Exploring the translational potential of influencing macrophage polarization in the tumor microenvironment in small cell lung cancer with a Boolean control network model

PhD thesis booklet

Gábor Szegvári

Semmelweis University Doctoral School
Károly Rácz Conservative medicine Division

Supervisor: Zoltán Lohinai, MD, Ph.D.
Official reviewers: János Fekete, Ph.D.
Tamás Korcsmáros, Ph.D.
Head of the Complex Examination Committee:
György Reusz, MD, D.Sc.
Members of the Complex Examination Committee:
László Szabó, MD, Habil.
Tamás Micsik, MD, Ph.D.

Budapest
2024

1. Introduction

1.1. Small Cell Lung Cancer

Small-cell lung cancer (SCLC) is an aggressive malignancy of great concern, due to its high incidence of metastases and low survival rates (the median is barely above 1 year, even with therapy). In recent years the importance of taking into account the differences between observable subtypes and the parameters of the tumor microenvironment (TME) has come to attention.

Subtypes of SCLC are assigned based mainly on available biomarkers, and gene expression patterns. E.g. according to the levels of neuroendocrine genes, high expression (NE-high) and low expression (NE-low) types are defined. The clinical relevance of subgroups needs further exploration as they represent phenotypes with potentially unique biological and clinical behavior.

Cancer cells can only exist if they are able to modulate both their immediate surroundings (the tumor nest) and the adjoining tissue (the stroma). They affect tissue structure and angiogenesis, as well as the infiltrating immune cells, creating a unique TME. Characterization of the immune components (density and (sub)types of cells, cytokine milieu) associated

with the tumor might help in understanding sensitivity and resistance to immunotherapies and in developing novel treatments of SCLC.

T lymphocytes are often in the focus of studies, and both $CD4^+$ helper and $CD8^+$ cytotoxic cells play an important role in the TME. However, the organization of the immune response is more complex, and there is a great deal to be learned from the examination of the role of myeloid cells in tumors, a topic not yet explored in SCLC. Macrophages are especially interesting, as they can act both in support of and against tumor progression. This is due to the remarkable level of post-differentiation plasticity they exhibit.

1.2. Macrophage polarization

Immune cells have to react to a wide variety of antigens and physiological and pathological environments and are thus capable of serving a multitude of functions. In particular, macrophages can direct or adjust the adaptive immune response with cytokine production and as antigen presenting cells (APC) and can undergo a process called polarization. In response to their surroundings, they dedicate themselves to a proinflammatory or a tissue regenerative modality. Polarization is dynamic, and the possible states form a continuous spectrum. The results of polarization are subject to modification based on

incoming cues, like cytokines, damage- and pathogen-associated molecular patterns and cell-cell interactions. In the TME, we can often find factors that aim to influence polarization in favor of the tumor. This milieu can give rise to so-called Tumor-Associated Macrophages (TAM), which mostly resemble regenerative-focused cells, but also have some unique traits.

Due to the interactions of immune cells, the effects of polarization can have a wider reach than just macrophages: they impact, among others, T-cell function and differentiation, and angiogenesis. We hypothesize that a therapy aimed at repolarizing macrophages could effectively alleviate the immunosuppressive nature of the TME.

The actual implementation of therapeutic repolarization presents unique challenges. Creating a model of the TME *in vitro* requires the setup or preservation of the complex interactions of multiple cell types. Here, I present an *in silico* model system to examine the influence of extracellular factors and therapeutic intervention on macrophage function. In addition, in order to incorporate the latest research data obtained, our group has also developed an online platform to aid researchers in drug repurposing for highly aggressive malignancies.

2. Objectives

➤ Characterize the immune cell population (especially TAM and MDSC) of the SCLC TME, according to neuroendocrine subtypes and localization in tumor compartments to identify influential cell types suitable for modulation.

