

DOI: <https://doi.org/10.60797/jbg.2026.31.1>

EDN: KTFNUH

**DUAL DECONVOLUTION OF CANINE MAMMARY TUMORS REVEALS ROBUST CD4<sup>+</sup> T CELL ENRICHMENT VIA EPIC AND CIBERSORTX**

Research article

**Herath Mudiyansele R.<sup>1,\*</sup>**<sup>1</sup>ORCID : 0009-0000-7982-0690;<sup>1</sup>Independent researcher, Stockholm, Sweden

\* Corresponding author (ruwiniherath92[at]gmail.com)

**Abstract**

Canine mammary tumors (CMTs) are an excellent model for human breast cancer because of their spontaneous origin and immunological similarity. This study applied dual immune deconvolution — EPIC and CIBERSORTx — to bulk RNA-seq data (GSE136197) from 48 tumor and normal canine mammary samples to define the immune tumor microenvironment. Tumor tissues showed a large, significant enrichment of CD4<sup>+</sup> T cells (Cohen's  $d = -0,83$ ,  $FDR < 0,01$ ), detectable as early as carcinoma in situ. CIBERSORTx further resolved resting and activated memory CD4<sup>+</sup> T cells and T follicular helper cells, highlighting functional diversity. These results reveal conserved CD4<sup>+</sup> T-cell-driven immune remodeling across dogs and humans and demonstrate that combined EPIC–CIBERSORTx deconvolution is a cost-effective way to generate high-resolution immune profiles that can guide comparative oncology and immunotherapy development.

**Keywords:** canine mammary tumour, immune deconvolution, CD4<sup>+</sup> T cells, EPIC, CIBERSORTx, comparative oncology.

**ДВОЙНАЯ ДЕКОНВОЛЮЦИЯ ОПУХОЛЕЙ МОЛОЧНОЙ ЖЕЛЕЗЫ У СОБАК ВЫЯВЛЯЕТ ЗНАЧИТЕЛЬНОЕ ОБОГАЩЕНИЕ CD4<sup>+</sup> Т-КЛЕТОК С ПОМОЩЬЮ EPIC И CIBERSORTX**

Научная статья

**Херат Мудиянселге Р.<sup>1,\*</sup>**<sup>1</sup>ORCID : 0009-0000-7982-0690;<sup>1</sup>Независимый исследователь, Стокгольм, Швеция

\* Корреспондирующий автор (ruwiniherath92[at]gmail.com)

**Аннотация**

Опухоли молочной железы собак (ОМС) являются отличной моделью рака молочной железы человека благодаря своему спонтанному происхождению и иммунологическому сходству. В данном исследовании проведена двойная иммунная деконволюция — EPIC и CIBERSORTx — данных транскриптомного секвенирования (GSE136197) из 48 образцов опухолевых и нормальных тканей молочной железы собак для определения иммунного микроокружения опухоли. В опухолевых образцах выявлено значительное и статистически достоверное обогащение CD4<sup>+</sup> Т-клеток (коэновское  $d = -0,83$ ,  $FDR < 0,01$ ), которое прослеживается уже на стадии карциномы in situ. Анализ CIBERSORTx дополнительно позволил выделить покоеющиеся и активированные клетки памяти CD4<sup>+</sup> Т, а также Т-фолликулярные хелперы, что подчеркивает функциональное разнообразие. Полученные результаты свидетельствуют о сохранённом механизме иммунного ремоделирования с участием CD4<sup>+</sup> Т-клеток у собак и человека и демонстрируют, что комбинированная деконволюция EPIC–CIBERSORTx является экономичным способом получения высокодетализированных иммунных профилей, способных служить основой для сравнительной онкологии и разработки иммунотерапии.

**Ключевые слова:** опухоль молочной железы у собак, иммунная деконволюция, Т-клетки CD4<sup>+</sup>, EPIC, CIBERSORTx, сравнительная онкология.

**Introduction**

Canine mammary tumors (CMTs) are the most common neoplasms in intact female dogs and share key characteristics with human breast cancer, including spontaneous tumor development, hormonal dependence, and similar age at onset and histological subtypes [1]. Molecular parallels include comparable expression patterns of estrogen and progesterone receptors, epidermal growth factor receptors, p53, Ki-67, and cyclooxygenases, supporting CMTs as useful models in comparative oncology [2]. Prognostic markers such as lymph node involvement, tumor size, and clinical stage also align across species [3].

