



OPEN ACCESS

EDITED BY

Russell C. Rockne,
Beckman Research Institute, United States

REVIEWED BY

Marco Vanoni,
University of Milano-Bicocca, Italy
David Linnaeus Gibbs,
Institute for Systems Biology (ISB), United States

*CORRESPONDENCE

Jonatan Blais,
✉ jonatan.blais.med@ssss.gouv.qc.ca

RECEIVED 28 November 2023

ACCEPTED 01 February 2024

PUBLISHED 01 March 2024

CITATION

Blais J and Jeukens J (2024), Pareto task inference analysis reveals cellular trade-offs in diffuse large B-Cell lymphoma transcriptomic data.

Front. Syst. Biol. 4:1346076.

doi: 10.3389/fsysb.2024.1346076

COPYRIGHT

© 2024 Blais and Jeukens. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Pareto task inference analysis reveals cellular trade-offs in diffuse large B-Cell lymphoma transcriptomic data

Jonatan Blais^{1*} and Julie Jeukens²

¹Oncology Research Axis, Centre de Recherche du CHU de Québec-Université Laval, Québec City, QC, Canada, ²Department of Laboratory Medicine, CHU de Québec-Université Laval, Québec City, QC, Canada

One of the main challenges in cancer treatment is the selection of treatment resistant clones which leads to the emergence of resistance to previously efficacious therapies. Identifying vulnerabilities in the form of cellular trade-offs constraining the phenotypic possibility space could allow to avoid the emergence of resistance by simultaneously targeting cellular processes that are involved in different alternative phenotypic strategies linked by trade-offs. The Pareto optimality theory has been proposed as a framework allowing to identify such trade-offs in biological data from its prediction that it would lead to the presence of specific geometrical patterns (polytopes) in, e.g., gene expression space, with vertices representing specialized phenotypes. We tested this approach in diffuse large B-cell lymphoma (DLCLB) transcriptomic data. As predicted, there was highly statistically significant evidence for the data forming a tetrahedron in gene expression space, defining four specialized phenotypes (archetypes). These archetypes were significantly enriched in certain biological functions, and contained genes that formed a pattern of shared and unique elements among archetypes, as expected if trade-offs between essential functions underlie the observed structure. The results can be interpreted as reflecting trade-offs between aerobic energy production and protein synthesis, and between immunotolerant and immune escape strategies. Targeting genes on both sides of these trade-offs simultaneously represent potential promising avenues for therapeutic applications.

KEYWORDS

pareto theory, transcriptomics, lymphoma, oncology, archetypes, trade-offs, optimality, systems biology

Introduction

Oncology has made tremendous progress over the last few decades ([Cancer Progress Timeline | ASCO](#)). New treatment modalities have been developed and existing ones have been refined, combined and optimized to the extent that clinical outcome, including overall survival, has improved significantly for many tumour types (e.g., [Arnold et al., 2019](#)). Yet, in many cases, despite these advances, and although response to therapies is often initially good for many types of cancer, resistance develops and leads to relapse/refractory tumours ([Vasan et al., 2019](#)). For too many patients, a cancer diagnosis is still synonymous with premature death and cancer still represents one of the main causes of mortality and morbidity in human populations.

With the development of “omics” sciences, the last few decades have also seen the accumulation of a huge amount of publicly available data, especially DNA and RNA sequences, as well as data about numerous signalling pathways and cellular processes. We are therefore facing a situation where the amount of available data is no longer the main limiting factor in advancing cancer therapy. Although, the quality of data in terms of spatio-temporal correlations with respect to relevant pathophysiological processes may still remain out-of-reach of currently available resources and/or technological possibilities. In any case, the availability of such “big data” has yet to propel cancer research to dramatic new therapeutic advances or to cure cancer. Indeed, immunotherapy is probably the most significant innovation in cancer treatment of the last decade and its development does not directly derive from the omics revolution but rather from a more conventional incremental increase in our knowledge of immunology (Paucek et al., 2019).

Arguably, one of the main reasons explaining this inability to take full advantage of the information contained within these public databases stems from the lack of a theoretical framework, organizing principles and methodology allowing to target only their most biologically relevant elements (thus turning data into information). Such a theoretical framework should provide a way to understand the most crucial and essential tasks that malignant cells must perform in order to survive and the genes used to perform these tasks, with the ultimate goal of selectively eliminating malignant cells.

Typically, cancer is understood to be the result of the accumulation of random mutations in genes involved in processes such as cell cycle control, apoptosis and DNA repair, eventually producing the malignant phenotype recognized as cancer (Sonnenschein and Soto, 2020). This view, often called the “somatic”, “reductionist” or “somatic mutation theory (SMT)” does not offer any specific guiding principle to navigate omics data (Selvarajoo and Giuliani, 2023). Under this classical paradigm, basic cancer research aims at providing increasingly detailed descriptions of various genetic alterations and their impact on cellular pathways, but with no reference to a deeper organising theoretical framework to increase our understanding of the phenomenon and thus becomes an essentially descriptive exercise (Monti et al., 2022).

In order to solve what is currently recognized as one of the main problems in cancer therapy, namely, the development of resistance to previously effective treatments (Bukowski et al., 2020; Saleh and Elkord, 2020), an alternative approach might be to try gaining insight into cancer cell vulnerabilities by looking at the fitness of different genotypes through analysis of the genomic fitness landscapes. In principle, this exercise should allow the prediction of cancer evolution under various circumstances. The genomic fitness landscape allows not only the determination of the selection coefficient and gradient associated with any particular allele, but also the identification of critical biological functions in the malignant phenotype and the prediction of response to therapy. In practice however, this is a daunting task as it requires within patient serial measurements at single-cell genomic resolution (Salehi et al., 2021), which is currently out of reach for most research teams.

Interestingly, however, an approach from systems biology may alleviate the need to characterize tumors’ fitness landscapes to identify biologically critical tasks and functions performed by

tumor cells. This idea relies on the fact that living organisms and cells need to optimize multiple tasks simultaneously and inevitably face trade-offs in doing so. Mathematically, it can be shown that such constraints imply that the data will form polytopes (n-dimensional generalization of 3-dimensional polyhedra: geometrical object with flat polygonal faces, straight edges and sharp vertices) in trait space (e.g., gene expression space) where vertices of the polytope each represent the optimal specialist phenotype for performing a given task (Hart et al., 2015). Known as Pareto task inference (implemented in the ParTI package in MatLab available at: [Pareto Task Inference \(ParTI\) method | Uri Alon \(weizmann.ac.il\)](#)), this idea comes from the general Pareto optimality theory that has been used in fields like engineering and economics for many decades in order to solve multi-task optimization problems involving trade-offs between different objectives. Briefly, ParTI allows the use of algorithms to find the best fitting simplex (the simplest polytope) on the data after dimensionality reduction by principal component analysis (PCA), and to test the statistical significance of the simplex by means of the t-ratio test. This test computes the ratio between the area encompassed by the fitted polytope to the area defined by the convex hull of the data (the t-ratio) and then repeat this process after randomizing the data in order to obtain an empirical null distribution for the t-ratio. Pareto task inference has so far been applied to a few biological datasets, including solid tumor gene expression (Hausser et al., 2019; Hausser and Alon, 2020). Most importantly, by circumventing the need to reconstruct the fitness landscape, this method has the potential to directly unveil the most relevant cellular processes for therapeutic intervention by allowing one to infer the fitness consequences of disrupting the cell’s ability to carry out specific tasks.

Although some authors have raised the possibility that the ParTI algorithm might produce spurious results due to phylogenetic, ancestral or other population structure (Sun and Zhang, 2020; Mikami and Iwasaki, 2021), Adler et al. (Adler et al., 2022) tested whether such problems could affect cancer gene expression data and found no evidence that this was the case. The risk of false-positives due to “p-hacking” through data preprocessing has also been raised (Sun and Zhang, 2020). Amid these concerns, rather than speculating on possible artefactual polytopes or on the possibility of statistical false-positives, a more fruitful way of addressing the validity of the ParTI method is to derive specific predictions of the Pareto optimization theory besides the geometrical prediction of polytopes in trait space. In particular, the biological interpretability and underlying logic of the results obtained from that method appear to be of particular relevance in that regard.

For instance, if the most specialized cellular phenotypes (the “archetypes”) identified as the vertices of the polytope in gene expression space are the results of artefactual data structure unrelated to biological specialization and optimization, one would not expect to find statistically significant and distinct enrichment in particular biological functions at these archetypes. Moreover, if polytope formation results from optimization trade-offs, as per the theory, the various tasks defining the archetypes should all be considered important, if not essential, for cell survival. Indeed, trade-offs imply that some of these tasks cannot be simply eliminated completely from the cell’s functional repertoire,

otherwise there should be no trade-offs constraining the optimization process. Hence, archetypes represent different solutions of proportional allocations of resources to these essential functions. It is therefore likely that different archetypes would show different proportions of a number of shared functions/genes, reflecting these different solutions. Alternatively, a pattern where all vertices are characterized by a unique and distinct set of functions and gene-transcripts would not be easily interpretable as the results of trade-offs constraining the cell's ability to maximize certain functions to their full genetic potential since each vertex would represent a unique and distinct functional domain. Thus, if the existence of the archetypes identified by the ParTI algorithm is the results of optimization trade-offs as predicted by the Pareto optimization theory, we should expect these archetypes to be statistically enriched in specific biological functions, and to be characterized by varying proportions of some shared tasks/genes in different combinations.

Because of its potential as a useful theoretical principle allowing meaningful analysis of omics and other biological data, and thus to provide important information about both basic biology and therapeutically actionable knowledge, the Pareto task inference approach should be carefully and independently evaluated and validated. So far, the ParTI package has been used on a few solid tumor datasets but, to our knowledge, never on hematological malignancies. Because of their cell type (malignant lymphocytes) and phylogenetic (from clonal proliferation) high level of homogeneity, data from lymph node biopsies of lymphoma patients are particularly well suited to study the validity of this method. In order to test the above predictions, we analyzed transcriptomic data from diffuse large B-cell lymphoma (DLBCL) samples for 481 patients using ParTI code. Archetypes identified by the algorithm were then tested for functional enrichment and compared in terms of gene composition and representation. Results were ultimately evaluated for their interpretability, potential biological meaning and usefulness for generating hypotheses.

