UNIVERSITY OF CALIFORNIA SAN DIEGO

Extrachromosomal DNA tumors are immune-cold, not due to reduced antigenicity or MHC loss, but due to an impaired immune response to clonal MHC-II neoantigens.

A thesis submitted in partial satisfaction of the
requirements for the degree Master of Science

In

Computer Science

By

Krithika Iyer

Committee in charge:

Professor Hannah Carter, Chair
Professor Vineet Bafna, Co-Chair
Professor Melissa Gymrek

2026

The thesis of Krithika Iyer is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego
2026

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my advisor, Dr. Hannah Carter, for her invaluable guidance, encouragement, and support throughout this work. Her insight and mentorship have greatly shaped both this research and my growth as a researcher.

I am also thankful to all the members of the Carter Lab for their support, discussions, and encouragement along the way. Being part of such a collaborative and inspiring group has been a truly meaningful experience.

I am deeply thankful to my parents, brother, aunt and uncle, for their constant love, support, and encouragement throughout this journey.

Finally, I want to thank my closest friends for their constant encouragement, and for making this journey more memorable.

To all of you, thank you for being part of this journey.

ABSTRACT OF THE THESIS

Extrachromosomal DNA tumors are immune-cold, not due to reduced antigenicity or MHC loss, but due to an impaired immune response to clonal MHC-II neoantigens

By

Krithika Iyer

Master of Science in Degree Computer Science

University of California San Diego, 2026

Professor Hannah Carter, Chair

Professor Vineet Bafna, Co-Chair

Extrachromosomal DNA (ecDNA) is associated with aggressive tumor biology, but its role in tumor immune escape remains incompletely understood. We combined ecDNA annotations, RNA-seq data, somatic mutations, HLA presentation metrics, tumor purity estimates, and immune deconvolution results from a TCGA pan-cancer cohort. Our goal was to determine if ecDNA-positive tumors are immune-cold and to understand the mechanism behind this phenotype. Once adjusted for cancer type and tumor purity, tumors with ecDNA showed decreased cytolytic activity, a weaker CD8 signal, and reduced inferred cytotoxic immune infiltration, and these effects became more severe with higher ecDNA loads. Conversely, ecDNA-positive tumors did not reveal significant decreases in PHBR-based antigenicity metrics

or strong binder fractions, which contradicts the notion of a broad loss of neoantigens. Although HLA-A and B2M showed selective decreases, the broader MHC-I transcriptional program was not coordinately suppressed. The most significant mechanistic finding was the sustained positive link between clonal MHC-II neoantigen load and overall cytolytic activity. However, this association was notably weakened in tumors containing ecDNA. These results indicate that ecDNA tumors are immune-cold due to a reduced ability to translate clonal MHC-II neoantigenicity into effective cytolytic immune activation, rather than a deficiency in antigenicity or widespread MHC loss.

Introduction

Extrachromosomal DNA (ecDNA) has emerged as an important structural feature of many cancer genomes. Unlike traditional chromosomal amplifications, ecDNA drives massive oncogene copy number gains, altered gene regulation, and tumor heterogeneity (Wu, Sihan, et al.). Pan-cancer analyses have shown that ecDNA-associated amplifications are linked to worse clinical outcomes across multiple tumor types (Kim et al.). The development of computational tools such as AmpliconArchitect has made it possible to identify circular amplicons at scale and study their biological consequences in large patient cohorts (Deshpande et al.).

Most work on ecDNA has focused on its role in oncogene amplification and tumor evolution (Turner et al.). These studies have established ecDNA not merely as a genomic abnormality but as a driver of unique evolutionary patterns and improved adaptability. However, a critical gap remains: we still do not fully understand how the presence of ecDNA shapes the tumor-immune interface.

A fundamental question in cancer immunology is how tumors escape immune surveillance despite carrying somatic alterations capable of generating neoantigens, successfully evading immune surveillance. The classical immunoediting framework suggests that immune pressure should eliminate these immunogenic clones. Under this model, neoantigen burden, antigen presentation, and immune infiltration are tightly linked features of tumor immunogenicity, and prior pan-cancer work showed that predicted neoantigen load can correlate with local cytolytic activity (Rooney et al.).

However, antigenicity alone does not generate an effective anti-tumor response. For a tumor to be recognized, peptides must be processed and presented via Major Histocompatibility Complex (MHC) pathways. MHC class I presentation is essential for direct CD8+ T-cell recognition, whereas MHC class II-associated neoantigens and CD4+ T-cell responses can support helper cell priming, cytokine production, and downstream CD8-dependent antitumor immunity (Marty et al. 2017, Alspach et al.). Prior work has shown that MHC-II genotype constrains tumor evolution and that neoantigen-specific CD4+ T cells can mediate potent anti-tumor effects, including in settings where tumors themselves are not strongly MHC-II positive (Marty et al. 2018).

Recent pan-cancer work specifically implicated ecDNA in tumor immune response. In TCGA-based analyses, ecDNA-containing tumors were supported to show lower immune infiltration, reduced cytotoxic activity, and decreased expression of both MHC class I and MHC class II antigen presentation genes, suggesting that ecDNA may promote immune evasion through impaired antigen presentation (Wu, Tao, et al.). This model is biologically plausible, but it leaves several unresolved questions. It is not yet clear whether ecDNA-associated immune suppression is primarily explained by reduced antigenicity, broad collapse of MHC class I machinery, altered immune cell recruitment, or a more selective defect in how neoantigen information is converted into cytotoxic immune response.

In this study, the term immune-cold is used to describe tumors with reduced productive anti-tumor immune engagement, reflected by lower cytolytic activity, lower CD8-association immune signal, and reduced abundance of cytotoxic immune populations. This usage is narrower than the broader literature, where immune cold states may include both immune desert tumors with minimal infiltration and immune excluded tumors in which immune cells are present but functionally prevented from providing an effective response. Because our analyses

are based on bulk molecular data, we use the term to denote reduced anti-tumor immune engagement rather than a specific architecture.

These unresolved questions are especially important because pan-cancer analyses can be confounded by differences in tumor lineages in baseline immune infiltration, purity, stromal content, and transcriptional state. Distinguishing whether ecDNA marks a truly immune-cold state within a tumor type, rather than reflecting differences in tumor composition across cancer types, requires integrated analysis with adjustment for lineage and purity.

This study integrated ecDNA annotations, RNA expression levels, somatic mutation data, HLA presentation metrics, neoantigen summaries, and immune deconvolution outputs from a large TCGA cohort. We first asked whether ecDNA-positive tumors are reproducibly immune-cold after accounting for cancer type and tumor purity. We then tested whether this phenotype could be explained by reduced global antigenicity or a coordinated loss of the MHC class I transcriptional program. Finally, we examined whether ecDNA alters the coupling between neoantigen burden and immune response, with particular focus on clonal MHC-II neoantigens because of their potential relevance to helper cell priming and broader anti-tumor immunity.

Our hypothesis is that ecDNA-positive tumors are immune cold not because they lack neoantigens or undergo broad MHC-I loss, but because they are impaired in converting clonal MHC-II neoantigenicity into effective cytolytic immune activation. Under this model, ecDNA-associated immune escape would reflect a defect in antigen to response coupling within helper/priming and effector immune programs rather than a simple reduction in antigen generation. This framework naturally extends prior ecDNA immune escape work while testing a

more specific mechanistic explanation for why ecDNA bearing tumors may remain poorly inflamed despite preserved antigenic potential.

Dataset and Cohort

1.1 Overview

This study integrates genomic, transcriptomic, and immunogenomic data from The Cancer Genome Atlas (TCGA) to investigate immune correlates of extrachromosomal DNA (ecDNA) in human tumors (Weinstein et al.). The study's analysis framework combined ecDNA annotations derived from AmpliconArchitect, RNA sequencing expression, HLA binding predictions, tumor purity estimates, and sample-level clinical and metadata annotations.