➤ Analyze the gene expression patterns of SCLC to identify molecular targets in tumor subtypes according to immune cell infiltration and neuroendocrine (NE) marker expression status (low vs high).

➤ Build an in silico model of tumor associated macrophage polarization in response to cytokines and other small molecule transmitters in the tumor microenvironment to facilitate cancer research and therapy development in the field.

➤ Identify elements of signal transduction in macrophages as potential drug targets in order to modulate the polarization process, and counter or diminish the immunosuppressive effect of macrophages in the TME.

➤ Evaluate, in silico, the potential of combination therapies versus the single-target approach in macrophage repolarization.

➤ Create a user friendly and collaborative platform collating and presenting the latest drug target information to enhance SCLC drug target validations and drug repurposing.

3. Methods

3.1. Tissue sample processing

Samples were acquired from surgical resections of 32 patients with histologically confirmed limited-stage SCLC. After fixation in formalin, 5 micron slices were created, and the areas of cancer determined. $CD45^+$, $CD3^+$, $CD68^+$ and $CD163^+$ cells were immunohistochemically stained and counted with the cell counter plugin of ImageJ, by two independent observers. Classification of overall immune infiltration levels of samples was based on the number of $CD45^+$ and $CD3^+$ cells.

We measured the RNA expression levels of 2560 oncologically relevant genes using a specialized biomarker panel. This data was processed using R software packages to identify potential molecular targets. Expression levels of neuroendocrine genes were used as a basis for the NE-high or NE-low classification of samples.

3.2. In silico model

For the creation of the *in silico* model system, data was collected from freely available sources. The basis of the network was laid using protein-protein interaction (PPI) databases, including KEGG and STRING. This was then expanded manually and cross-checked with a systematic search of primary

research articles, including transcription factor (TF) – gene interactions, for which no database suitable for our purposes was accessible.

From this data a Boolean control network (BCN) was constructed. Components are represented by nodes that have two states: active and inactive. Their interactions are encoded as directed edges connecting the nodes. We derive the gene expression pattern induced by a given set of extracellular factors by letting their signal propagate through the system in discrete steps. Following a synchronous update scheme, the state of each node is determined based on the nodes directing edges at it. This process is iterated until an attractor is reached, in other words, a stable loop of system states (possibly consisting of a single one). Outputs are designated to be furthering an M1-like or M2-like polarization and an index score is calculated from them to place the given state of the cell on a spectrum between these two extremes.

Therapeutic intervention is simulated by setting a particular node or a combination of two of them to inactive. The effects of these perturbations are analyzed using the polarization index mentioned above. A synergy index is also calculated to highlight combinations of particular interest.

3.3. Online data collation platform

Our platform collects and organizes information from multiple databases. It reads a list of queries from either a Google Sheet document or a TSV file, then accesses the clue.io REST API to find matches. It automatically searches the FDALabel service and the website of the European Medicines Agency (EMA) and filters the results to provide direct links to relevant drug labels. Using the UniProt ID, data about the target is collected through the UniProt REST API, including Gene Ontology (GO) terms from the molecular function and cellular component categories, and references to STRING and Reactome. Using the compound name found in clue.io, a search request is sent to the PubMed database. Results are filtered for randomized controlled trials, clinical trials, reviews, systematic reviews and meta-analyses and the top 3 according to the best match algorithm of PubMed are recorded. This is then presented to the user through HTML. A more in-depth explanation of underlying functions is available at https://cycle20.github.io/EZCancerTarget/index.html.

4. Results

4.1. TME analysis

Based on cell densities, we observe that the immune infiltration of SCLC is lower than that of other lung cancers.

Both primary tumors and lymph node metastases were classified into the NE-high and NE-low categories. In our analysis, only 56% of infiltration-high tumors were classified NE-low, and only 87.5% of infiltration-low tumors as NE-high. Thus, while there does seem to be some connection between the two classifications, we conclude that all 4 subgroups should be handled as separate.