Materials and methods

Gene expression data

Data were obtained through the National Cancer Institute's Genome Data Commons Data Portal ([GDC \(cancer.gov\)](https://gdc.cancer.gov)). Gene expression quantification expressed as transcripts per millions (TPM) from RNASeq was available for 517 DLBCL patient-lymph node biopsy samples (cases.disease_type in ["mature b-cell lymphomas"] and cases.primary_site in ["lymph nodes"]) from projects NCICCR-DLBCL, TCGA-DLBC, and CTSP-DLBCL1. Of these, 481 (93%) were from the NCICCR-DLBCL project, while only 29 (5.6%) and 7 (1.4%) were from the TCGA-DLBC and CTSP-DLBCL1 projects respectively. In order to ensure maximum homogeneity in sampling and sequencing protocol and hence between-sample comparability of data ([Zhao et al., 2020](#)), we analyzed the 481 cases from the NCICCR-DLBCL project. All 481 samples were fresh frozen lymph node biopsy samples, 96.3% collected prior to any treatment. The same RNASeq protocol was applied to all samples ([Schmitz et al., 2018](#)). We kept all informative

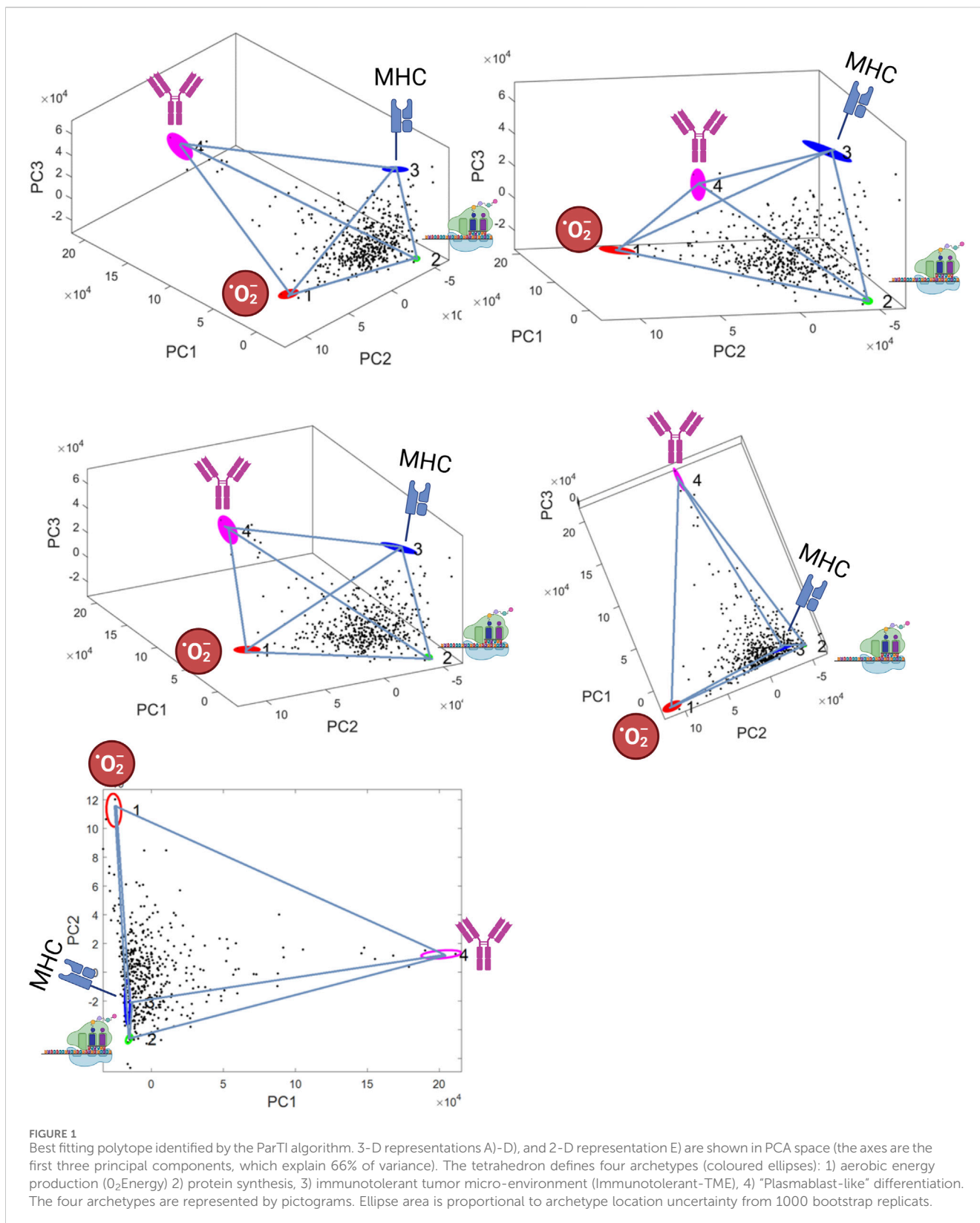
genes/isoforms (22 250), defined as having standard deviation and variance ≥ 1 TPM, for downstream analyses. To avoid p-hacking issues, we refrained from performing any preprocessing transformations/manipulations of data and used the raw TPM gene expression quantification. TPM is considered a better unit of RNA abundance than RPKM/FPKM since it respects the invariant-average property (the average TPM is a constant equal to 10^6 divided by the number of annotated transcripts) and is proportional to the average RNA molar concentration (rmc). It has thus been adopted by the latest computational algorithms for transcript quantification such as RSEM, Kallisto and Salmon ([Zhao et al., 2020](#)). The resulting data matrix had 481 sample lines by 22 250 genes/isoforms columns.

PCA and pareto task inference analyses

Because the ParTI code package relies on principal component analysis (PCA) for dimensionality reduction, as proposed by Vieira ([Vieira, 2012](#)), we first tested the validity of performing PCA on the data by computing the Ψ and ϕ statistics, the number of significant principal components, as well as the number of genes/isoforms with significant correlations with each of the principal components, by permutations and bootstrapping using PCAtest in R ([Camargo, 2022](#)).

Following Mikami and Iwasaki ([Mikami and Iwasaki, 2021](#)), the simplex best fitting the data was determined by the SDVMM algorithm ([Chan et al., 2012b](#)). Five such fitting algorithms are available in ParTI. The SISAL algorithm is not recommended for datasets of less than about 1000 data points (or more precisely of less than 10^N points, with N being the number of PCA dimensions), and will estimate the archetypes outside of the convex hull of the data, while potentially leaving important points outside of the fitted polytope and t-ratio test analysis (PartiCode homepage: [Pareto Task Inference \(ParTI\) method | Uri Alon \(weizmann.ac.il\)](#) accessed 11/03/2023), which may generate false positive or false negative results of the t-ratio test. The MVA and MVE algorithms, which also locate the archetypes outside of the convex hull, could be used with smaller datasets of <1000 points, but are not robust to noise and outliers. Contrary to the previous algorithms, the PCHA algorithm estimate the archetypes within the data but also suffer from a susceptibility to noise and outliers. Thus, we agree with Mikami and Iwasaki ([Mikami and Iwasaki, 2021](#)) that the SDVMM algorithm is the only option combining a strict data constraint with a robustness to noise and outliers and seems clearly preferable for statistical testing. The default value of eight PCA dimensions, as per the ParTI's example file "exampleCancerRNAseq.m", was kept as input. Using ParTI's discrete and continuous attribute functions, we also tested whether archetypes were associated with three clinical correlates available in the dataset, namely, tumor cell of origin (COO) subtype (which was available for all samples), progression free survival (PFS) in years (which was available for 48% of the samples), and the international prognostic index (IPI) score (which was available for 74% of the samples).

Archetype coefficients in original gene expression space were then extracted from ParTI's output and genes ordered according to coefficient value. Genes with the highest positive and negative coefficient values for each archetype (genes with the highest PCA



loadings at the archetype location and thus defining the archetype in terms of most positively/negatively correlated expression values) were identified by plotting ordered coefficient values and visually identifying the abrupt change in slope between the vast majority of

genes with coefficients close to zero and the few genes with highly positive (highly positively correlated) or negative (highly negatively correlated) values, and taking this inflexion point as cut-off. These archetype-defining gene lists were then tested for functional

enrichment using Gene Ontology tools (GO aspect: biological process) (Ashburner et al., 2000; Thomas et al., 2022; Aleksander et al., 2023). Functions with a Fisher's exact test enrichment p -value <0.05 (after adjustment for false-discovery rate (FDR)) were considered significantly enriched at the archetype. The composition of significantly enriched functions was compared and the level of shared vs. unique functions among archetype analyzed using Circos (Krzywinski et al., 2009).

Data availability

The dataset analysed for this study is publicly available on the Genome Data Commons Data Portal at [GDC \(cancer.gov\)](https://gdc.cancer.gov). The code needed to replicate the ParTI analysis from these data is provided as [Supplementary Table S1](#).

Results

PCA validation

PCAtest was run for 100 permutations and bootstrapping replicates and estimated an empirical Ψ value of 51848193.065, with a max null $\Psi = 979604.238$ and a min null $\Psi = 978848.881$, for a p -value <0.00001 . Likewise, the ϕ value was estimated at 0.337, with a max null $\phi = 0.047$ and min null $\phi = 0.047$, for a p -value <0.00001 . These results indicate highly significant non-random correlations in the data, justifying the use of PCA and confirming that the analysis is biologically meaningful. Moreover, the analysis showed that the first eight principal components (PC) explained 80% of the total original variance (from 1.8% to 33% for individual PC). Finally, the number of genes with significant correlations with each of these eight principal components ranged from 2830 to 7546. Together, these results strongly support the use of PCA on this dataset.

Polytope fitting

The elbow method applied by the PartiCode package suggested the presence of four archetypes, thus defining a tetrahedron (3-d simplex) ([Figure 1](#)). The identified polytope was highly significant with the t-ratio test indicating a p -value <0.00001 . The four vertices represented archetypes of 1) aerobic energy production (O₂Energy), 2) protein synthesis, 3) Immunotolerant tumor micro-environment (Immunotolerant-TME), and 4) Immune-escape via plasmablast-like differentiation ("Plasmablast-like"). A detailed description of the functional analysis characterizing these archetypes follows.

Archetype-defining genes selection

For each of the four archetypes, ordered coefficients were plotted and revealed a clear pattern of abrupt change from coefficients close to zero for the vast majority of genes, to a few genes showing strongly positive or negative coefficients ([Supplementary Figure S1](#)). Archetype-defining genes were therefore selected by considering a horizontal line (slope = 0) for genes with a coefficient close to zero

and a line with a slope of 1 (vertical) for genes with strongly positive or negative coefficients. These lines were extrapolated and the line bisecting these right-angled extrapolated horizontal and vertical lines at 45° was determined. The mid-point of this 45-degree line, when positioned so it just touched the data, was considered as the inflexion point (elbow point) and used as the coefficient cut-off to select the positive and negative archetype defining genes ([Supplementary Figure S2](#)).

