



**Russian-Armenian (Slavonic) University**

*Printed by the decision of the  
Academic Council of RAU*

**V E S T N I K**  
**OF RUSSIAN-ARMENIAN (SLAVONIC) UNIVERSITY**  
**(SERIES: PHYSICAL-MATHEMATICAL**  
**AND NATURAL SCIENCES)**

**RAU University Press**  
**No. 1/2024**

**Российско-Армянский (Славянский) университет**

*Печатается по решению  
Ученого Совета РАУ*

**В Е С Т Н И К**  
**РОССИЙСКО-АРМЯНСКОГО (СЛАВЯНСКОГО)**  
**УНИВЕРСИТЕТА**  
**(СЕРИЯ: ФИЗИКО-МАТЕМАТИЧЕСКИЕ**  
**И ЕСТЕСТВЕННЫЕ НАУКИ)**

**Издательство РАУ**  
**№ 1/2024**

**Главный редактор:** академик НАН РА, д.ф.-м.н., проф. **Казарян Э.М.**

**Зам. главного редактора:** к.ф.-м.н., д.филос.н., проф. **Аветисян П.С.**

**Ответственный секретарь:** к.х.н. **Шагинян Р.С.**

**Редакционная коллегия:**

**Математика и информатика**

*Р.Г. Арамян* – д.ф.-м.н., проф., Российско-Армянский университет (РАУ), Институт Математики НАН РА

*Д.Г. Асатрян* – д.т.н., проф., Институт проблем информатики и автоматизации НАН РА

*Г.Г. Казарян* – д.ф.-м.н., проф., Российско-Армянский университет (РАУ)

*О.В. Бесов* – член-корр.РАН, д.ф.-м.н., проф., Математический институт им. В.А. Стеклова РАН (Россия)

*В.И. Буренков* – д.ф.-м.н., проф., Математический институт им. В.А. Стеклова РАН (Россия)

*А.Г. Сергеев* – академик РАН, д.ф.-м.н., проф., Математический институт им. В.А. Стеклова РАН (Россия)

*А.И. Аветисян* – академик РАН, д.ф.-м.н., проф., Институт системного программирования им. В.П. Иванникова РАН (Россия)

*В.Ш. Меликян* – член-корр. НАН РА, д.т.н., проф., Российско-Армянский университет (РАУ), Synopsys

**Биологические и химические науки**

*А.А. Аракелян* – д.б.н., проф., Российско-Армянский университет (РАУ), Институт Молекулярной биологии НАН РА

*В.И. Мурунец* – д.б.н., проф., Московский государственный университет (МГУ) (Россия)

*А.А. Оганесян* – к.б.н., доц., Российско-Армянский университет (РАУ)

*Р.В. Захарян* – к.б.н., Российско-Армянский университет (РАУ)

*Г.Г. Данагулян* – член-корр. НАН РА, д.х.н., проф., Российско-Армянский университет (РАУ)

*К.Б. Назарян* – д.б.н., проф., Институт Молекулярной биологии НАН РА

*Л.М. Епископосян* – д.б.н., проф., Институт Молекулярной биологии НАН РА

*А.Б. Киракосян* – д.б.н., проф., Массачусетский технологический институт (США)

**Физико-технические науки**

*Д.А. Фирсов* – д.ф.-м.н., проф., СПбПУ им. Петра Великого (Россия)

*А.В. Папоян* – член-корр. НАН РА, д.ф.-м.н., проф., Институт физических исследований НАН РА

*С.Г. Петросян* – член-корр. НАНРА, д.ф.-м.н., проф., Российско-Армянский университет (РАУ)

*А.А. Саркисян* – д.ф.-м.н., проф., Российско-Армянский университет (РАУ)

*Е.Ш. Мамасакхисов* – д.ф.-м.н., проф., Российско-Армянский университет (РАУ)

*В.Г. Аветисян* – д.т.н., проф., Российско-Армянский университет (РАУ)

Журнал основан в 2003 году и входит в перечень периодических изданий, зарегистрированных ВАК РА и РИНЦ

**Editor-in-Chief:** Academician of NAS RA, D.Sc. in Physics and Mathematics,  
Prof. **Kazaryan E.M.**

**Deputy Editor-in-chief:** Ph.D. in Mathematics, D. Sc. in Philosophy, Prof. **Avetisyan P.S.**

**Executive secretary:** Ph.D. in Science (Chemistry) **Shaginyan R.S.**

**Editorial team**

**Mathematics and Computer Science**

**R.G. Aramyan** – D. Sc. (Mathematics), Prof., Russian-Armenian University (RAU), Institute of Mathematics NAS RA

**D.G. Asatryan** – D. Sc. (Technical Sciences), Prof., Institute of Informatics and Automation Problems of NAS RA

**G.G. Kazaryan** – D. Sc. (Mathematics), Prof., Russian-Armenian University (RAU)

**O.V. Besov** – corresponding member of RAS, D. Sc. (Mathematics), Prof., Mathematical Institute Named after V.A. Steklov RAS (Russia)

**I.A. Burenkov** – D. Sc., Prof., Mathematical Institute Named after V.A. Steklov RAS (Russia)

**A.G. Sergeev** – Academician of the Russian Academy of Sciences, D. Sc. (Mathematics), Prof., Mathematical Institute Named after V.A. Steklov RAS (Russia)

**A.I. Avetisyan** – Academician of the Russian Academy of Sciences, D. Sc. (Mathematics), Prof., Institute of System Programming Named after V.P. Ivannikov RAS (Russia)

**V.Sh. Melikyan** – corresponding member of NAS RA, D. Sc. (Technical Sciences), Prof., Russian-Armenian University (RAU), Synopsys

**Biological and Chemical Sciences**

**A.A. Arakelyan** – D. Sc. (Biology), Prof., Institute of Molecular Biology NAS RA

**I.A. Muronets** – D. Sc. (Biology), Prof., Moscow State University (Russia)

**A.A. Hovhannisyan** – Ph.D. in Science (Biology), Russian-Armenian University (RAU)

**R.V. Zakharyan** – Ph.D. in Science (Biology), Russian-Armenian University (RAU)

**G.G. Danagulyan** – corresponding member of NASRA, D. Sc. (Chemistry), Prof., Russian-Armenian University (RAU)

**K.B. Nazaryan** – D.Sc. (Biology), Prof., Institute of Molecular Biology NAS RA

**L.M. Episkoposyan** – D.Sc. (Biology), Prof., Institute of Molecular Biology NAS RA

**A.B. Kirakosyan** – D.Sc. (Biology), Prof., Massachusetts Institute of Technology (USA)

**Physical and Technical Sciences**

**D.A. Firsov** – D. Sc. (Physics), Prof., SPbPU Named after Peter the Great (Russia)

**A.V. Papoyan** – corresponding member of NAS RA, D.Sc. (Physics), Prof., Institute of Physical Research NAS RA

**S.G. Petrosyan** – corresponding member of NAS RA, D. Sc. (Physics), Prof., Russian-Armenian University (RAU)

**A.A. Sargsyan** – D.Sc. (Physics), Prof., Russian-Armenian University (RAU)

**E.S. Mamasakhlisov** – D.Sc. (Physics), Prof., Russian-Armenian University (RAU)

**V.G. Avetisyan** – D. Sc. (Technical Sciences), Prof., Russian-Armenian University (RAU)

The journal founded in 2003 is included in the list of periodicals registered by HAC RA and RSCI

## СОДЕРЖАНИЕ

### МАТЕМАТИКА И ИНФОРМАТИКА

<b>Манукян А.</b> Обращение Мура-Пенроуза трехдиагональных косозермитовых матриц с нулевыми диагональными элементами .....	8
<b>Арамян Р.Г., Апинян Л.А.</b> Восстановление эллипса по распределениям хорд по направлениям .....	17
<b>Алтунян В.Н.</b> Сравнительный анализ инструментов для расчета скорректированного симметрией RMSD в молекулярном докинге .....	25
<b>Фаградян А., Багдасарян Т., Мелконян В., Саакян В., Киракосян Л., Оганесян О., Саргсян С., Дарбинян А.</b> Сравнение и улучшение методов обнаружения изменений .....	37

### БИОЛОГИЧЕСКИЕ НАУКИ

<b>Мелентьева Д.М., Зернов Н.А., Камарян В.С., Унанян Л.С., Попугаева Е.А.</b> Влияние модуляторов TRPC6 на актиновый цитоскелет <i>in vitro</i> .....	44
<b>Паронян А.К.</b> Сравнительный анализ программ предсказания структур белков на примере 14-3-3 .....	56
<b>Давитавян С.С.</b> Мульти-омикс исследование опухоли молочной железы .....	61
<b>Котлов Н.Ю.</b> Валидация прогностической силы подтипов микроокружения крупноклеточной лимфомы при лечении обинутозумабом в первой линии .....	67

# CONTENTS

## MATHEMATICS AND COMPUTER SCIENCE

<b>Manukyan A.</b> The Moore-Penrose Inverse of Tridiagonal Skew-Hermitian Matrices with Zero Diagonal Elements .....	8
<b>Aramyan R., Apinyan L.</b> Reconstruction of an Ellipse from the Distributions of Chords in Directions.....	17
<b>Altunyan V.</b> Comparative Analysis of Symmetry Corrected RMSD Calculation Tools in Molecular Docking .....	25
<b>Fahradyan A., Baghdasaryan T., Melkonyan V., Sahakyan V., Kirakosyan L., Hovhannisyanyan O., Sargsyan S., Darbinyan A.</b> Comparing and Improving Change Detection Methods .....	37

## BIOLOGICAL SCIENCES

<b>Melenteva D., Zernov N., Ghamaryan V., Hunanyan L., Popugaeva E.</b> Effect of TRPC6 Modulators on Actin Cytoskeleton in vitro .....	44
<b>Paronyan A.</b> A Comparative Analysis of Protein Structure Prediction Programs Using the Example of 14-3-3 .....	56
<b>Davitavyan S.</b> Multi-omics portrayal of breast cancers .....	61
<b>Kotlov N.Yu.</b> Lymphoma Microenvironment Subtypes Prognostic power Validation in frontline Obinutuzumab treated patients with Diffuse large B-cell Lymphoma.....	67

# МАТЕМАТИКА И ИНФОРМАТИКА

DOI 10.24412/1829-0450-2024-1-8-16  
УДК 519.6

Поступила: 21.03.2024г.  
Сдана на рецензию: 26.03.2024г.  
Подписана к печати: 04.04.2024г.

## THE MOORE-PENROSE INVERSE OF TRIDIAGONAL SKEW-HERMITIAN MATRICES WITH ZERO DIAGONAL ELEMENTS

*A. Manukyan*

*Yerevan State University  
avetiq.manukyan@ysu.am*

### ABSTRACT

The present paper is devoted to deriving closed form expressions for the Moore-Penrose inverse of tridiagonal skew-Hermitian matrices with zero diagonal elements. Using the formulas obtained, a numerical algorithm with optimal computational efforts is constructed.

**Keywords:** Moore-Penrose inverse, skew-Hermitian matrix, tridiagonal matrix.

**Introduction.** Let us denote by  $A^*$  the conjugate transpose of a matrix  $A$  with complex elements (see [1], for instance). The matrix  $A$  is said to be a *skew-Hermitian* if  $A^* = -A$ . Skew-Hermitian and, in particular, skew-symmetric matrices have many applications in various fields, such as statistical analysis, signal processing, computational mathematics, machine learning and others [2].

As follows from the definition, the diagonal elements of a skew-Hermitian matrix are either purely imaginary numbers or zeros. In this work we will discuss the case when all diagonal elements are zero. To find the Moore-Penrose inverse for tridiagonal matrices, we will use an approach developed in [3, 4] for skew-symmetric matrices. We are talking about generalized inversion, since tridiagonal skew-Hermitian matrices of odd order are singular (it is easy to verify this). Recall that for a  $m \times n$  matrix  $A$  the Moore-Penrose inverse  $A^+$  is the unique  $n \times m$  matrix that satisfies the following four conditions [1,5]:

$$AA^+A = A, A^+AA^+ = A^+, (A^+A)^* = A^+A, (AA^+)^* = AA^+. \quad (1)$$

If  $A$  is a square nonsingular matrix, then  $A^+ = A^{-1}$ . Thus the Moore-Penrose inverse generalizes ordinary matrix inversion. Note that the Moore-Penrose inverse of a skew-Hermitian matrix is also skew-Hermitian [1, 5].

Let us consider a tridiagonal matrix



$$A = \begin{bmatrix} 0 & a_1 & & & \\ -\bar{a}_1 & 0 & a_2 & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & -\bar{a}_{n-2} & 0 & a_{n-1} \\ & & & -\bar{a}_{n-1} & 0 \end{bmatrix}, \quad (2)$$

where  $n \geq 3$ . Note that throughout this article  $\bar{z}$  stands for the complex conjugate of the complex number  $z$ . We assume that  $a_i \neq 0$  for all  $i = 1, 2, \dots, n - 1$ . This requirement is not restrictive, since if some of the overdiagonal elements of  $A$  are equal to zero, the problem of computing the Moore-Penrose inverse is decomposed into several similar problems for matrices of lower order.

We will consider separately the matrices of even and odd orders.

**The matrix of even order.** Let  $n = 2m$ . Then from (2) we have

$$A = \begin{bmatrix} 0 & a_1 & & & \\ -\bar{a}_1 & 0 & a_2 & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & -\bar{a}_{2m-2} & 0 & a_{2m-1} \\ & & & -\bar{a}_{2m-1} & 0 \end{bmatrix}. \quad (3)$$

According to above assumption about the overdiagonal elements, this matrix is nonsingular. It follows from the easily proven equality

$$\det A = (|a_1| \cdot |a_3| \cdots |a_{2m-1}|)^2. \quad (4)$$

Note that, as follows from the formula (4), for the matrix  $A$  to be nonsingular, it is sufficient that only overdiagonal elements in the odd rows be nonzero.

Thus, in our case  $A^+ = A^{-1}$ . Finding the inverse matrix will be based on a special decomposition of the matrix  $A$ . To do this, we introduce into consideration bidiagonal matrix

$$B = \begin{bmatrix} -\bar{a}_1 & a_2 & & & \\ & -\bar{a}_3 & a_4 & & 0 \\ & & \ddots & \ddots & \\ 0 & & & -\bar{a}_{2m-3} & a_{2m-2} \\ & & & & -\bar{a}_{2m-1} \end{bmatrix} \quad (5)$$

of order  $m$ . Next, we define the following matrices:

$$P = [p_{ij}]_{m \times 2m}, \quad p_{ij} = \begin{cases} 1, & \text{if } j = 2i \\ 0, & \text{if } j \neq 2i \end{cases}, \quad i = 1, 2, \dots, m \quad (6)$$

and

$$Q = [q_{ij}]_{m \times 2m}, \quad q_{ij} = \begin{cases} 1, & \text{if } j = 2i - 1 \\ 0, & \text{if } j \neq 2i - 1 \end{cases}, \quad i = 1, 2, \dots, m. \quad (7)$$

Then the matrix  $A$  can be written as follows:

$$A = P^T B Q - (P^T B Q)^*. \quad (8)$$

The following easily verified properties of the matrices  $P$  and  $Q$  defined in (6) and (7) hold:

$$P P^T = I_m, \quad Q Q^T = I_m, \quad P Q^T = 0, \quad P^T P + Q^T Q = I_{2m} \quad (9)$$

(the subscript of the identity matrix indicates its order).

Proceeding from (8) and using properties (9), one can directly verify that

$$A^{-1} = Q^T B^{-1} P - (Q^T B^{-1} P)^*. \quad (10)$$

So, the problem of finding the inverse for the matrix  $A$  from (3) is reduced to the problem of inverting the matrix  $B$  defined in (5).

The inverse of nonsingular matrix  $B$  can be easily found.

**Lemma 1.** The elements of the matrix  $B^{-1} = [w_{ij}]_{m \times m}$  are as follows: for the indices  $j = 1, 2, \dots, m$  we have:

$$w_{ij} = -\frac{1}{a_{2j-1}} \prod_{s=i}^{j-1} r_s, \quad i = 1, 2, \dots, j-1; \quad w_{jj} = -\frac{1}{a_{2j-1}}; \quad (11)$$

$$w_{ij} = 0, \quad i = j+1, j+2, \dots, m,$$

where

$$r_s \equiv \frac{a_{2s}}{a_{2s-1}}, \quad s = 1, 2, \dots, m-1.$$

You can verify the validity of formulas (11) by directly checking.

Based on the formulas (11), the elements of the upper triangular matrix  $B^{-1}$  can be calculated by the following recurrence relations: for the indices  $j = 1, 2, \dots, m$  we find:

$$w_{jj} = -\frac{1}{a_{2j-1}}; \quad w_{ij} = r_i w_{i+1j}, \quad i = j-1, j-2, \dots, 1. \quad (12)$$

It can be readily seen that calculations by formulas (12) requires

$$A_{ops}^{(even)} = \frac{1}{2} m^2 + O(m)$$



of size  $m \times m + 1$ . Next, we define the following matrices:

$$F = [f_{ij}]_{m \times 2m+1}, \quad f_{ij} = \begin{cases} 1, & \text{if } j = 2i \\ 0, & \text{if } j \neq 2i \end{cases}, \quad i = 1, 2, \dots, m \quad (15)$$

and

$$G = [g_{ij}]_{m+1 \times 2m+1}, \quad g_{ij} = \begin{cases} 1, & \text{if } j = 2i - 1 \\ 0, & \text{if } j \neq 2i - 1 \end{cases}, \quad i = 1, 2, \dots, m + 1. \quad (16)$$

Then the matrix  $A$  can be written as follows:

$$A = F^T B G - (F^T B G)^*. \quad (17)$$

The following easily verified properties of the matrices  $F$  and  $G$  hold:

$$F F^T = I_m, \quad G G^T = I_{m+1}, \quad G F^T = 0, \quad F^T F + G^T G = I_{2m+1} \quad (18)$$

(the subscript of the identity matrix indicates its order).

Taking into account properties (18), it can be proved that

$$A^+ = G^T B^+ F - (G^T B^+ F)^*. \quad (19)$$

To do this, it is enough to show that the matrix (19) satisfies the conditions (1).

So, the problem of finding the Moore-Penrose inverse for matrix  $A$  from (13) is reduced to a similar problem for matrix  $B$  defined in (14). An approach to derive the Moore-Penrose inverse of the matrix  $B$  is based on the well-known formula

$$B^+ = \lim_{\varepsilon \rightarrow +0} (B^* B + \varepsilon I_{m+1})^{-1} B^* \quad (20)$$

(see [5], for instance). At first we find the inverse matrix  $(B^* B + \varepsilon I_{m+1})^{-1}$ . For this purpose, we use the computational procedure developed in the paper [6]. Then the elements of the matrix  $(B^* B + \varepsilon I_{m+1})^{-1} B^*$  are calculated and a character of their dependence on the parameter  $\varepsilon$  is revealed. Thereafter, according to the equality (20), passing to the limit when  $\varepsilon \rightarrow +0$ , we will arrive at closed form expressions for the elements of the matrix  $B^+$ . This sequence of actions, applicable to skew-symmetric matrices, is detailed in paper [4]. We only need to repeat all the arguments, taking into account the difference between skew-symmetric and skew-Hermitian matrices. As a result we get the following statement.

**Lemma 2.** Let  $B$  be a bidiagonal matrix given in (14) and  $a_i \neq 0$ ,  $1 \leq i \leq 2m$ . Then the elements of the Moore-Penrose inverse  $B^+ = [w_{ij}]_{m+1 \times m}$  are as follows: for  $1 \leq j \leq m$ ,

$$w_{ij} = - \frac{(\prod_{s=1}^{i-1} \frac{1}{r_s}) \sum_{k=j}^m (\prod_{s=j}^k \frac{1}{r_s}) (\prod_{s=k+1}^m \bar{r}_s)}{\bar{a}_{2j-1} \sum_{k=0}^m (\prod_{s=1}^k \frac{1}{r_s}) (\prod_{s=k+1}^m \bar{r}_s)}, \quad i = 1, 2, \dots, j \quad (21)$$

and

$$w_{ij} = \frac{(\prod_{s=i}^m r_s) \sum_{k=1}^j (\prod_{s=1}^{k-1} r_s) (\prod_{s=k}^{j-1} \frac{1}{r_s})}{\bar{a}_{2j-1} \sum_{k=0}^m (\prod_{s=1}^k \frac{1}{r_s}) (\prod_{s=k+1}^m r_s)}, \quad i = j + 1, \dots, m + 1, \quad (22)$$

where

$$r_s \equiv \frac{a_{2s}}{a_{2s-1}}, \quad s = 1, 2, \dots, m.$$

As in paper [4], based on the expressions (21) and (22), below we suggest a numerical procedure to compute the elements of the matrix  $B^+ = [w_{ij}]_{m+1 \times m}$ .