The distribution of $CD33^+$ (myeloid), $CD68^+$ (macrophage) and $CD163^+$ (M2-like macrophage) cells shows a particular pattern in our samples. First, compared to the tumor nest, in the stroma overall immune infiltration is significantly higher, and it has $CD33^+$ cells. Second, in primary NE-low tumors, compared to NE-high samples we can detect a significantly higher density of $CD68^+$ and $CD33^+$ cells in the stroma, and a significantly higher density of $CD68^+$ and $CD163^+$ cells in the tumor nest.

Comparing our measurements with previous results from our group, we can also conclude that myeloid cells show an

association with overall immune infiltration markedly different from $CD3^+$ (T lymphocyte) cells. CD3 expression correlates strongly with that of CD45 (a pan-leukocyte marker) in all subtypes and compartments. CD68 shows lower correlation that is more pronounced in the nest compartment of primary tumor sites. CD163 also associates with CD45 more in primary tumors, but its correlation often considerably exceeds that of CD68, in particular in the stroma of primary tumors. This supports that macrophage infiltration and polarization are influenced by different factors, and both should be examined for a comprehensive understanding of their function in the TME.

We have identified potential molecular targets, based on RNA expression data. Between the patterns in primary sites and metastases there are very few genes showing a considerable difference. Based on the level of infiltration a handful of genes get highlighted, notably CD70, CXCL9, CXCR2, FCGR1A, GZMA, HLA-B, MMP7, ITGAM, IFI27, and YBX3 in infiltration-high cases and CDH2, CHGA, FGF5, GRP, IL9, INS, ISL1, NCAM1, and SYP in infiltration-low cases. Based on NE status we can see a higher number of genes, many with higher fold change and significance. In NE-low: ANXA1, CD44, CD70, CXCR2, FCGR1A, HLA-B, IFI27, ITGAM, ITGB6, ITGB4, KRT5, MYC, MMP7, YBX3. In NE-high:

CDH2, CHGA, FGF5, GRP, ISL1, NCAM1, NKX2, SOX3, SYP. Genes in the NE-high group are mainly neural or neuroendocrine differentiation factors. The NE-high and infiltration-low, and NE-low and infiltration-high lists show considerable overlap.

4.2. Network model of macrophage polarization

I have constructed a network of 106 nodes and 217 edges, representing the processes responsible for early events in macrophage polarization. I have verified it based on reported experiments linking inputs to output points (satisfying 94.5% of criteria, 52 out of 55). General network parameters including degree distribution, and betweenness centrality also support the model's validity. The network's degree distribution is close to scale-free, a quality observed in most biological molecular networks. Betweenness centrality is higher for well-recognized central actors of the pathways involved.

Using this model, I explore all 511 combinations of 9 different inputs, both with and without outside perturbations. 67 inner nodes are used as targets of inhibition for therapeutic purposes. Targeting a single one and all 2211 combinations of two are both simulated, and their impact evaluated.

I observe that the system is fairly robust to the inhibition of a single node. It can have some effect on the polarization

status, but is not generally sufficient to create a shift from one end of the spectrum to the other (i.e. repolarization). With only one target node inhibited at a time, from a total of 34,237 instances (67 targets x 511 input combinations), we see a considerable change in the polarization index in 3000 (8.76%), but only 4 (0.01%) show repolarization. With the parallel inhibition of two nodes, out of 1,129,821 total instances, 528 are capable of repolarization. While this is still a low number in a relative sense (0.05% of all instances and 0.315% of those creating a considerable polarization shift), it is a substantial amount in regards to identifying new therapeutic options.

Examining the combinations in more detail, I found that in most instances, one part of an effective combination has little to no effect when inhibited alone. Repolarization to an M2-like state is achieved by targeting NFAT5, a node with negligible effect if inhibited alone, and another node with a higher individual impact. The pairs show a high synergy, indicating convergent pathways being affected. Repolarization to an M1-like state involves inhibition of STAT6 in most cases, but with the inclusion of another target, we can see the effect in a higher portion of input combinations. The other proteins most prominent in combinations are JAK1 and JAK3, with Tyk2, STK4 and Sp1 also present in multiple instances. These pairs

also show low synergy, implying the involvement of parallel pathways.

