml b) Impact of NK cell quantity on survival probabilities for all patient's tumor average z-score > -0.1 or < -0.1 c) relationship between NK cell quantity and survival probabilities in patients with low AFP levels d) relationship between NK cell quantity and survival probabilities in patients with high AFP levels

Conclusions

Targeting GPC3 with NK-cell engager like NY-303 show the potential to improve survival outcomes.

- Glypican-3 (GPC3) and Alpha-Fetoprotein (AFP) expression was identified as a significant factor influencing survival. Higher GPC3 and AFP levels were generally associated with shorter survival times, but this adverse effect was mitigated by high NK cell infiltration. This interaction underscores the potential of targeting GPC3 in combination with NK cell-based therapies to improve survival outcomes.
- The integration of clinical, genomic, ppi data, transcriptomic data and graph features with machine learning models refines the prediction of HCC patient survival, demonstrating the value of a comprehensive approach to cancer prognosis and enhancing the understanding of the tumor microenvironment.
- Models are built without knowing the applied therapy on the patients. Further improvement can be reached by considering treatment like NY-303 (a GPC3 targeting NK engager), and by considering more features or collecting more external data sources.

References

1. Newman AM, etal. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019 Jul;37(7):773-782
2. Capurro M, Wanless IR, Sherman M, et al. Glypican-3: a novel serum and histochemical marker for hepatocellular carcinoma. Gastroenterology. 2003;125(1):89–97
3. Vivier E , etal. Innate or adaptive immunity? The example of natural killer cells. Science. 2011 Jan 7;331(6013):44-9
4. Szklarczyk D, etal. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023 Jan 6;51(D1):D638-D646.
5. Zaki M.J. Scalable algorithms for association mining. IEEE transactions on knowledge and data engineering, 2000, 12(3), 372-390
6. Arulanandam A, etal. Derivation and Preclinical Characterization of CYT-303, a Novel NKp46-NK Cell Engager Targeting GPC3. Cells. 2023; 12(7):996. https://doi.org/10.3390/cells12070996