**Algorithm MPinverse  $B^+$**

1. Input elements  $a_1, a_2, \dots, a_{2m}$  of the matrix  $B$ .
2. Calculate the quantities  $r_s$ :

$$r_s \equiv a_{2s}/\bar{a}_{2s-1}, \quad s = 1, 2, \dots, m.$$

3. Calculate the quantities  $\tilde{\mu}_i$ :

$$\tilde{\mu}_{m+1} = 1; \quad \tilde{\mu}_i = r_i \tilde{\mu}_{i+1}, \quad i = m, m-1, \dots, 1.$$

4. Calculate the quantities  $\tilde{v}_i$ :

$$\tilde{v}_1 = 1; \quad \tilde{v}_{i+1} = \tilde{v}_i / r_i, \quad i = 1, 2, \dots, m.$$

5. Calculate the quantities  $R_i$ :

$$R_m = r_m / |a_{2m}|^2;$$

$$R_i = r_i (|a_{2i+1}|^2 R_{i+1} + \tilde{\mu}_{i+1}) / |a_{2i}|^2, \quad i = m-1, m-2, \dots, 1.$$

6. Calculate the quantities  $S_i$ :

$$S_1 = 1/|a_1|^2;$$

$$S_i = (|a_{2i-2}|^2 S_{i-1} + \tilde{v}_{i-1}) / |a_{2i-1}|^2 r_{i-1}, \quad i = 2, 3, \dots, m.$$

7. Calculate the quantity

$$\sigma = |a_1|^2 R_1 + \tilde{\mu}_1.$$

8. Calculate the elements of matrix matrix  $B^+$ :

$$w_{ij} = \begin{cases} -a_{2j-1} \tilde{v}_i \bar{R}_j / \sigma, & i = 1, 2, \dots, j \\ a_{2j-1} \tilde{\mu}_i \bar{S}_j / \sigma, & i = j + 1, j + 2, \dots, m + 1 \end{cases}; \quad j = 1, 2, \dots, m.$$

9. Output the matrix  $B^+ = [w_{ij}]_{m+1 \times m}$ .

**End Algorithm**

Direct calculations show that numerical implementation of the algorithm **MPinverseB<sup>+</sup>** requires  $m^2 + O(m)$  complex arithmetical operations. By this fact, the algorithm may be considered as an optimal one.

The elements of the matrix  $B^+$  obtained in Lemma 2 are, in the same time, the elements of the matrix  $A^+$  for skew-Hermitian matrix (13). Indeed, let us return to the equalities (17) and (19). We introduce the notation

$$G^T B^+ F \equiv V = [v_{\alpha\beta}]_{2m+1 \times 2m+1}.$$

According to formula (19),

$$A^+ = V - V^*. \quad (23)$$

We also introduce an intermediate notation

$$G^T B^+ \equiv U = [u_{\alpha\beta}]_{2m+1 \times m}.$$

Then  $V = UF$ . Using definitions (15) and (16) of the matrices  $F$  and  $G$ , respectively, consider the following four options for the arrangement of the elements in rows and columns of the matrix  $V$ .

- *Elements at the intersection of odd rows and odd columns of the matrix  $V$ .*

For  $i = 1, 2, \dots, m + 1$  and  $j = 1, 2, \dots, m + 1$ :

$$v_{2i-1 \ 2j-1} = \sum_{k=1}^m u_{2i-1 \ k} f_{k \ 2j-1} = 0. \quad (24)$$

- *Elements at the intersection of odd rows and even columns of the matrix  $V$ .*

For  $i = 1, 2, \dots, m + 1$  and  $j = 1, 2, \dots, m$ :

$$v_{2i-1 \ 2j} = \sum_{k=1}^m u_{2i-1 \ k} f_{k \ 2j} = u_{2i-1 \ j} = \sum_{k=1}^{m+1} g_{k \ 2i-1} w_{kj} = w_{ij}. \quad (25)$$

- *Elements at the intersection of even rows and odd columns of the matrix  $V$ .*

For  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, m + 1$ :

$$v_{2i \ 2j-1} = \sum_{k=1}^m u_{2i \ k} f_{k \ 2j-1} = 0. \quad (26)$$

- *Elements at the intersection of even rows and even columns of the matrix  $V$ .*

For  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, m$ :

$$v_{2i \ 2j} = \sum_{k=1}^m u_{2i \ k} f_{k \ 2j} = u_{2i \ j} = \sum_{k=1}^{m+1} g_{k \ 2i} w_{kj} = 0. \quad (27)$$

Having expressions (24)-(27) and proceeding from representation (23) of the matrix  $A^+$ , we arrive at the following statement.

**Theorem 2.** Let  $A$  be a skew-Hermitian matrix of order  $2m + 1$  given in (13). It is assumed that  $a_i \neq 0, 1 \leq i \leq 2m$ . The elements of the Moore-Penrose inverse  $A^+ = [z_{\alpha\beta}]_{2m+1 \times 2m+1}$  are as follows:

$$z_{2i-1, 2j} = w_{ij}, \quad z_{2j, 2i-1} = -\overline{w_{ij}}, \quad i = 1, 2, \dots, m + 1, \quad j = 1, 2, \dots, m,$$

where  $w_{ij}$  are the elements of the  $B^+$  obtained in Lemma 2. The remaining elements of the matrix  $A^+$  are equal to zero.

As in the case of even matrix order, below we give the layout of the elements of matrix  $A^+$  for the value  $m = 3$ :

$$A^+ = \begin{bmatrix} 0 & w_{11} & 0 & w_{12} & 0 & w_{13} & 0 \\ -\overline{w_{11}} & 0 & -\overline{w_{21}} & 0 & -\overline{w_{31}} & 0 & -\overline{w_{41}} \\ 0 & w_{21} & 0 & w_{22} & 0 & w_{23} & 0 \\ -\overline{w_{12}} & 0 & -\overline{w_{22}} & 0 & -\overline{w_{32}} & 0 & -\overline{w_{42}} \\ 0 & w_{31} & 0 & w_{32} & 0 & w_{33} & 0 \\ -\overline{w_{13}} & 0 & -\overline{w_{23}} & 0 & -\overline{w_{33}} & 0 & -\overline{w_{43}} \\ 0 & w_{41} & 0 & w_{42} & 0 & w_{43} & 0 \end{bmatrix}.$$

To summarize, in conclusion we emphasize two main results of this work. First, we obtained closed form expressions for the elements of the Moore-Penrose inverse of tridiagonal skew-Hermitian matrices with zero diagonal. Secondly, on the basis of the obtained formulas, a numerical algorithm which is optimal in terms of computational costs was constructed.

### REFERENCES

1. Horn R. and Ch. Johnson. Matrix Analysis, Cambridge University Press, 1986.
2. Balonin N.A. and Sergeev M.B. Special matrices: pseudoinverse, orthogonal, Hadamard and Cretan, Polytechnika Publ., Saint-Petersburg, 2019 (in Russian).
3. Hakopian Yu.R., Manukyan A.H. and Mikaelyan H.V. The Moore-Penrose inverse of tridiagonal skew-symmetric matrices. I, Proceedings of the Yerevan State University, 57 (1) (2023), 1–8.
4. Hakopian Yu.R., Manukyan A.H. and Mikaelyan H.V. The Moore-Penrose inverse of tridiagonal skew-symmetric matrices. II, Proceedings of the Yerevan State University, 57 (2) (2023), 31–43.
5. Ben-Israel A. and T.N.E. Greville. Generalized Inverses: Theory and Applications, Springer, New-York, 2003.
6. Hakopian Yu.R. and Manukyan A.H. Analytical inversion of tridiagonal Hermitian matrices, Mathematical Problems of Computer Science, 58 (2022), 7–19.

**ОБРАЩЕНИЕ МУРА-ПЕНРОУЗА ТРЕХДИАГОНАЛЬНЫХ  
КОСОЭРМИТОВЫХ МАТРИЦ С НУЛЕВЫМИ ДИАГОНАЛЬ-  
НЫМИ ЭЛЕМЕНТАМИ**

*А. Манукян*

*Ереванский государственный университет*

**АННОТАЦИЯ**

Настоящая статья посвящена выводу явных выражений для элементов обратной матрицы Мура-Пенроуза в случае трехдиагональных косоэрмитовых матриц с нулевыми диагональными элементами. На основе полученных формул получен оптимальный, в смысле объема вычислительных затрат, численный алгоритм.

**Ключевые слова:** обращение Мура-Пенроуза, косоэрмитовая матрица, трехдиагональная матрица.



DOI 10.24412/1829-0450-2024-1-17-24  
УДК 519.218.5

Поступила: 20.02.2024г.  
Сдана на рецензию: 26.02.2024г.  
Подписана к печати: 19.05.2024г.

## ВОССТАНОВЛЕНИЕ ЭЛЛИПСА ПО РАСПРЕДЕЛЕНИЯМ ХОРД ПО НАПРАВЛЕНИЯМ

*Р.Г. Арамян, Л.А. Апинян*

*Российско-Армянский (Славянский) университет  
apinlev00@gmail.com, rafikaramyan@yahoo.com*

### АННОТАЦИЯ

Задача нахождения информации или полного восстановления выпуклых тел по распределениям характеристик  $k$ -мерных сечений – одна из основных задач стохастической томографии. Данная работа посвящена восстановлению эллипса на плоскости через распределения случайных хорд по направлениям. Предлагается алгоритм восстановления.

**Ключевые слова:** эллипс, выпуклость, функция распределения, опорная функция, восстановление эллипса.

### 1. Введение

Геометрические характеристики выпуклого тела  $K$ , инвариантные относительно Евклидовых движений, могут быть вычислены статистическими методами. Решение таких задач выходит далеко за рамки теоретического интереса.

Эта область была значительно развита в течение последних десятилетий, когда требовалось все большее количество реальных приложений. В частности, восстановление выпуклого тела по его случайным сечениям является центральной задачей стохастической томографии.

Восстановление выпуклых тел (компактных множеств) является главной задачей геометрической томографии. Геометрическая томография – это математическая область, которая фокусируется на проблемах восстановления однородных (часто выпуклых) объектов из томографических данных (это могут быть рентгеновские лучи, проекции, сечения, функции яркости или ковариограммы). По словам Ричарда Гарднера, «Геометрическая томография посвящена поиску информации о геометрическом объекте из данных о его проекциях (тенях) на плоскостях или сечениях плоскостями». Геометрическая томография в рамках математики в основном связана с выпуклой геометрией, интегральной геометрией, геометрией Минковского, функциональным ана-

лизом и комбинаторикой, а также вне математики, с компьютерной томографией, медицинской томографией, дискретной томографией, электронной микроскопией, компьютерным зрением и распознаванием образов.

Стохастическая томография посвящена поиску информации о геометрическом объекте из случайных данных о его проекциях (тенях) на плоскостях или распределения случайных сечений  $k$ -мерными плоскостями.

### **Постановка задачи**

Задача состоит в восстановлении эллипса с помощью функции распределения длин хорд по направлениям. В данной работе мы определим несколько геометрических понятий и рассмотрим возможность восстановления эллипса с помощью функции распределения длин хорд по направлению.

Проблему восстановления выпуклых тел по распределению длин случайных хорд впервые рассмотрел Вильгельм Бляшке [1]. Пусть задана выпуклая фигура. Случайная прямая (с инвариантной мерой), пересекающая фигуру, задает случайную хорду  $\chi(g)$ . В. Бляшке поставил задачу: определяется ли выпуклое тело по распределению длины случайных хорд с точностью до параллельных переносов и отражений? Позже Колин Л. Маллоус и Джон Морис Кларк опровергали вышесказанное [2], построив два неконгруэнтных 12-угольника с одинаковыми распределениями длин хорд.

После была поставлена вторая гипотеза: восстанавливается ли выпуклая фигура по распределению длин случайных хорд по направлению с точностью до параллельных переносов и отражений? Также эту проблему рассматривал Жорж Матерон. В 1975г. он ввел понятие ковариограммы, показал связь между ней и функцией распределения длины случайной хорды по направлению [6], а в 1986г. предложил гипотезу [4]: ковариограмма однозначно определяет выпуклое тело в  $\mathbb{R}^n$ , с точностью до параллельных переносов и отражений. В 1993г. Вернер Нагель доказал эту гипотезу для всех выпуклых многоугольников [3]. Матерон также предложил положительный ответ при  $n = 2$ , который был доказан в 2007 г. Аверковым и Г. Бьянки [4]. Габриэль Бьянки построил контрпримеры в  $\mathbb{R}^n$  для каждого  $n \geq 4$  и показал, что задача ковариограммы в общей постановке имеет отрицательный ответ [5]. Для  $n = 3$  проблема пока не решена. Но несмотря на то, что задача имеет положительный ответ при  $n = 2$ , алгоритм восстановления выпуклых фигур с помощью распределения длин хорд по направлению до сих пор остается открытым.

Цель данной статьи – найти какие-то характеристики эллипса от его распределений хорд и выяснить способ восстановления эллипса через его распределения длин хорд по направлениям.

## 2. Предварительные понятия

### 2.1. Координаты прямой на плоскости

Пространство прямых на плоскости  $\mathbb{R}^2$  будем обозначать через  $G$ . Прямую  $g \in G$  параметризуем парой координат  $(p, \phi)$ ,  $p \geq 0$ ,  $0 \leq \phi < 2\pi$ , то есть полярными координатами основания перпендикуляра опущенного на прямую  $g$  из начала координат  $O$ . Введем в  $G$  инвариантную относительно евклидовых движений, меру  $dg = dpd\phi$ .

### 2.2. Распределение длины хорды по направлению

Пусть  $D$  – компактное множество в Евклидовом пространстве  $\mathbb{R}^2$ .  $G$  пространство прямых в  $\mathbb{R}^2$ ,  $g \in G$ ,  $(p, \phi)$  – полярные координаты прямой  $g$ , где  $p \geq 0$ ,  $\phi \in S^1$ ,  $S^1$  – единичная окружность на плоскости с центром в начале координат. Прямая  $g$  пересекает множество  $D$  и образует хорду, которую обозначим  $\chi(g)$  (Рис. 1).

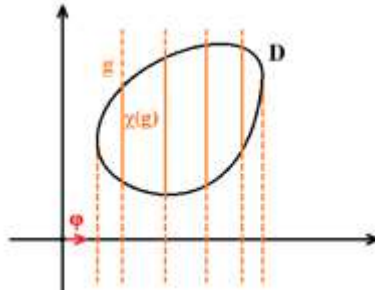


Рисунок 1.

Зафиксируем направление  $\phi$ . Рассмотрим все прямые в этом направлении. Каждая такая прямая однозначно определяется ее пересечением с осью  $Ox$ .

Функция распределения длины хорды по направлению – это

$$F_\phi(t) = P\{|\chi(g_\phi)| < t : \chi(g_\phi) = g_\phi \cap D\},$$

где  $g_\phi$  имеет направление  $\phi$ .

## 3. Основные результаты

Здесь рассмотрен эллипс на  $\mathbb{R}^2$  и выявлена его функция распределения длин хорд по направлению, описана опорная функция и представлен способ восстановления.

### 3.1. Опорная функция эллипса

Опорная функция тела  $K$  – это  $\text{Sup}_{x \in K} \langle \vec{\varphi}, \vec{x} \rangle \equiv h(\varphi)$ , где  $\vec{\varphi}$  есть единичный вектор в направлении  $\varphi$ . Известно, что опорная функция выпуклого тела  $K$  однозначно задает  $K$ . Пусть  $K$  эллипс с полуосями  $a$  и  $b$ :

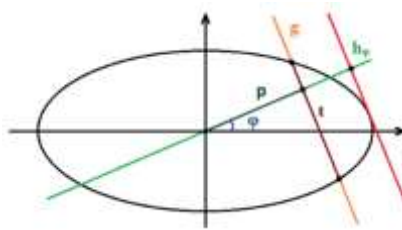
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Известно, что опорная функция эллипса имеет следующий вид:

$$h(\varphi) = \sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi}$$

### 3.2. Функция распределения длин хорд для эллипса по направлению

Для эллипса  $K$  обозначим через  $t$  длину хорды  $(A, B)$  отсекаемой произвольной прямой  $g$  брошенной по направлению  $\varphi$ , обозначим через  $p$  длину перпендикуляра, проведенного от центра эллипса до хорды  $t$ . Найдем функцию распределения длины хорды  $K$  по направлению  $\varphi$ , то есть  $F_\varphi(t)$ .



Из определения функции распределения имеем:

$$F_\varphi(t) = \frac{h(\varphi) - p}{h(\varphi)}$$

Чтобы найти координаты  $A = (x_1, y_1)$  и  $B = (x_2, y_2)$ , рассмотрим следующую систему:

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \\ x \cos \varphi + y \sin \varphi = p \end{cases}$$

Заметим, что

$$t = |AB| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Подставляя выражение для  $x$  из второго уравнения системы в первое, получим

$$\frac{\left(\frac{p - y \sin \varphi}{\cos \varphi}\right)^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Откуда получим

$$(y_2 - y_1)^2 = \frac{4p^2 b^4 \sin^2 \varphi}{h^4(\varphi)} - \frac{4b^2 p^2 - 4b^2 a^2 \cos^2 \varphi}{h^2(\varphi)}$$

Аналогичными образом получим

$$(x_2 - x_1)^2 = \frac{4p^2 a^4 \cos^2 \varphi}{h^4(\varphi)} - \frac{4a^2 p^2 - 4b^2 a^2 \sin^2 \varphi}{h^2(\varphi)}$$

Так как

$$t^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 = \frac{4p^2}{h^2(\varphi)} \left( \frac{a^4 \cos^2 \varphi - b^4 \sin^2 \varphi}{h^2(\varphi)} - a^2 - b^2 \right) + \frac{4a^2 b^2}{h^2(\varphi)}$$

Окончательно получается

$$\left( t^2 - \frac{4a^2 b^2}{h^2(\varphi)} \right) = - \frac{4p^2 a^2 b^2}{h^4(\varphi)}$$

Отсюда и получается выражение для p

$$p = h(\varphi)^2 \sqrt{\frac{1}{h(\varphi)^2} - \frac{t^2}{4a^2 b^2}}$$

Подставив его в  $F_\varphi(t)$  и сделав соответствующие упрощения, был получен окончательный вид для функции распределения длин хорд по направлению:

$$F_\varphi(t) = 1 - \sqrt{1 - \frac{t^2 h(\varphi)^2}{4a^2 b^2}}$$

Отсюда можно выразить  $h(\varphi)$  через  $F_\varphi(t)$

$$h(\varphi) = \frac{2ab}{t} \sqrt{1 - (1 - F_\varphi(t))^2}$$

Получается следующая лемма:

**Лемма.** Для любого эллипса и для любого направления  $\varphi$  выражение

$$\frac{2S}{\pi t} \sqrt{1 - (1 - F_\varphi(t))^2}$$

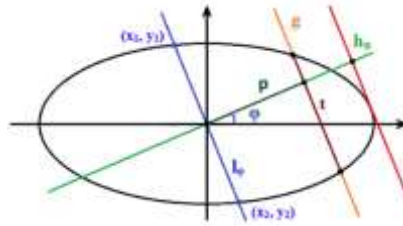
не зависит от  $t$ .

Обратное утверждение мы представляем как гипотезу:

**Гипотеза.** Если для некой фигуры и любого направления  $\varphi$  выражение  $\frac{2S}{\pi t} \sqrt{1 - (1 - F_\varphi(t))^2}$  не зависит от  $t$ , то эта фигура является эллипсом.