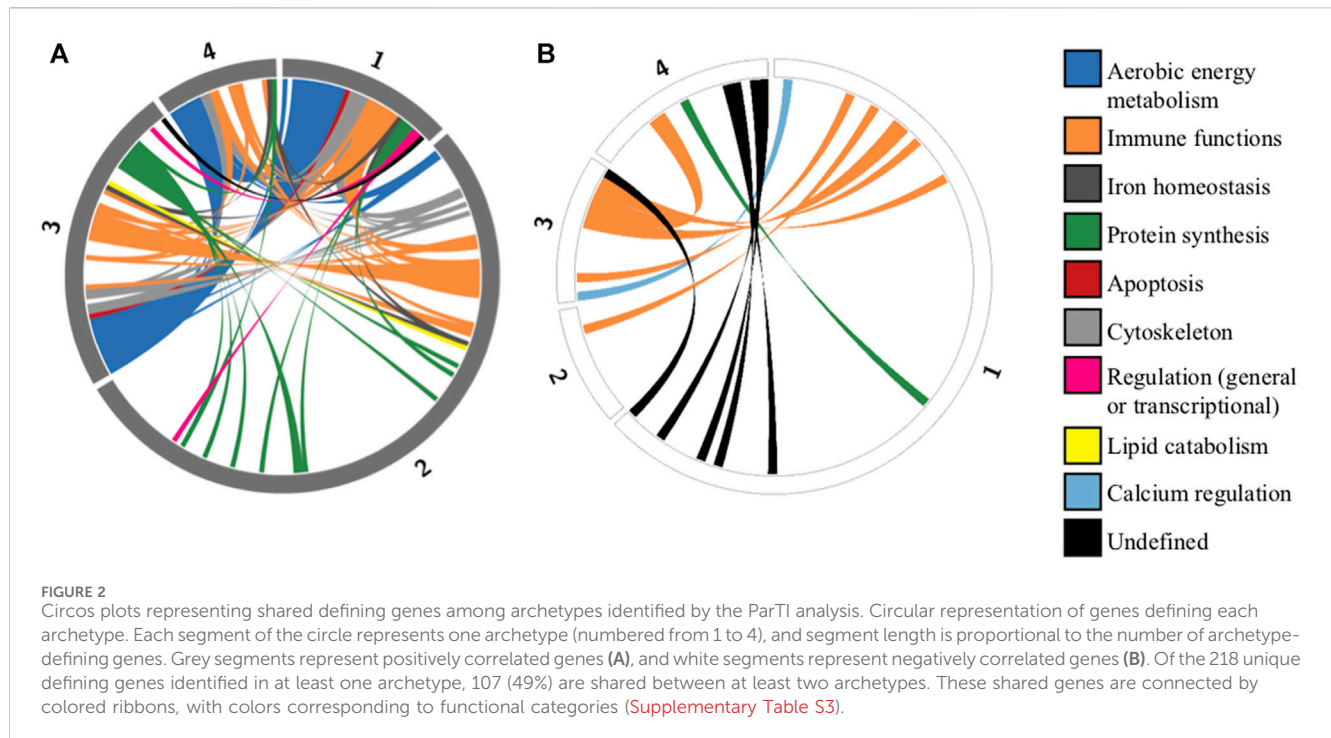
Enrichment analysis

The above procedure identified between 23 and 132 archetype-defining genes with positive loadings, and between 12 and 85 archetype-defining genes with negative loadings depending on the archetype ([Supplementary Table S2](#)). For each archetype, both positive and negative gene lists were submitted to the Panther database for enrichment analysis. All eight archetype-defining gene lists (one positive and one negative loading list for each of the four archetypes) were significantly enriched (after FDR adjustment) in some functions. The 10% functions with the most significant enrichment p -value for each of the eight lists are presented in [Supplementary Table S3](#). GO terms associated with each individual archetype-defining genes are listed in [Supplementary Datasheet S1](#).

Archetype characterization

The first archetype was positively enriched in aerobic energy production (three most significant GO terms: oxidative phosphorylation/aerobic respiration/aerobic electron transport chain) and negatively enriched in protein translation (three most significant GO terms: cytoplasmic translation/translation/peptide biosynthetic process). The second archetype was positively enriched in protein translation (three most significant GO terms: cytoplasmic translation/translation/peptide biosynthetic process) and negatively enriched in energy production and immune functions (three most significant GO terms: oxidative phosphorylation/adaptive immune response/cellular respiration). The third and fourth archetype were positively enriched in a mixture of aerobic energy production and immune functions, and negatively enriched in immune functions. The 10% lowest FDR-adjusted p -values of the enrichment analysis were very small, ranging from 0.0025 to 5.67×10^{-95} , indicating that, from a functional point of view, archetype-defining genes represented highly non-random genomic sub-samples. As predicted, a significant proportion (49%) of archetype-defining genes were shared between at least two archetypes, albeit in different combinations and proportions ([Figure 2](#)).

Archetype 2 and 3 were significantly associated with the COO "germinal center B-cell" (GCB) and "unclassified" subtypes respectively (p -value <0.001 and <0.01 ; hypergeometric distribution), and those associations remained significant after the Benjamini–Hochberg correction for multiple tests. Archetype 3 was also significantly associated with the IPI "intermediate" score (p -value = 0.031; hypergeometric distribution), although the p -value did not remain significant after Benjamini–Hochberg correction. Likewise, archetype 3 was significantly associated with a shorter PFS



(median difference = -3.2 years; p -value = 0.018; Mann-Whitney), while archetype 2 and 4 were marginally associated with a longer PFS (median difference = 3.1 and 3.8 years; p -value = 0.056 and 0.085; Mann-Whitney respectively) but these associations became non-significant after Benjamini-Hochberg correction. As mentioned previously, the sample sizes and (therefore statistical power) for testing archetypes association with IPI and PFS were reduced by 26% and 52% respectively.

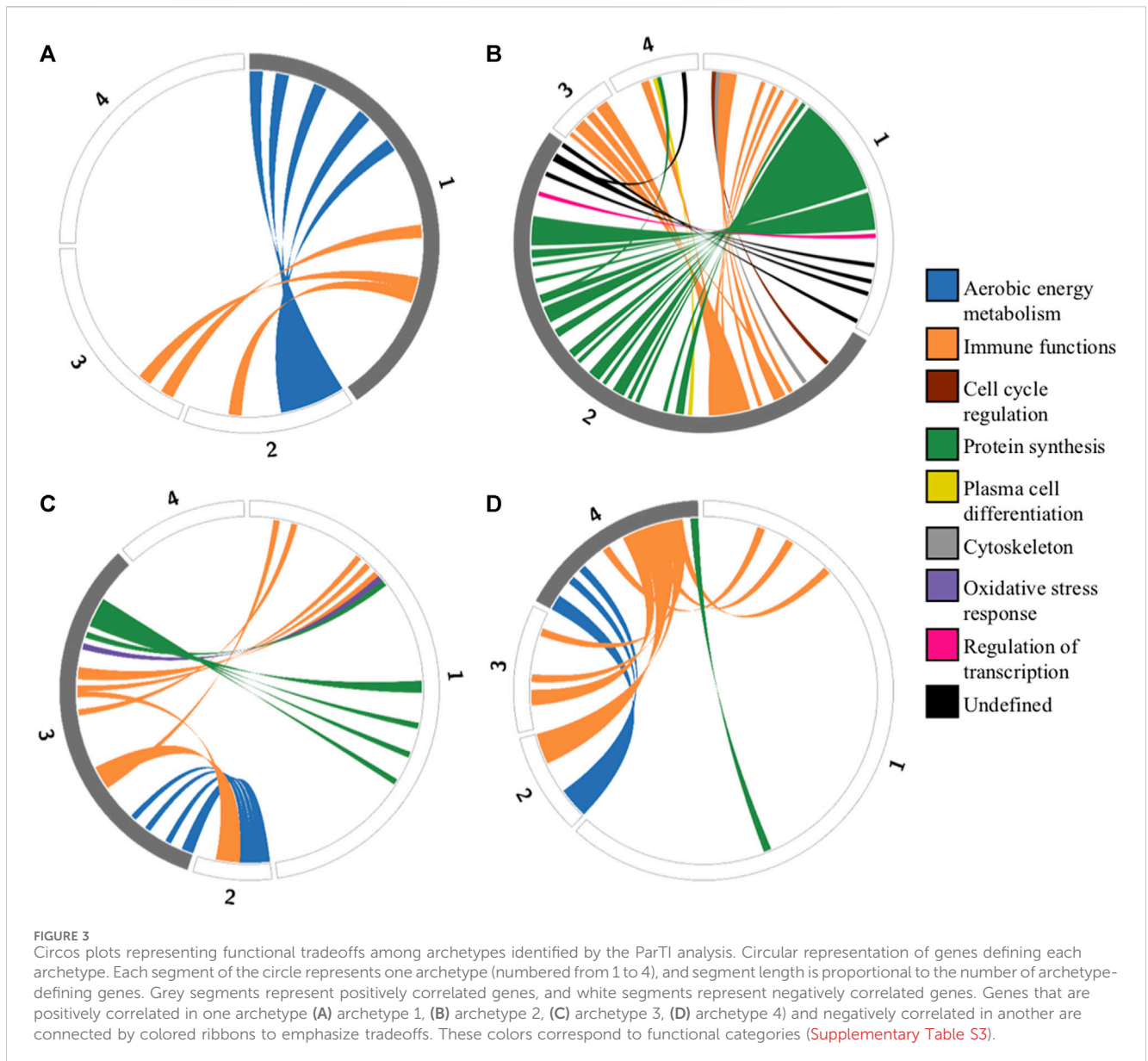
Biological interpretability

Archetype 1 and 2: metabolic trade-offs between aerobic energy production vs. protein synthesis

One of the most striking features of the functional enrichment analysis was the contrast between archetype 1, which showed specialization for oxidative phosphorylation (OXPHOS) at the expense of protein synthesis, and archetype 2 which was specialized in protein synthesis at the expense of both OXPHOS and immune functions (Figures 3A, B). A detailed analysis of the genes with expression positively correlated with archetype 2 and negatively correlated with archetype 1 revealed 34 genes involved in protein synthesis. These genes consisted of various ribosomal subunits and two translation elongation factors, as well as CHCHD2. Interestingly, it was recently shown that in stress conditions produced by carbonyl cyanide *m*-chlorophenylhydrazone (CCCP) treatment, a decoupling agent known to induce oxidative stress (Park et al., 2015), CHCHD2 knockdown triggered the integrative stress response (ISR) in cultured HeLa cells (Ruan et al., 2022). The role of the ISR is to maintain homeostasis under stress conditions, including

oxidative stress, and is known to slow down protein synthesis via the phosphorylation of eIF2 α (Bilen et al., 2022). Indeed, the ISR was recently observed to inhibit protein synthesis in the context of oxidative stress from mitochondria-derived production of reactive oxygen species (ROS) in a cardiac ischemia/reperfusion model (Zhang et al., 2021). Because of the high rate of energy production via OXPHOS and the associated ROS generated, archetype 1 cells may downregulate CHCHD2 in order to restore homeostasis via the ISR, at the expense of protein synthesis. In contrast, archetype 2 cells may overexpress CHCHD2 in order to benefit from its anti-apoptotic effect (Jiang et al., 2022). Recently, the importance of CHCHD2 in cancer and its potential as drug target was highlighted by Gundamaraju et al. (Gundamaraju et al., 2020). Consistent with the hypothesis of a driving role for increased ROS production in archetype 1, the expression of ROMO1 was negatively correlated with this archetype. ROMO1 is known to increase ROS production (Amini et al., 2019) and thus might be downregulated by archetype 1 as an adaptation to compensate for the high ROS generated by energy production via the electron transport chain.