The final integrated cohort comprised 2471 primary tumor samples spanning multiple TCGA cancer types (Weinstein et al.). Among these, 1040 tumors were classified as ecDNA positive based on criteria described below. We performed individual downstream analyses on subsets determined by the availability of RNA expression data, tumor purity, neoantigens, or pathway data, as not all samples had complete data for every modality.

1.2 ecDNA Annotations

ecDNA status was determined using structural variant reconstruction generated by AmpliconArchitect (AA) (Deshpande et al.). It identifies focal, high-copy DNA amplifications and reconstructs their architecture to distinguish circular extrachromosomal amplicons from linear chromosomal amplifications.

We obtained ecDNA annotations from the Amplicon Repository, which provides AA reconstructions for TCGA tumors. Tumors were classified as ecDNA-positive if at least one circular amplicon meeting established structural criteria was identified, including evidence of circular structure, focal high-copy amplification, and rearrangement architecture consistent with

extrachromosomal topology. Tumors without such circular amplicons were classified as ecDNA-negative.

For each tumor, we derived the following ecDNA features:

- Binary ecDNA status - positive or negative
- ecDNA burden - number of circular amplicons detected per tumor
- Structural complexity metrics
- Amplified interval length

Within the TCGA-unfiltered cohort of 2471 tumors, 1040 were classified as ecDNA-positive and 1431 as ecDNA-negative.

1.3 RNA sequencing and gene expression data

Gene expression data were obtained from TCGA RNA sequencing datasets accessed through UCSC Xena and quantified using RSEM and normalized to transcripts per million (TPM) (Goldman et al.). Gene models were defined using GENCODE v43 annotations (Frankish et al.). Expression values were log-transformed prior to downstream analyses.

From gene-level expression data, we derived multiple immune-related signatures:

- Cytolytic activity score (CYT), which is the geometric mean of GZMA and PRF1 expression.
- CD8 expression levels
- Antigen Presentation Machinery (APM) score based on expression of HLA-A, HLA-B, HLA-C, B2M, TAP1, TAP2 and NLRC5
- STING pathway effector signatures
- Chronic STING pathway activation scores

- Pathway programs related to IFNγ signalling, MHC-II

These signatures were analyzed both continuously and after normalization within cancer types where appropriate.

1.4 Somatic Mutation Data

Somatic mutation data were obtained from harmonized TCGA mutation calls generated from whole-exome sequencing. Mutation annotations included genomic coordinates, gene assignments, coding consequence classification, and variant allele frequency (VAF) used in downstream immunogenomic analyses.

For neoantigen and presentation analyses, we focused on non-synonymous coding mutations, since these are the variants most likely to generate altered peptides with potential HLA presentation. Variants were further annotated with predicted peptide-HLA binding metrics and patient-level neo-antigen presentation summaries.

Tumor barcodes were restricted to primary tumor samples (TCGA sample type 01) to ensure consistency across genomic and transcriptomic analyses.

1.5 HLA Typing and Neo-antigen Metrics

Patient-specific HLA class I and class II alleles were inferred using established computation pipelines applied to TCGA sequencing data. Predicted peptide-HLA binding metrics were used to compute Patient Harmonic Best Rank (PHBR) scores for each mutation (Marty et al. 2017, Marty et al. 2018).

Two primary presentation metrics were analyzed:

- PHBR-I (HLA class I presentation potential)

- PHBR-II (HLA class II presentation potential)

Additional neo-antigen metrics included:

- Strong binder fractions
- clonal strong binder fractions
- Patient-level summaries of antigen presentation capacity

These metrics were used to test whether ecDNA-positive tumors differed in global antigenicity and to evaluate whether ecDNA altered the relationship between neoantigen burden and downstream immune response.

1.6 Tumor Purity, Immune Deconvolution, and Additional Covariates

Tumor purity estimates were incorporated where available to account for variation in non-tumor cellular mix across samples (Carter et al.). Purity was included as a covariate in multivariable models because highly pure tumors can show lower apparent immune cell-associated expression due to reduced immune and stromal content.

To provide orthogonal measures of immune infiltration beyond bulk transcriptional signatures, we also incorporated immune deconvolution outputs, including:

- CIBERSORT-derived immune cell fractions (Newman et al.)
- MCP-counter derived immune and stromal abundance estimates (Becht et al.)

1.7 Cohort Integration

Samples were included in the final cohort if they satisfied all of the following:

1. Available ecDNA structural annotations
2. Available RNA sequencing expression data
3. Available somatic mutation data with HLA binding annotation
4. Primary tumor designation

Standardizing all datasets with TCGA barcodes was crucial for maintaining consistent patient-level linkage between genomic, transcriptomic, and structural data. The resulting integrated dataset comprised 2471 tumors spanning multiple cancer types and served as the basis for all downstream statistical analyses.

To account for tumor heterogeneity and cancer variability, immune expression metrics were evaluated in both raw form and after cancer type normalization. These regression models incorporated covariates including cancer types, tumor mutation burden, and tumor purity (where available).

Python was used to implement reproducible computational pipelines for all statistical analyses.

Methodology

The core hypothesis of this study is that tumors characterized by ecDNA exhibit an immune-cold phenotype. To test this hypothesis, we tested whether this phenotype is explained by reduced antigenicity or broad loss of antigen-presentation machinery, or instead reflects impaired translation of clonal MHC-II neoantigen burden into a productive anti-tumor immune response.

Two primary biological questions motivated this analysis. First, do ecDNA-positive tumors exhibit reduced cytotoxic immune activity and reduced T-cell associated infiltration? Second, if ecDNA-positive tumors are immune cold, is this phenotype explained by reduced antigenicity, impaired antigen presentation machinery, or altered coupling between neoantigen burden and immune response?

To answer these questions, we analyzed RNA-seq derived immune signatures, immune deconvolution outputs, HLA presentation metrics, ecDNA burden measures, and multivariable regression models adjusted for cancer type and tumor purity

2.1 RNA Processing and Normalization

We obtained gene expression data from the TCGA RNA sequencing dataset, quantified as transcripts per million (TPM) and accessed through UCSC Xena (Goldman et al.). Expression values were log-transformed as:

$$E_{log} = log_2(TPM + 1)$$

Log-transformed values were used to compute all immune signatures. To account for gene scale differences, we standardized each gene across samples using Z-score normalization:

$$Z_{g,i} = \frac{E_{g,i} - \mu_g}{\sigma_g}$$

Where $E_{g,i}$ denotes the log-transformed expression of gene g in sample i, μ_g and $\sigma_g$represent the mean and standard deviation of gene g across the cohort. In secondary analysis, we normalized the immune signatures by cancer type to control for tumor-specific baseline infiltration differences.

2.2 Immune Signature construction

2.2.1 Cytolytic Activity Score (CYT)

We quantified cytolytic activity using the established Rooney et al. immune cytolytic activity signature. (Rooney et al.) Rather than computing the geometric mean on TPM values, we used the mean of z-scored log-transformed expression values:

$$CYT = \frac{Z_{GZMA,i} * Z_{PRF1,i}}{2}$$

where, $GZMA$ is Granzyme A and where $PRF1$ is Perforin.

This score reflects the cytotoxic T-cell effector function. Higher CYT scores correlate with an improved anti-tumor immune response.

2.2.2 CD8 T-cell Infiltration Score

The CD8 T-cell abundance was approximated using the expression of canonical CD8 markers:

$$CD8_i = \frac{Z_{CD8A,i} + Z_{CD8B,i}}{2}$$

The score captures CD8 T-cell infiltration without considering their effector activation state.

2.2.3 Antigen Presentation Machinery (APM) score

To evaluate antigen presentation capacity, we defined two related transcriptional summaries. First, a core APM score was calculated from major MHC class I surface presentation genes:

$$APM_i = \frac{Z_{HLA-A,i} + Z_{HLA-B,i} + Z_{HLA-C,i} + Z_{B2M,i}}{4}$$

Second, for a more comprehensive evaluation of MHC class I pathway expression, we also constructed an expanded MHC-I composite using HLA-A, HLA-B, HLA-C, B2M, TAP, TAP2, NLRC5.