### 3.3. Восстановление неповернутого эллипса

Мы предлагаем следующий алгоритм для восстановления неповернутого эллипса



Для восстановления необходимы 2 самые длинные хорды в двух не коллинеарных направлениях. Самую длинную хорду в направлении  $\varphi$  обозначим  $l_\varphi$ . Найти ее можно из данной системы:

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \\ x \cos \varphi + y \sin \varphi = 0 \end{cases}$$

$$l_\varphi = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$y_{1,2} = \pm \frac{abc \cos \varphi}{\sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi}}$$

$$x_{2,1} = \pm \frac{abs \sin \varphi}{\sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi}}$$

Откуда получаем, что

$$l_\varphi = \frac{2ab}{\sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi}}$$

Остаётся решить следующую систему выразив  $a$  и  $b$  через эти хорды и их направления

$$\begin{cases} l_{\varphi_1} = \frac{2ab}{\sqrt{a^2 \cos^2 \varphi_1 + b^2 \sin^2 \varphi_1}} \\ l_{\varphi_2} = \frac{2ab}{\sqrt{a^2 \cos^2 \varphi_2 + b^2 \sin^2 \varphi_2}} \end{cases}$$

Окончательное выражение для полуосей выглядят так:

$$a = \frac{l_{\varphi_1} l_{\varphi_2} \cos \varphi_1}{2} \cdot \sqrt{\frac{\sin^2 \varphi_2 - \operatorname{tg}^2 \varphi_1}{l_{\varphi_1}^2 \cos^2 \varphi_1 - l_{\varphi_2}^2 \cos^2 \varphi_2}}$$

$$b = \frac{l_{\varphi_1} l_{\varphi_2} \sin \varphi_1}{2} \cdot \sqrt{\frac{\cos^2 \varphi_2 - \operatorname{ctg}^2 \varphi_1}{l_{\varphi_1}^2 \sin^2 \varphi_1 - l_{\varphi_2}^2 \sin^2 \varphi_2}}$$

### 3.4. Восстановление повернутого эллипса

Мы предлагаем следующий алгоритм для восстановления эллипса повернутого на угол  $\psi$ . Аналогичным образом составляется следующая система:

$$\begin{cases} l_{\varphi_1} = \frac{2ab}{\sqrt{a^2 \cos^2(\varphi_1 - \psi) + b^2 \sin^2(\varphi_1 - \psi)}} \\ l_{\varphi_2} = \frac{2ab}{\sqrt{a^2 \cos^2(\varphi_2 - \psi) + b^2 \sin^2(\varphi_2 - \psi)}} \\ l_{\varphi_3} = \frac{2ab}{\sqrt{a^2 \cos^2(\varphi_3 - \psi) + b^2 \sin^2(\varphi_3 - \psi)}} \end{cases}$$

Решив систему аналогичным образом, мы получим  $a$ ,  $b$  и  $\psi$  в терминах  $\varphi_1, \varphi_2, \varphi_3, l_{\varphi_1}, l_{\varphi_2}, l_{\varphi_3}$ . Получаются следующие выражения для параметров эллипса:

$$\psi = \frac{\arctg\left(\frac{C_2}{C_1 - C_3}\right)}{2}$$

$$a = \sqrt{\frac{2}{C_1 + C_3 - \sqrt{(C_1 - C_3)^2 + C_2^2}}}$$

$$b = \sqrt{\frac{2}{C_1 + C_3 + \sqrt{(C_1 - C_3)^2 + C_2^2}}},$$

где

$$C_1 = \frac{4}{l_{\pi}^2},$$

$$C_3 = \frac{4}{l_0^2},$$

$$C_2 = C_1 + C_3 - \frac{8}{l_{\pi}^2}$$

### Заключение

Таким образом мы показали, способ восстановления произвольного эллипса (повернутого) по его распределению случайных хорд по трем направлениям. Вопрос восстановления произвольной выпуклой фигуры по распределению случайных хорд по направлениям открыт

### ЛИТЕРАТУРА

1. *Luis A. Santalo*. Integral Geometry and Geometric Probability, Addison-Wesley, Reading, MA (2004).
2. *Mallows C. L., Clark J. M.C.* Linear-intercept distributions do not characterize plane sets. J. Appl. Probab. 7, 240–244 (1970).
3. *Nagel W.* Orientation-dependent chord length distributions characterize convex polygons. J. Appl. Probab. 30, 730–736 (1993).

4. *Averkov G., Bianchi G.* Confirmation of Matheron's conjecture on the covariogram of a planar convex body, *J. Eur. Math. Soc.*, in press, arXiv:0711.0572 [math.MG].
5. *Bianchi G.* The covariogram determines three-dimensional convex polytopes. *Advances in Mathematics* 220 (2009), 1771–1808; arXiv:0805.1605v1 [math.MG].
6. *Matheron G.* *Random Sets and Integral Geometry.* Willey, New York (1975).

## RECONSTRUCTION OF AN ELLIPSE FROM THE DISTRIBUTIONS OF CHORDS IN DIRECTIONS

***R. Aramyan, L. Apinyan***

*Russian-Armenian(Siavonic) University*

### ABSTRACT

The problem of finding information or completely reconstructing convex bodies from the distributions of characteristics of  $k$ -dimensional sections is one of the main problems of stochastic tomography. This work is devoted to the reconstruction of an ellipse on a plane through the distribution of random chords in directions. A reconstruction algorithm is proposed.

**Keywords:** ellipse, convexity, distribution function, support function, ellipse restoration.



DOI 10.24412/1829-0450-2024-1-25-36  
УДК 004.052.2

Поступила: 13.05.2024г.  
на рецензию: 13.05.2024г.  
Подписана к печати: 20.05.2024г.

## COMPARATIVE ANALYSIS OF SYMMETRY CORRECTED RMSD CALCULATION TOOLS IN MOLECULAR DOCKING

*V. Altunyan*

*Yerevan State University  
altunyanv@gmail.com  
ORCID: 0009-0002-0021-3789*

### ABSTRACT

The accurate calculation of RMSD, especially with symmetry correction, is vital for reliable molecular comparisons in computational chemistry. This study conducts a comparative analysis of three advanced RMSD calculation tools to assess their accuracy, failure rates, and computational efficiency across varied molecular scenarios. Utilizing two large datasets, each comprising approximately 2000 molecular structures, the tools are evaluated in the context of molecular docking simulations. These datasets are sourced from well-established molecular databases and include a range of complexities to challenge the tools under test. The research extends to include special molecular cases, designed to expose strengths and weaknesses of each tool, providing a comprehensive overview of their performance. This analysis aims to guide computational chemists in selecting the most appropriate RMSD calculation tools for their research needs, enhancing the accuracy and reliability of molecular modeling studies.

**Keywords:** symmetry corrected RMSD, molecular docking.

### Introduction

Root Mean Square Deviation (RMSD) is a widely used metric in computational chemistry for quantifying the similarity between molecular structures, particularly in the context of molecular docking and structural biology. The precision of RMSD calculations directly influences the reliability of molecular modeling studies, affecting interpretations and decisions in drug design and other applications. However, RMSD calculations can be particularly challenging when dealing with symmetric molecules, as traditional RMSD often fails to account for intrinsic molecular symmetries, potentially leading to misleading results.

Recent advancements have led to the development of symmetry corrected RMSD calculation tools that address these issues by considering molecular symmetries in their algorithms. This paper focuses on tools which use graph isomorphism for finding symmetries, specifically: *DockRMSD* [1], *spyRMSD*[2] and

*obrms* (from OpenBabel [3]), which represent significant improvements in the accuracy and efficiency of RMSD calculations.

Previous studies have begun to explore comparisons between these RMSD calculation tools, focusing on their performance in terms of accuracy, speed, and reliability. Notably, references [1, 2] present comparisons between DockRMSD (with generalized bonds) and *obrms*, as well as between *spyRMSD* and *obrms*, evaluating their equivalence in results and computational speed. These comparisons showed that all three tools yielded the same RMSDs, indicating a high level of consistency. However, these studies did not discuss instances where the tools fail or produce inaccurate results. This omission highlights a gap in the current understanding of these tools reliability across more diverse or challenging molecular scenarios.

To address this gap, this study employs two large datasets, each comprising approximately 2000 molecular structures. The first dataset is randomly sampled from the refined set of PDBBind [4], widely recognized for its relevance in protein-ligand binding studies. The second dataset is composed of randomly selected complexes fetched from the Protein Data Bank (PDB)<sup>1</sup>, including complexes that present challenging ligands, which are known to test the limits of computational tools in terms of accuracy and efficiency. These datasets provide a robust basis for evaluating the performance of each RMSD tool under standardized as well as challenging conditions. Additionally, special molecular cases are introduced to test each tool's ability to handle complex scenarios that challenge their computational and algorithmic limits.

By systematically comparing these tools across multiple datasets and specialized cases, this study aims to elucidate the strengths and limitations of each RMSD calculation method. The insights gained from this research will not only enhance the understanding of current RMSD calculation capabilities but also aid in the continuous improvement and development of these essential tools in computational chemistry.

### **RMSD Calculation Tools**

*DockRMSD* [1] is an open-source tool specifically engineered to calculate RMSD for symmetric molecules. It utilizes graph isomorphism to ensure accurate atom mapping, addressing the common challenge in traditional RMSD calculations where symmetrical structures are misaligned due to arbitrary atom ordering. *DockRMSD* applies graph isomorphism to ensure that each atom is precisely paired with its corresponding counterpart, even in complex symmetrical molecules. Distinctive features of this tool include its implementation in C, which makes it excep-

---

<sup>1</sup> Protein Data Bank (rcsb.org)

tionally lightweight and free from dependencies. Additionally, *DockRMSD* considers bond types too, and only if no isomorphism is found, it assumes that all bonds are of the same type.

*SpyRMSD* [2] is a Python-based tool designed to perform symmetry corrected RMSD calculations, thereby enhancing the accuracy of molecular similarity assessments. While the tool also supports translational and rotational symmetries, our focus will primarily be on its capability regarding graph isomorphism. Distinctive features of *SpyRMSD* include its implementation in Python, which allows for seamless integration into existing computational workflows<sup>2</sup>. Additionally, the tool is noted for its reliability, as it leverages well-known libraries such as *NetworkX* [5] to manage graph isomorphisms effectively.

*Obrms* is part of the OpenBabel [3] suite, a powerful tool for chemical data handling, which includes functionality for calculating RMSD. *Obrms* leverages OpenBabel's extensive capabilities in format conversion and molecular manipulation to facilitate RMSD calculations even for complex and large molecular systems. While *obrms* is primarily written in C++ for high performance, it is also accessible via a Python wrapper provided by the OpenBabel suite, enhancing its versatility. For the purposes of this study, however, we will utilize the standalone executable version of the tool.

## Evaluation Methodology

This study evaluates several key aspects of RMSD calculation tools: accuracy, failure rate, and computational efficiency. The comparative analysis of RMSD calculation tools in this study is based on two selected datasets: *PDBBind Refined* and *PDB Random*, each comprising approximately 2000 molecular structures. These datasets were chosen to provide a broad spectrum of molecular complexities and scenarios, ensuring a comprehensive evaluation of the RMSD calculation tools involved.

*PDBBind Refined* dataset is derived from the PDBBind database's refined set. PDBBind provides a curated collection of protein-ligand complexes with experimentally measured binding affinities, widely utilized in the validation and benchmarking of molecular docking algorithms. The refined set specifically contains high-quality entries that have been rigorously validated for structural and resolution integrity. It also provides carefully processed molecular structure files. In case of uncertainty in bond types, bonds are marked as aromatic in MOL2 files, so only in uncertain cases *DockRMSD* will proceed with generalized bond types (*Figure 1*). On the left side the crystal structure (with N1 and N2 atoms) is PDBBind processed molecule, and on the right side the crystal structure (with N1 and N2 atoms) is

---

<sup>2</sup> <https://github.com/RMeli/spyrmsd/tree/develop>

generated from complex PDB file and ligand SMILES. There is also a generated conformation in both pictures. In first case DockRMSD will generalize bonds and produce RMSD of 3.0293143 Å, which is correct thing to do in this case, while in the second case it'll find atom mapping N1□N2□ and N2□N1□ which will cause RMSD of 3.1660216 Å.

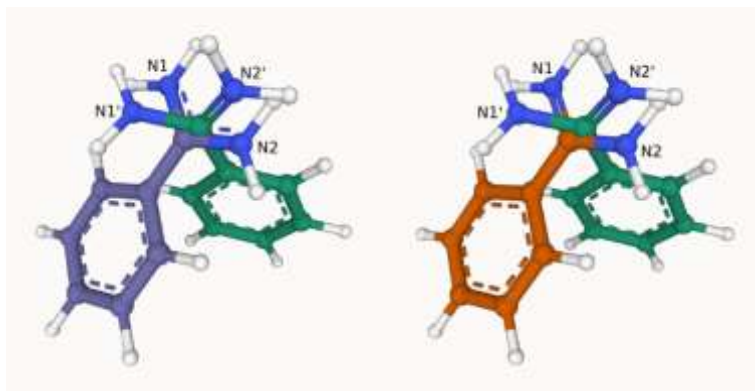


Figure 1. BENZAMIDINE molecule.

*PDB Random* dataset consists of randomly selected complexes from the Protein Data Bank, an expansive repository of three-dimensional structural data of biological macromolecules. This dataset is enriched with complexes that include challenging ligands, known for testing the robustness and adaptability of computational tools.

### Dataset preparation

For the first dataset, we randomly selected 2000 complexes from a pool of approximately 5000 available complexes and conducted molecular docking using SMINA software [6]. This procedure was designed to generate a variety of distinct conformations for each complex. In the case of the second dataset, we created a pool of roughly 10,000 complexes using their 4-letter PDB IDs and subsequently extracted the small molecules found within these complexes. A filtering process was applied to isolate protein-ligand pairs, ensuring that each pair featured unique PDB IDs for both the complex and the ligand. This selection strategy primarily aims to ensure the uniqueness of ligands in the final dataset. While two complexes with different PDB IDs may share the same protein, the uniqueness of ligand PDB IDs guarantees that the ligands are distinct. Ultimately, we selected 2000 such complexes and performed molecular docking in a similar manner to the first dataset, aiming to generate the necessary conformations.

Molecular docking using SMINA resulted in SDF files including multiple conformations of the same molecule, which were converted into MOL2 format and written into separate files using OpenBabel toolbox. Details about the final datasets can be found in *Table 1*. Additionally, these datasets will be made publicly available to facilitate further research and validation by other researchers.

*Table 1. Summary of generated datasets*

Name	Initial complexes	Successfully docked	Total conformations
<i>PDBBind Refined</i>	2000	1996	12574
<i>PDB Random</i>	2000	1983	16422

### **Evaluation of RMSD Tools**

As previously discussed, this study evaluates several key aspects of RMSD calculation tools: accuracy, failure rate, and computational efficiency. Given that RMSD calculations are typically used to evaluate docking results involving around 10 conformations per complex, it is crucial that these calculations are performed quickly. To address this, we have implemented a 10-second timeout for each RMSD calculation task in our study, establishing this cutoff to ensure timely processing and to highlight cases where the duration may be problematic. If a tool exceeds this time limit, it is classified as a failure. However, it is important to distinguish between failures due to time constraints and those resulting from algorithmic shortcomings. This differentiation will allow us to more accurately assess the practical limitations and strengths of each tool under test.

Conformations were processed in terms of RMSD calculation in parallel. The following versions of software were used:

- DockRMSD – built from the source, available in website<sup>3</sup>.
- spyRMSD – v0.7.0.
- obrms – OpenBabel v3.1.1.

During processing of each complex, sequentially DockRMSD, spyRMSD and obrms were run on all the docking generated conformations of given complex, then results were finalized, and processing of a single complex was completed. Calculations were run on a machine with AMD EPYC 9654 96-core processor, using 1/3 of its processors, while no other jobs were running.

---

<sup>3</sup> <https://zhanggroup.org/DockRMSD/>

## Results

### Failure rate

By failure we mean a job which doesn't give any numeric output within the timeout period. This may be caused of very long running time or runtime errors of tool in case of which program just aborts. As illustrated in *Figure 2*, DockRMSD exhibits the highest failure rates, with approximately 5% in the *PDB Random* dataset and around 30% in the *PDBBind Refined* dataset. While spyRMSD also encountered several failures, all were due to timeout issues. In contrast, failures in DockRMSD were caused by segmentation faults in all but two complexes from the PDB Random dataset. Those two complexes (comprising a total of 15 conformations) are similar to each other and represent special cases where DockRMSD struggles, often operating for hours before successfully identifying the optimal isomorphism.

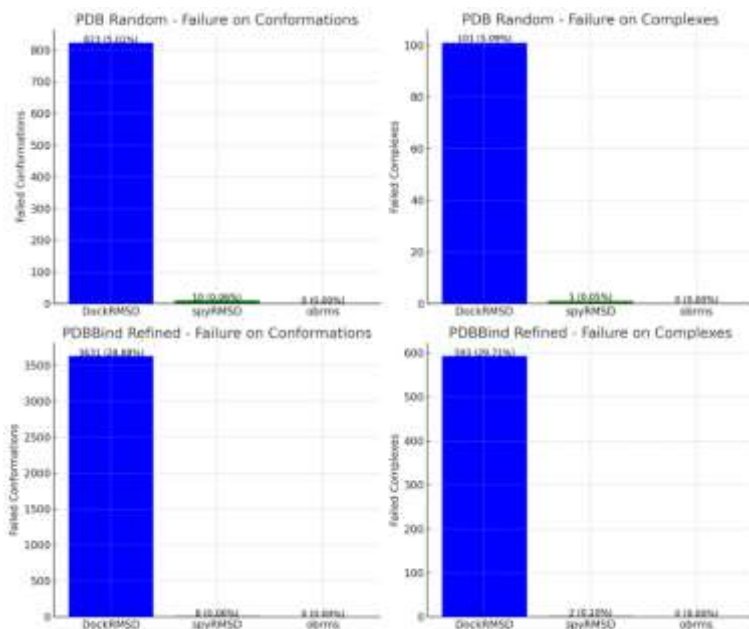


Figure 2. Failure rate comparison.

### Accuracy

In this section, we focus on examining the agreement between various RMSD calculation tools. This assessment is particularly crucial as there currently exists no dataset with symmetry corrected RMSDs calculated in an established, accurate manner. By comparing the outputs of these tools, particularly using obrms as a ref-

erence standard, we aim to identify the degree of consistency across different implementations of core algorithms. As noted in the beginning DockRMSDs approach slightly differs from the other two. For comparison of core algorithms of our tools, a modification is made to DockRMSDs approach to match those of spyRMSD and obrms by generalizing the bond types of input molecules immediately.

Upon analysis, it was found that all tools consistently agreed on the outputs for all successfully processed conformations across both datasets. The mean absolute error (MAE) between spyRMSD and obrms was approximately  $10^{-5}\text{\AA}$ , highlighting a high degree of precision. In contrast, DockRMSD displayed a slightly higher MAE of approximately  $5\times 10^{-5}\text{\AA}$ . This variation is attributed to DockRMSD's rounding of output values to the third decimal place.

These results confirm that the core algorithms of the tools are fundamentally aligned, with no significant algorithmic discrepancies noted in DockRMSD apart from the consideration of bond types. To further investigate, we will analyze the performance of DockRMSD without any modifications to its algorithm, specifically focusing on cases where the results deviate significantly from the established reference standard.

We begin with the *PDBBind Refined* dataset, known for addressing issues related to ambiguous bond types in MOL2 files. Our focus will be on complexes with docked conformations where the DockRMSD results deviate from the obrms reference by more than 0.1 Å. Notable examples from this dataset include complexes **2FLE**, **3A2O**, **6CSP** and **1EBW**. An in-depth examination of two such instances will be provided.

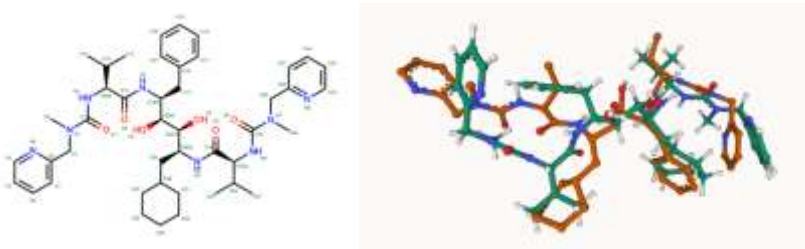


Figure 3. Ligand with PDB ID **AI** and docked position for it in complex **2FLE**.

Figure 3 illustrates the ligand of complex **2FLE**, providing both 2D visualization and 3D visualization along with some docked pose. The 2D representation clearly indicates that the molecule is not symmetric. However, if we were to hypothetically consider the lower ring as aromatic, the molecule could be perceived as symmetric. This is relevant because the docking algorithm almost perfectly aligned the two pairs of rings on the right side of the molecule, as shown in the figure.

Despite this near alignment, the docking results reveal a critical issue. One of the pairs consists of an aromatic ring and a non-aromatic ring and considering them





results. Such insights are crucial for refining docking algorithms and improving the precision of molecular simulations.

In the context of the *PDB Random* dataset, it's observed that there are over 200 complexes where the results from DockRMSD significantly differ from the reference. This variance is largely attributed to the less processed nature of the complexes within this dataset compared to those in the *PDBBind Refined* dataset.

### Runtime

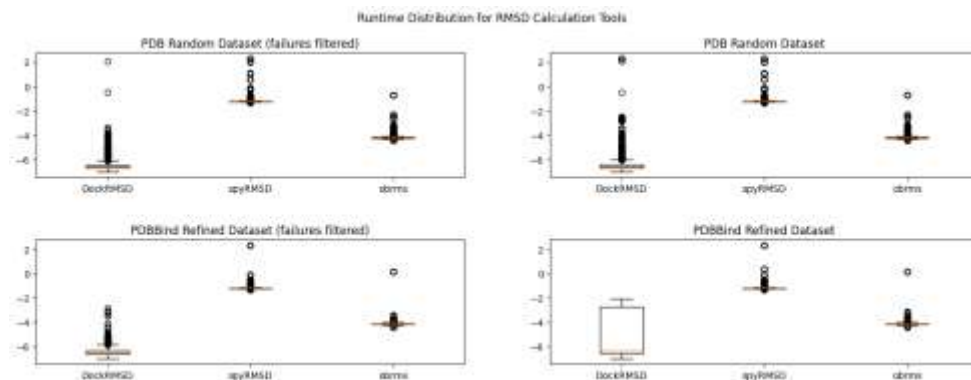


Figure 5. Runtime distribution (log(seconds)).