Another noticeable contrast between archetypes 1 and 2 was the positive correlation of the expression of CKS2 and TPT1 with archetype 2, while the expression of these two genes were negatively correlated with archetype 1. CKS2 and TPT1 have both been implicated in cancer, including as potential therapeutic targets (Acunzo et al., 2014; You et al., 2015; Lee et al., 2022). TPT1 encodes the translationally controlled tumor protein (TCTP) and CKS2 encodes the cyclin-dependent kinase regulatory subunit 2, which is known to bind and be necessary for the activity of the cyclin B1-CDK1 protein kinase, an essential factor for cells to progress past the G2 phase of the cell cycle (Wang et al., 2021). Interestingly, both TPT1 and CKS2 have also been associated with



OXPHOS-induced oxidative stress (Lucibello et al., 2011; Jonsson et al., 2019), as well as with cell-cycle regulation and with protein synthesis/degradation (Bommer and Telerman, 2020; Grey et al., 2023). In an hematopoietic cells (HSC) mouse model, CKS2 knockout was associated with an accelerated cell cycle (Grey et al., 2018) (which may contribute to the malignant phenotype) but also with an increase ROS production (Grey et al., 2023). Moreover, the same study also found that CKS2 was involved in proteostasis of HSC, which may be related to the trade-off in protein synthesis identified for archetype 1 (Grey et al., 2023). The role of TCTP in protein synthesis regulation is known to involve its interaction with elongation factors eEF1A and eEF1B (Bommer and Telerman, 2020). The fact that both eEF1A1 and eEF1B2 followed the same correlation of expression patterns between archetypes 1 and 2 as TPT1 and CKS2 supports this interpretation.

In general, TPT1/TCTP is thought to protect cells against apoptosis and oxidative stress (Bommer, 2017). It might thus seem surprising that its expression should be negatively correlated with archetype 1 given the proposed trade-offs involving increased OXPHOS-induced ROS and ISR activation. However, while mild oxidative stress was found to upregulate TCTP, strong oxidative stress was found to downregulate its expression (Lucibello et al., 2011), and both CKS2 and TPT1 were downregulated in a butyrate resistant cell line conferring tumorigenesis, apoptosis and stress resistance in a colon adenocarcinoma model (López de Silanes et al., 2004). Thus, the level of oxidative stress in archetype 1 might reach levels associated with reduced TCTP expression.

TCTP and CKS2 expression patterns are linked through the TCTP/CDC25C/CDK1 pathway, which was shown to be dysregulated in hepatocellular carcinoma (Chan et al., 2012a).

Metabolic trade-offs

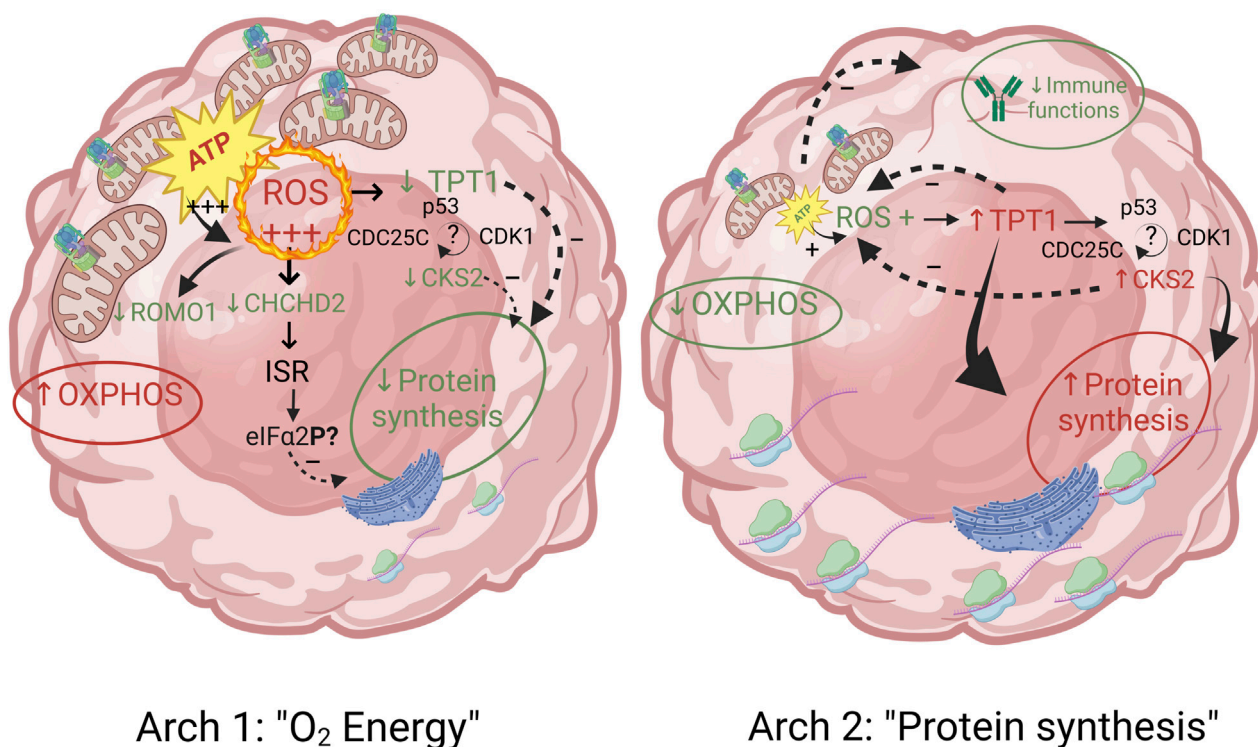


FIGURE 4

Metabolic trade-offs between archetypes 1 and 2 are characterized by genes involved in energy production, protein synthesis and immune functions. The proposed model of mechanistic relations underlying these trade-offs between aerobic energy production and protein synthesis involves a complex interplay between ROS production and its inhibitory effect on protein synthesis, and minimal energy requirement for protein synthesis and immune functions. Archetype defining functions and genes are shown in red (positively correlated) and green (negatively correlated). Functions and genes in black are hypothetically inferred and not part of archetype defining genes.

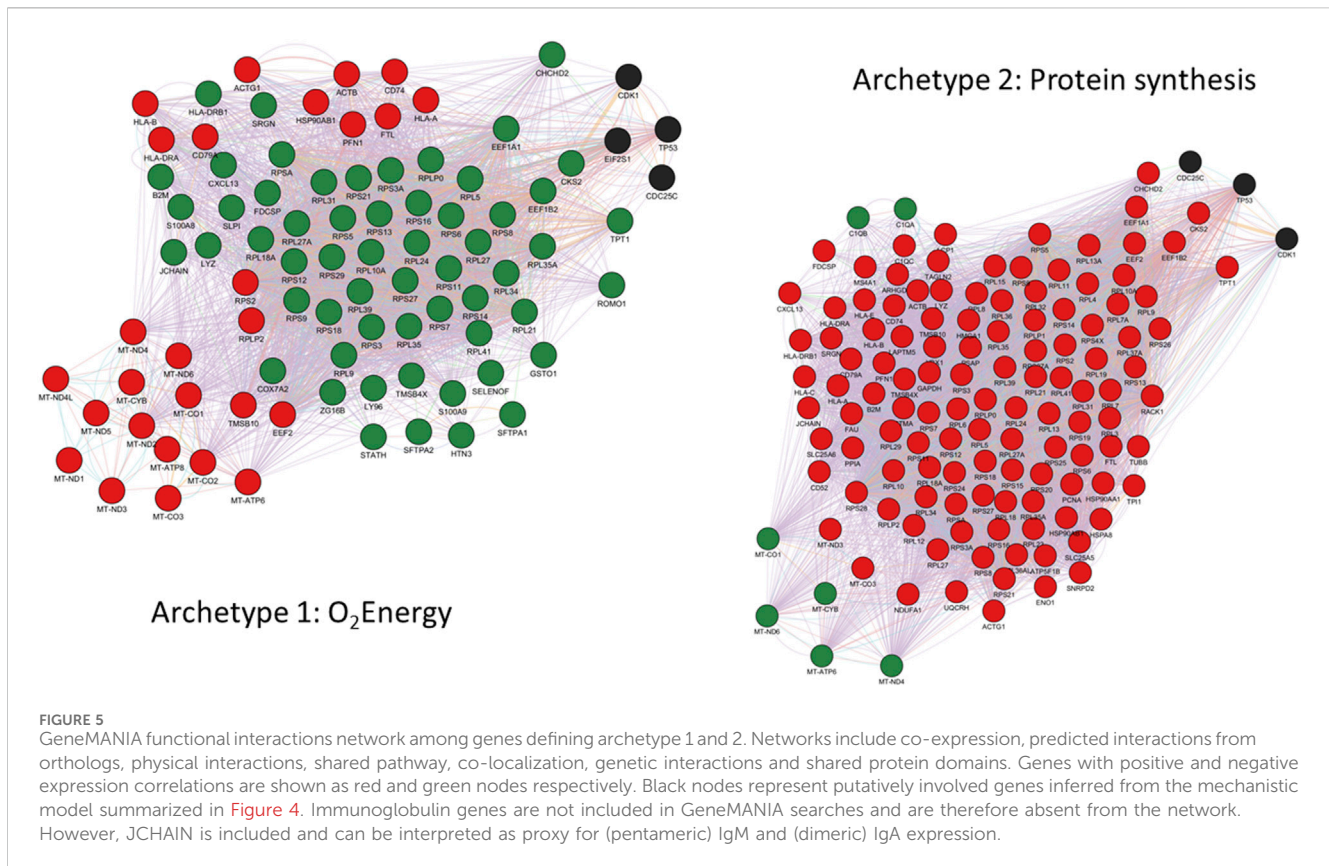
Overexpression of TPT1 has been associated with reduced CDK1 activity via ubiquitin-proteasome degradation of CDC25C, which is necessary for the dephosphorylation and activation of CDK1 (Chan et al., 2012a). By contrast, CKS2 is thought to promote CDK1 expression (You et al., 2015) and to be required for its function. Thus, the reciprocal correlation patterns seen between archetype 1 and 2 might reflect the need to compensate for the reduced activation of CDK1 by CDC25C from the increased expression of TPT1 by an increase in CKS2 expression and vice versa.