4.3. Online tool for target based drug repurposing

We built an online platform that gathers and organizes information on existing molecular entities associated with biological targets to streamline drug repurposing efforts, which is especially important in translational research into more scarcely explored diseases, like SCLC and other malignancies. Targets acting as input for the platform are given with HUGO and UNiProtKB ID-s. A default list of targets is provided, but it can be extended or replaced by the user. A user-defined label can also be attached for categorization purposes. Entries without an associated drug registered are automatically filtered out. Users are able to browse results per target, with all known drugs fit for repurposing listed, along with their mechanism of action, clinical status, and relevant links to our sources. In addition, information providing biological context is made available for each target to aid in the interpretation of results from larger input libraries. The platform is available at https://github.com/cycle20/EZCancerTarget.

5. Conclusions

➤ $CD33^+$ MDSC in the stroma, and $CD163^+$ M2-like TAM cells in the tumor nest are significantly more prevalent in NE-low SCLC, compared to NE-high, establishing an immunosuppressive TME with potential future therapeutic applications.

➤ In primary tumors, leukocyte numbers correlate with both macrophage markers (CD68 and CD163) in the nest, but in the stroma only with CD163, i.e. M2-like cells, implying distinct mechanisms of colonization in different tumor compartments.

➤ Interventions aimed at tipping the balance of immune infiltration and suppression are expected to be particularly effective in CD68- and CD163-high subsets of NE-low tumors. We identified CD70, ANXA1, FCGR1A, ITGB6, MMP7, YBX3 and CXCR2 as potential targets for this purpose.

➤ NE-low and infiltration-high tumors showed highly similar expression profiles at both primary sites and lymph node metastases. Smaller subgroups of NE-high, but phenotypically infiltration-high and NE-low, but phenotypically infiltration-low tumors exist, with distinct expression profiles.

➤ Our in silico model can replicate experimentally observed cell behavior in its scope and shows the expected characteristics

of a complete network encompassing the selected pathways involved in macrophage polarization.

➤ When aiming to affect repolarization, interventions designed against a single target are suboptimal compared to treatments simultaneously targeting two intracellular actors. Individual targets in combination therapies might not show a significant effect on cellular state on their own.

➤ When attempting to repolarize cells into an M1-like state, our model highlights STAT6, JAK1 and JAK3, and to a lesser extent Tyk2, STK4 and Sp1 as potential targets. When trying to push the system toward an M2-like state, NFAT5 emerges as a point of interest.

➤ We created a platform for drug repurposing in oncopharmacology that is accessible and expandable by the international research community, intended to aid research into SCLC and other highly aggressive malignancies.

6. Bibliography of the candidate's publications

Related to the thesis

➤ Szegvári et al..: Effective Reversal of Macrophage Polarization by Inhibitory Combinations Predicted by a Boolean Protein-Protein Interaction Model, Biology (Basel) 2023

➤ Dóra et al.: Characterization of Tumor-Associated Macrophages and the Immune Microenvironment in Limited-Stage Neuroendocrine-High and -Low Small Cell Lung Cancer, Biology (Basel) 2021

➤ Dóra et al.: EZCancerTarget: an open-access drug repurposing and data-collection tool to enhance target validation and optimize international research efforts against highly progressive cancers. BioData Mining, 2022

Other publications

➤ Horváth et al.: Effect of allosteric inhibition of non-muscle myosin 2 on its intracellular diffusion, Scientific Reports 2020

➤ Bátora et al.: Subcellular dissection of a simple neural circuit: functional domains of the Mauthner-cell during habituation, Frontiers in Neural Circuits 2021