Figure 5 displays the runtime distributions of the RMSD calculation tools across both datasets, including those datasets without failed conformations. Despite the frequent failures of DockRMSD, filtering of such conformations influenced only DockRMSD tools runtime distributions. The failed runs of DockRMSD generally lasted between 0.06 and 0.08 seconds, which can be attributed to the try-catch block mechanism implemented in the wrapper used for these computations. This setup was responsible for managing segmentation faults, leading to minimal runtime accumulation even in failure scenarios. Consequently, these failures represent mere noise within the overall runtime distribution of DockRMSD, suggesting that the more relevant insights are derived from the graphs on the left, which exclude these anomalous data points. Mean and standard deviations of runtime distributions can be found in Table 2.

Table 2. Summary of runtimes of tools on each of datasets (all failures filtered)

Name	DockRMSD		spyRMSD		obrms	
	Mean	Std	Mean	Std	Mean	Std
<i>PDBBind Refined</i>	0.00167	0.00119	0.30955	0.29248	0.01663	0.03368
<i>PDB Random</i>	0.00208	0.06322	0.31170	0.31179	0.01593	0.01271

From the table we can see that *PDB Random* dataset was “harder” compared to *PDBBind Refined* for DockRMSD and spyRMSD in terms of runtime, while obrms showed better running times and more stability in runtime.

The big increase in DockRMSD standard deviation from *PDBBind Refined* to *PDB Random* is because of timeout failures. For some conformations of timeout failure complexes DockRMSD managed to find optimal isomorphism within timeout window, but the runtime of for those conformations affected the runtime distribution. By filtering such conformations too, DockRMSD also shows stability on both datasets.

Overall results on both datasets are similar in general and not assuming special cases that we intentionally added to the datasets, our tools showed no slight changes in runtime distributions.

We won't compare spyRMSD against other tools in term of runtime, instead we already compared its performance on two datasets, and it showed consistency. Let's now compare tools DockRMSD and obrms with each other.

DockRMSD outperformed obrms nearly 10x both in terms of mean runtime and standard deviation. Similar comparisons have been done like in [1], but for our datasets we had conformations for which DockRMSD was slower than obrms (not the special cases). More specifically about 10 jobs in *PDB Random* and 4 jobs in *PDBBind Refined* datasets. Overall DockRMSD works better than obrms if it doesn't crush.

### *Special cases*

There are special types of molecules and their structures for which DockRMSD will run for very long period. Let's have a look at ligand with PDB ID **PE3** which is ligand corresponding to complex with PDB ID **1XWV** from *PDB Random* dataset.

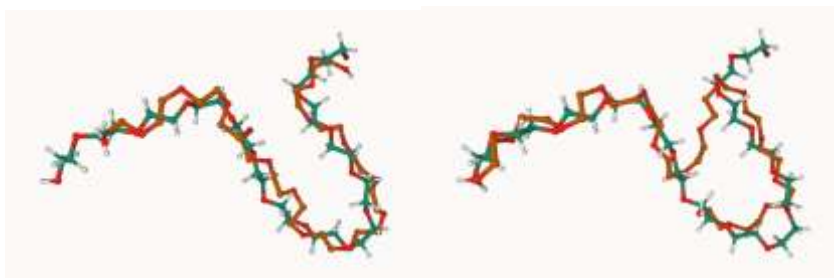


Figure 6. Ligand with PDB ID **PE3** and docked positions for it in complex **1XWV**.

For the example on the left DockRMSD runtime was 0.017 seconds, while in case of the example on the right execution time didn't fit into 10 second window. The problem comes from DEE criterion not reaching for long time in case of the

example on the right. This is also caused by the fact that tool makes its initial correspondence possibilities space artificially large. In this case assuming the implementation of DockRMSD, it'll consider more than 10 potential matches for each heavy atom, while increasing depth of tree the tool constructs for each atom will eventually lead to 2 potential matches in chain like symmetric molecules. Similar structure was also present in dataset: the ligand with PDB ID **33O** as part of complex with PDB ID **5YJ6** from *PDB Random* dataset, which just like this molecule but 3 atoms shorter from one end.

## Conclusions

This paper has explored several notable RMSD calculation tools that utilize graph isomorphism to account for molecular symmetry. While each tool has demonstrated distinct advantages, they also exhibit specific limitations under various scenarios. These tools are among the most recognized in the field for their application of graph isomorphism principles in symmetry corrected RMSD calculations.

Despite their robustness and foundational research, the tools have encountered challenges in certain conditions, which were highlighted through comprehensive evaluations. Addressing these limitations presents a significant opportunity for future research. Potential areas for development include refining existing algorithms to enhance their accuracy and reliability. Additionally, there is room for innovation in the creation of new RMSD calculation tools which may use features that organic compound graphs have.

The datasets used in this study are publicly available to ensure the reproducibility of our results and to facilitate further research in the field. They can be accessed at <https://altunyanv.github.io/rmsd-analysis-paper>.

## REFERENCES

1. Bell E.W., Zhang Y. DockRMSD: an open-source tool for atom mapping and RMSD calculation of symmetric molecules through graph isomorphism // *Journal of Cheminformatics* 11, 40 (2019): <https://doi.org/10.1186/s13321-019-0362-7>.
2. Meli R., Biggin P.C. spyrmsd: symmetry-corrected RMSD calculations in Python. *Journal of Cheminformatics* 12, 49 (2020): <https://doi.org/10.1186/s13321-020-00455-2>.
3. O'Boyle N.M., Banck M., James C.A., Morley C., Vandermeersch T., Hutchison G.R. Open Babel: An open chemical toolbox // *Journal of Cheminformatics* 3, 33 (2011): <https://doi.org/10.1186/1758-2946-3-33>.
4. Su M., Yang Q., Du Y., Feng G., Liu Z., Li Y., Wang R. Comparative Assessment of Scoring Functions: The CASF-2016 Update // *Journal of Chemical Information and Modeling*, 2019 Feb 25;59(2): 895–913: <https://doi.org/10.1021/acs.jcim.8b00545>.
5. Aric A. Hagberg, Daniel A. Schult and Pieter J. Swart. Exploring network structure, dynamics, and function using NetworkX // *Proceedings of the 7th Python in Science Conference (SciPy2008)*, Aug 2008. PP. 11–15.

6. Koes D.R., Baumgartner M.P., Camacho C.J. Empirical scoring with smina from the CSAR 2011 benchmarking exercise // Journal of Chemical Information and Modelling, 53 (2013). PP. 1893–1904.

### СРАВНИТЕЛЬНЫЙ АНАЛИЗ ИНСТРУМЕНТОВ ДЛЯ РАСЧЕТА КОРРЕКТИРОВАННОГО СИММЕТРИЕЙ RMSD В МОЛЕКУЛЯРНОМ ДОКИНГЕ

*В.Н. Алтунян*

*Ереванский государственный университет*

*ORCID: 0009-0002-0021-3789*

#### АННОТАЦИЯ

Точный расчет RMSD, особенно с учетом коррекции симметрии, имеет важное значение для надёжного сравнения молекул в вычислительной химии. В данном исследовании проводится сравнительный анализ трех современных инструментов для расчета RMSD с целью оценки их точности, частоты сбоев и вычислительной эффективности в различных молекулярных сценариях. Используя два крупных набора данных, каждый из которых содержит примерно 2000 молекулярных структур, инструменты анализируются в контексте молекулярного докинга. Эти наборы данных взяты из известных молекулярных баз данных и включают в себя разнообразие структур, что позволяет оценить инструменты в испытательных условиях. Исследование также охватывает специальные молекулярные случаи, разработанные для выявления сильных и слабых сторон каждого инструмента, предоставляя всесторонний обзор их производительности. Этот анализ направлен на помощь вычислительным химикам в выборе наиболее подходящих инструментов для расчета RMSD, что повышает точность и надежность исследований молекулярного моделирования.

**Ключевые слова:** RMSD с коррекцией симметрии, молекулярный докинг.

DOI 10.24412/1829-0450-2024-1-37-43  
УДК 004.02

Поступила: 24.05.2024г.  
Сдана на рецензию: 25.05.2024г.  
Подписана к печати: 30.05.2024г.

## COMPARING AND IMPROVING CHANGE DETECTION METHODS

*A. Fahradyan<sup>1</sup>, T. Baghdasaryan<sup>2</sup>, V. Melkonyan<sup>1</sup>, V. Sahakyan<sup>1</sup>,  
L. Kirakosyan<sup>1</sup>, O. Hovhannisyan<sup>1</sup>, S. Sargsyan<sup>1</sup>, A. Darbinyan<sup>1</sup>*

<sup>1</sup>*Russian-Armenian (Slavonic) University*

<sup>2</sup>*Yerevan State University*

*vahagn.melkonyan@student.rau.am*

*vardan.sahakyan@student.rau.am*

### ABSTRACT

This paper delves into a comparative analysis of various Change Detection (CD) methods on aerial image pairs known for their promising outcomes. By incorporating adjustments, like Image Matching, Sliding Window, and our proposed method, the Gaussian Sliding Window, these methods were trained on distinct CD datasets to attain the most effective solution for the proposed problem.

**Keywords:** Change Detection, Transformers, Image Matching, UAV, Aerial Images.

### Introduction

Change Detection is a fundamental task in remote sensing and computer vision. It involves the identification of disparities between pairs of images captured at the same geographical location but at different temporal instances. These changes can range from identifying alterations in buildings, the movement of vehicles or military equipment, and deforestation, to assessing damage caused by natural disasters. Each of them presents a unique facet of the change detection problem.

An effective CD model excels at recognizing these significant changes filtering out irrelevant alterations at the same time. The latter are caused by factors, like seasonal variations, shadows, atmospheric conditions, and lighting changes. However, tackling this task is challenging due to limited datasets and the problem that a one-size-fits-all approach doesn't work.

Presently, cutting-edge CD methods primarily rely on Deep Convolutional Neural Networks (DCNN). These networks are favored for their ability to extract valuable features. To improve their performance, recent research has focused on expanding their ability to understand large-scale context across space and time. This perception has led to the development of models with advanced features like stacked convolutional layers, dilated convolutions, and attention mechanisms.

In this paper, we aim to conduct a comprehensive analysis of various CD network-based methods like ChangeFormer [1], BIT-CD [2], and ChangerEx [3]. These methods were trained on distinct CD datasets and subjected to thorough evaluation. Additionally, we introduce enhancements to augment their performance, which will be discussed in detail in the Proposed Method section.

Furthermore, it's crucial to note that the success of current deep-supervised CD methods is closely tied to the availability of large annotated datasets. However, annotating aerial images is a costly process, especially for bi-temporal comparisons. It demands an expert examination of pixel-level differences.

To provide context, the images we exploit in this paper are captured from Unmanned Aerial Vehicles (UAVs). They are equipped with cameras and gimbals to stabilize camera vibrations.

With this study, we seek to provide a holistic understanding of the capabilities and limitations of these approaches, providing insights into their performance across different scenarios and dataset conditions.

### **Related Work**

In our study, we explore network-based methods, such as ChangeFormer [1], BIT-CD [2], and ChangerEx [3]. These networks were trained on a range of well-known datasets, such as S2Looking [4], SYSU-CD [5], CDD-Dataset [6], LEVIR-CD [7], and SECOND (SEmantic Change detectiON Dataset) [8] and on their merged version.

The problem with the existing methods lies in their limitations when it comes to detecting changes across diverse datasets. These methods often struggle to achieve consistent and accurate results across various data types. When these networks are trained on one specific dataset, their performance tends to suffer once they are applied to a different dataset. Additionally, a common requirement in network-based approaches is that the input images must be of the same size, which further restricts their adaptability. This can be a significant drawback since real-world scenarios often involve the need to detect changes in images of varying shapes and sizes.

### **Proposed Method**

In this research paper, we address the challenge of change detection thoroughly, going beyond just comparing images. We acknowledge certain practical complexities often overlooked in the research landscape, specifically:

#### ***A. Image Matching***

Our first consideration revolves around addressing potential disparities in perspective within image pairs. These differences in perspectives can occur due to some factors. For example, the gimbal not being perfectly stable, GPS deviations, or variations in the flight routes. We assume that the common area shared by the

image pair is substantial enough for meaningful analysis. To address this, we employ traditional image-matching techniques. The latter includes keypoint detection, keypoint matching, and the derivation of homographic transformations to align one image with the perspective of the other, as shown in Figure 1. Well-established algorithms like SIFT [9], SURF [10], ORB [11], and even neural network-based solutions like SuperGlue [12] are available. In our case, we opt for ORB [11] due to its simplicity and effectiveness.

### ***B. Training with Cropped Images***

In our proposed method, we address the challenge of effectively training neural networks on larger images by employing a cropping strategy. We support cropping these images into smaller, standardized sizes, such as  $256 \times 256$  pixels, before feeding them into the networks for training. This approach is particularly beneficial for aerial or satellite imagery, where large images may contain subtle details or changes that could be overlooked by the network when processed in its entirety. By segmenting the images into smaller regions, we ensure that the network receives input where the region of interest is not too small.

### ***C. Testing Methods***

During our testing phase, we encounter challenges with large images on multiple fronts. Firstly, using neural networks directly on these images demands a lot of computing power and memory. Secondly, finding labeled datasets with large images is tough. Also, our network must handle images of different sizes.

To handle these challenges, we use a neural network trained on  $256 \times 256$  size images. Then we test that network using the **Resized, Cropped, Sliding Window**, and **Gaussian Sliding Window** techniques. The Resized technique is quite simple, it resizes the image pair into  $256 \times 256$  images before feeding them to the network. The rest of the methods are visualized in Figure 2.

In Figure 2(c), the proposed method Gaussian Sliding Window works in the following way. It employs the classic sliding window, but the entries are weights generated from a 2D Gaussian bell. The proposed method helps align and handle big images better for change detection.

### ***D. Change Detection Process***

To provide a comprehensive understanding of our change detection methodology, we present the entire process here. In our approach, we start with two pre-change and post-change images (Figure 1), which are initially processed. One of these images is aligned with the perspective of the other (Figure 1(c)), as discussed in the Image Matching section.

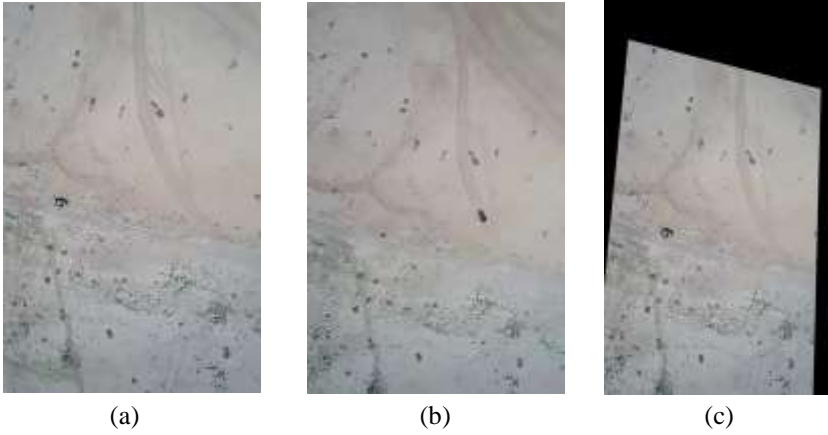


Figure 1. (a) pre-change image (b) post-change image (c) pre-change image aligned with the perspective of the post-change image.

Next, to deal with the challenges posed by large images, we introduce different testing techniques. Instead of directly using neural networks on large images, we divide them into smaller  $256 \times 256$  pixel regions with the methods specified in section Testing Methods. This process ensures that our network can effectively analyze images of varying sizes while avoiding object fragmentation issues.

We then employ three neural networks, Change Former [1], BIT-CD [2] and ChangerEx [3] to generate change maps. The use of testing methods (except Resize) results in multiple  $256 \times 256$  change maps, which are then combined into a single comprehensive change map (Figure 3).

Figure 3 illustrates the entire change detection process, showcasing the interplay of Image Matching, Testing Methods, and the generation of the final change map.

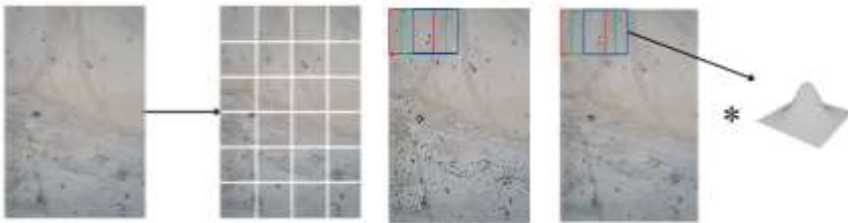


Figure 2. Testing Methods. (a) Crop Method (b) Sliding Window Method (c) Gaussian Sliding Window Method, where \* is the convolution operator.



## Results

### A. Experimental setup

We conducted our experiments using several CD key datasets mentioned in “Related Work”. Networks were trained on Ubuntu 22.04, Intel i7-11700F CPU, and Nvidia GeForce RTX 3060 GPU.

### B. Experimental Results

From the results presented in Table 1, it becomes evident that the models trained on one dataset fail to perform on another one, that is why we propose to train the models on a merged dataset, which consists of the 5 datasets, which significantly improves the F1 scores. Also, the best results are obtained when the weights obtained from the training on the merged dataset are used and the perspectives of bi-temporal images are aligned. On average the proposed Gaussian Sliding Window approach performs better than other methods.

Table 1. Test method comparison on 3 models and 6 datasets based on F1 scores.

Train Dataset	Test Dataset	ChangeFormer				ChangerEx				BIT			
		Resize	Crop	SW Average	Gaussian SW	Resize	Crop	SW Average	Gaussian SW	Resize	Crop	SW Average	Gaussian SW
SYSU	SYSU	71.42	71.51	71.61	<b>71.66</b>	<b>74.87</b>	73.33	72.18	71.29	<b>77.09</b>	77.08	77.07	77.07
	SECOND	39.74	41.32	42.55	<b>43.47</b>	<b>45.19</b>	42.7	40.79	39.31	43.08	45.93	46.25	<b>46.39</b>
	S2Looking	<b>3.6</b>	2.77	2.65	2.6	<b>4.65</b>	4.33	4.06	3.84	1.54	4.65	4.64	<b>4.68</b>
	CDD-Dataset	<b>29.67</b>	29.16	28.78	28.47	32.5	32.54	32.55	<b>32.56</b>	<b>31.6</b>	31.02	30.63	30.36
	LEVIR-CD	<b>16.71</b>	15.93	15.43	15.06	<b>18.74</b>	17.8	17.18	16.69	14.25	18.4	18.7	<b>18.77</b>
SECOND	SYSU	53.83	53.9	53.96	<b>54</b>	<b>54.31</b>	54.06	53.88	53.74	54.61	54.66	54.68	<b>54.69</b>
	SECOND	<b>67.11</b>	66.67	66.37	66.15	<b>67.54</b>	67.42	67.41	67.43	63.24	66.98	67.63	<b>67.91</b>
	S2Looking	<b>14</b>	13	12.47	12.1	8.75	15.72	16.1	<b>16.25</b>	<b>16</b>	14.83	14.38	14.1
	CDD-Dataset	21.88	22.07	22.22	<b>22.34</b>	<b>22.01</b>	21.86	21.75	21.66	<b>22.26</b>	22.26	22.26	22.26
	LEVIR-CD	<b>25.87</b>	25.48	25.21	25	22.87	24.14	25.14	<b>25.88</b>	14.99	22.36	22.73	<b>22.89</b>
S2Looking	SYSU	<b>18.51</b>	18.22	17.98	17.79	21.91	21.97	21.99	<b>22</b>	<b>20.74</b>	19.91	19.32	18.88
	SECOND	<b>36.23</b>	35	33.89	33.01	<b>41.24</b>	39.2	37.68	36.6	37.85	41.88	42.23	<b>42.42</b>
	S2Looking	<b>58.75</b>	57.52	56.66	56.01	25.46	55.63	56.78	<b>57.27</b>	56.75	57.7	58.47	<b>59.02</b>
	CDD-Dataset	<b>23.74</b>	23.27	22.85	22.48	28.19	28.21	28.21	<b>28.22</b>	<b>26.57</b>	25.53	24.81	24.28
	LEVIR-CD	<b>28.87</b>	28.4	27.99	27.66	24.86	26.53	27.88	<b>28.94</b>	11.28	23.99	24.71	<b>24.95</b>
CDD Dataset	SYSU	16.01	16.35	16.63	<b>16.86</b>	14.72	15.11	15.39	<b>15.6</b>	14.08	14.11	14.12	<b>14.13</b>
	SECOND	22.15	22.41	22.5	<b>22.57</b>	15.84	19.1	21.01	<b>22.33</b>	14.6	<b>15.36</b>	15.16	15.08
	S2Looking	<b>14.04</b>	13.71	13.42	13.16	10.11	14.08	14.44	<b>14.54</b>	<b>14.5</b>	14.35	14.2	14.08
	CDD-Dataset	<b>89.18</b>	88.8	88.48	88.22	<b>92.94</b>	91.49	90.45	89.66	95.07	<b>95.08</b>	<b>95.08</b>	<b>95.08</b>
	LEVIR-CD	<b>22.68</b>	20.56	18.93	17.62	20.82	21.93	22.49	<b>22.82</b>	12.79	20.64	20.81	<b>20.89</b>