The tumor immune microenvironment (TIME), composed of immune and stromal cells, plays a critical role in breast cancer progression and treatment response [4]. Subtypes like HER2-positive and triple-negative breast cancer show greater tumor-infiltrating lymphocytes (TILs), while luminal subtypes have lower immune infiltration and poorer outcomes [5]. Specific immune cells, such as cytotoxic lymphocytes, regulatory T cells, mast cells, and tumor-associated macrophages, influence immune escape and therapeutic resistance [6]. TIME subclassification frameworks (for example, inflamed, immune-desert, and immune-excluded) have been developed in human oncology to guide immunotherapy strategies [4], but their application to CMTs remains limited.

Computational deconvolution of bulk RNA-seq enables estimation of immune cell fractions based on transcriptomic data, offering a scalable approach to TIME profiling [7]. EPIC (Estimating the Proportions of Immune and Cancer cells) and CIBERSORTx [8] are such methods, validated across tissues and even across species, particularly for abundant immune



populations like CD4<sup>+</sup> and CD8<sup>+</sup> T cells, B cells, macrophages, and cancer-associated fibroblasts [7]. Benchmarking studies have shown that deconvolution tools like EPIC, CIBERSORT, and MCP-counter perform well with high-quality reference signatures, though limitations exist for under-characterized species or cell types [9].

This study uses EPIC to profile immune infiltration in publicly available CMT transcriptomic data. The aim of the work is to determine whether CD4<sup>+</sup> T cell abundance, a key immune correlate in human breast cancer, is also a distinguishing feature in canine tumors, thus contributing to cross-species immuno-oncology insights.

### Research methods and principles

#### Sample collection and RNA sequencing:

This study re-analyzed raw RNA-seq data from 48 canine mammary gland samples (normal tissue, atypia-normal, carcinoma in situ, and invasive carcinoma) generated by Graim et al. [10] (GSE136197). In the original work, tumor and matched normal tissues were obtained at surgical excision or diagnostic biopsy from privately owned dogs with informed owner consent and institutional ethical approval. Immediately after resection, tissues were flash-frozen in liquid nitrogen and stored at -80 °C until processing. Total RNA was extracted using TRIzol reagent, and integrity was confirmed by Bioanalyzer (RIN ≥ 7). Poly(A)-enriched, strand-specific cDNA libraries were prepared with the Illumina TruSeq RNA Sample Preparation Kit and sequenced on an Illumina HiSeq 2500 platform to a depth of ~50 million paired-end reads per sample.

#### Metadata curation and tumor quality control:

Sample-level metadata, including histological classification, versus normal status, and sequencing run information, were obtained from the original publication and the associated GEO records. All annotations were cross-checked and harmonized as follows: (i) sample identifiers were matched between the supplementary tables of Graim et al. and the GEO sample sheets; (ii) histological labels were standardized to four categories (normal, atypia-normal, carcinoma in situ, invasive carcinoma); (iii) any records with incomplete or ambiguous histology were excluded (none met this criterion); and (iv) sequencing metrics such as total read count and mapping rate were screened to confirm data completeness. The final curated metadata table was merged with immune cell fraction outputs for downstream statistical analysis. The data were then reshaped into long format using the `pivot_longer` function from the `tidyr` package, which facilitated efficient plotting and subgroup comparisons.

#### Data preprocessing for immune deconvolution:

Raw read counts were transformed into Transcripts Per Million (TPM) to standardize expression values across samples and ensure compatibility with the algorithms used. Because EPIC and CIBERSORTx were originally developed for human transcriptomes, Ensembl canine gene identifiers were mapped to their corresponding human orthologs using the `biomaRt` package, retaining only one-to-one orthologs to preserve downstream cell-type annotation fidelity and avoid ambiguities from gene duplication or divergence.