In addition, TCTP and CKS2 were both found to exhibit reciprocal repression with p53. In the case of TPT1, p53 is repressed via TCTP ubiquitin-mediated degradation of p53 while p53 directly represses TPT1 transcription (Amson et al., 2012; Acunzo et al., 2014). Likewise, CKS2 expression was found to be repressed by p53 (Rother et al., 2007), while the overexpression of CKS2 was associated with reduced p53 protein abundance in gastric cancer (Kang et al., 2009). However, the mechanism by which TPT1 and CKS2/CDK1 are repressed by p53 has been questioned and might be due to the indirect DREAM pathway rather than direct interaction with p53 (Fischer et al., 2014). Interestingly, CDC25C was also found to be repressed by p53 (Liu et al., 2020) and thus complex TPT1/CDC25C/CKS2/CDK1/p53 interactions might be

behind the TPT1-CKS2 opposite correlation pattern seen in these two archetypes. A schematic and synthetic representation of the hypothetical model outlined above for the metabolic trade-offs suggested by the analysis of archetypes 1 and 2's defining genes is provided on Figure 4. The full network of functional interactions among genes characterizing archetype 1 and 2 was created using the GeneMANIA app in Cytoscape v.3.10.1 (Montejo et al., 2010) and is shown on Figure 5.

Archetype 3 and 4: immune evasion trade-offs between immune-tolerant tumor micro-environment vs. immune escape via plasmablast-like differentiation

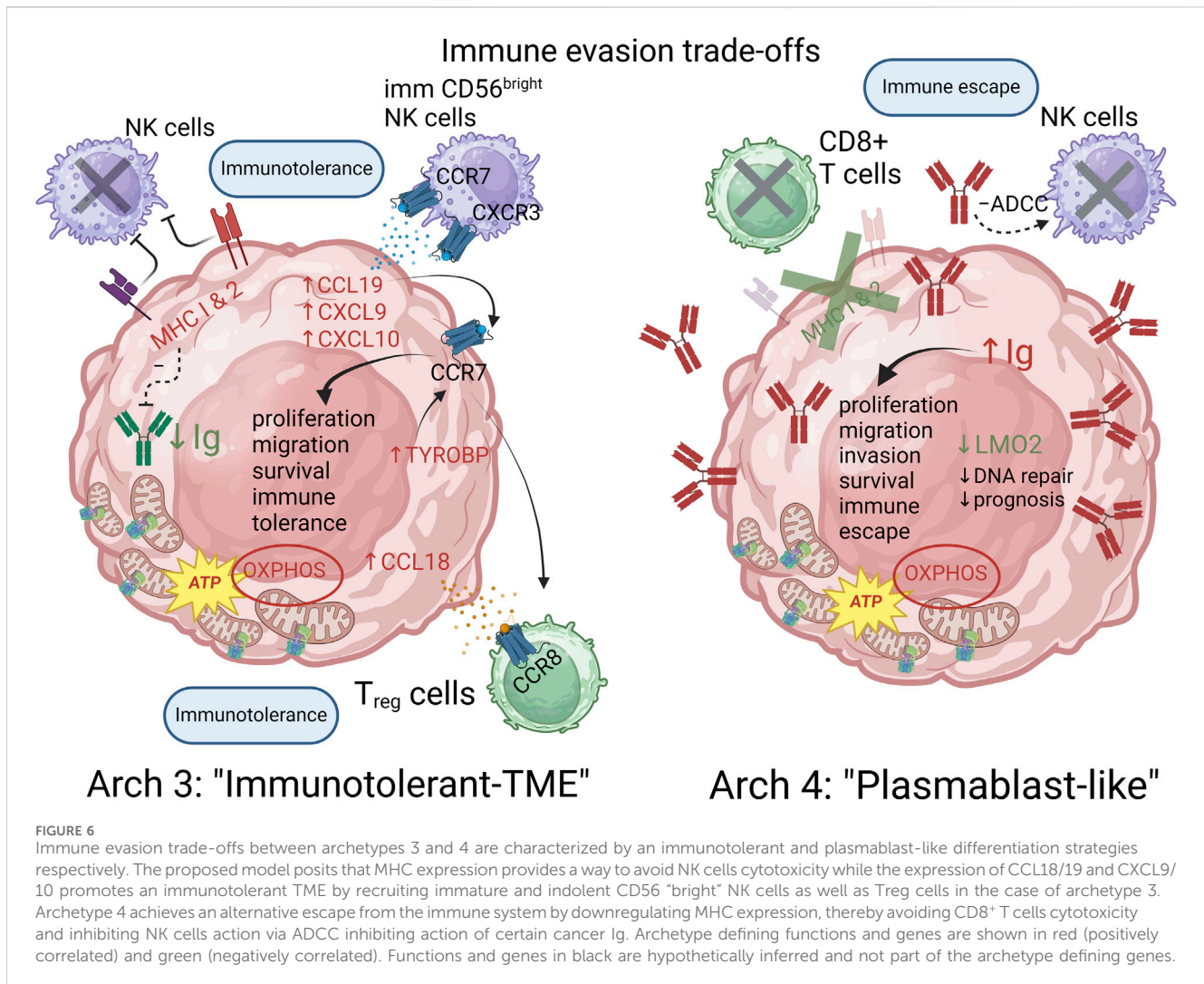
A breakdown of the genes characterizing archetype 3 and 4 was also conducted (Figures 3C, D). Many of these genes turned out to be known for their involvement in cancer in general and/or lymphoma in particular. For instance, LMO2 expression was negatively correlated with archetype 4, while IGHM expression was positively correlated with this archetype. This expression pattern has previously been associated with the activated B-cell (ABC) DLBCL COO subtype (Blenk et al., 2007). LMO2 expression reduces double-strand break DNA repair mechanisms and has been associated with a better prognosis in DLBCL patients



treated with poly(ADP-ribose) polymerase (PARP) inhibitors (Parvin et al., 2019). Interestingly, the expression of HLA-A,B,C,E, B2M and HLA-DRA/DRB1 was positively correlated with archetype 3, while expression of immunoglobulins (Ig) IGHM, IGHV4-34, IGHV5-51, IGKV3-20, IGLC2, IGLV1-47, IGLV3-1, IGLV3-19, IGLV3-21, JCHAIN were negatively correlated. In contrast, the expression of several Ig genes was positively correlated with archetype 4, while only the invariant HLA-DRA was positively correlated with the archetype. This pattern would be consistent with an immune system escape from MHC loss (de Charette and Houot, 2018) via a partial plasmablast cell differentiation pathway (Wilkinson et al., 2012) in archetype 4. Such a strategy would also be consistent with the many pro-tumorigenic effect of cancer derived Ig that have been identified, including proliferation, migration, invasion, survival, and immune evasion through inhibitory effect on antibody-dependent cell-cytotoxicity (ADCC) from NK cells (Cui et al., 2021). This picture is also consistent with archetype 4 resembling the ABC subtype of the COO classification (Wilkinson et al., 2012; van der Meeren et al., 2018; Takahara et al., 2023).

However, since Ig are common neo-antigens in B-cell malignancies and Ig-derived neoantigen presentation by MHC is a general phenomenon in lymphomas including DLBCL (Khodadoust et al., 2019), archetype 3 may limit the production of Ig in the context of retained MHC expression to avoid displaying neo-antigen Ig to the immune system (Han et al., 2022a). Moreover, archetype 3 was also positively correlated with the expression of CCL18, CCL19, CXCL9 and CXCL10. CCL18 is known to increase

the proliferation of B-cell lymphoma (Korbecki et al., 2020) and may contribute to immune evasion from its effect on immune surveillance mediated by macrophages and dendritic cells and by simultaneously favoring T cell-tolerance (Korbecki et al., 2020; Cardoso et al., 2021; Kidani et al., 2022). CCL19 directs B-cell migration after activation via antigen binding and is known to be upregulated in both GCB and ABC DLBCL subtypes. Recently, autocrine CCR7-CCL19 signaling was proposed to significantly contribute to lymphomagenesis under malignant conditions via a stronger activation of the survival pathways (Uhl et al., 2022). More generally, CCR7 signalling upon binding to its ligands (CCL19/21) is associated with many pro-tumorigenic effects in hematological malignancies, including migration, proliferation, survival and immune evasion (Cuesta-Mateos et al., 2021). Interestingly, increased expression of CCL19, CXCL9 and CXCL10 is associated with recruitment of immature CD56^{bright} NK cells with low perforin content in the tumor microenvironment (TME), which is thought to protect tumor cells from NK cells (Castriconi et al., 2018). Additionally, MHC expression is a well-known way for tumor cells to avoid immune detection and destruction by NK cells (Castriconi et al., 2018). This correlation pattern involving MHC/CXCL9-10/CCL18-19 might thus represent the signature of an alternative immune evasion strategy for archetype 3, distinct from the one displayed by archetype 4. Finally, both archetypes were positively correlated with the expression of CD74, a gene involved in B-cell differentiation, proliferation and survival (Zhao et al., 2019). An integrative schematic summary of the immune evasion trade-offs suggested by the above analysis of archetypes