Train Dataset	Test Dataset	ChangeFormer				ChangerEx				BIT			
		Resize	Crop	SW Average	Gaussian SW	Resize	Crop	SW Average	Gaussian SW	Resize	Crop	SW Average	Gaussian SW
LEVIR-CD	SYSU	<b>5.51</b>	5.2	4.95	4.74	3.22	4.42	5.26	<b>5.89</b>	<b>1.39</b>	<b>1.39</b>	1.38	1.38
	SECOND	13.77	16.04	17.63	<b>18.86</b>	12.32	12.77	12.94	<b>13.1</b>	11.77	<b>12.69</b>	12.54	12.49
	S2Looking	<b>6.05</b>	5.79	5.57	5.41	4.1	5.18	5.71	<b>6.09</b>	0.76	<b>4.11</b>	4.1	<b>4.11</b>
	CDD-Dataset	<b>25.89</b>	25.78	25.68	25.6	24.86	25.4	25.77	<b>26.04</b>	23.95	23.97	23.97	<b>23.98</b>
	LEVIR-CD	<b>87.38</b>	86.44	85.88	85.44	86.76	87.23	87.63	<b>87.89</b>	<b>87.38</b>	86.44	85.88	85.44
Merged	SYSU	66.86	67.55	68.12	<b>68.59</b>	<b>72.48</b>	69.66	67.58	65.99	<b>76.55</b>	<b>76.55</b>	<b>76.55</b>	<b>76.55</b>
	SECOND	53.16	53.95	54.64	<b>55.2</b>	<b>63.42</b>	58.7	55.41	53.05	59.13	63.98	64.82	<b>65.17</b>
	S2Looking	36.28	36.52	36.74	<b>36.94</b>	<b>51.11</b>	44.72	40.02	36.43	18.76	49.25	51.01	<b>51.65</b>
	CDD-Dataset	<b>55.77</b>	55.24	54.81	54.44	<b>57.02</b>	56.77	56.58	56.42	57.15	57.28	57.32	<b>57.35</b>
	LEVIR-CD	71.41	72.53	73.5	<b>74.26</b>	<b>83.83</b>	77.98	74.34	71.82	18.83	82.21	84.09	<b>84.76</b>

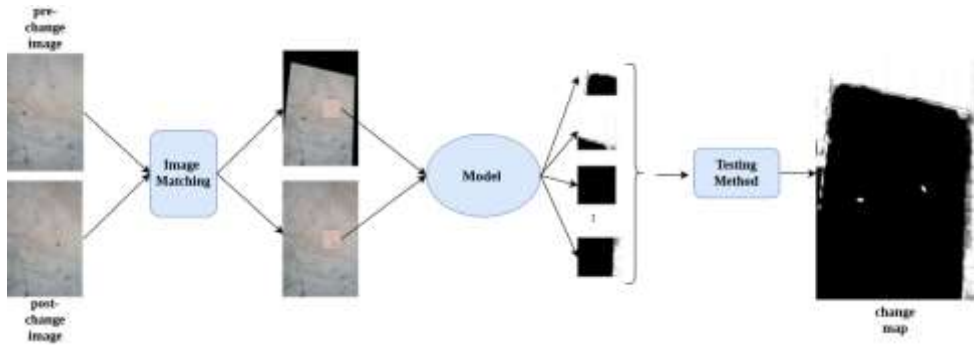


Figure 3. Change Detection Process Overview. Pre-change and post-change images are brought to the same perspective, after which with the chosen test method the images are processed and fed to the model as input. The model outputs the final change map.

## Conclusion

In conclusion, this research has provided a deep analysis of CD methods and introduced practical improvements to tackle issues with diverse datasets and large images to enhance accuracy in CD for aerial images.

## Acknowledgement

This work was supported by the Science Committee of RA (Research project № 23EDP-2B001).

## REFERENCES

1. *Bandara W.G.C. and Patel V.M.* (2022). July. A transformer-based Siamese network for change detection. In IGARSS 2022–2022// IEEE International Geoscience and Remote Sensing Symposium (PP. 207–210). IEEE.
2. *Chen H., Qi Z. and Shi Z.*, 2021. Remote sensing image change detection with transformers // IEEE Transactions on Geoscience and Remote Sensing, 60. PP. 1–14.
3. *Fang S., Li K. and Li Z.* 2023. Changer: Feature interaction Is what you need for change detection. // IEEE Transactions on Geoscience and Remote Sensing.
4. *Shen L., Lu Y., Chen H., Wei H., Xie D., Yue J., Chen R., Lv S., Jiang B.* S2looking: A satellite side-looking dataset for building change detection. Remote Sensing 13(24), 5094 (2021).
5. *Shi Qian, Mengxi Liu, Shengchen Li, Xiaoping Liu, Fei Wang, and Liangpei Zhang.* “A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection” // IEEE transactions on geoscience and remote sensing, 60 (2021): 1–16.
6. *Lebedev M., Yu V. Vizilter, Vygolov O., Vladimir A. Knyaz and A. Yu Rubis.* “Change detection in remote sensing images using conditional adversarial networks” // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42 (2018): 565–571.
7. *Hao Chen and Zhenwei Shi,* “A spatial-temporal attention-based method and a new dataset for remote sensing image change detection”, Remote Sensing, vol. 12, no. 10. PP. 1662, 2020.
8. *Kunping Y., Xia G.-S., Liu Z., Bo Du, Yang W., Pelillo M. and Zhang L.* “Semantic change detection with asymmetric Siamese networks” // arXiv preprint arXiv:2010.05687 (2020).
9. *Lowe, David G.* “Object recognition from local scale-invariant features” // In Proceedings of the seventh IEEE International Conference on Computer Vision, vol. 2 PP. 1150–1157. Ieee, 1999.
10. *Lowe David G.* “Object recognition from local scale-invariant features” // In Proceedings of the seventh IEEE International Conference on Computer Vision, vol. 2. PP. 1150–1157. Ieee, 1999.
11. *Bay H., Andreas E., Tuytelaars T. and Gool Van L.* “Speeded-up robust features (SURF)” // Computer vision and image understanding 110, no. 3 (2008): 346–359.
12. *Rublee Et., Rabaud V., Konolige K. and Bradski G.* “ORB: An efficient alternative to SIFT or SURF” // In 2011 International conference on computer vision. PP. 2564–2571. Ieee, 2011.
13. *Paul-Edouard S., DeTone D., Malisiewicz T. and Rabinovich A.* (2020) “Superglue: Learning feature matching with graph neural networks” // In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. PP. 4938–4947.

СРАВНЕНИЕ И УЛУЧШЕНИЕ МЕТОДОВ ОБНАРУЖЕНИЯ  
ИЗМЕНЕНИЙ

*А. Фаградян<sup>1</sup>, Т. Багдасарян<sup>2</sup>, В. Мелконян<sup>1</sup>, В. Саакян<sup>1</sup>, Л. Киракосян<sup>1</sup>,  
О. Оганесян<sup>1</sup>, С. Саргсян<sup>1</sup>, А. Дарбинян<sup>1</sup>*

## АННОТАЦИЯ

Данная статья посвящена сравнительному анализу различных методов обнаружения изменений (CD), известных своими многообещающими результатами. Внедряя такие коррективы, как сопоставление изображений, скользящее окно и наш предложенный метод – Гауссово скользящее окно, эти методы были обучены на различных наборах данных CD, чтобы достичь наиболее эффективного решения для предложенной проблемы.

**Ключевые слова:** обнаружение изменений, Трансформеры, Сопоставление изображений, БПЛА, аэрофотоснимки.

DOI 10.24412/1829-0450-2024-1-44-55  
УДК: 577.25

Поступила: 07.03.2024г.  
Сдана на рецензию: 07.03.2024г.  
Подписана к печати: 19.03.2024г.

## ВЛИЯНИЕ МОДУЛЯТОРОВ TRPC6 НА АКТИНОВЫЙ ЦИТОСКЕЛЕТ IN VITRO

*Д.М. Мелентьева<sup>1</sup>, Н.А. Зернов<sup>1</sup>, В.С. Камарян<sup>2</sup>,  
Л.С. Унанян<sup>1,2</sup>, Е.А. Попугаева<sup>1,\*</sup>*

<sup>1</sup>*Санкт-Петербургский политехнический университет Петра Великого,  
г. Санкт-Петербург, Россия*

<sup>2</sup>*Российско-Армянский (Славянский) университет, г. Ереван, Армения  
\* lena.porugaeva@gmail.com*

### АННОТАЦИЯ

В данных экспериментах оценивался эффект модуляторов канала плазматической мембраны TRPC6, оказываемый на клеточный актин. В качестве агонистов применялись соединения GSK1702934A, C20, Z12-30 в качестве блокатора SKF96365. Эксперименты проводились на первичной культуре фибробластов, а также на нейронах первичной гиппокампальной культуры.

**Ключевые слова:** TRPC6, нейродегенерация, актиновый цитоскелет, нейрон.

### Введение

TRPC6 канал – это один из белков семейства TRPC (“transient receptor potential canonical”), канал плазматической мембраны, состоящий из 4 субъединиц, формирующих ион-проницаемую пору [1, 2]. Настоящее исследование направлено на изучение модуляторов TRPC6 в контексте изучения кальциевой теории патогенеза болезни Альцгеймера (БА). Ранее показано значительное снижение экспрессии TRPC6 у пациентов с БА [3], а также продемонстрировано улучшение пространственной памяти у мышей с гиперэкспрессией TRPC6 [4, 5]. Предполагается, что TRPC6 участвует в работе различных регуляторных механизмов, в том числе и депо-управляемого [6], и рецептор-управляемого входа кальция [7], что делает данный канал гибким регуляторным инструментом кальциевого баланса в нейронах с возможностью фармакологической коррекции его работы [8, 9].

Однако требования, предъявляемые к фармакологическим агентам, включают не только успешную модуляцию TRPC6 канала, но и набор фармакокинетических свойств. Поэтому продолжается поиск новых структур, которые будут соответствовать вышеперечисленным требованиям. Ранее было обнаружено соединение 51164, производное пиперазина, демонстрирующее

нейропротекторные свойства “*in vitro*” [10]. На HEK293T клетках было показано, что 51164 активирует кальциевый вход через TRPC6 [10]. Также 51164 в концентрации 10 нМ способно восстанавливать структуры грибовидных шипиков в условиях амилоидной токсичности, а в концентрации 100 нМ исправляет дефицит долговременной потенции у 6-месячных мышей 5xFAD [10]. Кроме того, для 51164 был обнаружен альтернативный протекторный механизм – на первичной культуре мышечных фибробластов впервые было показано, что 51164 в концентрации 1 мкМ, препятствует деполимеризации актина [11].

Однако 51164 демонстрирует неудовлетворительный профиль биодоступности. В этой связи с помощью методов биоинформатики и на основе фармакофора 51164 был определен новый кандидат из группы пиперазинов – соединение Z12-30 [12]. В данной работе исследовалось влияние на организацию актина соединения Z12-30 – также производного пиперазина. В работе исследовалось влияние на актин Z12-30 и некоторых ранее известных модуляторов TRPC6.

## Материалы и методы

### □ Химические соединения

3-(3-,4-Dihydro-6,7-dimethoxy-3,3-dimethyl-1-isoquinolinyl)-2H-1-benzopyran-2-one (C20) получен из Tocris (Tocris, Bristol, UK, # 6875); Цитохалазин-D (Tocris, #1233); SKF96365 (Tocris, #1147); GSK1702934A (Sigma, #SML2323).

### □ Первичные культуры мышечных фибробластов

Первичная культура фибробластов была получена по протоколу, описанному ранее [11]. Коротко: фибробласты выделялись из хвостиков новорожденных (1–3 день) мышей линии FVB. Небольшие участки хвостов измельчались в холодном буфере (1% 10× CMF-HBSS (Gibco, Grand Island, NY, USA; #14185), 1% Pen Strep (Gibco, Grand Island, NY, USA; #15140)). Проводилась ферментация в растворе папаина (Worthington, Columbus, USA; #LK003176) в течении 1,5 часов при 37 °C, затем промывка (дважды) раствором ДНКазы I (Sigma, St Louis, MO, USA; #DN-25). Рост клеток происходил в среде для культивации DMEM (Gibco, Grand Island, NY, USA; #41965) с 10% FBS (Gibco, Grand Island, NY, USA; #10500), 1% PEST, (Gibco, Grand Island, NY, USA; #15140) при 37 °C в 5% CO<sub>2</sub> инкубаторе. На 2–3 пассаже клетки пересаживались на 15 мм покровные стекла (Fisher brand microscope cover glass, #052212-9), предварительно обработанные 0.1 мг/мл поли-D-лизин (poly-d-lysine, Sigma, #SLBN9056V). На следующие сутки клетки инкубировались с 1 мкМ Z12-30 и/или 0,125 мг/мл цитохалазина D, затем отмывались в растворе фосфатного буфера PBS (VWR life science, #0015C192) и фиксировались в 4% параформальдегиде (Across organics, #B0150852) 15–20 минут при комнатной температуре, пермеабелизовывались раствором 0,1% Triton X-100

(Helicon, #3616C357) в течении 10 минут. В следующие 10 минут клетки окрашивались 3–5 мкг/мл родамин-фаллоидином (Invitrogen by Thermo Fisher Scientific, #1475357).

□ Первичные гиппокампальные культуры.

Выделение первичной гиппокампальной культуры проводилось по протоколу, описанному ранее [11, 13]. Коротко: гиппокампальные культуры выделялись из новорожденных мышей (0–2 день) линии FVB (Jackson Laboratory, 001800). Вырезание гиппокампов осуществлялось в холодном буфере (1% 10× CMF-HBSS (Gibco, Grand Island, NY, USA; #14185), 1% Pen Strep (Gibco, Grand Island, NY, USA; #15140), 16 mM HEPES (Sigma, St Louis, MO, USA; #H3375), 10 mM NaHCO<sub>3</sub> (Sigma, St Louis, MO, USA; #S5761); pH=7.2). После 30 минут ферментации при 37°C в растворе папаина (Worthington, Columbus, OH, USA; #LK003176) гиппокампы обрабатывались раствором ДНКазы I (Sigma, St Louis, MO, USA; #DN-25). Клетки рассивались на покровные стекла, обработанные 0.1 мг/мл поли-D-лизин (poly-d-lysine, Sigma, #SLBN9056V) в 24-луночный планшет. Клетки культивировались при 37°C в 5% CO<sub>2</sub> инкубаторе в среде Neurobasal (Gibco, Grand Island, NY, USA; #10888) с 1% FBS (Gibco, Grand Island, NY, USA; #10500), 2% 50xB27 (Gibco, Grand Island, NY, USA; #17504), и 0.05 mM L-глутамином (Gibco, Grand Island, NY, USA; #250030). Через 24 часа культивации клеточная культура инкубировалась с одним из модуляторов TRPC6, затем фиксировалась по протоколу, аналогичному для фиксации первичной культуры фибробластов с окраской родамин-фаллоидином.

□ Иммуноцитохимия.

Для визуализации микротрубочек нейронов фиксированная в 4% параформальдегиде культура блокировалась 30 минут при комнатной температуре в 5% растворе BSA в PBS (Sigma, #SLBZ6631), пермеабелизовывались 10 минут раствором 0,1% Triton X-100. Затем проводилась окраска мышечными моноклональными антителами anti-MAP2A (Sigma, #2500681), растворенными 1:1000 в 2,5% растворе BSA в PBS, 2 часа при комнатной температуре. Окраска вторичными антителами Alexa Fluor 488 goat anti-mouse antibody (Thermo Fisher Scientific, #R37120), растворенными 1:2000 в 2,5% растворе BSA 1 час при комнатной температуре. Для одновременной визуализации нейронального актина клетки окрашивались родамин-фаллоидином в течение 10 минут. После каждого этапа окраски препарат 3 раза промывался раствором PBS по 2 минуты.

□ Конфокальная микроскопия.

Микроскопия препаратов осуществлялась с использованием конфокального микроскопа Thorlabs (ThorImLS1 software 5 (Newton, NJ, USA)) с увеличением 40x (LUMPlanFL N, 40x/0.80 water, OLYMPUS), разрешением 0,116 мкм/пиксель для культуры фибробластов и с увеличением 100x (UPlanSApo,

100x/1.40 Oil, OLYMPUS), разрешением 0,046 мкм/пиксель для культуры гиппокампальных нейронов.

□ Анализ организации актиновых тяжей в клеточных культурах

Для анализа изображений фибробластов использовался метод вычисления фрактальной размерности, алгоритм “box-counting method”, описанный ранее [11, 14]. В ходе метода на предварительно бинаризованное изображение накладывается квадратная сетка и подсчитывается относительное число ячеек, содержащих элементы изображения, затем частота разбиения сетки увеличивается, и относительное число ячеек подсчитывается вновь; процедура повторяется. Фрактальной размерностью называется предел полученной последовательности чисел. Это значение может характеризовать организацию клеточного актина: в клетках с деполимеризованным под действием внешних факторов актином (актин в виде отдельных точек) фрактальная размерность будет ниже, чем для нормальных клеток. Данный анализ выполнялся в ПО “ImageJ” – с использованием плагина “FracLac”.

Для оценки нейрональных конусов роста вычислялся средний уровень интенсивности флуоресценции (данные светимости конусов роста из 16-битных изображений, деленные на число пикселей внутри выделенной области конуса роста с вычетом значений фоновой интенсивности).

## Результаты

□ Производное пиперазина поддерживает актиновые филаменты в полимеризованной форме в культуре клеток мышинных фибробластов.

Было сделано предположение о наличии у соединения Z12-30 способности поддерживать актин в полимеризованной форме. Данный эффект был продемонстрирован на первичной культуре клеток мышинных фибробластов. Выбор культуры для исследования обусловлен морфологическими свойствами фибробластов – хорошо развитым актиновым цитоскелетом, в норме заполняющим все цитоплазматическое пространство [15].

Для выявления эффекта Z12-30 клетки мышинных фибробластов в течение 15 минут подвергались действию 0,125 мг/мл цитохалазина D (токсин, разрушающий фибриллярный актин (F-актин) до его мономеров (G-актина)) вместе с Z12-30 в концентрации 10 мкМ или эквивалентным объемом DMSO. Затем, чтобы визуализировать клеточный актин, фибробласты фиксировались и окрашивались родамин-фаллоидином. Изображение, полученные на конфокальном микроскопе, анализировались с использованием метода измерения фрактальной размерности. Данный метод позволяет количественно оценить организацию актиновой структуры клетки [14].

На Рис. 1 представлены репрезентативные изображения фибробластов и результаты количественного анализа. Клетки, инкубированные с цитохалазином D (Цит.D), отличались меньшим значением фрактальной размерности в

сравнении с контрольной группой клеток [ $1.75 \pm 0.01$  ( $n = 30$ , по 10 клеток из 3 различных культур) в контрольной группе – против группы клеток, обработанных Цит.Д  $1.45 \pm 0.03$  ( $n = 30$ ),  $p = 6.2 \cdot 10^{-11}$ , тест Крускала-Уоллиса, с последующим апостериорным тестом Данна]. Инкубация с Цит.Д, совместно с Z12-30, привела к увеличению фрактальной размерности в сравнении с группой клеток, подвергшихся воздействию только с Цит.Д [ $1.61 \pm 0.02$  ( $n = 30$ ) под действием Цит.Д + Z12-30, против группы под действием только с Цит.Д  $1.45 \pm 0.03$  ( $n = 30$ ),  $p = 0.0006$ ].

Этот результат доказывает, что производное пиперазина Z12-30 способно препятствовать деполимеризации актина через механизмы, не связанные с сигнальным путем активации TRPC6.