The expanded composite was computed as the mean of z-scored expression across these genes. This allowed us to distinguish selective downregulation of individual genes from coordinated suppression of the broader MHC-I transcriptional program.

2.2.4 Exploratory ERAP1/ERAP2 Antigen-Processing Score

As an exploratory measure of antigen processing, we evaluated expression of ERAP1 and ERAP2, two endoplasmic reticulum aminopeptidases involved in peptide trimming before MHC class I loading. Because ERAP1/2 can both facilitate and destroy presentable peptides depending on substrate and context, their expression was interpreted as a marker of altered peptide trimming state rather than a direct measure of impaired antigen presentation. An ERAP score was computed as the mean of standardized ERAP1 and ERAP2 expression value.

$$ERAP_i = \frac{Z_{ERAP1,i} + Z_{ERAP2,i}}{2}$$

2.2.5 Additional Immune and Pathway scores

We additionally analyzed several immune context programs derived from predefined gene sets including:

- STING effector signature
- Chronic STING activation signature
- IFNγ program
- MHCII program
- Recruitment program
- TGF$\boldsymbol{\beta}$_ECM program

These pathway-level scores were used to assess whether ecDNA-associated immune suppression mapped more closely to antigen presentation, recruitment, or innate immune signaling contexts.

2.3 ecDNA annotation and burden quantification

Tumors were classified as ecDNA-positive if at least one circular amplicon was identified by AmpliconArchitect reconstruction.

In addition to binary ecDNA status, ecDNA burden was defined as the number of circular amplicons identified per tumor. This helped evaluate the dose-response relationships between structural amplification and immune phenotypes.

2.4 HLA Presentation and Neoantigen Metrics

For each patient, HLA class I and class II presentation summaries were derived from predicted peptide-HLA binding data. We analyzed both patient-level and mutation-level neoantigen metrics, including:

- mean PHBR-I
- mean PHBR-II
- strong MHC-I binder fraction
- strong MHC-II binder fraction
- strong clonal MHC-I binder fraction
- clonal strong MHC-II neoantigen burden

This metrics were used for two related purposes. First, they were used to test whether ecDNA-positive tumors differed in overall antigenicity or HLA presentation potential. Second, they were used to determine whether ecDNA altered the relationship between neoantigen burden and downstream immune response.

2.5 Immune Deconvolution Analyses

To validate expression-signature-based findings using orthogonal methods, we analyzed inferred immune cell composition from bulk transcriptomic data using both CIBERSORT (Newman et al.) and MCP-counter outputs (Becht et al.).

CIBERSORT-derived cell fractions were used to quantify relative abundance of immune cell populations including CD8 T cells, T follicular helper cells, NK-cell states, macrophage states, dendritic cells, and other leukocyte subsets. MCP-counter scores were used as an

independent estimate of immune and stromal population abundance, including cytotoxic lymphocytes, NK cells, T cells, and myeloid dendritic cells.

These analyses were used to determine whether the ecDNA-associated immune-cold phenotype was reproducible across multiple orthogonal frameworks rather than driven by a single RNA signature.

2.6 Statistical Analysis

2.6.1 Unadjusted Group comparison

Differences in immune signature scores between ecDNA-positive and ecDNA-negative were evaluated using Wilcoxon rank sum. This non-parametric approach was used due to the normal distribution of immune scores and the presence of outliers.

2.6.2 Dose Response Analysis

Associations between ecDNA burden and immune signatures were evaluated using Spearman correlation and linear regression models. Scatter plots with regression overlays were generated to visualize trends between ecDNA burden status and CYT score, and CD8 score.

2.6.3 Multivariable Regression Adjustment

To assess whether observed immune differences were independent of mutation burden and tumor purity, multivariable linear regression models were fit of the form:

$$ImmuneScore_i = \beta_0 + \beta_1(ecDNA_i) + \beta_2(TMB_i) + \beta_3(Purity_i) + \epsilon_i$$

where,

- $ImmuneScore_i$ represents CYT or CD8 score
- $ecDNA_i$ is the binary ecDNA status
- TMB_i represents tumor mutation burden

- $Purity_i$ represents tumor purity

This model tests whether ecDNA status independently predicts immune suppression.

2.6.4 Antigenicity-to-Immune-Response Coupling Analyses

To test whether ecDNA altered the relationship between neoantigen burden and immune response, we fit interaction models of the form:

$$ImmuneScore_i = \beta_0 + \beta_1 N_i + \beta_2 ecDNA_i + \beta_3 (N_i * ecDNA_i) + \beta_4 Purity_i + \sum_k \gamma_k CancerType_{ik} + \epsilon_i$$

where,

- N_i denotes neoantigen-related quantities such as mutation burden, strong MHC-I binder burden, or clonal strong MHC-II neoantigen burden
- β_1 is the baseline relationship between neoantigen burden and immune activity
- β_3 tests whether this relationship is weakened or altered in ecDNA-positive tumors

2.6.5 Robustness Analyses of Clonal MHC-II Coupling

Since clonal MHC-II neoantigen burden depends on threshold values, robustness analyses were performed across multiple PHBR-II thresholds and cancer cell fraction thresholds. To investigate which immune context might explain the ecDNA associated weakening of clonal MHC-II neoantigen-to-CYT coupling, pathway attenuation analyses were performed by sequentially adding pathway scores to the base interaction model. These included:

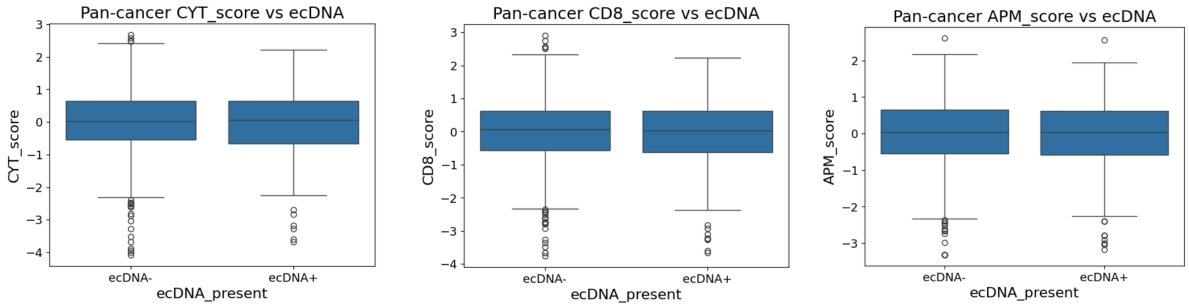
- IFNγ program
- MHCII program
- Recruitment program
- TGF$\boldsymbol{\beta}$_ECM program

Results

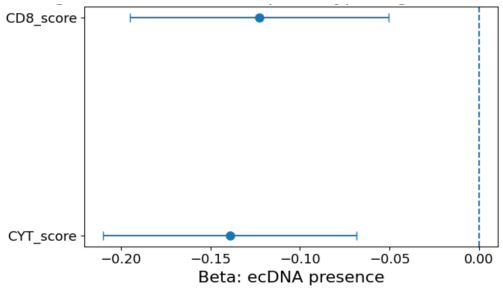
3.1 ecDNA-positive tumors exhibit an immune-cold phenotype within tumor type

We first assessed whether ecDNA-positive tumors differ from ecDNA-negative tumors in RNA-derived immune activity signatures. In an unadjusted pan-cancer comparison, median CYT, CD8, and APM scores were not significantly different between groups. As shown in Fig. 1(a), raw pan-cancer distributions of CYT, CD8, and APM substantially overlap between ecDNA-positive and ecDNA-negative tumors. This indicates that cross-cancer baseline differences obscure ecDNA-associated immune effects in unadjusted analyses and motivates cancer type-adjusted modeling. Thus, at the raw pan-cancer level, ecDNA-positive tumors were not globally more inflamed or more antigen-presenting. The distributions of key structural, immune, and neoantigen-related variables used in downstream analyses are shown in Supplementary Figure S1.