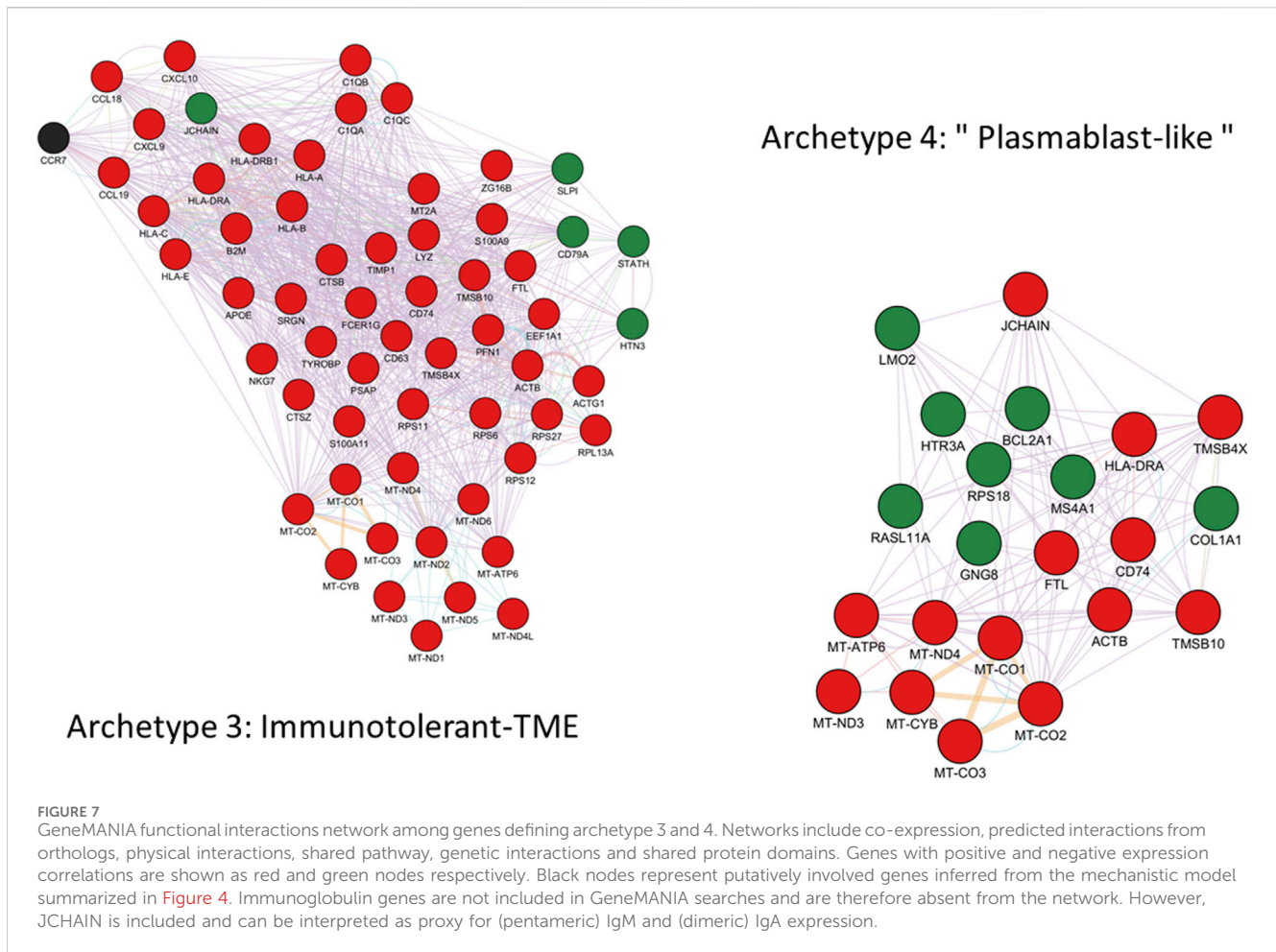


3 and 4 is provided on [Figure 6](#). The full network of functional interactions among genes characterizing archetype 3 and 4 is shown on [Figure 7](#).

Comparisons with differential expression studies

By looking at some of the genes defining each archetype, trade-offs and expression patterns noted by other studies using different approaches could be detected. Besides the cases already highlighted, the proportion of archetype-defining genes found in nine comparative DLBCL or non-Hodgkin lymphoma gene expression studies ([Monti et al., 2005](#); [Dybkaer et al., 2015](#); [Michaelsen et al., 2018](#); [Davies et al., 2019](#); [Tripodo et al., 2020](#); [Kotlov et al., 2021](#); [Steen et al., 2021](#); [de Groot et al., 2022](#); [Rapier-Sharman et al., 2022](#)) was assessed. Overall, 23% of the archetype defining genes identified by ParTI were part of the significantly differently expressed genes reported by these studies. Of note, the recently comparative transcriptomic study of [Rapier-Sharman et al.](#)

([Rapier-Sharman et al., 2022](#)) compared RNASeq gene expression data from 322 samples, including 134 B-cell lymphoma samples (of which 123 were LBCL/DLBCL) and 188 healthy B-cell controls. Among the 20 most differently expressed genes between B-lymphoma and normal control samples, 7 (35%) were found among the archetype defining genes identified here by the ParTI algorithm (CXCL9, CXCL13, C1QA, C1QB, C1QC, CCL18, CCL19). An additional three (15%) archetype defining genes identified by ParTI were among the 20 genes showing the most significant differences in the presence of splice variants (APOE, COL1A1 and RPL5). Although the tasks and trade-offs identified by ParTI are not necessarily expected to match differently expressed genes between malignant and normal cells, the fact that a significant proportion of genes showing differential expression between normal and lymphoma cells in differential expression studies are also found among archetype-defining genes identified by the ParTI approach provides additional and convincing evidence that the identified archetypes are related to the malignant process and phenotype. As such, it increases confidence that the Pareto



optimality theoretical framework is uncovering biologically meaningful and interpretable information at the scale of systems organization, with potential therapeutic relevance.

Discussion

Uncovering cancer cell vulnerabilities in the form of trade-offs represents a promising avenue to avoid the emergence of treatment resistance. The presence of trade-offs implies that certain tasks cannot be completely avoided by the cells, and yet cannot be simultaneously optimized to the maximum level allowable in principle by the genetic potential. The genes underlying these trade-offs are thus attractive therapeutic targets that could make resistance more difficult to acquire for the malignant cells. The Pareto task inference approach predicts that whenever such trade-offs exist, they should produce detectable geometrical structures in the data. We further predicted that besides geometrical patterns in trait space, if the genes located at the vertices of the polytope identified by the algorithm do indeed represent phenotypic optima (archetypes), as predicted by the theory, rather than artefacts unrelated to phenotypic optimization, these genes

should be significantly enriched in biological functions and characterized by patterns of different combinations of a certain proportion of shared tasks/genes.

The data analyzed here do indeed confirm those three predictions at a high level of statistical significance. The t-ratio test empirically calculates the probability of observing a polytope providing as good or better fit to the data as compared to the best possible fit defined as the convex hull. To this end, the data are randomized and the best fitting polytope and its ratio to the convex hull for each replicate data set are re-estimated. The value of the observed ratio is then compared to the distribution of ratios from the simulated randomized data to derive its probability. This test strongly supports the presence of a polytope defined by 4 vertices (tetrahedron) in this transcriptomic dataset.

According to the theory, the vertices of this polytope should represent optimal specialist phenotypes in trait space. Thus, these archetypes are predicted to be significantly enriched in certain particular functions, some of which being either different or performed by different genes, and others shared among different archetypes. FDR-adjusted Fisher's exact-tests on archetype defining gene lists clearly show that these genes are non-random genomic sub-samples, being statistically highly significantly enriched in particular functions. Some of the

most significantly enriched functions were different among archetypes, although the third and fourth archetypes showed substantial overlap in broadly defined functions, but little overlap in the genes underlying these functions within each archetype (the overall percentage of positively and negatively correlated genes unique to either archetype in pairwise comparison was 72.5%). The fact that different archetypes are statistically significantly enriched in different functions and/or gene combinations rules out the possibility that the observed functional enrichment is simply reflecting the general “B-lymphocyte” phenotype and is rather consistent with each archetype representing distinct optimization solutions. Given the typical complete effacement of lymph node architecture and extensive infiltration by malignant B-lymphocytes in DLBCL, these specialized functional sub-categories are likely to reflect, at least partly, various B-cell malignant phenotypic strategies, although some contribution from other cell types from the TME such as dendritic cells, macrophages or T-lymphocytes cannot be completely excluded. Nevertheless, even in the unlikely possibility that the archetypes represent cell-type specializations rather than tumor-cell specializations, Pareto theory’s predictions (polyhedron structure of the data, functional enrichment at the vertices, shared/modular structure of the functions/genes characterizing the phenotypes) would remain confirmed, albeit at a different level of cellular differentiation. Indeed, this theory scales to different levels of biological organization and has been applied to test for phenotypic optimization in various kinds of biological entities from cells to different animal species (e.g., Tandler et al., 2015). Thus, this possibility would not change the general conclusion that the theory is valid and that the archetypes identified represent trade-offs in the functions defining those archetypes.

However, besides these different functional characteristics, archetypes also displayed significant proportions of shared genes. As predicted if archetypes result from optimization in the face of trade-offs, these shared elements were distributed in different combinations and proportions among different archetypes. This pattern is consistent with archetypes having to reconcile various functional constraints by shuffling certain genetic toolkits, turning on and off the expression of genes in a way that preserves certain core functions and functional combinations, at the expense of other less essential and potentially dispensable elements. Whether the archetypes identified by the ParTI approach represent irreversible commitment to certain phenotypic pathways or phenotype through which cells can cycle sequentially remains an open question. The former possibility could thus represent stages in malignant progression, while the latter would include the possibility that different archetype might represent temporary adaptations to transient intra or extra-cellular environmental circumstances. Our analysis also confirmed that the archetypes identified by the Pareto approach are biologically interpretable and can be used to generate hypotheses about possible mechanisms underlying the identified correlation patterns. Thus, taken together, these results broadly confirm the

predictions of the Pareto optimality theory as applied to these transcriptomic data.

If trade-offs can be identified by the Pareto task inference approach, this could open the possibility to exploit these results to develop therapeutic strategies tailored to minimize the risk of resistance. Thus, the metabolic and immune evasion trade-offs suggested by the data may represent therapeutic opportunities that deserve further study. The first trade-off supports recent findings suggesting an important role for mitochondria, OXPHOS and ROS in tumorigenesis (Ghosh et al., 2020; Liu and Shi, 2020; Vasan et al., 2020), including in relation with the stress responses induced by increased ROS production and their impact on macromolecule synthesis (Jin et al., 2022). The second trade-off is in line with recent suggestions that in follicular lymphomas, loss of MHCII may be selectively acquired in cells that have accumulated immunogenic mutations in their idiotype sequences in order to avoid displaying Ig neoantigens to T-cells (Han et al., 2022b). In that study, MHCII expressing lymphoma cells were associated with a TME rich in a CD4⁺ T-cell population with a high cytotoxicity expression profile signature (CD4_{CTL}), while the reverse was observed for cells expressing low levels of MHCII (Han et al., 2022b). The pattern revealed by ParTI is also consistent with earlier findings that MHC loss can occur through a partial plasmablastic phenotypic differentiation which could be associated with high levels of Ig production (Wilkinson et al., 2012). As mentioned previously, the retention of MHC associated with the expression of CXCL9/10 and CCL18/19 in archetype 3 might represent an alternative immune evasion strategy to avoid NK-cells detection (from the expression of MHC and the action of CXCL9/10 and CCL19) and promote T-cell tolerance from CCL18 (Korbecki et al., 2020; Cardoso et al., 2021; Seliger and Koehl, 2022). Uncovering the best way to exploit such trade-offs in energy production, protein synthesis and immune evasion, will require additional detailed studies focused at testing the effect of disrupting the function of specific archetype defining genes on both side of the trade-offs simultaneously.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: Genome Data Commons Data Portal at GDC (cancer.gov).