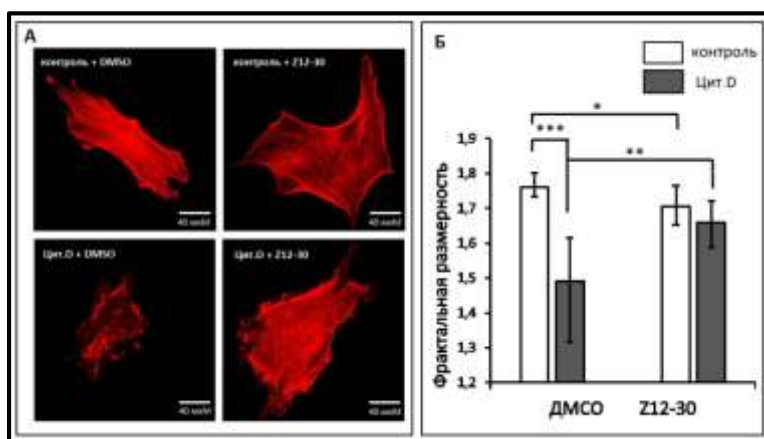


Рисунок 1. Исследование влияния Z12-30 на организацию F-актина в первичной культуре фибробластов мыши. А. Репрезентативные конфокальные изображения клеток первичной культуры мышечных фибробластов, окрашенных родамин-фаллоидином. Длина масштабного отрезка 40 мкм. Б. Количественный анализ степени заполненности клеток первичной культуры фибробластов актиновыми тяжами, выполненный с использованием метода вычисления фрактальной размерности. Результаты представлены как медиана  $\pm$  верхний/нижний квартиль. Вид распределения данных проверен с использованием теста Шапиро-Уилка. При статистическом анализе использовался тест Крускала-Уоллиса, Post-hoc анализ-тест Данна,  $*P < 0.05$ ,  $**P < 0.005$ ,  $***P < 0.0001$  ( $n = 30$  для каждой группы).

- Действие блокатора TRPC6 приводит к перераспределению нейронального актина и разрушению актина в нейрональных конусах роста.

Существуют модуляторы TRPC6, которые предположительно влияют на актин, используя непосредственно сигнальный путь, включающий работу TRPC6 канала. Согласно литературным данным в нейронах имеет место каскад реакций, запускаемых TRPC6, приводящий к активации генов, регулирующих синаптическую пластичность, среди которых нейротрофный фактор мозга ("brain derived neurotrophic factor", BDNF) [16, 17]. Показано также, что



стимуляция входа кальция через TRPC6 приводит к росту дендритов [18] и управляет движением конусов роста [19]. Основой такой динамики морфологических признаков является изменение структуры актинового цитоскелета [20].

Для выявления роли TRPC каналов в регуляции состояния нейронального актина использовалось соединение SKF96365. SKF96365 – это неселективный блокатор TRPC и потенциал-зависимых натриевых каналов [21]. В эксперименте соединение добавлялось в среду для культивации нейронов в концентрации 10 мкМ, т.к. показано, что такая концентрация полностью блокирует вход кальция [22].

Эффект от действия SKF96365 оценивался по конфокальным снимкам нейронального актина в первичной культуре гиппокампальных нейронов (div1), окрашенной родамин-фаллоидином. Актиновый цитоскелет в нейронах представляет собой скопления G-актина в теле нейрона (наблюдаемые в виде отдельных точек) и звездообразные структуры F-актина на концах отростков нейронов – конусы роста. Состояние конусов роста является очень изменчивым и отражает внутриклеточные процессы на ранних стадиях развития нейронов [23–25].

Изменения в нейрональном цитоскелете были зафиксированы после 5-часового воздействия 10 мкМ SKF96365. Блокатор TRPC вызвал перераспределение актина: заполнение молекулами актина тела нейрона и редукцию конусов роста (Рис. 2). По данным литературы, вышеописанное явление ретроградного транспорта актина противопоставляется полимеризации актина и нейритогенезу, росту нейрональных отростков [26, 27].

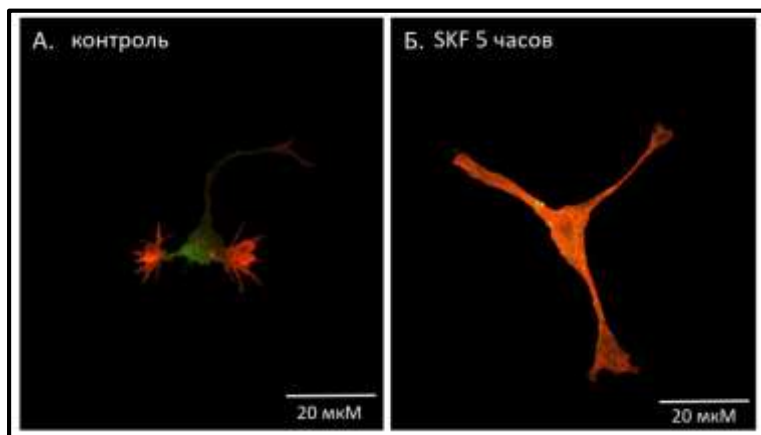


Рисунок 2. Актиновый и микротрубочковый цитоскелет в нейронах первичной гиппокампальной культуры. Флуоресцентные изображения показывают распределение F-актина (красный, окраска родамин-фаллоидином) и микротрубочек (зеленый, окраска антителами anti-MAP2A) в зафиксированных

нейронах, подвергнутых иммунофлуоресцентному окрашиванию. А. Нейрон без химического воздействия. Б. Нейрон после воздействия 10 мкМ SKF96365 в течение 5 часов.

Количественная оценка влияния SKF96365 заключалась в анализе средней интенсивности флуоресценции конусов роста, окрашенных родамин-фаллоидином. Инкубирование первичной гиппокампальной культуры с SKF96365 в течение 5 часов значимо снижало интенсивность флуоресценции [ $5018 \pm 117$  ( $n=15$ ) под действием SKF96365, против контрольной группы  $9148 \pm 102$  ( $n=15$ ),  $p = 0.0001$ , Т-тест Стьюдента]. Блокатор TRPC каналов привел к деполимеризации тяжелой F-актина и деградации конусов роста (Рис. 3).

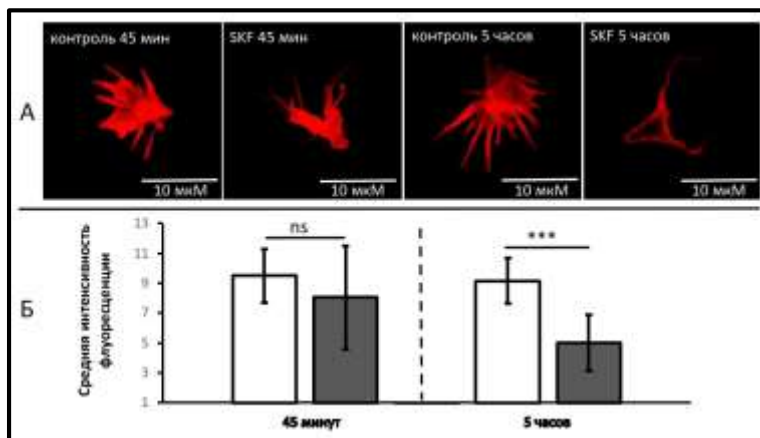


Рисунок 3. Действие блокатора TRPC каналов SKF96365 на состояние актина в нейрональных конусах роста. А. Репрезентативные конфокальные изображения клеток первичной гиппокампальной культуры, окрашенных родамин-фаллоидином после воздействия 10 мкМ SKF96365 через 45 минут и 5 часов, а также контрольной группы клеток. Б. Количественный анализ состояния конусов роста нейронов, выполненный путем измерения средней интенсивности флуоресценции конусов роста. Результаты представлены как среднее значение  $\pm$  SD (стандартное отклонение). Вид распределения и равенство дисперсий проверены с использованием теста Колмогорова-Смирнова и теста Бартлетта соответственно. При статистическом анализе использовался Т-тест с поправкой Уэлча, \*\*\* $P < 0,0001$ , ns - non significant ( $n=15$  для каждой группы).

- Изменения в организации актина конусов роста под действием положительных модуляторов TRPC6 не наблюдались.

Далее, в аналогичных условиях эксперимента были протестированы положительные модуляторы TRPC6 канала: GSK1702934A [28], C20 [29] и Z12-30 [12].

Ожидалось в результате анализа флуоресценции конусов роста зафиксировать значимое увеличение числа тяжелой F-актина в нейритах. Это предположение основывалось на данных об увеличении продукции актина в результате действия активатора TRPC6 [30].

Однако в описываемых в данной работе экспериментальных условиях выявить эффект от применения положительных модуляторов TRPC6 не удалось [ $1.00 \pm 0.17$  ( $n = 35$ ) для контрольной группы,  $0.98 \pm 0.18$  ( $n = 43$ ) для GSK1702934A,  $1.09 \pm 0.22$  ( $n = 35$ ) для C20,  $0.97 \pm 0.17$  ( $n = 33$ ) для Z12-30, при сравнении каждой группы с контрольной  $P > 0.05$ , Т-тест Стьюдента] (Рис. 4).

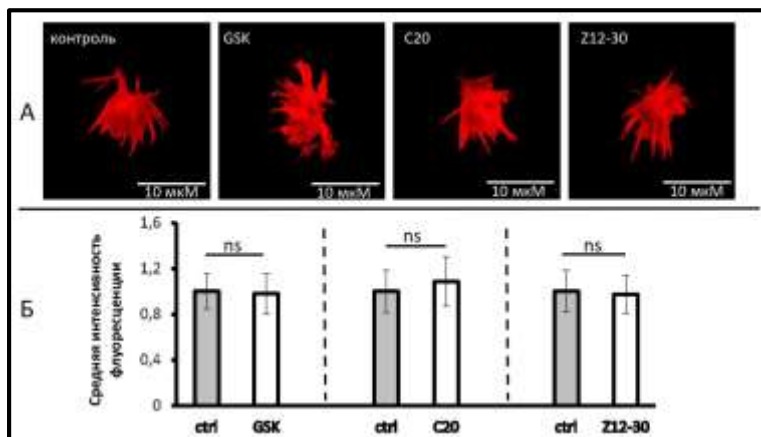


Рисунок 4. Действие активаторов TRPC6 каналов на состояние актина в нейрональных конусах роста. А. Репрезентативные конфокальные изображения клеток первичной гиппокампальной культуры, окрашенных родамин-фаллоидином через 5 часов после воздействия соединениями GSK1702934A, C20, Z12-30 в концентрации 10 мкМ. Б. Количественный анализ состояния конусов роста нейронов, выполненный путем измерения средней интенсивности флуоресценции конусов роста, нормированной на контрольные значения. Результаты представлены как среднее значение  $\pm$  SD (стандартная ошибка среднего). Вид распределения и равенство дисперсий проверены с использованием теста Шапиро-Уилка и теста Бартлетта, соответственно. При статистическом анализе использовался однофакторный анализ ANOVA, Post-hoc анализ- тест Шидака, *ns* - non significant  $P > 0,05$  ( $n$  от 33 до 43 для каждой группы).

## Обсуждение результатов

Настоящая статья описывает возможные механизмы воздействия модуляторов TRPC на актиновый цитоскелет. Данные механизмы, вероятно, могут затрагивать как кальций-зависимые каскады белковых взаимодействий, так и непосредственное воздействие химических агентов на молекулы актина.

Клетки первичной культуры фибробластов, ввиду отличий в экспрессии TRPC6 от нейронов [31, 32], характеризуются отличным сигнальным путем регуляции организации актина. Так фибробласты могут служить удобной моделью для изучения альтернативных механизмов влияния на клеточный актин, которые характерны для соединений класса производных пиперазина. Основываясь на данном исследовании Z12-30 и ранее полученных данных для структуры 51164 о способности поддерживать актин в полимеризованной

форме, можно сделать вывод, что такое свойство является общим для группы производных пиперазина.

Следует рассмотреть данное свойство с точки зрения цитотоксичности. С одной стороны, выраженным негативным эффектом обладают химические агенты, разрушающие актиновые филаменты [33, 34]. С другой, актин-связывающие соединения, увеличивающие фракцию полимеризованной формы актина, также цитотоксичны и применяются как противоопухолевые препараты [35]. Такая точка зрения позволяет рассматривать пиперазины как основу для поиска и разработки нового класса противоопухолевых препаратов, поскольку стабилизаторы актина способствуют ингибированию пролиферации и нарушению миграции раковых клеток [36]. При этом установленная в настоящем исследовании актин-стабилизирующая концентрация Z12\_30 (10 мкМ) на порядок больше концентраций, используемых в “in vivo”- экспериментах. Для выявления цитотоксических свойств при различных концентрациях требуются дальнейшие эксперименты.

В данной статье впервые был продемонстрирован “in vitro”-эффект агонистов TRPC6 на нейронах первичной гиппокампальной культуры. В это же время, полученные результаты воздействия блокаторов свидетельствуют о том, что вход кальция через TRPC является необходимым для формирования конусов роста, а блокировка данных каналов приводит к фатальным изменениям клеточного сигналинга, разрушению цитоскелета и гибели нейронов. Однако, ввиду наличия у SKF96365 способности активировать TRPC и потенциал-зависимые натриевые каналы, нельзя однозначно заключить какой из кальциевых проницаемых каналов играет первостепенную роль в процессах полимеризации актина в незрелых первичных культурах гиппокампа.

Исследование выполнено за счет гранта Российского научного фонда № 20-75-10026: <https://rscf.ru/project/20-75-10026/>».

#### ЛИТЕРАТУРА

1. *Dietrich A., Gudermann T.* Trpc6 // *Handb Exp Pharmacol.* 2007.10.1007/978-3-540-34891-7\_7 № 179. СС. 125–41.
2. *Trebak M., Vazquez G., Bird G.S., Putney J.W., Jr.* The TRPC3/6/7 subfamily of cation channels // *Cell Calcium.* 2003. Т. 33, № 5–6. СС. 451–61.
3. *Chen J. M., Li Q.W., Liu J.S., Jiang G.X., Liu J.R., Chen S.D., Cheng Q.* TRPC6 mRNA levels in peripheral leucocytes of patients with Alzheimer's disease and mild cognitive impairment: A case-control study // *Prog Neuropsychopharmacol Biol Psychiatry.* 2019. Т. 92. СС. 279–284.
4. *Zhou J., Du W., Zhou K., Tai Y., Yao H., Jia Y., Ding Y., Wang Y.* Critical role of TRPC6 channels in the formation of excitatory synapses // *Nat Neurosci.* 2008. Т. 11, № 7. СС. 741–3.
5. *Gualdani R., Gailly P.* How TRPC Channels Modulate Hippocampal Function // *Int J Mol Sci.* 2020. Т. 21, № 11.

6. Zhang H., Sun S., Wu L., Pchitskaya E., Zakharova O., Fon Tacer K., Bezprozvanny I. Store-Operated Calcium Channel Complex in Postsynaptic Spines: A New Therapeutic Target for Alzheimer's Disease Treatment // *J Neurosci*. 2016. Т. 36, № 47. СС. 11837–11850.
7. Gudermann T., Hofmann T., Mederos y Schnitzler M., Dietrich A. Activation, subunit composition and physiological relevance of DAG-sensitive TRPC proteins // *Novartis Found Symp*. 2004. Т. 258. СС. 103–18; discussion 118–22, 155–9, 263–6.
8. Wang H., Cheng X., Tian J., Xiao Y., Tian T., Xu F., Hong X., Zhu M. X. TRPC channels: Structure, function, regulation and recent advances in small molecular probes // *Pharmacol Ther*. 2020. Т. 209. СС. 107–497.
9. Zernov N., Popugaeva E. Role of Neuronal TRPC6 Channels in Synapse Development, Memory Formation and Animal Behavior // *Int J Mol Sci*. 2023. Т. 24, № 20.
10. Popugaeva E., Chernyuk D., Zhang H., Postnikova T.Y., Pats K., Fedorova E., Poroikov V., Zaitsev A. V., Bezprozvanny I. Derivatives of Piperazines as Potential Therapeutic Agents for Alzheimer's Disease // *Mol Pharmacol*. 2019. Т. 95, № 4. СС. 337–348.
11. Zernov N., Ghamaryan V., Makichyan A., Melenteva D., Hunanyan L., Popugaeva E. Piperazine Derivative Stabilizes Actin Filaments in Primary Fibroblasts and Binds G-Actin In Silico // *Curr Issues Mol Biol*. 2022. Т. 44, № 11. СС. 5191–5208.
12. Попугаева Е.А. Мелентьева Д.М., Зернов Н.И., Камарян В.С., Унаниян Л.С. Безпро-званный И.Б. Производные дизамещенных пиперазинов, агонисты trpc6 каналов, дем-онстрируют нейропротекторные свойства в *in vitro* модели болезни Альцгеймера, 2022.
13. Chernyuk D., Zernov N., Kabirova M., Bezprozvanny I., Popugaeva E. Antagonist of neuronal store-operated calcium entry exerts beneficial effects in neurons expressing PSEN1DeltaE9 mutant linked to familial Alzheimer disease // *Neuroscience*, 2019. Т. 410. СС. 118–127.
14. Revittser A., Selin I., Negulyaev Y., Chubinskiy-Nadezhdin V. The analysis of F-actin structure of mesenchymal stem cells by quantification of fractal dimension // *PLoS One*. 2021. Т. 16, № 11.
15. Khan M., Gasser S. Generating Primary Fibroblast Cultures from Mouse Ear and Tail Tissues // *J Vis Exp*. 2016.10.3791/53565 № 107.
16. Gibon J., Deloulme J. C., Chevallier T., Ladeveze E., Abrous D. N., Bouron A. The antidepressant hyperforin increases the phosphorylation of CREB and the expression of TrkB in a tissue-specific manner // *Int J Neuropsychopharmacol*. 2013. Т. 16, № 1. СС. 189–98.
17. Middei S., Spalloni A., Longone P., Pittenger C., O'Mara S. M., Marie H., Ammassari-Teule M. CREB selectively controls learning-induced structural remodeling of neurons // *Learn Mem*. 2012. Т. 19, № 8. СС. 330–6.
18. Tai Y., Feng S., Ge R., Du W., Zhang X., He Z., Wang Y. TRPC6 channels promote dendritic growth via the CaMKIV-CREB pathway // *J Cell Sci*. 2008. Т. 121, № Pt 14. СС. 2301–7.
19. Li Y., Jia Y.C., Cui K., Li N., Zheng Z.Y., Wang Y.Z., Yuan X.B. Essential role of TRPC channels in the guidance of nerve growth cones by brain-derived neurotrophic factor // *Nature*. 2005. Т. 434, № 7035. СС. 894–8.
20. Tai Y., Feng S., Du W., Wang Y. Functional roles of TRPC channels in the developing brain // *Pflugers Arch*. 2009. Т. 458, № 2. СС. 283–9.
21. Chen K.H., Liu H., Yang L., Jin M.W., Li G.R. SKF-96365 strongly inhibits voltage-gated sodium current in rat ventricular myocytes // *Pflugers Arch*. 2015. Т. 467, № 6. СС. 1227–36.

22. *Leuner K., Kazanski V., Muller M., Essin K., Henke B., Gollasch M., Harteneck C., Muller W.E.* Hyperforin—a key constituent of St. John's wort specifically activates TRPC6 channels // *FASEB J.* 2007. T. 21, № 14. CC. 4101–11.
23. *Braini C., Bugnicourt G., Villard C.* Neuronal growth from a volume perspective // *Phys Biol.* 2021. T. 18, № 1. CC.
24. *Ahnert-Hilger G., Holtje M., Grosse G., Pickert G., Mucke C., Nixdorf-Bergweiler B., Boquet P., Hofmann F., Just I.* Differential effects of Rho GTPases on axonal and dendritic development in hippocampal neurones // *J Neurochem.* 2004. T. 90, № 1. CC. 9–18.
25. *Henley J., Poo M.M.* Guiding neuronal growth cones using Ca<sup>2+</sup> signals // *Trends Cell Biol.* 2004. T. 14, № 6. CC. 320–30.
26. *Rosner H., Moller W., Wassermann T., Mihatsch J., Blum M.* Attenuation of actinomyosinII contractile activity in growth cones accelerates filopodia-guided and microtubule-based neurite elongation // *Brain Res.* 2007. T. 1176. CC. 1–10.
27. *Gomez T. M., Letourneau P.C.* Actin dynamics in growth cone motility and navigation // *J Neurochem.* 2014. T. 129, № 2. CC. 221–34.
28. *Klein M.G., Kovacs L., Simon B.J., Schneider M.F.* Decline of myoplasmic Ca<sup>2+</sup>, recovery of calcium release and sarcoplasmic Ca<sup>2+</sup> pump properties in frog skeletal muscle // *J Physiol.* 1991. T. 441. CC. 639–71.
29. *Zernov N., Veselovsky A.V., Poroikov V.V., Melentjeva D., Bolshakova A., Popugaeva E.* New Positive TRPC6 Modulator Penetrates Blood-Brain Barrier, Eliminates Synaptic Deficiency and Restores Memory Deficit in 5xFAD Mice // *Int J Mol Sci.* 2022. T. 23, № 21.
30. *Heiser J.H., Schuwald A.M., Sillani G., Ye L., Muller W.E., Leuner K.* TRPC6 channel-mediated neurite outgrowth in PC12 cells and hippocampal neurons involves activation of RAS/MEK/ERK, PI3K, and CAMKIV signaling // *J Neurochem.* 2013. T. 127, № 3. CC. 303–13.
31. *Kunert-Keil C., Bisping F., Kruger J., Brinkmeier H.* Tissue-specific expression of TRP channel genes in the mouse and its variation in three different mouse strains // *BMC Genomics.* 2006. T. 7. C. 159.
32. *Wu X., Babnigg G., Villereal M.L.* Functional significance of human trp1 and trp3 in store-operated Ca(2+) entry in HEK-293 cells // *Am J Physiol Cell Physiol.* 2000. T. 278, № 3. CC. 526–36.
33. *Spector I., Braet F., Shochet N.R., Bubb M.R.* New anti-actin drugs in the study of the organization and function of the actin cytoskeleton // *Microsc Res Tech.* 1999. T. 47, № 1. CC. 18–37.
34. *Ryu H.J., Kim J. E., Yeo S.I., Kim D.W., Kwon O.S., Choi S.Y., Kang T.C.* F-actin depolymerization accelerates clasmotodendrosis via activation of lysosome-derived autophagic astroglial death // *Brain Res Bull.* 2011. T. 85, № 6. CC. 368–73.
35. *Duncan K.L., Duncan M.D., Alley M.C., Sausville E.A.* Cucurbitacin E-induced disruption of the actin and vimentin cytoskeleton in prostate carcinoma cells // *Biochem Pharmacol.* 1996. T. 52, № 10. CC. 1553–60.
36. *Foerster F., Braig S., Chen T., Altmann K.H., Vollmar A.M.* Pharmacological characterization of actin-binding (-)-doliculide // *Bioorg Med Chem.* 2014. T. 22, № 18. CC. 5117–22.