#### Statistical analysis:

All statistical analyses were performed in R version 4.2.3 using the packages `tidyverse`, `phatmap`, `FSA`, and `effsize`. Because immune cell fractions were not normally distributed, group-wise comparisons were carried out with the Kruskal–Wallis test, followed by Dunn's post hoc pairwise tests. P-values were adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR) procedure, and statistical significance was defined as FDR-adjusted  $p < 0,05$ . Effect sizes were calculated with Cohen's  $d$  and interpreted as small ( $\geq 0,2$ ), medium ( $\geq 0,5$ ), or large ( $\geq 0,8$ ). All figures and heatmaps were generated directly from these R workflows to ensure full reproducibility.

Immune cell proportions were estimated using two computational deconvolution tools: EPIC (Estimating the Proportions of Immune and Cancer cells) and CIBERSORTx. The EPIC R package was applied using its default reference signature, which estimates the abundance of major immune and stromal cell types, including CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, B cells, natural killer (NK) cells, macrophages, cancer-associated fibroblasts (CAFs), and endothelial cells. The category “otherCells”, which includes uncharacterised stromal or malignant cells, was also retained to provide a comprehensive overview of cellular heterogeneity. In parallel, CIBERSORTx was used via its web-based interface, leveraging the LM22 signature matrix that defines 22 human hematopoietic cell types. The TPM-normalised, human ortholog-converted matrix was uploaded without batch correction, as all samples originated from a single dataset. CIBERSORTx was run in relative mode, and immune cell fraction estimates were downloaded for integration with EPIC results.

To visualize immune infiltration patterns, z-score scaled heatmaps were generated using the `phatmap` package, with scaling applied across rows to highlight differences in immune cell abundance between histological subtypes. Boxplots were created using `ggplot2` to compare the distributions of specific immune cell populations, particularly CD4<sup>+</sup> T cells, across different histological groups. To quantify the magnitude of immune differences, effect sizes were calculated using the `cohen.d` function from the `effsize` package, with thresholds for interpretation set at 0,2 (small), 0,5 (medium), and 0,8 (large). Particular attention was given to CD4<sup>+</sup> T cells due to their established relevance in human breast cancer immunology.

Given the heterogeneous distribution and small sample sizes within each subgroup, non-parametric statistical testing was used throughout. Kruskal–Wallis tests were employed for overall comparisons between histology groups, followed by Dunn's post hoc tests using the `dunnTest` function from the `FSA` package to identify pairwise differences. To control the false discovery rate, all  $p$ -values were adjusted using the Benjamini–Hochberg procedure. All code and analytical workflows were implemented in Jupyter Lab with an R kernel and are available upon request to ensure reproducibility.

#### Data availability:

The full analysis pipeline, including code scripts for EPIC deconvolution, ortholog mapping, and figure generation, is available at the following GitHub repository: <https://github.com/ruwinimadhu/Immune-Deconvolution-GSE136197>. The processed data derived from GSE136197 and human ortholog mappings are included in the repository for reproducibility.

The data used for EPIC analysis were downloaded from the NCBI Gene Expression Omnibus (GEO) under accession number GSE136197 [10]. The dataset is available under a CC BY-NC 4.0 license.

### Funding & ethics disclosure:

This research did not receive any specific grant from funding agencies. No new animal experiments were conducted. The analysis was performed on publicly available data.

### Main results

Computational deconvolution of RNA-seq data from 48 canine mammary tissue samples using EPIC revealed a distinct reshaping of the tumor immune microenvironment (TIME) between tumor and normal samples. A consistent increase in CD4<sup>+</sup> T cell infiltration was observed across all histological subtypes in tumor tissues. This increase was statistically significant (Cohen's  $d = -0,83$ ,  $p < 0,01$ ), representing a large effect size and indicating a biologically meaningful elevation in helper T cell recruitment or expansion (Fig. 1). Although CD8<sup>+</sup> T cells showed a modest upward trend, differences in macrophages, B cells, and NK cells were less pronounced (Cohen's  $d < 0,5$ ) and did not reach statistical significance.

These findings were visually reinforced by heatmap clustering of row-wise z-scores *To further dissect*, which demonstrated clear separation between tumor and normal samples, primarily driven by differences in CD4<sup>+</sup> and CD8<sup>+</sup> T cell levels. Statistical comparisons using the Kruskal–Wallis test confirmed that CD4<sup>+</sup> T cell proportions varied significantly across histological groups ( $p < 0,01$ ), and Dunn's post hoc analysis pinpointed significant differences between carcinoma and normal tissue. Notably, CD4<sup>+</sup> T cell enrichment was already evident in early lesions, such as carcinoma in situ, suggesting that immune remodeling is an early event in canine mammary tumor progression.