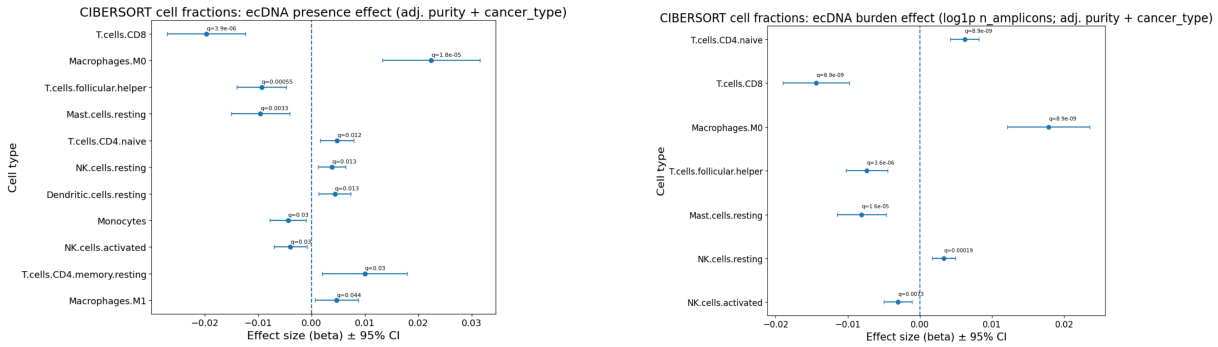
However, this changed after accounting for cancer type baseline differences. After adjustment for cancer type, ecDNA positivity was significantly associated with lower CYT and CD8 scores (CYT: beta = -0.216, $p = 1.11 \times 10^{-6}$; CD8: beta = -0.182, $p = 5.06 \times 10^{-5}$), whereas APM was not significantly reduced (beta = -0.064, p = 0.170). As shown in Fig 1(b), ecDNA-positive tumors are relatively immune-cold within the tumor rather than globally immune-cold across the entire pan cancer cohort. This indicates that the dominant ecDNA-associated signal is reduced cytotoxic immune activity rather than broad suppression of antigen-presentation gene transcription.



(a)



(b)



(c)

Fig 1. ecDNA-positive tumors are immune-cold after adjustment for tumor type and purity; (a) Raw pan-cancer immune signature distributions show limited separation by ecDNA status; (b) Cancer-type adjusted models reveal lower cytolytic activity and CD8 signal in ecDNA positive tumors; (c) Orthogonal immune deconvolution confirms reduced cytotoxic infiltration and altered immune composition in ecDNA-positive tumors.

We next tested whether the signal remained after controlling for tumor purity, a critical factor because higher purity tumors typically contain fewer non-tumor immune cells. In the purity adjusted cohort (n=1833), ecDNA presence remained significantly associated with lower CYT score (beta = -0.147, $p = 2.71 \times 10^{-4}$, q = 0.0011) and lower CD8 score (beta = -0.111, $p = 7.10 \times 10^{-3}$, q = 0.014). In contrast, neither APM score nor STING effector score showed significant suppression in the same purity adjusted models. These results indicate that ecDNA positive tumors are not different because of tissue composition or cancer type mixture; rather, after controlling for those factors, ecDNA presence and ecDNA burden are associated with reduced cytotoxic immune activity.

To validate the immune-cold phenotype using an orthogonal framework, we analyzed inferred immune cell fractions from deconvolution data in a larger subset of tumors (n= 2292 with CIBERSORT estimates), adjusting for cancer type and tumor purity. As shown in Figure 1(c), ecDNA-positive tumors had significantly lower CD8 T-cell fractions both as a binary presence effect and as a burden-dependent effect (presence: beta = -0.0197, $q = 4.0 \times 10^{-6}$; burden: beta = -0.0144, $q = 8.86 \times 10^{-9}$). Beyond CD8 T-cells, ecDNA-positive tumors also showed reduced T follicular helper signal (presence: beta = -0.00934, $q = 5.47 \times 10^{-4}$; burden: beta = -0.00736, $q = 3.55 \times 10^{-6}$) and reduced activated NK-cell fractions (presence: beta = -0.00394, q = 0.030; burden: beta = -0.00304, q = 0.0073), together with increased Macrophage M0 fractions (presence: beta = 0.0224, $q = 1.8 \times 10^{-5}$; burden: beta = 0.0178, $q = 8.86 \times 10^{-9}$).

These shifts are consistent with a less cytotoxic and less productive anti-tumor immune microenvironment and reinforce that the ecDNA associated phenotype is not limited to a single RNA derived score. These models showed that ecDNA-positive tumors had significantly lower CD8 T-cell infiltration both as a binary presence effect and as a burden dependent effect.

3.2 ecDNA associated Immune-cold phenotype is not explained by global antigenicity loss or coordinated MHC-I shutdown

We next tested whether ecDNA-positive tumors are less immunogenic simply because they contain either fewer predicted neoantigens or neoantigens with weaker predicted HLA presentation. To do this, we analyzed patient-level HLA binding and neoantigen summary metrics in the subset with PHBR and strong binder annotations (n=1686). ecDNA presence was not significantly associated with mean PHBR-I (beta=0.019, p=0.74), mean PHBR-II (beta = 0.030, p=0.92), strong MHC-I binder fraction (beta= -0.0031, p=0.64), strong MHC-II binder fraction (beta = 0.013, p=0.71), or a strong clonal MHC-I binder fraction (beta= 0.0077, p=0.39). As shown in Fig. 2(a), ecDNA status was not significantly associated with broad shifts in PHBR-I, PHBR-II, or strong binder fractions, arguing against a model of global antigenicity loss.

These results contradict against the idea that ecDNA-positive tumors are immune cold because they are globally less antigenic or less capable of generating predicted peptide HLA binders. Instead, ecDNA-positive tumors appear to preserve broad antigenicity metrics despite reduced immune activity.

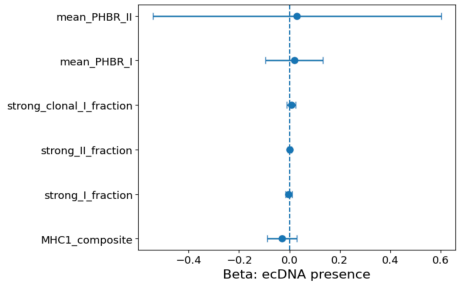
We then tested whether ecDNA disrupts the relationship between mutation-derived antigenicity and immune response. Across tumors, mutation burden and neoantigen burden remained positively associated with CYT and to CD8 signal. The log transformed mutation count was positively associated with CYT (beta = 0.162, p= 1.00×10^{-12}) and log transformed strong MHC-I binder count was also positively associated with CYT (beta = 0.164, p= 3.49×10^{-13}). ecDNA did not significantly weaken these relationships for total mutation burden or strong MHC-I binders. Thus, the general mutation to CYT and MHC-I neoantigen to CYT relationships

remain intact in ecDNA positive tumors, contradicting the classical MHC-I focused immunoediting model.

Because prior pan-cancer work reported decreased antigen presentation gene expression in ecDNA containing tumors, we evaluated MHC-I pathway expression in more detail. In cancer type adjusted gene level models, ecDNA positivity was associated with reduced expression of HLA-A (beta= -0.133, p=0.038) and B2M (beta = -0.143, p= 0.010), while HLA-B, HLA-C, TAP1, TAP2, and NLRC5 were not significantly associated with ecDNA status. Similarly, higher ecDNA burden was associated with lower HLA-A (beta = -0.026 per amplicon, p= 0.030) and lower B2M (beta = -0.0.035, p= 0.012), with weaker or absent effects for the remaining genes.