## EFFECT OF TRPC6 MODULATORS ON ACTIN CYTOSKELETON IN VITRO

*D. Melenteva<sup>1</sup>, N. Zernov<sup>1</sup>, V. Ghamaryan<sup>1,2</sup>, L. Hunanyan<sup>1,2</sup>, E. Popugaeva<sup>1,\*</sup>*

*<sup>1</sup>Peter the Great St.Petersburg Polytechnic University, St.Petersburg, Russia*

*<sup>2</sup>Russian-Armenian (Slavonic)University, Yerevan, Armenia*

### ABSTRACT

During experiments, the effect of TRPC6 plasma membrane channel modulators on cellular actin was evaluated. Compounds GSK1702934A, C20, Z12-30 were used as agonists of TRPC6, and SKF96365 were used as blocker. Experiments were performed on primary tip tail fibroblasts culture, as well as on primary hippocampal culture.

**Keywords:** TRPC6, neurodegeneration, actin cytoskeleton, neuron.

DOI 10.24412/1829-0450-2024-1-56-60  
УДК 577

Поступила: 22.03.2024г.  
Сдана на рецензию: 25.03.2024г.  
Подписана к печати: 09.04.2024г.

## СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОГРАММ ПРЕДСКАЗАНИЯ СТРУКТУР БЕЛКОВ НА ПРИМЕРЕ 14-3-3.

*А.К. Паронян*<sup>1,2</sup>

<sup>1</sup>Российско-Армянский (Славянский) университет

<sup>2</sup> Институт молекулярной биологии НАН РА

*adrine.paronyan@rau.am*

ORCID: 0009-0005-9315-6606

### АННОТАЦИЯ

Изучение третичных структур белков очень важно для понимания молекулярных механизмов действия, особенно если данный белок играет ключевую роль в патогенезе заболеваний. На данный момент имеется большой спектр программных пакетов и серверов, позволяющие смоделировать полноразмерные структуры белков, исходя из аминокислотной последовательности. В данной работе были рассмотрены три наиболее используемые и точные программы по предсказанию структур белков: “AlphaFold2”, “Rosetta”, “I-TASSER”.

**Ключевые слова:** моделирование структур белков, “in silico”, “AlphaFold2”, “Rosetta”, “I-TASSER”.

### Введение

Моделирование третичных структур белков является актуальной проблемой для решения задач по дизайну лекарственных препаратов, а также для выяснения механизмов действия белков при патогенезе. Для проведения рентгенографического анализа необходимо вырастить кристаллы белка, поместить один из них под рентгеновский луч и собрать данные дифракции луча, обработав которые можно определить структуру. Некоторые белки легко кристаллизуются, а другие – нет, и исследователи не могут определить их структуру [1]. На данный момент ученые создают и конструируют программы и алгоритмы, которые решают проблему моделирования структур белков, для которых не известны их кристаллографические структуры. Целью данного исследования являлось сравнительный анализ программ моделирования структур белков на примере 14-3-3ε, для выявления наилучшей полноразмерной модели.

Изоформы белка 14-3-3 были первыми белками, идентифицированными как специфически связывающие фосфосерин/фосфотреонин содержащие мо-



тивы [2]. 14-3-3 участвуют в передаче клеточных сигналов, регуляции развития клеточного цикла, нейродегенеративных процессах, апоптозе, канцерогенезе, аутофагии, а также в процессе репликации вирусов [4].

В клетках млекопитающих имеется семь изоформ белка 14-3-3, каждая из которых обозначается греческой буквой ( $\beta$ -бета,  $\gamma$ -гамма,  $\delta$ -дельта,  $\epsilon$ -эпсилон,  $\zeta$ -дзета,  $\theta/\tau$ -тета/тау,  $\eta$ -ета) [5]. Изоформа 14-3-3 $\epsilon$  является наиболее консервативным членом семейства 14-3-3. 14-3-3 $\epsilon$  сильно экспрессируется в головном мозге и взаимодействует с множеством белков, которые могут играть критическую роль в функциях мозга, и общепризнано, что 14-3-3 $\epsilon$  связан с множеством неврологических заболеваний [6].

### Материалы и методы

Для оценки качества программ по моделированию третичных структур существует конкурс CASP (“Critical Assessment of protein Structure Prediction”), который проводится раз в два года и результатом которого является выявление наилучшей смоделированной модели. В связи с этим, в контексте данного исследования были рассмотрены три наиболее распространенных программы по моделированию структур белков. Сервер I-TASSER представляет собой онлайн-платформу, реализующую основанные алгоритмы предсказания структуры и функции белка [7]. Было разработано множество методов для создания белковых структур “ab initio”. Одним из подходов является метод сборки на основе фрагментов, идея Боуи и Айзенберга [8]. В Rosetta использовались библиотеки фрагментов из трех и девяти остатков [9]. AlphaFold2 является программой искусственного интеллекта, разработанная “DeepMind”, которая выполняет прогнозирование структуры белка и разработана как система глубокого обучения [10].

Оценочными критериями достоверности полученной модели принято считать несколько, из них самыми распространенными являются: RMSD (среднеквадратичное отклонение положений атомов), GDT-TS (является критерием сходства между двумя белковыми структурами), TM-score (метрика для оценки топологического сходства белковых структур). Для оценки TM-score и GDT-TS был использован веб-сервер от Zhang

Lab [11]. Для оценки RMSD была использована программа ICM-Pro 3.8-7 [12]. В качестве контрольной модели для проверки была использована доступная кристаллографическая структура из базы данных PDB (PDB ID: 7C8E). Так как для своего нормального функционирования 14-3-3 должен быть в димерной форме, необходимо провести белок-белковый докинг, что было сделано с помощью программы ICM-Pro 3.8-7 [13].

## Результаты и обсуждение

На Рис. 1. указаны полученные смоделированные модели. В случае “AlphaFold2” программа дает возможность создавать комплексы сразу без необходимости дальнейших манипуляций, которые были сделаны после полученных моделей “Rosetta” и “I-TASSER”, а именно белок-белковый докинг с преобразование Фурье (FFT) [13].

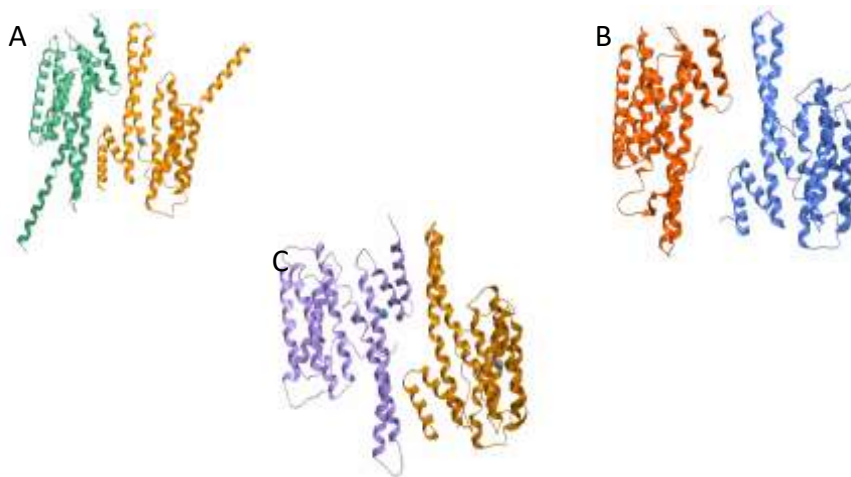


Рисунок 1. Полученные модели. А: 14-3-3ε полученная с помощью “Rosetta”; Б: 14-3-3ε полученная с помощью “AlphaFold2”; В: 14-3-3ε полученная с помощью “I-TASSER”.

Все структуры были суперимпозированы по отношению к кристаллографической структуре PDB ID: 7C8E и получены оценки моделирования, которые указаны в Табл. 1.

Таблица 1. Оценочные значения полученных структур.

Структура Оценка	TM score	RMSD	GDT-TS
<b>14-3-3ε I-TASSER</b>	0.7998	4.4472	0.6473
<b>14-3-3ε Rosetta</b>	0.9830	1.8173	0.9168
<b>14-3-3ε AlphaFold2</b>	0.3881	5.4922	0.2217

Исходя из Табл., можно заметить, что наилучшие результаты по оценкам у программы “Rosetta”, где TM score 0.9830, RMSD 1.8173 и GDT-TS 0.9168. Но стоит отметить и тот факт, что мы рассматривали структуру в димерной

форме, если рассмотреть по мономерам, то наилучший результат по значению RMSD у программы “AlphaFold2” и равняется 1.6343 (TM-score 0.8788, а GDT-TS 0.8422, т.е. по этим двум критериям у “Rosetta” значения немного выше).

Полученные результаты дают нам возможность, проведя сравнительный анализ, выявить наилучшую димерную модель 14-3-3ε, что в дальнейшем позволит проводить моделирование структур остальных изоформ. Несмотря на то, что у “AlphaFold2” результаты за последние года самые высокие и достоверные, стоит учитывать тот факт, что при моделировании и другими программами возможно получение хороших результатов, и как показал анализ на отдельных мономерах, у “AlphaFold2” и “Rosetta” получение оценки очень близки.

#### ЛИТЕРАТУРА

1. Ronda L., Bruno S., Bettati S., Storici P. and Mozzarelli A. (2015). From protein structure to function via single crystal optical spectroscopy *Front. Mol. Biosci.* 2.
2. Obsil T. and Obsilova V. (2011). Structural basis of 14-3-3 protein functions // *Seminars in Cell & Developmental Biology.* 22. 663–72.
3. Shimada T., Fournier A.E. and Yamagata K. (2013). Neuroprotective Function of 14-3-3 Proteins in Neurodegeneration // *BioMed Research International.* 2013. 1–11.
4. Fu H., Subramanian R.R. and Masters S.C. (2000). 14-3-3 Proteins: Structure, Function, and Regulation // *Annu. Rev. Pharmacol. Toxicol.* 40. 617–47.
5. Aghazadeh Y., Zirkin B.R. and Papadopoulos V. (2015). Pharmacological regulation of the cholesterol transport machinery in steroidogenic cells of the testis // *Vitam Horm.* 98. 189–227.
6. Yin S.-J., Lee J.-R., Kwak H., Lee B.-N., Han J.-W., Hahn M.-J., Park Y.-D. and Yang J.-M. (2020). Functional study of 14-3-3 protein epsilon (YWHAЕ) in keratinocytes: microarray integrating bioinformatics approaches // *Journal of Biomolecular Structure and Dynamics.* 38. 2633–49.
7. Yang J., Yan R., Roy A., Xu D., Poisson J., Zhang Y. (2015) The I-TASSER Suite: Protein structure and function prediction // *Nature Methods*, 12: 7–8.
8. Bowie J.U. and Eisenberg D. (1994). An evolutionary approach to folding small alpha-helical proteins that uses sequence information and an empirical guiding fitness function // *Proc. Natl. Acad. Sci. U.S.A.* 91 4436–40.
9. Rohl C.A., Strauss C.E.M., Misura K.M.S. and Baker D. (2004). Protein Structure Prediction Using Rosetta // *Methods in Enzymology*, vol 383 (Elsevier). PP. 66–93.
10. Jumper J., Evans R., Pritzel A., Green T., Figurnov M., Ronneberger O. et al. (2021). Highly accurate protein structure prediction with AlphaFold // *Nature.* 596 583–9.
11. Zhang Y., Skolnick J. (2004). Scoring function for automated assessment of protein structure template quality, *Proteins*, 57: 702–710.
12. Cardozo T., Totrov M. and Abagyan R. (1995). Homology modeling by the ICM method // *Proteins* 23(3): 403–414.
13. Totrov M. and Abagyan R. (1994). Detailed ab initio prediction of lysozyme-antibody complex with 1.6 Å accuracy. *Nat. Struct. Biol.* 1. 259–263.

## A COMPARATIVE ANALYSIS OF PROTEIN STRUCTURE PREDICTION PROGRAMS USING THE EXAMPLE OF 14-3-3.

*A. Paronyan*<sup>1,2</sup>

<sup>1</sup>*Russian-Armenian (Slavonic) University*

<sup>2</sup>*Institute of Molecular Biology NAS RA*

### ABSTRACT

Identifying the tertiary structures of proteins is very important for understanding the molecular mechanisms of action, especially if a given protein plays a key role in disease pathogenesis. A wide range of software packages and servers are currently available to model the full-length structures of proteins based on the amino acid sequence. In this paper, we reviewed the three most common programs for predicting protein structures: AlphaFold2, Rosetta, and I-TASSER.

**Keywords:** protein structure modeling, in silico, prediction methods, AlphaFold2, Rosetta, I-TASSER.

DOI 10.24412/1829-0450-2024-1-61-67  
УДК 515.112

Поступила: 12.04.2024г.  
Сдана на рецензию: 12.04.2024г.  
Подписана к печати: 19.04.2024г.

## MULTI-OMICS PORTRAYAL OF BREAST CANCERS

*S. Davitavyan*<sup>1,2</sup>

<sup>1</sup>*Institute of Biomedicine and Pharmacy, Russian-Armenian University*

<sup>2</sup>*Research Group of Bioinformatics, Institute of Molecular Biology NAS RA*  
*surendavitavyan98@gmail.com*

### ABSTRACT

The factors that influence tumor development are diverse and ambiguous. Consequently, there is a necessity for the development of new data-mining approaches that will contribute to efficient disease diagnostics, prognostics, and monitoring of disease courses.

This study is aimed to identify patterns of genetic changes that provide changes in expression in tumor samples. Multiomics data for analyzing the main manifestations of genetic shifts were obtained from the GDC database. We used a multilayer self-organizing map (ML-SOM) algorithm for clustering and dimension reduction. This approach allows for distinguishing the sets of factors associated with cancer, and on the other hand, it provides information on the distribution of multilayer genetic data among tumor samples.

**Keywords:** breast cancer, multi-omics data, self-organizing maps.

### Introduction

Breast cancer is the most common cancer worldwide with 7.8 million women alive as of the end of 2020 who had received a diagnosis within the previous five years. Women in diverse age groups are at risk of receiving a breast cancer diagnosis, moreover, the development of this disease is directly correlated with aging [1]. The implication of new approaches is becoming imperative because of the diverse course across the different subtypes of ductal cancer. Receptor-based molecular classification methods continue to be applied for modern clinical purposes.

Nevertheless, new classification approaches have been developed. One of the examples is transcriptomics-based PAM50 classification which uses expression signals from 50 genes to classify breast cancer[2]. Multi “-omics” data has revealed new ways to understand the covert relationships between genomic features and serves as a basis for developing machine learning models capable of comprehensive analysis of big biological data.

In our previous studies, we showcased the advantages of self-organizing map pipelines in analyzing and describing multi-layer datasets[3]. Low-grade gliomas

study [4], B-cell lymphomas study[5], and others prove the credibility of the SOM method to depict the origins and further progression of the illness.

To unveil the cornerstone of the Breast Carcinoma development process, we have used multi-SOM machine learning algorithms on gen-, transcriptomic, and epigenetic data.

## **Materials and methods**

### ***Study Datasets***

In this study, we used available-omic datasets of the TCGA-BRCA project [6]. Total, RNA-seq counts, microarray promoter and gene body methylation, microarray CNV, and SNV were obtained for 996 samples.

### ***Data preprocessing***

RNA-seq counts were subjected to library size normalization and converted to log counts using variant stabilization transformation.

Promoter methylation data was converted from betta to m values.

CNVs were normalized by adding small numbers to avoid constant values.

SNVs were summarized by genes.

Also, we have divided samples into subtypes by PAM50 gene signature classification, which separates breast cancer into Basal (N=161), LumA (N=480), LumB (N=199), HER2E (N= 74), normal-like (N= 32), True Normal (N= 46), and CLOW (N= 4) groups.

### ***Integrated analysis of cancer molecular features with multi-layer SOM***

To conduct a comprehensive integrative analysis of -omics datasets of breast cancer, we employed a refined multi-layer self-organizing maps (ml-SOM) approach, extending our prior research. The self-organizing map (SOM) algorithm is a neural network-based technique for dimensionality reduction and clustering. Numerous studies have demonstrated its effectiveness in grouping genes based on their expression profiles into co-expressed modules. Furthermore, the SOM implementation in the “oposSOM” package is enriched with robust function mining capabilities, facilitating the assignment of biological functions to gene clusters. This capability efficiently reduces the high-dimensional gene space into numerous differentially expressed functional modules. As a result, this approach enables a seamless transition from analyzing individual genes to conducting systems-level analyses while preserving the integrity of the original information.

In this approach, we organized all the datasets into distinct layers and trained them collectively on a single SOM grid, similar to a classical single-layer SOM (sl-SOM) [7].

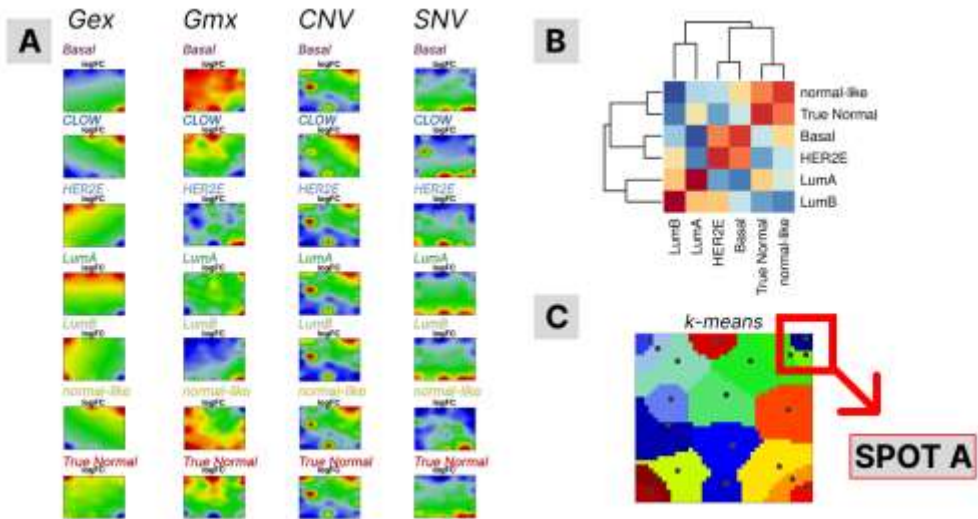
### **Analysis of survival**

Survival analysis (overall survival, disease-free survival, disease-specific survival, and progression-free survival) was performed using the Cox proportional

hazards regression using *contsurvplot* R package [8]. The model includes survival as a dependent variable and group information, and spot -omic profiles as predictors.

## Results

The ml-SOM algorithm not only highlighted the differences between subtypes, as we expected, but showed relationships between data categories as well. (Figure 1A). By primary visual comparison of group portraits, we have identified specific spots for subtypes, meanwhile, we have observed several common clusters between these groups. To confirm the last hypothesis, we have conducted a correlational analysis of the expressional levels for each subtype (Figure 1B). The analysis indicates the similarity of True Normal and normal-like, Basal and HER2E groups. There was a weak correlation between LumA and Basal, as well as LumB and normal-like groups.