To further dissect the composition of the CD4<sup>+</sup> compartment, CIBERSORTx was applied using the LM22 signature matrix. Compared to EPIC, CIBERSORTx provided higher granularity by identifying specific T cell subsets within the CD4<sup>+</sup> population. The dominant subsets in tumor samples included both resting and activated CD4<sup>+</sup> memory T cells. Additionally, a subset of tumors showed detectable levels of T follicular helper (Tfh) cells. Regulatory T cells (Tregs) were also present in several tumor samples, although typically at low abundance (Fig. 2).

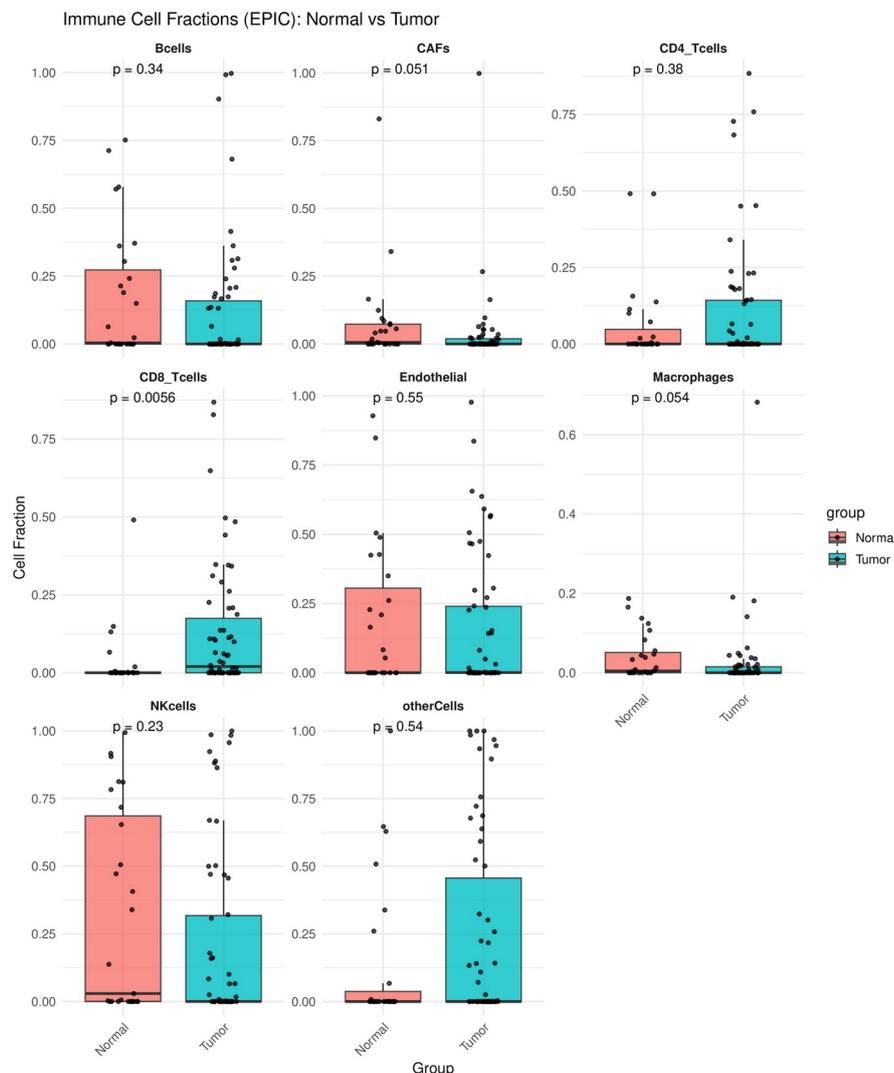


Figure 1 - Immune cell fraction in normal vs tumour tissues estimated using EPIC deconvolution

DOI: <https://doi.org/10.60797/jbg.2026.31.1.1>

Note: several immune cell types showed a significant difference between groups. Notably, CD4<sup>+</sup> T cells were significantly enriched in tumour tissues