However, these selective gene level associations did not translate into strong suppression of the full MHC-I machinery. In the purity adjusted models, the composite MHC-I score was not significantly associated with ecDNA presence (beta = -0.0043, p= 0.917) or burden (beta = -0.0298, p= 0.257). Likewise, the broader APM score was not significantly reduced in the main purity adjusted models (presence beta = -0.0086, p= 0.846; burden beta = -0.0276, p= 0.331).

Thus, ecDNA positive tumors do show evidence of selective suppression of particular MHC-I related genes, especially HLA-A and B2M, but this doesn't extend to broad collapse of the MHC-I transcriptional program once tumor purity and cancer type are taken into account. This distinction suggests that the immune cold phenotype in ecDNA positive tumors is not primarily explained by global MHC-I loss, but instead may reflect a more selective pathway modulation or defects downstream of antigen generation.

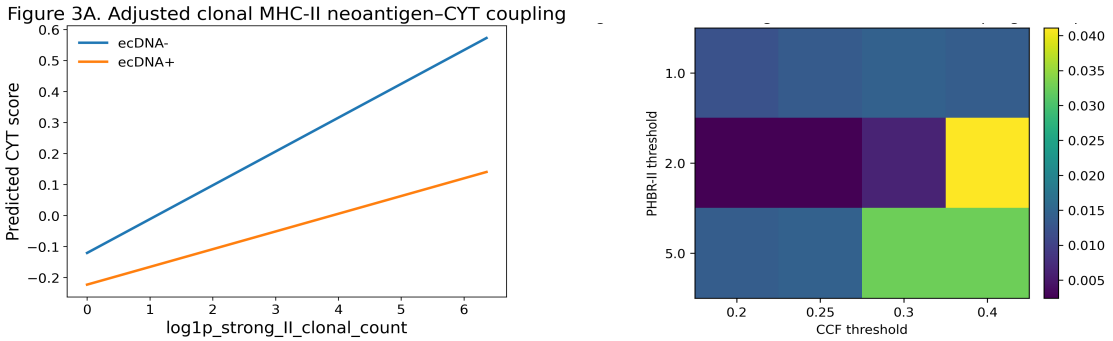


(a)

Fig 2. ecDNA associated immune suppression is not explained by global antigenicity loss or coordinated MHC-I shutdown; (a) Predicted antigenicity and global MHC-I related presentation metrics are largely preserved in ecDNA positive tumors.

3.3 Selective decoupling of clonal MHC-II neoantigenicity from Cytolytic Response is a key mechanistic feature of ecDNA associated immune suppression

Given that overall antigenicity is preserved, we investigated further whether ecDNA alters how neoantigen burden translates into effective immune response. The strongest signal emerged for clonal MHC-II neoantigens. In ecDNA-negative tumors, higher clonal MHC-II neoantigen burden was associated with substantially higher predicted CYT, whereas this slope was markedly attenuated in ecDNA-positive tumors (Fig 3(a)). This indicates that although clonal MHC-II neoantigens remain immunologically relevant across tumors, ecDNA-positive tumors derive less cytolytic immune activation from the neoantigen burden. We noticed that while a higher count of clonal MHC-II presentable neoantigens is generally associated with increased cytolytic activity (CYT) across tumors (beta = 0.105, $p = 8.69 \times 10^{-8}$), this critical immunological result is significantly attenuated in the presence of ecDNA positive tumors.

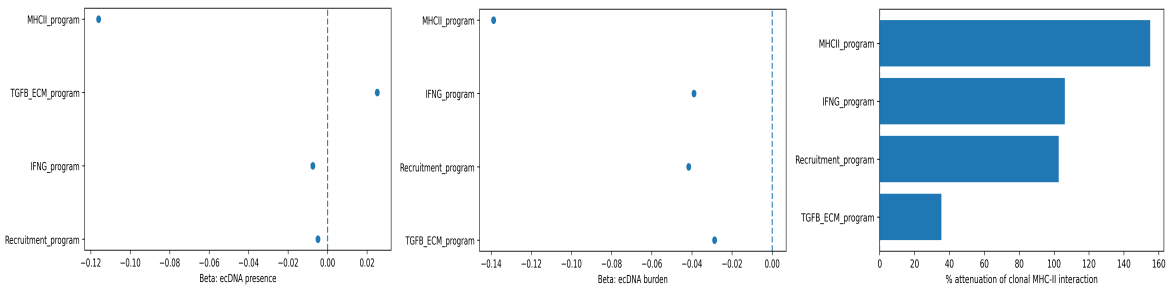


(a) (b)

Fig 3. ecDNA weakens the conversion of clonal MHC-II neoantigen burden into cytolytic activity; (a) ecDNA attenuates the association between clonal MHC-II neoantigen burden and predicted cytolytic activity; (b) Clonal MHC-II decoupling remains robust across PHBR-II and clonality thresholds.

This result indicates that although clonal MHC-II neoantigens remain immunologically relevant overall, ecDNA-positive tumors derive less cytolytic immune activation from that neoantigen burden. In contrast, clonal MHC-I neoantigen coupling did not show a statistically significant ecDNA specific weakening (clonal strong MHC-I x ecDNA interaction on CYT: beta = -0.0564, p= 0.0546, q=0.109). Similarly, the clonal MHC-II interaction was not significant for CD8 score alone (beta = -0.0428, p = 0.122), suggesting that the strongest effect is on cytolytic immune activation rather than on bulk CD8 abundance alone.

This result points to a more specific mechanistic model - ecDNA positive tumors do not appear to evade immunity by eliminating neoantigens, but rather by weakening the conversion of clonal MHC-II associated neoantigenicity into productive cytolytic response.



(a) (b) (c)

Fig 4. The clonal MHC-II decoupling signal maps to priming related immune programs; (a) ecDNA presence is associated most strongly with reduced MHC-II related immune program activity; (b) Increasing ecDNA burden is associated with progressive lower MHC-II program activity; (c) MHC-II IFNγ, and recruitment programs strongly attenuate the clonal MHC-II decoupling signal.

To better understand what might account for the weakened clonal MHC-II to CYT relationship in ecDNA positive tumors, we performed attenuation analyses by adding immune context pathway programs into the interaction model. Among the pathway programs tested, the MHCII program showed the clearest negative association with ecDNA presence (Fig 4(a)). By contrast, IFNγ and recruitment programs showed weaker or near-null baseline ecDNA associations, and TGF$\boldsymbol{\beta}$ECM did not show a consistent pattern suggestive of a simple global stromal exclusion model. ecDNA burden produced an even stronger negative association with MHCII program than binary ecDNA status alone (Fig. 4 (b)), supporting a dose dependent relationship between increasing ecDNA load and impaired MHC-II associated immune context.

In the base model, the clonal MHC-II x ecDNA interaction term was negative (beta= -0.0875, p= 1.11 x 10^{-3}). When the immune pathway was added, the interaction term was nearly eliminated (beta = 0.0054, p=0.60). Inclusion of MCH-II reversed the interaction coefficient, indicating very strong overlap between the clonal MHC-II coupling signal and MHC-II related transcriptional state. As shown in Fig 3(b), the negative ecDNA x clonal MHC-II interaction remained significant across a range of PHBR-II and cancer cell fraction thresholds. This

demonstrates that the decoupling signal is not an artifact of a single threshold choice and strengthens the interpretation that impaired response to clonal MHC-II neoantigens is a stable property of ecDNA-positive tumors. Adding the MHC-II program, IFNγ program, or Recruitment program to the interaction model substantially attenuated the ecDNA-associated clonal MHC-II decoupling signal, whereas TGF$\boldsymbol{\beta}$_ECM produced only partial attenuation.