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants’ legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

JB: Conceptualization, Formal Analysis, Investigation, Methodology, Visualization, Writing—original draft. JJ: Formal Analysis, Investigation, Methodology, Visualization, Writing—review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We thank Uri Alon and his research group for useful advices on running the PartiCode package.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Acunzo, J., Baylot, V., So, A., and Rocchi, P. (2014). TCTP as therapeutic target in cancers. *Cancer Treat. Rev.* 40, 760–769. doi:10.1016/j.ctrv.2014.02.007
- Adler, M., Tendler, A., Hausser, J., Korem, Y., Szekely, P., and Bossel, N., (2022). Controls for phylogeny and robust analysis in Pareto task inference. *Mol. Biol. Evol.* 39, msab297. doi:10.1093/molbev/msab297
- Aleksander, S. A., Balhoff, J., Carbon, S., Cherry, J. M., Drabkin, H. J., and Ebert, D., (2023). The gene ontology knowledgebase in 2023, 224. *Genetics*. doi:10.1093/genetics/iyad031
- Amini, M. A., Talebi, S. S., and Karimi, J. (2019). Reactive oxygen species modulator 1 (ROMO1), a new potential target for cancer diagnosis and treatment. *Chonnam Med. J.* 55, 136–143. doi:10.4068/cmj.2019.55.3.136
- Amson, R., Pece, S., Lespagnol, A., Vyas, R., Mazzarol, G., and Tosoni, D., (2012). Reciprocal repression between P53 and TCTP. *Nat. Med.* 18, 91–99. doi:10.1038/nm.2546
- Arnold, M., Rutherford, M. J., Bardot, A., Ferlay, J., Andersson, T. M., and Myklebust, T., (2019). Progress in cancer survival, mortality, and incidence in seven high-income countries 1995–2014 (ICBP SURVMARK-2): a population-based study. *Lancet Oncol.* 20, 1493–1505. doi:10.1016/S1470-2045(19)30456-5
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., and Cherry, J. M., (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. doi:10.1038/75556
- Bilen, M., Benhammouda, S., Slack, R. S., and Germain, M. (2022). The integrated stress response as a key pathway downstream of mitochondrial dysfunction. *Curr. Opin. Physiology* 27, 100555. doi:10.1016/j.cophys.2022.100555
- Blenk, S., Engelmann, J., Weniger, M., Schultz, J., Dittrich, M., and Rosenwald, A., (2007). Germinal center B cell-like (GCB) and activated B cell-like (ABC) type of diffuse large B cell lymphoma (DLBCL): analysis of molecular predictors, signatures, cell cycle state and patient survival. *Cancer Inf.* 3, 420. doi:10.1177/117693510700300004
- Bommer, U. A. (2017). The translational controlled tumour protein TCTP: biological functions and regulation. *Results Probl. Cell Differ.* 64, 69–126. doi:10.1007/978-3-319-67591-6_4
- Bommer, U. A., and Telerman, A. (2020). Dysregulation of TCTP in biological processes and diseases. *Cells* 9. doi:10.3390/cells9071632
- Bukowski, K., Kciuk, M., and Kontek, R. (2020). Mechanisms of multidrug resistance in cancer chemotherapy. *Int. J. Mol. Sci.* 21, 3233. doi:10.3390/ijms21093233
- Camargo, A. (2022). PCATest: testing the statistical significance of Principal Component Analysis in R. *PeerJ* 10, e12967. doi:10.7717/peerj.12967
- Cardoso, A. P., Pinto, M. L., Castro, F., Costa, A. M., Marques-Magalhães, A., and Canha-Borges, A., (2021). The immunosuppressive and pro-tumor functions of

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fsysb.2024.1346076/full#supplementary-material>

SUPPLEMENTARY FIGURE S1

Typical distribution of gene expression coefficients at archetype locations. Most genes have coefficients very close to zero (slope close to zero), with a small subset of genes having strongly negative or positive coefficients (slope close to one).

SUPPLEMENTARY FIGURE S2

Archetype defining genes selection method. The intersection of horizontal (slope close to zero) and vertical (slope close to one) lines was bisected at 45° and the mid-point of this 45° bisecting line was used as the cutoff coefficient value for archetype defining genes.

CCL18 at the tumor microenvironment. *Cytokine Growth Factor Rev.* 60, 107–119. doi:10.1016/j.cytogfr.2021.03.005

Castriconi, R., Carrega, P., Dondero, A., Bellora, F., Casu, B., and Regis, S., (2018). Molecular mechanisms directing migration and retention of natural killer cells in human tissues. *Front. Immunol.* 9, 2324. doi:10.3389/fimmu.2018.02324

Chan, T. H., Chen, L., and Guan, X. Y. (2012a). Role of translationally controlled tumor protein in cancer progression. *Biochem. Res. Int.* 2012, 369384. doi:10.1155/2012/369384

Chan, T. H., Liou, J. Y., Ambikapathi, A., Ma, W. K., and Chi, C. Y. (2012). "Fast algorithms for robust hyperspectral endmember extraction based on worst-case simplex volume maximization", in: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 1237–1240. 25–30 March 2012, Kyoto, Japan,

Cuesta-Mateos, C., Terrón, F., and Herling, M. (2021). CCR7 in blood cancers - review of its pathophysiological roles and the potential as a therapeutic target. *Front. Oncol.* 11, 736758. doi:10.3389/fonc.2021.736758

Cui, M., Huang, J., Zhang, S., Liu, Q., Liao, Q., and Qiu, X. (2021). Immunoglobulin expression in cancer cells and its critical roles in tumorigenesis. *Front. Immunol.* 12, 613530. doi:10.3389/fimmu.2021.613530

Davies, A., Cummin, T. E., Barrans, S., Maishman, T., Mamot, C., and Novak, U., (2019). Gene-expression profiling of bortezomib added to standard chemoimmunotherapy for diffuse large B-cell lymphoma (REMoDL-B): an open-label, randomised, phase 3 trial. *Lancet Oncol.* 20, 649–662. doi:10.1016/S1470-2045(18)30935-5

De Charette, M., and Houot, R. (2018). Hide or defend, the two strategies of lymphoma immune evasion: potential implications for immunotherapy. *Haematologica* 103, 1256–1268. doi:10.3324/haematol.2017.184192

De Groot, F. A., De Groen, R. a.L., Van Den Berg, A., Jansen, P. M., Lam, K. H., Mutsaers, P., et al. (2022). Biological and clinical implications of gene-expression profiling in diffuse large B-cell lymphoma: a proposal for a targeted BLYM-777 consortium panel as part of a multilayered analytical approach. *Cancers (Basel)* 14, 1857. doi:10.3390/cancers14081857

Dybkaer, K., Bogsted, M., Falgreen, S., Bodker, J. S., Kjeldsen, M. K., Schmitz, A., et al. (2015). Diffuse large B-cell lymphoma classification system that associates normal B-cell subset phenotypes with prognosis. *J. Clin. Oncol.* 33, 1379–1388. doi:10.1200/JCO.2014.57.7080

Fischer, M., Steiner, L., and Engeland, K. (2014). The transcription factor p53: not a repressor, solely an activator. *Cell Cycle* 13, 3037–3058. doi:10.4161/15384101.2014.949083