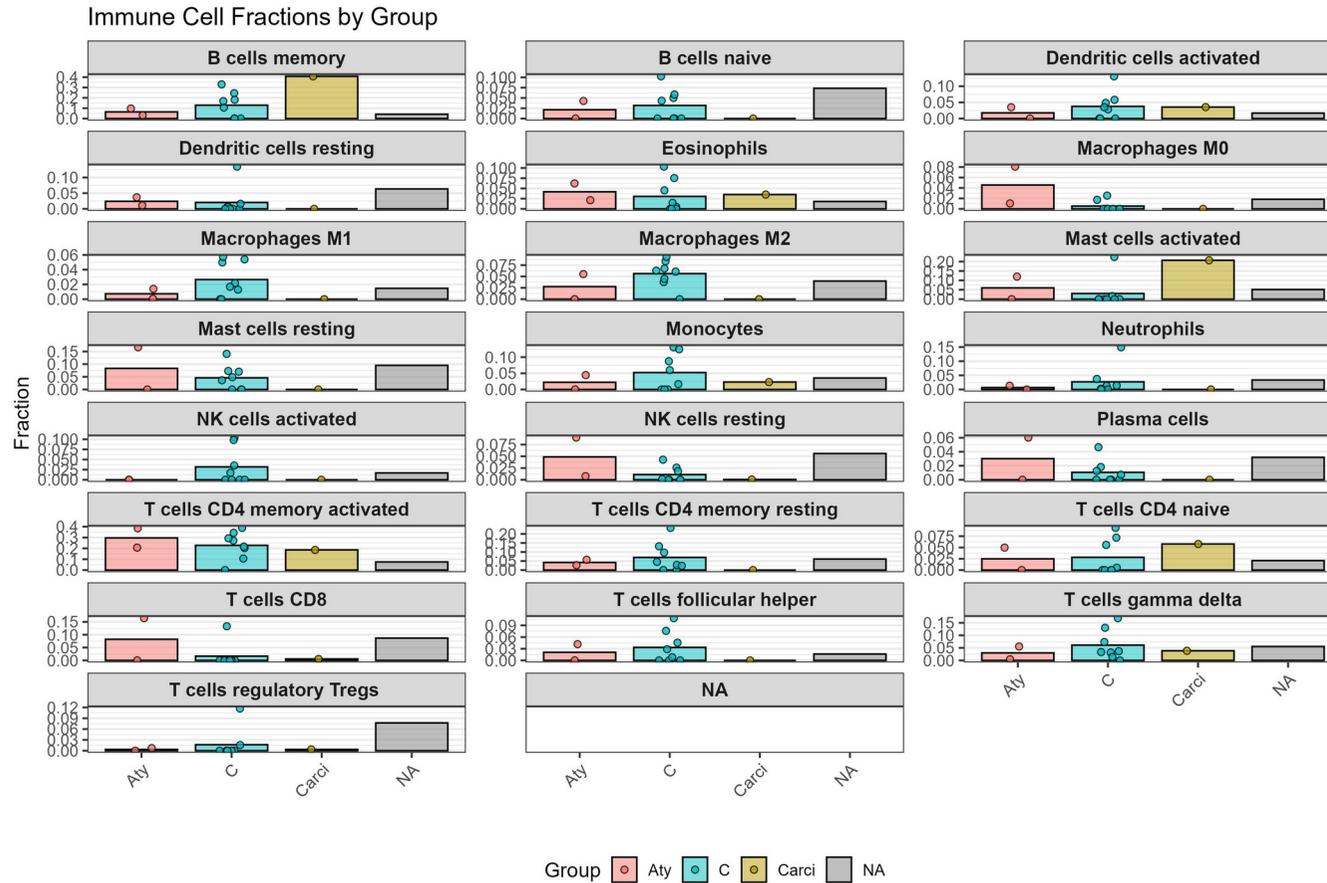


Figure 2 - CD4<sup>+</sup> T cell subset composition across histological subtypes of canine mammary tumours as estimated by CIBERSORTx  
 DOI: <https://doi.org/10.60797/jbg.2026.31.1.2>