These patterns suggest that the ecDNA associated failure to translate clonal MHC-II neoantigens into cytolytic activity is embedded within broader immune context programs related to helper cell priming, IFNγ associated activation, and recruitment. We also tested whether STING signalling could explain the immune cold phenotype. In the purity-adjusted STING cohort (n=1871), ecDNA presence was not associated with acute STING effector score (beta = 0.0055, p=0.88), nor was ecDNA burden (beta = -0.0197, p=0.41). Moreover, adding STING score as a covariate did not attenuate the ecDNA effect on CYT or CD8; if anything, the ecDNA coefficients remained similar or slightly stronger. Thus, although STING effector activity is strongly positively associated with CYT and CD8 overall, it does not explain the ecDNA associated immune suppression in these models. These analyses support that ecDNA associated immune suppression is more closely related to impaired antigen to response coupling in the effector immune context than to acute STING deficiency.

Finally, we tested whether the immune cold phenotype in ecDNA positive tumors could be attributed to enrichment of damaging mutations in canonical immune evasion genes. We did not observe enrichment of damaging alterations in antigen-presentation machinery genes or JAK/STAT pathway genes in ecDNA positive tumors.

3.4 Overall Interpretation

Overall, these results refine the current model of ecDNA-associated immune escape. Consistent with prior pan-cancer work, ecDNA-positive tumors in our integrated TCGA analysis are relatively immune-cold: after adjustment for cancer type and tumor purity, they show reduced cytolytic activity, reduced CD8 signal, and lower inferred CD8 T-cell infiltration, with stronger immune suppression as ecDNA burden increases.

However, unlike earlier work suggesting broad suppression of antigen-presentation genes, our adjusted analyses do not support a coordinated global shutdown of the MHC-I transcriptional program. Instead, we observe a more selective pattern: HLA-A and B2M show evidence of reduction in some models, but the full MHC-I composite and broader APM score are not robustly suppressed after stringent adjustment. Likewise, global neoantigen presentation metrics are largely preserved, and classical MHC-I-focused immunoediting signals are not evident.

The strongest mechanistic signal is instead that clonal MHC-II neoantigen burden remains positively associated with CYT overall, but this relationship is significantly weakened in ecDNA-positive tumors. This supports the hypothesis that the ecDNA-positive tumors are immune-cold not because they lack antigenicity or undergo broad MHC-I loss, but because they are impaired in converting clonal MHC-II neoantigenicity into effective cytolytic immune activation.

Discussion

Our results provide a refined perspective on the current model of ecDNA-associated immune escape. Similar to findings from earlier pan-cancer investigations, the immune-cold phenotype of ecDNA-positive tumors persisted even after rigorous adjustments for cancer type and tumor purity (Wu, Tao, et al.). This state is characterized by diminished cytolytic activity, a reduced CD8-associated signal, and decreased inferred cytotoxic immune cell infiltration. Significantly, these findings were reproducible across RNA-derived immune signatures and orthogonal immune deconvolution analyses, supporting that the phenotype reflects a true within-tumor type immune deficit rather than differences in lineage composition alone.

However, we found that this immune-cold phenotype cannot be explained by reduced global antigenicity or a broad shutdown of MHC-I antigen presentation machinery. ecDNA-positive tumors did not show major differences in PHBR-I, PHBR-II, or strong binder fractions, and classical MHC-I focused immunoediting signals were not evident. While some models suggested selective negative associations for HLA-A and B2M, the comprehensive MHC-I composite and APM scores did not show significant suppression following strict adjustments. Together, these findings argue against a model in which ecDNA tumors evade immunity primarily by losing neoantigens or globally disabling MHC-I expression.

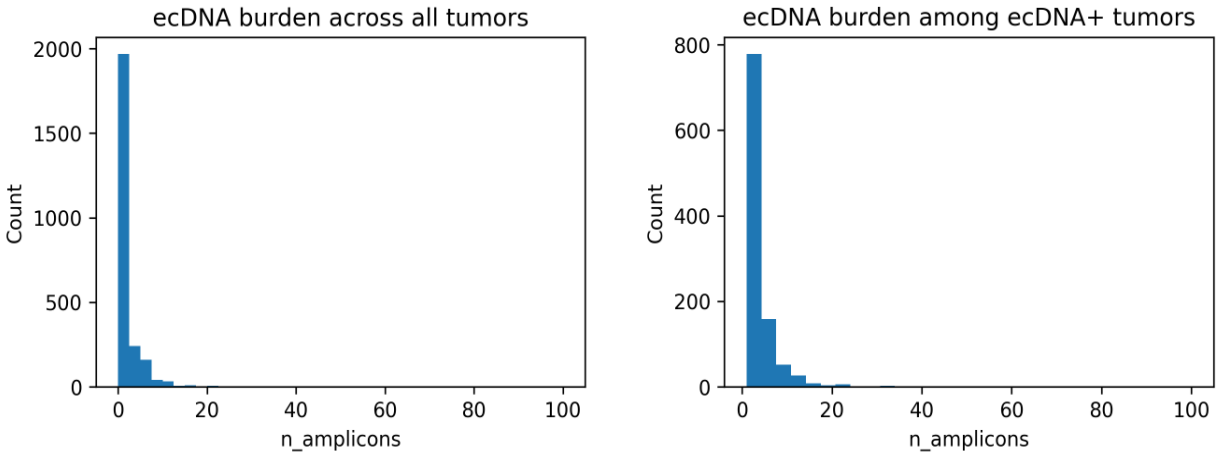
An exploratory analysis suggested that part of the ecDNA-associated immune-cold phenotype overlaps with broader aneuploidy-associated immune suppression. ecDNA-positive tumors showed higher SCNA burden, both for binary ecDNA presence and increasing ecDNA burden. When SCNA was added to the CYT model, the ecDNA effect was substantially attenuated, suggesting that some of the reduction in cytolytic activity in ecDNA-positive tumors may reflect shared biology with highly aneuploid tumors. However, PERK and RIDD UPR

branch scores were not significant and did not explain the ecDNA effect, indicating that the specific UPR readouts tested here do not account for the ecDNA-associated immune-cold phenotype. Importantly, the ecDNA-associated attenuation of clonal MHC-II neoantigen-to-CYT coupling remained stable after adjustment for SCNA and UPR variables, supporting the interpretation that this disconnect is not simply a byproduct of general aneuploidy-associated immune suppression (Supplementary Figure S2).

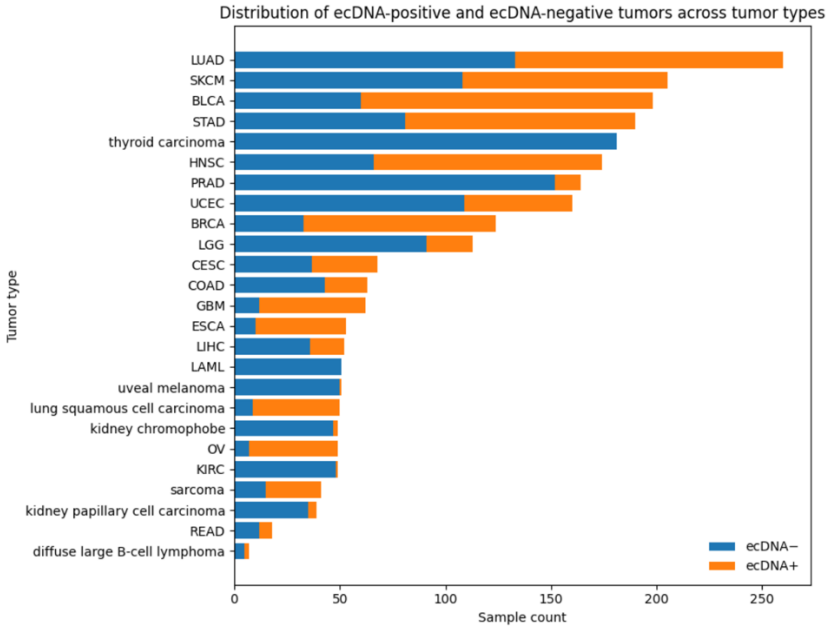
We also explored NeoPrecis-based mutation immunogenicity scores as an extension beyond peptide-HLA binding alone. However, NeoPrecis-based MHC-II summaries did not reproduce the ecDNA-associated negative interaction observed with binding-based clonal MHC-II burden, potentially reflecting reduced sample size and more limited MHC-II training data for this framework. Accordingly, the main conclusions of this study are based on mutations predicted to be subject to immune surveillance rather than on a broader TCR-reactivity model.