- Ghosh, P., Vidal, C., Dey, S., and Zhang, L. (2020). Mitochondria targeting as an effective strategy for cancer therapy. *Int. J. Mol. Sci.* 21, 3363. doi:10.3390/ijms21093363
- Grey, W., Atkinson, S., Rix, B., Casado, P., Ariza-Mcnaughton, L., and Hawley, C., (2023). The Cks1/Cks2 proteostasis Axis is crucial to maintain hematopoietic stem cell function. *Hemisphere* 7, e853. doi:10.1097/HS9.0000000000000853
- Grey, W., Ivey, A., Milne, T. A., Haferlach, T., Grimwade, D., and Uhlmann, F., (2018). The Cks1/Cks2 axis fine-tunes Mll1 expression and is crucial for MLL-rearranged leukaemia cell viability. *Biochim. Biophys. Acta Mol. Cell Res.* 1865, 105–116. doi:10.1016/j.bbamcr.2017.09.009
- Gundamaraju, R., Lu, W., and Manikam, R. (2020). CHCHD2: the power house's potential prognostic factor for cancer? *Front. Cell Dev. Biol.* 8, 620816. doi:10.3389/fcell.2020.620816
- Han, G., Deng, Q., Marques-Piubelli, M. L., Dai, E., Dang, M., and Ma, M. C. J., (2022a). Follicular lymphoma microenvironment characteristics associated with tumor cell mutations and MHC class II expression. *Blood Cancer Discov.* 3, 428–443. doi:10.1158/2643-3230.BCD-21-0075
- Han, G., Deng, Q., Marques-Piubelli, M. L., Dai, E., Dang, M., and Ma, M. C. J., (2022b). Follicular lymphoma microenvironment characteristics associated with tumor cell mutations and MHC class II expression. *Blood Cancer Discov.* 3, 428–443. doi:10.1158/2643-3230.BCD-21-0075
- Hart, Y., Sheftel, H., Hausser, J., Szekely, P., Ben-Moshe, N. B., and Korem, Y., (2015). Inferring biological tasks using Pareto analysis of high-dimensional data. *Nat. Methods* 12, 233–235. doi:10.1038/nmeth.3254
- Hausser, J., and Alon, U. (2020). Tumour heterogeneity and the evolutionary trade-offs of cancer. *Nat. Rev. Cancer* 20, 247–257. doi:10.1038/s41568-020-0241-6
- Hausser, J., Szekely, P., Bar, N., Zimmer, A., Sheftel, H., and Caldas, C., (2019). Tumor diversity and the trade-off between universal cancer tasks. *Nat. Commun.* 10, 5423. doi:10.1038/s41467-019-13195-1
- Jiang, T., Wang, Y., Wang, X., and Xu, J. (2022). CHCHD2 and CHCHD10: future therapeutic targets in cognitive disorder and motor neuron disorder. *Front. Neurosci.* 16, 988265. doi:10.3389/fnins.2022.988265
- Jin, P., Jiang, J., Zhou, L., Huang, Z., Nice, E. C., and Huang, C., (2022). Mitochondrial adaptation in cancer drug resistance: prevalence, mechanisms, and management. *J. Hematol. Oncol.* 15, 97. doi:10.1186/s13045-022-01313-4
- Jonsson, M., Fjeldbo, C. S., Holm, R., Stokke, T., Kristensen, G. B., and Lyng, H. (2019). Mitochondrial function of Cks2 oncoprotein links oxidative phosphorylation with cell division in chemoradioresistant cervical cancer. *Neoplasia* 21, 353–362. doi:10.1016/j.neo.2019.01.002
- Kang, M. A., Kim, J.-T., Kim, J. H., Kim, S.-Y., Kim, Y. H., and Yeom, Y. I., (2009). Upregulation of the cyclin kinase subunit Cks2 increases cell proliferation rate in gastric cancer. *J. Cancer Res. Clin. Oncol.* 135, 761–769. doi:10.1007/s00432-008-0510-3
- Khodadoust, M. S., Olsson, N., Chen, B., Sworder, B., Shree, T., and Liu, C. L., (2019). B-cell lymphomas present immunoglobulin neoantigens. *Blood* 133, 878–881. doi:10.1182/blood-2018-06-845156
- Kidani, Y., Nogami, W., Yasumizu, Y., Kawashima, A., Tanaka, A., Sonoda, Y., et al. (2022). CCR8-targeted specific depletion of clonally expanded Treg cells in tumor tissues evokes potent tumor immunity with long-lasting memory. *Proc. Natl. Acad. Sci. U. S. A.* 119. doi:10.1073/pnas.2114282119
- Korbecki, J., Olbromski, M., and Dziegiel, P. (2020). CCL18 in the progression of cancer. *Int. J. Mol. Sci.* 21, 7955. doi:10.3390/ijms21217955
- Kotlov, N., Bagaev, A., Revuelta, M. V., Phillip, J. M., Cacciapuoti, M. T., Antysheva, Z., et al. (2021). Clinical and biological subtypes of B-cell lymphoma revealed by microenvironmental signatures. *Cancer Discov.* 11, 1468–1489. doi:10.1158/2159-8290.CD-20-0839
- Krzywinski, M. I., Schein, J. E., Birol, I., Connors, J., Gascoyne, R., and Horsman, D., (2009). *Circos: an information aesthetic for comparative genomics. Genome Research*. 19. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Lee, H.-J., Song, K.-H., Oh, S. J., Kim, S., Cho, E., and Kim, J., (2022). Targeting TCTP sensitizes tumor to T cell-mediated therapy by reversing immune-refractory phenotypes. *Nat. Commun.* 13, 2127. doi:10.1038/s41467-022-29611-y
- Liu, K., Zheng, M., Lu, R., Du, J., Zhao, Q., and Li, Z., (2020). The role of CDC25C in cell cycle regulation and clinical cancer therapy: a systematic review. *Cancer Cell Int.* 20, 213. doi:10.1186/s12935-020-01304-w
- Liu, Y., and Shi, Y. (2020). Mitochondria as a target in cancer treatment. *MedComm* 1, 129–139. doi:10.1002/mco2.16
- López De Silanes, I., Olmo, N., Turnay, J., González De Buitrago, G., Pérez-Ramos, P., and Guzmán-Aránguez, A., (2004). Acquisition of resistance to butyrate enhances survival after stress and induces malignancy of human colon carcinoma cells. *Cancer Res.* 64, 4593–4600. doi:10.1158/0008-5472.CAN-04-0711
- Lucibello, M., Gambacurta, A., Zonfrillo, M., Pierimarchi, P., Serafino, A., and Rasi, G., (2011). TCTP is a critical survival factor that protects cancer cells from oxidative stress-induced cell-death. *Exp. Cell Res.* 317, 2479–2489. doi:10.1016/j.yexcr.2011.07.012
- Michaelsen, T. Y., Richter, J., Brøndum, R. F., Klapper, W., Johnsen, H. E., and Albertsen, M., (2018). A B-cell-associated gene signature classification of diffuse large B-cell lymphoma by NanoString technology. *Blood Adv.* 2, 1542–1546. doi:10.1182/bloodadvances.2018017988
- Mikami, T., and Iwasaki, W. (2021). The flipping t-ratio test: phylogenetically informed assessment of the Pareto theory for phenotypic evolution. *Methods Ecol. Evol.* 12, 696–706. doi:10.1111/2041-210x.13553
- Monti, N., Verna, R., Piombaro, A., Querqui, A., Bizzarri, M., and Fedeli, V. (2022). Paradoxical Behav. Oncog. Undermines Somat. Mutat. Theor. 12. *Biomolecules*
- Monti, S., Savage, K. J., Kutok, J. L., Feuerhake, F., Kurtin, P., and Mihm, M., (2005). Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. *Blood* 105, 1851–1861. doi:10.1182/blood-2004-07-2947
- Montejo, J., Zuberi, K., Rodriguez, H., Kazi, F., Wright, G., and Donaldson, S. L., (2010). GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* 26, 2927–2928. doi:10.1093/bioinformatics/btq562
- Park, J. S., Kang, D. H., and Bae, S. H. (2015). p62 prevents carbonyl cyanide m-chlorophenyl hydrazone (CCCP)-induced apoptotic cell death by activating Nrf2. *Biochem. Biophys. Res. Commun.* 464, 1139–1144. doi:10.1016/j.bbrc.2015.07.093
- Parvin, S., Ramirez-Labrada, A., Aumann, S., Lu, X., Weich, N., and Santiago, G., (2019). LMO2 confers synthetic lethality to PARP inhibition in DLBCL. *Cancer Cell* 36, 237–249. doi:10.1016/j.ccell.2019.07.007
- Paucuk, R. D., Baltimore, D., and Li, G. (2019). The cellular immunotherapy revolution: arming the immune system for precision therapy. *Trends Immunol.* 40, 292–309. doi:10.1016/j.it.2019.02.002
- Rapier-Sharman, N., Clancy, J., and Pickett, B. E. (2022). Joint secondary transcriptomic analysis of non-hodgkin's B-cell lymphomas predicts reliance on pathways associated with the extracellular matrix and robust diagnostic biomarkers. *J. Bioinform. Syst. Biol.* 5, 119–135. doi:10.26502/jbsb.5107040
- Rother, K., Dengl, M., Lorenz, J., Tschöp, K., Kirschner, R., and Mössner, J., (2007). Gene expression of cyclin-dependent kinase subunit Cks2 is repressed by the tumor suppressor p53 but not by the related proteins p63 or p73. *FEBS Lett.* 581, 1166–1172. doi:10.1016/j.febslet.2007.02.028
- Ruan, Y., Hu, J., Che, Y., Liu, Y., Luo, Z., and Cheng, J., (2022). CHCHD2 and CHCHD10 regulate mitochondrial dynamics and integrated stress response. *Cell Death Dis.* 13, 156. doi:10.1038/s41419-022-04602-5
- Saleh, R., and Elkord, E. (2020). Acquired resistance to cancer immunotherapy: role of tumor-mediated immunosuppression. *Seminars Cancer Biol.* 65, 13–27. doi:10.1016/j.semcancer.2019.07.017
- Salehi, S., Kabeer, F., Ceglia, N., Andronescu, M., Williams, M. J., Campbell, K. R., et al. (2021). Clonal fitness inferred from time-series modelling of single-cell cancer genomes. *Nature* 595, 585–590. doi:10.1038/s41586-021-03648-3
- Schmitz, R., Wright, G. W., Huang, D. W., Johnson, C. A., Phelan, J. D., and Wang, J. Q., (2018). Genetics and pathogenesis of diffuse large B-cell lymphoma. *N. Engl. J. Med.* 378, 1396–1407. doi:10.1056/NEJMoa1801445
- Seliger, B., and Koehl, U. (2022). Underlying mechanisms of evasion from NK cells as rationale for improvement of NK cell-based immunotherapies. *Front. Immunol.* 13, 910595. doi:10.3389/fimmu.2022.910595
- Selvarajoo, K., and Giuliani, A. (2023). Systems biology and omics approaches for complex human diseases. *Biomolecules* 13, 1080. doi:10.3390/biom13071080
- Sonnenschein, C., and Soto, A. M. (2020). Over a century of cancer research: inconvenient truths and promising leads. *PLoS Biol.* 18. doi:10.1371/journal.pbio.3000670
- Steen, C. B., Luca, B. A., Esfahani, M. S., Azizi, A., Sworder, B. J., and Nabet, B. Y., (2021). The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. *Cancer Cell* 39, 1422–1437. doi:10.1016/j.ccell.2021.08.011
- Sun, M., and Zhang, J. (2020). Rampant false detection of adaptive phenotypic optimization by ParTI-based Pareto front inference. *Mol. Biol. Evol.* 38, 1653–1664. doi:10.1093/molbev/msaa330
- Takahara, T., Nakamura, S., Tsuzuki, T., and Satou, A. (2023). The immunology of DLBCL. *Cancers (Basel)* 15, 835. doi:10.3390/cancers15030835
- Tendler, A., Mayo, A., and Alon, U. (2015). Evolutionary tradeoffs, Pareto optimality and the morphology of ammonite shells. *BMC Syst. Biol.* 9, 12. doi:10.1186/s12918-015-0149-z
- Thomas, P. D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L. P., and Mi, H. (2022). PANTHER: making genome-scale phylogenetics accessible to all. *Protein Sci.* 31, 8–22. doi:10.1002/pro.4218
- Tripodo, C., Zanardi, F., Iannelli, F., Mazzara, S., Vegliante, M., Morello, G., et al. (2020). A spatially resolved dark-versus light-zone microenvironment signature subdivides germinal center-related aggressive B cell lymphomas. *iScience* 23, 101562. doi:10.1016/j.isci.2020.101562
- Uhl, B., Prochazka, K. T., Pansy, K., Wenzl, K., Strobl, J., and Baumgartner, C., (2022). Distinct chemokine receptor expression profiles in *de novo* DLBCL, transformed follicular lymphoma, richter's trans-formed DLBCL and germinal center B-cells. *Int. J. Mol. Sci.* 23, 7874. doi:10.3390/ijms23147874
- Van Der Meer, L. E., Visser, L., Diepstra, A., Nijland, M., Van Den Berg, A., and Kluin, P. M. (2018). Combined loss of HLA I and HLA II expression is more common in

the non-GCB type of diffuse large B cell lymphoma. *Histopathology* 72, 886–888. doi:10.1111/his.13445

Vasan, K., Werner, M., and Chandel, N. S. (2020). Mitochondrial metabolism as a target for cancer therapy. *Cell Metab.* 32, 341–352. doi:10.1016/j.cmet.2020.06.019

Vasan, N., Baselga, J., and Hyman, D. M. (2019). A view on drug resistance in cancer. *Nature* 575, 299–309. doi:10.1038/s41586-019-1730-1

Vieira, V. (2012). Permutation tests to estimate significances on principal components analysis. *Comput. Ecol. Softw.* 2, 103–123.

Wang, Z., Zhang, M., Wu, Y., Yu, Y., Zheng, Q., and Li, J. (2021). CKS2 overexpression correlates with prognosis and immune cell infiltration in lung adenocarcinoma: a comprehensive study based on bioinformatics and experiments. *J. Cancer* 12, 6964–6978. doi:10.7150/jca.63625

Wilkinson, S. T., Vanpatten, K. A., Fernandez, D. R., Brunhoeber, P., Garsha, K. E., and Glinzmann-Gibson, B. J. (2012). Partial plasma cell differentiation as a mechanism

of lost major histocompatibility complex class II expression in diffuse large B-cell lymphoma. *Blood* 119, 1459–1467. doi:10.1182/blood-2011-07-363820

You, H., Lin, H., and Zhang, Z. (2015). CKS2 in human cancers: clinical roles and current perspectives (Review). *Mol. Clin. Oncol.* 3, 459–463. doi:10.3892/mco.2015.501

Zhang, G., Wang, X., Li, C., Li, Q., An, Y. A., and Luo, X. (2021). Integrated stress response couples mitochondrial protein translation with oxidative stress control. *Circulation* 144, 1500–1515. doi:10.1161/CIRCULATIONAHA.120.053125

Zhao, S., Molina, A., Yu, A., Hanson, J., Cheung, H., and Li, X. (2019). High frequency of CD74 expression in lymphomas: implications for targeted therapy using a novel anti-CD74-drug conjugate. *J. Pathol. Clin. Res.* 5, 12–24. doi:10.1002/cjp2.114

Zhao, S., Ye, Z., and Stanton, R. (2020). Misuse of RPKM or TPM normalization when comparing across samples and sequencing protocols. *Rna* 26, 903–909. doi:10.1261/rna.074922.120