*Note: boxplots show relative fractions of CD4<sup>+</sup> memory resting T cells, CD4<sup>+</sup> memory activated T cells, Tregs, and T follicular helper cells, highlighting functional diversity within the CD4<sup>+</sup> compartment across tumour progression stages. These patterns reflect early immune activation and possible immune escape mechanisms*



## Discussion

While [10] used the same GSE136197 cohort to model the molecular development and oncogenic drivers of canine mammary tumors, their focus remained on transcriptional trajectories and cancer-initiating pathways. In contrast, the present work interrogates the tumor immune microenvironment by applying a dual deconvolution strategy (EPIC and CIBERSORTx) to quantify CD4<sup>+</sup> T-cell enrichment and to resolve key CD4<sup>+</sup> subsets. This immune-centric analysis provides a higher-resolution view of lymphocyte composition and positions CD4<sup>+</sup> T cells as candidate biomarkers and therapeutic targets, extending the dataset's translational value toward comparative immuno-oncology and the design of future immunotherapy studies.

The elevation of CD4<sup>+</sup> T cells observed in this study aligns with findings in human breast cancer, where altered CD4<sup>+</sup> T cell infiltration correlates with immune evasion, therapy response, and tumour subtype [4], [11]. CD4<sup>+</sup> T cells play diverse roles in coordinating adaptive immunity via cytokine signaling, antigen presentation, and effector cell recruitment [12]. Their infiltration is often associated with both prognostic value and therapeutic responsiveness in human disease.

The identification of resting and activated memory CD4<sup>+</sup> T cells suggests a combination of immune surveillance and active response states. T follicular helper cells may indicate the formation of tertiary lymphoid structures, further supporting immune involvement in the tumour microenvironment. Although Tregs were infrequent, their presence hints at early immunosuppressive activity, potentially contributing to tumour immune escape. This functional heterogeneity of CD4<sup>+</sup> subsets reflects similar trends seen in human breast cancer, where the balance between effector and regulatory cells can shape disease progression and treatment outcomes.

By enabling this level of resolution, CIBERSORTx complements EPIC and strengthens the translational utility of immune deconvolution in comparative oncology. The conserved immune patterns between canine and human mammary tumours reinforce the value of canine models for immunotherapy research.

However, limitations must be acknowledged. Both EPIC and CIBERSORTx rely on human-derived reference signatures, which may introduce bias when applied to canine data. Moreover, they lack spatial context and cannot capture dynamic functional states such as T cell exhaustion. The absence of clinical metadata in this study also limits the ability to link immune features with prognosis or treatment outcomes.

Despite these limitations, this study demonstrates the feasibility and utility of applying dual-method immune deconvolution pipelines to veterinary cancer data. The consistent elevation of memory and activated CD4<sup>+</sup> T cell subsets suggests their potential as biomarkers of tumour immunity and as targets for future therapeutic strategies in canine mammary tumours.

## Conclusion

This study demonstrates a significant increase in CD4<sup>+</sup> T cell infiltration within the tumour immune microenvironment of canine mammary tumours (CMTs), as revealed by EPIC deconvolution of RNA-seq data. The consistent enrichment of CD4<sup>+</sup> T cells across histological subtypes — including early lesions such as carcinoma in situ—suggests that immune remodelling is an early and sustained feature of CMT progression. These findings position CD4<sup>+</sup> T cells as potential biomarkers and therapeutic targets in veterinary oncology, while also underscoring key immunological parallels between canine and human breast cancer. The use of computational deconvolution tools such as EPIC and CIBERSORTx highlights the feasibility of cost-effective immune profiling using existing transcriptomic datasets. To advance the clinical relevance of these findings, future studies should incorporate clinical outcome data, spatial immune profiling, and single-cell analyses to further elucidate the functional and prognostic roles of immune cells in CMTs.



## Дополнительные материалы

Дополнительные материалы доступны на онлайн-странице статьи.

### Благодарности

Автор выражает благодарность всем, кто предоставил общедоступный набор данных РНК-секвенирования, использованный в данном исследовании (GSE136197). Особая благодарность разработчикам EPIC и CIBERSORTx за предоставление свободного доступа к своим вычислительным инструментам. Автор также выражает благодарность Магнусу Обринку из SLU за руководство и рекомендации в ходе основополагающей экспериментальной работы с CRISPRi–CMT, которая вдохновила на создание данного вычислительного расширения. Данный проект не получал специального финансирования.