Instead, the strongest mechanistic signal was the attenuation of clonal MHC-II neoantigen-to-CYT coupling in ecDNA-positive tumors. While a higher clonal MHC-II neoantigen burden was associated with greater cytolytic activity overall, this relationship was significantly weaker in ecDNA-positive tumors and remained robust across threshold choices. Pathway analyses further suggested that this effect is linked to MHC-II, IFNγ, and recruitment-related immune programs rather than acute STING deficiency or simple stromal exclusion. Overall, these findings support the updated hypothesis that ecDNA tumors are immune-cold not because they lack antigenicity or undergo broad MHC loss, but because they are functionally impaired in converting clonal MHC-II neoantigenicity into effective cytolytic immune activation.

Appendix A. Supplementary Figures



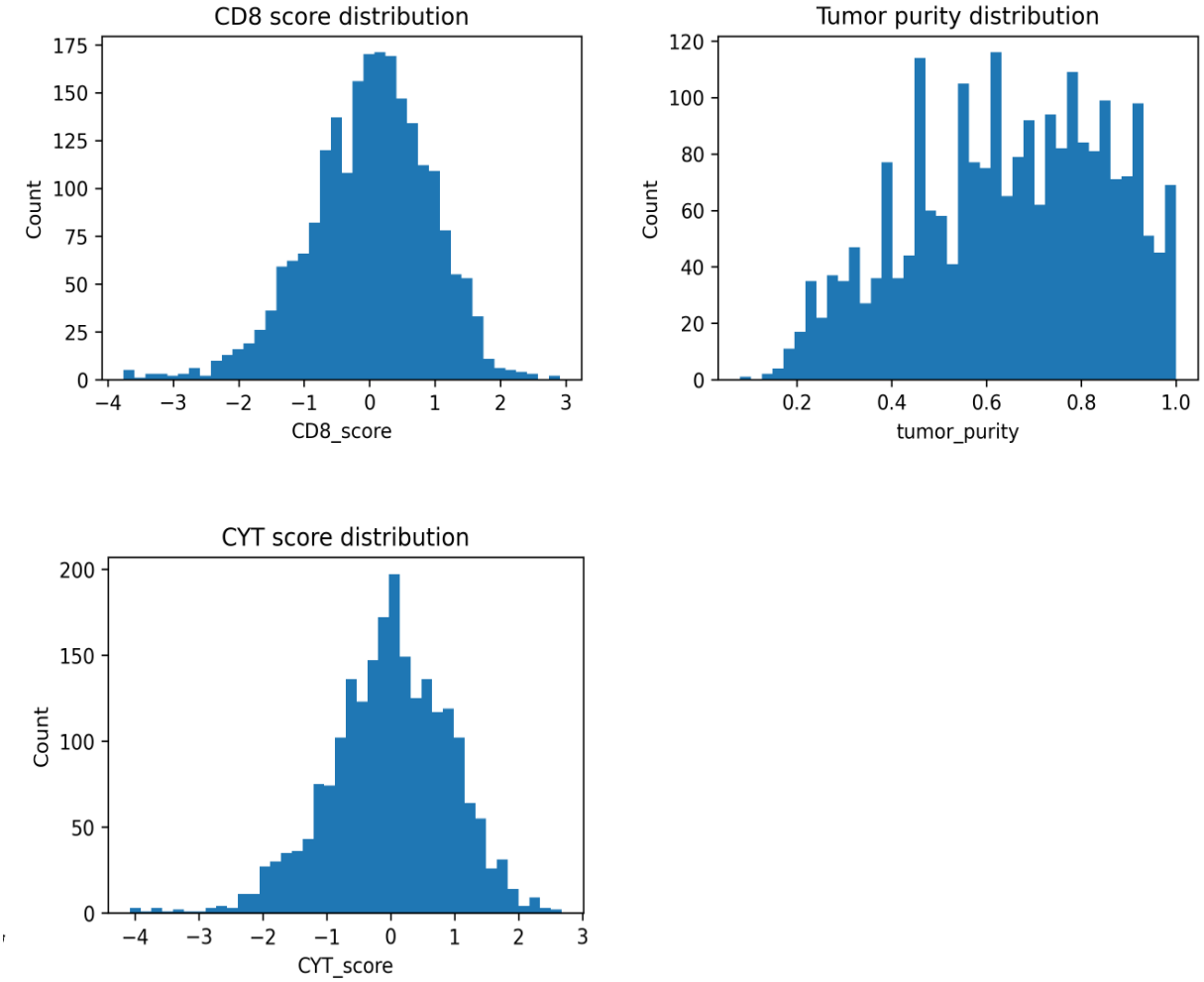
(a) (b)



(c)

Supplementary Figure S1. Cohort composition and distribution of key structural, immune, and neoantigen-related variables used in the study.

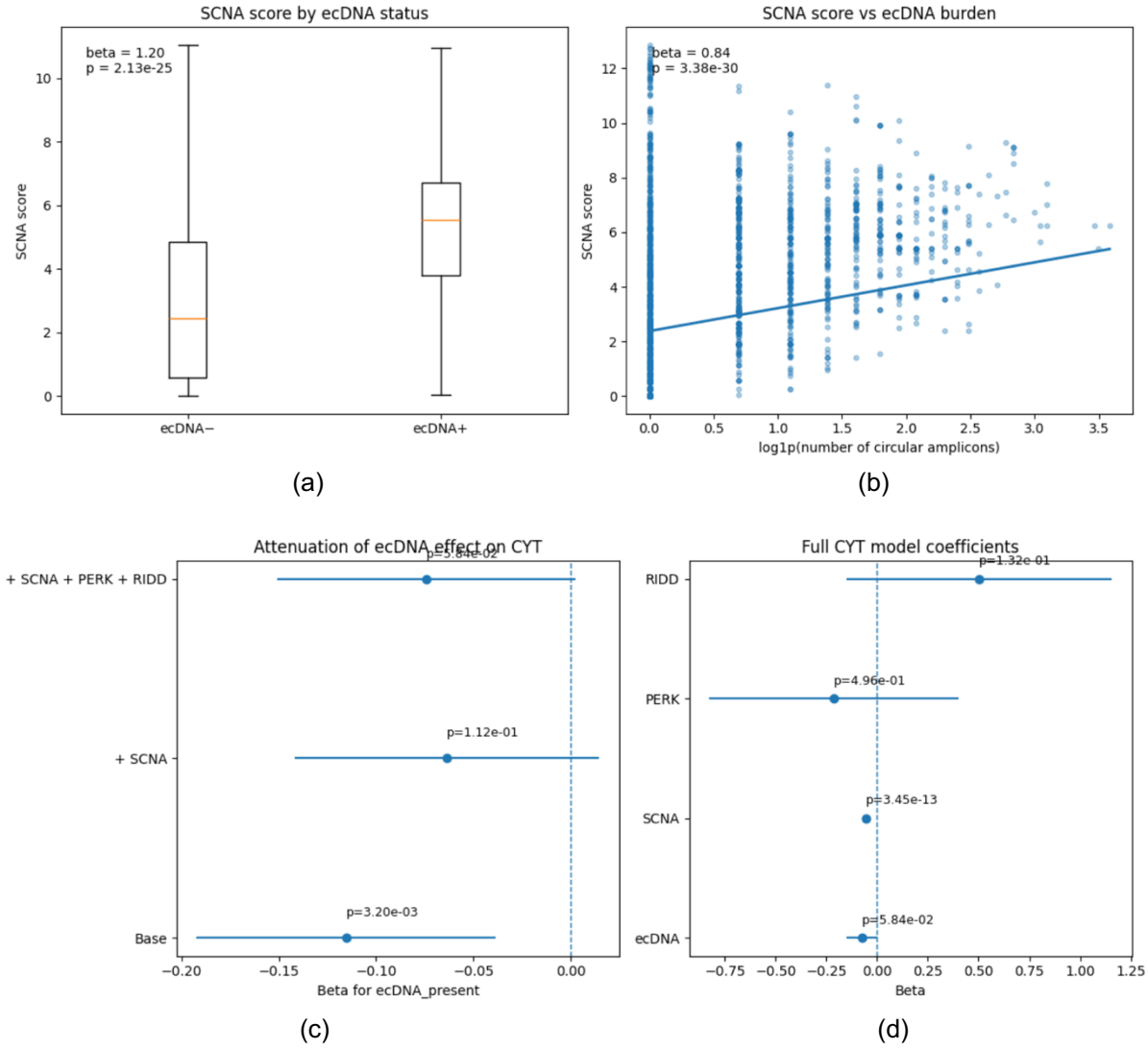
(a) Distribution of ecDNA burden across all tumors
(b) Distribution of ecDNA burdens across ecDNA-positive tumors only
(c) Distribution of ecDNA-positive and ecDNA-negative tumors across tumor types
(d) Distribution of CD8 score, Tumor purity distribution and CYT score. These plots illustrate the sparsity of ecDNA burden, heterogeneity across tumor types



(d)

Supplementary Figure S1. Cohort composition and distribution of key structural, immune, and neoantigen-related variables used in the study, Continued.

(a) Distribution of ecDNA burden across all tumors
(b) Distribution of ecDNA burdens across ecDNA-positive tumors only
(c) Distribution of ecDNA-positive and ecDNA-negative tumors across tumor types
(d) Distribution of CD8 score, Tumor purity distribution and CYT score. These plots illustrate the sparsity of ecDNA burden, heterogeneity across tumor types



Supplementary Figure S2. Partial overlap between ecDNA-associated immune suppression and SCNA burden

(a) SCNA score by ecDNA status
(b) SCNA score vs ecDNA burden
(c) Attenuation of ecDNA effect on CYT after adding SCNA / UPR covariates
(d) Full CYT model coefficients for ecDNA, SCNA, PERK, and RIDD

References

Alspach, E., Lussier, D.M., Miceli, A.P., Kizhvatov, I., DuPage, M., Luoma, A.M., Meng, W., Lichti, C.F., Esaulova, E., Vomund, A.N. and Runci, D., 2019. MHC-II neoantigens shape tumour immunity and response to immunotherapy. Nature, 574(7780), pp.696-701.

Becht, E., Giraldo, N.A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., Selves, J., Laurent-Puig, P., Sautès-Fridman, C., Fridman, W.H. and De Reynies, A., 2016. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome biology, 17(1), p.218.

Carter, S.L., Cibulskis, K., Helman, E., McKenna, A., Shen, H., Zack, T., Laird, P.W., Onofrio, R.C., Winckler, W., Weir, B.A. and Beroukhim, R., 2012. Absolute quantification of somatic DNA alterations in human cancer. Nature biotechnology, 30(5), pp.413-421.

Deshpande, V., Luebeck, J., Nguyen, N.P.D., Bakhtiari, M., Turner, K.M., Schwab, R., Carter, H., Mischel, P.S. and Bafna, V., 2019. Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. Nature communications, 10(1), p.392.

Adam Frankish, Sílvia Carbonell-Sala, Mark Diekhans, Irwin Jungreis, Jane E Loveland, Jonathan M Mudge, Cristina Sisu, James C Wright, Carme Arnan, If Barnes, Abhimanyu Banerjee, Ruth Bennett, Andrew Berry, Alexandra Bignell, Carles Boix, Ferriol Calvet, Daniel Cerdán-Vélez, Fiona Cunningham, Claire Davidson, Sarah Donaldson, Cagatay Dursun, Reham Fatima, Stefano Giorgetti, Carlos Garcıa Giron, Jose Manuel Gonzalez, Matthew Hardy, Peter W Harrison, Thibaut Hourlier, Zoe Hollis, Toby Hunt, Benjamin James, Yunzhe Jiang, Rory Johnson, Mike Kay, Julien Lagarde, Fergal J Martin, Laura Martínez Gómez, Surag Nair,

Pengyu Ni, Fernando Pozo, Vivek Ramalingam, Magali Ruffier, Bianca M Schmitt, Jacob M Schreiber, Emily Steed, Marie-Marthe Suner, Dulika Sumathipala, Irina Sycheva, Barbara Uszczynska-Ratajczak, Elizabeth Wass, Yucheng T Yang, Andrew Yates, Zahoor Zafrulla, Jyoti S Choudhary, Mark Gerstein, Roderic Guigo, Tim J P Hubbard, Manolis Kellis, Anshul Kundaje, Benedict Paten, Michael L Tress, Paul Flicek, GENCODE: reference annotation for the human and mouse genomes in 2023, Nucleic Acids Research, Volume 51, Issue D1, 6 January 2023, Pages D942–D949

Goldman, M.J., Craft, B., Hastie, M., Repečka, K., McDade, F., Kamath, A., Banerjee, A., Luo, Y., Rogers, D., Brooks, A.N. and Zhu, J., 2020. Visualizing and interpreting cancer genomics data via the Xena platform. Nature biotechnology, 38(6), pp.675-678.

Kim, H., Nguyen, N.P., Turner, K., Wu, S., Gujar, A.D., Luebeck, J., Liu, J., Deshpande, V., Rajkumar, U., Namburi, S. and Amin, S.B., 2020. Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. Nature genetics, 52(9), pp.891-897.

Pyke, R.M., Thompson, W.K., Salem, R.M., Font-Burgada, J., Zanetti, M. and Carter, H., 2018. Evolutionary pressure against MHC class II binding cancer mutations. Cell, 175(2), pp.416-428.

Marty, R., Kaabinejadian, S., Rossell, D., Slifker, M.J., van de Haar, J., Engin, H.B., de Prisco, N., Ideker, T., Hildebrand, W.H., Font-Burgada, J. and Carter, H., 2017. MHC-I genotype restricts the oncogenic mutational landscape. Cell, 171(6), pp.1272-1283.

Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M. and Alizadeh, A.A., 2015. Robust enumeration of cell subsets from tissue expression profiles. Nature methods, 12(5), pp.453-457.

Rooney, M.S., Shukla, S.A., Wu, C.J., Getz, G. and Hacohen, N., 2015. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell, 160(1), pp.48-61.

Turner, K.M., Deshpande, V., Beyter, D., Koga, T., Rusert, J., Lee, C., Li, B., Arden, K., Ren, B., Nathanson, D.A. and Kornblum, H.I., 2017. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. Nature, 543(7643), pp.122-125.

Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C. and Stuart, J.M., 2013. The cancer genome atlas pan-cancer analysis project. Nature genetics, 45(10), pp.1113-1120.

Wu, S., Turner, K.M., Nguyen, N., Raviram, R., Erb, M., Santini, J., Luebeck, J., Rajkumar, U., Diao, Y., Li, B. and Zhang, W., 2019. Circular ecDNA promotes accessible chromatin and high oncogene expression. Nature, 575(7784), pp.699-703.

Wu, T., Wu, C., Zhao, X., Wang, G., Ning, W., Tao, Z., Chen, F. and Liu, X.S., 2022. Extrachromosomal DNA formation enables tumor immune escape potentially through regulating antigen presentation gene expression. Scientific reports, 12(1), p.3590.