### Конфликт интересов

Не указан.

### Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

## Supplementary materials

Supplementary materials are available online on the article's webpage.

### Acknowledgement

The author gratefully acknowledges the original contributors of the publicly available RNA-seq dataset used in this study (GSE136197). Special thanks are extended to the developers of EPIC and CIBERSORTx for making their computational tools freely accessible. The author also wishes to thank Magnus Åbrink at SLU for supervision and guidance during the foundational CRISPRi–CMT experimental work, which inspired this computational extension. No specific funding was received for this project.

### Conflict of Interest

None declared.

### Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

## Список литературы на английском языке / References in English

1. Queiroga F.L. Canine Mammary Tumors as a Model to Study Human Breast Cancer: Most Recent Findings / F.L. Queiroga, T. Raposo, M.I. Carvalho [et al.] // *In Vivo*. — 2011. — Vol. 25. — № 3. — P. 455–465.
2. Abdelmegeed S. Canine mammary tumors as a model for human disease (Review) / S. Abdelmegeed, S. Mohammed // *Oncology Letters*. — 2018. — Vol. 15. — № 6. — P. 8195–8205. — DOI: 10.3892/ol.2018.8411.
3. Sleenckx N. Canine Mammary Tumours, an Overview / N. Sleenckx, H. De Rooster, E.J.B. Veldhuis Kroeze [et al.] // *Reproduction in Domestic Animals*. — 2011. — Vol. 46. — № 6. — P. 1112–1131. — DOI: 10.1111/j.1439-0531.2011.01816.x.
4. Binnewies M. Understanding the tumor immune microenvironment (TIME) for effective therapy / M. Binnewies, E.W. Roberts, K. Kersten [et al.] // *Nature Medicine*. — 2018. — Vol. 24. — № 5. — P. 541–550. — DOI: 10.1038/s41591-018-0014-x.
5. Otterlei Fjørtoft M. The Tumor Immune Microenvironment in Breast Cancer Progression / M. Otterlei Fjørtoft, K. Huse, I.H. Rye // *Acta Oncologica*. — 2024. — Vol. 63. — P. 359–367. — DOI: 10.2340/1651-226X.2024.33008.
6. Quail D.F. Breast Cancer Immunity: It is TIME for the Next Chapter / D.F. Quail, M. Park, A.L. Welm [et al.] // *Cold Spring Harbor Perspectives in Medicine*. — 2024. — Vol. 14. — № 2. — DOI: 10.1101/cshperspect.a041324.
7. *Bioinformatics for Cancer Immunotherapy: Methods and Protocols* / edited by S. Boegel // *Methods in Molecular Biology*. — New York : Springer US, 2020. — Vol. 2120. — DOI: 10.1007/978-1-0716-0327-7.
8. Newman A.M. Determining cell type abundance and expression from bulk tissues with digital cytometry / A.M. Newman, C.B. Steen, C.L. Liu [et al.] // *Nature Biotechnology*. — 2019. — Vol. 37. — № 7. — P. 773–782. — DOI: 10.1038/s41587-019-0114-2.
9. Sturm G. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology / G. Sturm, F. Finotello, F. Petitprez [et al.] // *Bioinformatics*. — 2019. — Vol. 35. — № 14. — P. i436–i445. — DOI: 10.1093/bioinformatics/btz363.
10. Graim K. Modeling molecular development of breast cancer in canine mammary tumors / K. Graim, D. Gorenshiteyn, D.G. Robinson [et al.] // *Genome Research*. — 2021. — Vol. 31. — № 2. — P. 337–347. — DOI: 10.1101/gr.256388.119.
11. Moorthy A. The Primary Tumor Immune Microenvironment Status Predicts the Response to Immunotherapy and Overall Survival in Breast Cancer / A. Moorthy, A. Quinn // *bioRxiv*. — 2020. — DOI: 10.1101/2020.07.03.186221.
12. Chu T. Cell type and gene expression deconvolution with BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology / T. Chu, Z. Wang, D. Pe'er [et al.] // *Nature Cancer*. — 2022. — Vol. 3. — № 4. — P. 505–517. — DOI: 10.1038/s43018-022-00356-3.