*Figure 1. Visualization of SOM group portraits (A). From left to right are group portraits of transcriptomics, methylation, copy number variations, and single nucleotide variations, respectively. Red spots belong to highly expressed clusters, which are the target for research. Correlations between groups in transcriptomic level (B). True Normal and normal-like, also Basal and HER2E subtypes show a high correlation. K-means defined gene clusters (C). Spot A differs not only between subtypes but also between data categories.*

Furthermore, gene clusters based on k-means rates were investigated (Figure 1C). This part of the study aimed to identify specific transcriptome-level changes across subtypes and suggest possible reasons for these changes by uncovering the influence of other categories of data. We have chosen Spot A because of its specificity in several

subtypes and different values among data categories. For unveiling the dependents of gene expression levels from methylation, CNV, and SNV data categories we have trained linear regression models. The methylation profile was significantly negatively correlated with expression levels in basal and normal-like groups, while a positive association with CNV was observed in lum A, lum B, and HER2 cancers. No significant trend was observed for SNV in all cancer subtypes (Figure 2).

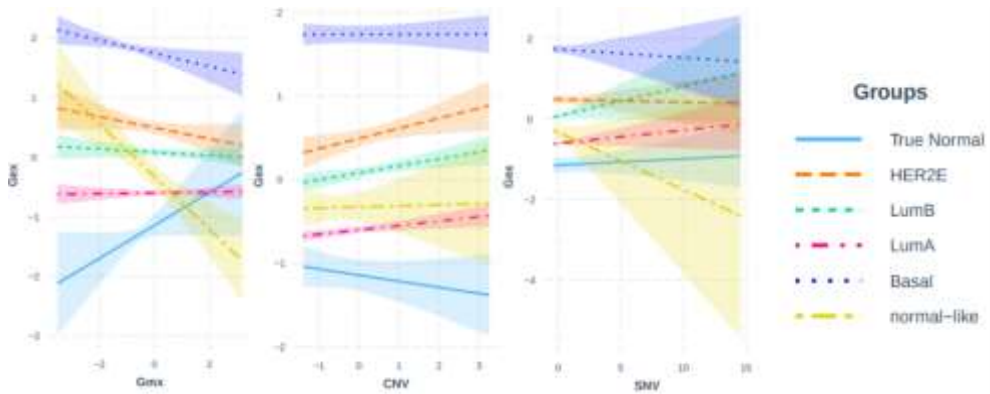


Figure 2. Representation of trained linear regression models. In the Y-axis are placed log transformed Gex values, from left to right X-axis refers to Gmx, CNV, and SNV data, and colors describe subtypes.

The cluster contains 114 genes, to analyze this number of genes we performed gene enrichment analysis using the online tool “Webgestalt” (Figure 3) [9]. We used KEGG pathways [10] to understand the likely phenotypic changes due to the mixing of changes at different levels in genes from spot A.

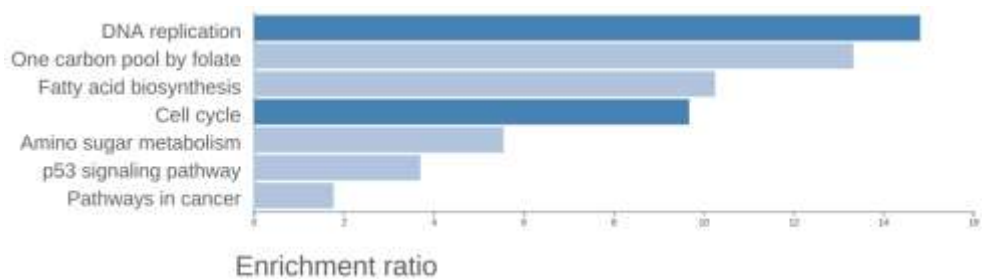


Figure 3. Gene enrichment analysis. KEGG pathways differed in genes from selected spots. The enrichment ratio shows the significance of the results.

The enrichment shows significance with the DNA replication[11], One carbon pool by folate[12], and Cell cycle pathways[13].



Also, survival analysis of samples showed that the LumA and LumB subtypes better survived than HER2E and Basal [14] (Figure 4).

Cross-checking the results of gene enrichment and survival analysis against literature sources showed the applicability of using the approach on multi-omics data.

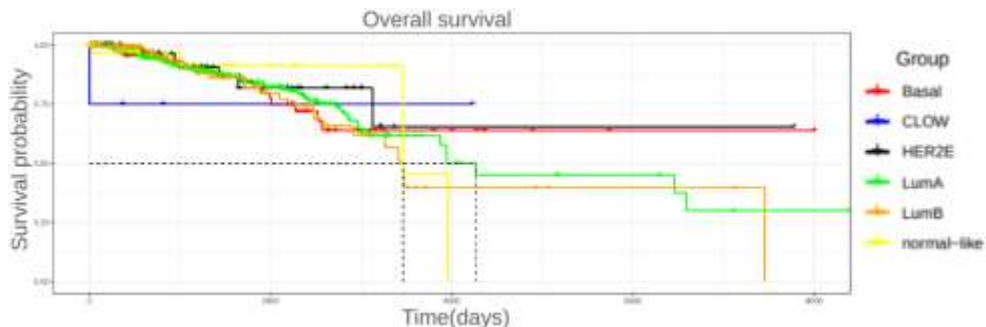


Figure 4. Survival plot of breast subtypes based on transcriptomic level. The colors refer to subtypes. Some of the subtypes do not provide sufficient significant results due to the small number of patients.

## Conclusions

Machine learning methods enable a comprehensive study of the nature of breast cancer, identifying the factors that determine the levels of gene expression.

The methylation profile exhibited a significant negative correlation with expression levels in basal and normal-like groups, while a positive association with CNV was observed in lum A, lum B, and HER2 cancers. No significant trend was observed for SNV in all cancer subtypes.

LumA and LumB subtypes have a better survival probability than other subtypes.

## REFERENCES

1. <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>
2. Chia Stephen K., et al. "A 50-gene intrinsic subtype classifier for prognosis and prediction of benefit from adjuvant tamoxifen" // *Clinical cancer research* 18.16 (2012): 4465-4472.
3. Löffler-Wirth H., Martin Kalcher and Hans Binder. "oposSOM: R-package for high-dimensional portraying of genome-wide expression landscapes on bioconductor". // *Bioinformatics* 31.19 (2015): 3225–3227.
4. Binder, Hans et al. "Integrated Multi-Omics Maps of Lower-Grade Gliomas". // *Cancers* vol. 14,11 2797. 4 Jun. 2022, DOI:10.3390/cancers14112797.
5. Hopp L., Nersisyan L., Löffler-Wirth H., Arakelyan A., & Binder H. (2015). Epigenetic Heterogeneity of B-Cell Lymphoma: Chromatin Modifiers // *Genes*, 6 (4), 1076–1112. <https://doi.org/10.3390/genes6041076>
6. *Cancer Genome Atlas Research Network et al.* "The Cancer Genome Atlas Pan-Cancer analysis project." // *Nature genetics* vol. 45,10 (2013): 1113–20. doi:10.1038/ng.2764

7. Löffler-Wirth H. et al. “oposSOM: R-package for high-dimensional portraying of genome-wide expression landscapes on bioconductor”. // *Bioinformatics* (Oxford, England) vol. 31,19 (2015): 3225-7. doi:10.1093/bioinformatics/btv342
8. Denz R.; Timmesfeld Ninaa. Visualizing the (Causal) Effect of a Continuous Variable on a Time-To-Event Outcome. *Epidemiology* 34(5): PP. 652–660, September 2023. | DOI: 10.1097/EDE.0000000000001630
9. Jing Wang, Suhas Vasaikar, Zhiao Shi, Michael Greer, Bing Zhang, WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit // *Nucleic Acids Research*, Volume 45, Issue W1, 3 July 2017, Pages W130–W137, <https://doi.org/10.1093/nar/gkx356>
10. Kanehisa Minoru et al. “KEGG for taxonomy-based analysis of pathways and genomes.” // *Nucleic acids research* vol. 51, D1 (2023): D587–D592. doi:10.1093/nar/gkac963
11. Rajamanickam S., Park J.H., Subbarayalu P. et al. Targeting aberrant replication and DNA repair events for treating breast cancers // *Commun Biol* 5, 493 (2022). <https://doi.org/10.1038/s42003-022-03413-w>
12. Xu Xinran and Jia Chen. “One-carbon metabolism and breast cancer: an epidemiological perspective”. *Journal of genetics and genomics = Yi chuan xue bao* vol. 36,4 (2009): 203–14. doi:10.1016/S1673-8527(08)60108-3
13. Thu K L et al. “Targeting the cell cycle in breast cancer: towards the next phase.” // *Cell cycle* (Georgetown, Tex.) vol. 17,15 (2018): 1871–1885. doi:10.1080/15384101.2018.1502567
14. Nguyen Chu Van et al. “Molecular classification predicts survival for breast cancer patients in Vietnam: a single institutional retrospective analysis.” // *International journal of clinical and experimental pathology* vol. 14,3 322–337. 1 Mar. 2021.

## МУЛЬТИ-ОМИКС ИССЛЕДОВАНИЕ ОПУХОЛИ МОЛОЧНОЙ ЖЕЛЕЗЫ

С.С. Давитавян<sup>1,2</sup>

<sup>1</sup>Институт Биомедицины и Фармации  
Российско-Армянский (Славянский) университет

<sup>2</sup>Исследовательская группа Биоинформатики,  
Институт Молекулярной Биологии НАН РА

### АННОТАЦИЯ

Факторы, влияющие на развитие опухоли, разнообразны и неоднозначны. Существует необходимость в разработке новых подходов к сбору данных, которые будут способствовать эффективной диагностике заболеваний, прогнозированию и мониторингу прогресса заболеваний. Целью данного исследования является выявление закономерностей генетических изменений, которые обуславливают изменения экспрессии в опухолевых образцах. Мульти-омиксные данные для анализа основных проявлений генетических изменений были получены из базы данных GDC. Мы использовали алгоритмы самоорганизующейся карты (ML-SOM) для кластеризации и уменьшения размерности данных. Такой подход позволяет выделить наборы факторов, связанных с опухолью и предоставляет информацию о распределении мульти-омиксных генетических данных среди образцов.

**Ключевые слова:** опухоль молочной железы, мульти-омиксные данные, самоорганизующиеся карты

DOI 10.24412/1829-0450-2024-1-67-74  
УДК 2788

Поступила: 03.05.2024г.  
Сдана на рецензию: 06.05.2024г.  
Подписана к печати: 14.05.2024г.

## ВАЛИДАЦИЯ ПРОГНОСТИЧЕСКОЙ СИЛЫ ПОДТИПОВ МИКРООКРУЖЕНИЯ КРУПНОКЛЕТОЧНОЙ ЛИМФОМЫ ПРИ ЛЕЧЕНИИ ОБИНУТУЗУМАБОМ В ПЕРВОЙ ЛИНИИ

*Н.Ю. Котлов*

*“Boston Gene Technologies” ООО  
nikita.kotlov@bostongene.com  
ORCID: 0000-0002-6393-7357*

### АННОТАЦИЯ

Стандартным лечением диффузной крупноклеточной лимфомы все еще остается комбинация иммуно-химиопрепаратов R-СНОР (ритуксимаб, циклофосфамид, доксирубицин, винкристин, прелнизон). При этом до 40% пациентов не достигают стабильной ремиссии. Обинутузумаб – моноклональное антитело против В-клеточного рецептора CD20, которое тестируется в клинических испытаниях в качестве альтернативы ритуксимабу в виде комбинации G-СНОР. Микроокружение лимфом играет значительную прогностическую роль при лечении R-СНОР в первой линии. В данном исследовании была подтверждена ассоциация подтипов LME-GC, LME-MS с лучшим прогнозом, чем у подтипа LME-DE при лечении комбинацией G-СНОР в первой линии на данных клинического испытания GOYA. **Ключевые слова:** диффузная крупноклеточная лимфома, микроокружение лимфом, прогностические биомаркеры.

### Введение

Диффузная крупноклеточная лимфома (DLBCL) является наиболее распространенной формой неходжкинских лимфом, и составляет 25–35% случаев. Частота заболеваемости в США и Восточной Европе оценивается в 5,5 человек на 100000. А пятилетняя выживаемость составляет около 65% [12]. Текущий стандарт лечения на первой линии – это комбинация ритуксимаба (моноклональное антитело против CD20) с химиотерапией (R-СНОР) для всех пациентов [10]. При таком лечении все еще существенный процент пациентов ~40% остается неизлечимым и умирает вследствие болезни. На текущий момент уже разработано более 10 разных терапий (включая AZA, CAR-T, Bites, IBN и т.д.), которые более эффективны в зависимости от некоторых молекулярных параметров пациентов [2, 3, 8, 9].

В 2000 году Ash Alizadeh и соавторы выделили два основных транскрипционных подтипа DLBCL: GCB – происходящие из клеток герминативного

центра, можно дальше поделить на центроциты и центробласты. ABC – происходящие из активированных В клеток или уже из вышедших из герминативной реакции. Пациенты с лимфомами типа GCB обычно имеют более благоприятный прогноз как в эпоху до использования ритуксимаба, так и в настоящее время. Также существует третий подтип СОО, известный как неклассифицированный (UNK) или тип 3, который лишен явных паттернов экспрессии генов, которые бы классифицировали его как GCB или ABC [1, 15].

Известно, что опухолевые В клетки активно взаимодействуют с другими клетками организма в лимфоузлах (опухолевым микроокружением) и часто влияют на композицию и поляризацию иммунных и стромальных клеток [11, 5]. Мы в 2021 году опубликовали 4 подтипа микроокружения DLBCL и продемонстрировали, что стандартная иммунохимиотерапия имеет различный успех в зависимости от подтипа, а также и то, что разные подтипы обогащены генетическими и транскрипционными изменениями, которые могут говорить о чувствительности на другие препараты [6].

Обинутузумаб, CD20 – направленное антитело, аналог ритуксимаба активно тестируется в качестве замены ритуксимаба на первой линии лечения [7, 13, 4]. По результатам рандомизированного исследования третьей фазы GOYA, обинутузумаб показал сравнимые результаты по OS/PFS с ритуксимабом [13, 14].

В данной работе мы задаемся вопросом, есть ли для G-СНОР аналогичная прогностическая ассоциация подтипов микроокружения с выживаемостью, как для R-СНОР.

## **Материалы и методы**

*Данные экспрессии генов.* Данные экспрессии генов исследования GOYA были получены из базы данных SRA (SRP183071) и пересчитаны из FASTQ файлов-методом, описанным ранее [6]. Анализ качества данных секвенирования РНК был проведен с помощью FastQC v0.11.5, FastQ Screen v0.11.1. Образцы, не прошедшие контроль качества [6], были исключены.

*Классификация подтипов.* Типирование подтипов СОО, LME было проведено методом из оригинальной статьи [6].

*Статистический анализ.* Анализ выживаемости был проведен с помощью теста “log rank”. Анализ пере-представленности процентов классов был проведен с помощью теста Фишера или теста хи-квадрат.

## Результаты и обсуждение

Для того чтобы проверить ассоциации подтипов ЛМЕ с выживаемостью пациентов на обинутузумаб-химии (G-СНОР) с ритуксимаб-химии (R-СНОР), мы использовали данные экспрессии генов из клинического испытания GOYA. После перерасчетов и контроля качества осталось 212 G-СНОР и 224 (R-СНОР) образца. Для сравнения также взяли Мета-когорту из исследования [6].

*Прогностическая сила СОО.* Для начала мы проверили, что состав анализируемой когорты ожидаем с точки зрения молекулярных параметров. Проценты СОО (GCB/ABC) был ожидаемый (50% и 55%) и не отличался между когортами G-СНОР и R-СНОР ( $p=0.5$ ), в то время как процент ABC был значимо повышен в подкогортах GOYA – относительно Мета когорты (0.0009 G-СНОР GOYA против Мета когорты, 0.03 R-СНОР GOYA против Мета когорты (Рис. 1).

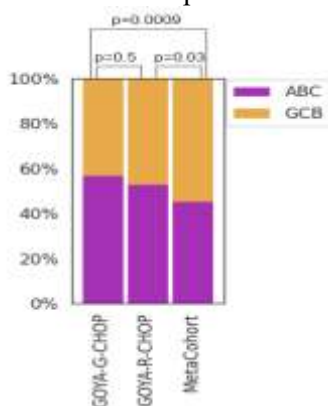


Рисунок 1. Распределение процентов СОО подтипов в подкогортах G-СНОР и R-СНОР исследования GOYA и Мета когорте. Показанные  $p$ -значения подсчитаны точным тестом Фишера.

Выживаемость без прогрессии GCB против ABC была значимо лучше в когорте G-СНОР ( $p=0.03$ ), в когорте R-чоп наблюдался тренд ( $p=0.1$ ). Во всех трех когортах выживаемость подтипа GCB была лучше, чем подтипа ABC (Рис. 2).

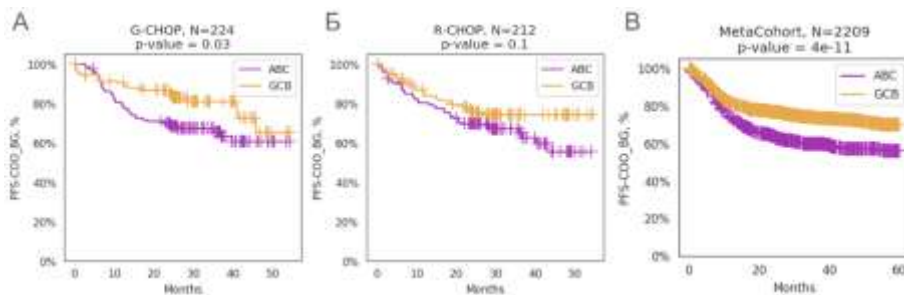


Рисунок 2. Сравнение выживаемости без прогрессии между группами ABC и GCB. А – когорта G-CHOP GOYA, Б – когорта R-CHOP GOYA, В – Мета-когорта. В заголовках представлены р-значения лог-ранк теста.

*Прогностическая сила подтипов LME.* Далее, каждому образцу был назначен LME подтип по методу из оригинальной статьи [6]. В когорте G-CHOP GOYA было больше LME-IN и LME-DP (63.8%), чем в когорте R-CHOP GOYA (54.2%,  $p=0.08$ ) и значительно больше, чем ожидаемое из Мета когорты (51.8%,  $p=0.002$ ). В когорте R-CHOP GOYA проценты подтипов значительно не отличались от ожидаемых из Мета когорты ( $p=0.6$ ).

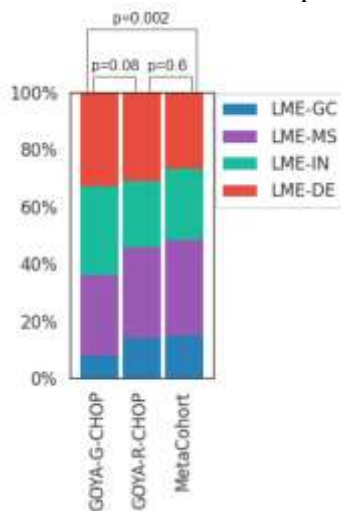


Рисунок 3. Распределение процентов LME подтипов в подкогортах G-CHOP и R-CHOP исследования GOYA и Мета-когорте. Показанные р-значения подсчитаны тестом хи-квадрат.

Ожидаемый прогноз отличается для каждого подтипа в зависимости от подтипа микроокружения (Рис. 4В, Рис. 5В). В когорте G-CHOP мы обнаружили ожидаемую ассоциацию выживаемости без прогрессии и общей выживаемости: подтипы LME-GC, LME-IN, LME-MS, по сравнению с подтипом

LME-DP, имели значимо лучшую общую выживаемость ( $p=0.002$ ) и выживаемость без прогрессии ( $p=0.03$ ) (Рис. 4А, Рис. 5А). Неожиданно, но в когорте пациентов, лечимых ритуксимабом, не было обнаружено ожидаемого расхождения кривых выживаемости ( $p=0.6$ ,  $p=0.3$ ) (Рис. 4Б, Рис. 5Б).

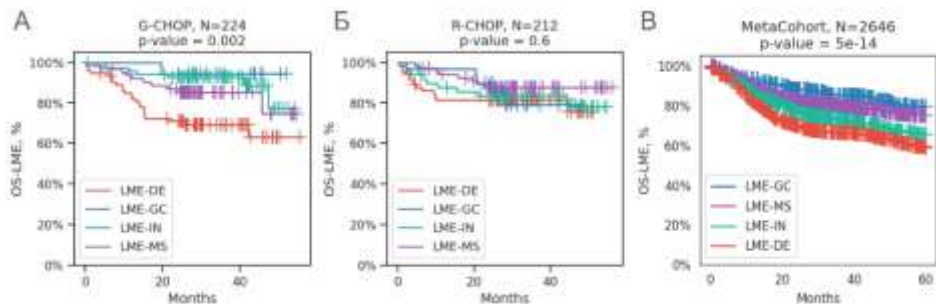


Рисунок 4. Сравнение общей выживаемости между группами LME. А – когорта G-CHOP GOYA, Б – когорта R-CHOP GOYA, В – Мета когорта. В заголовках представлены p-значения четырехстороннего лог-ранк теста.

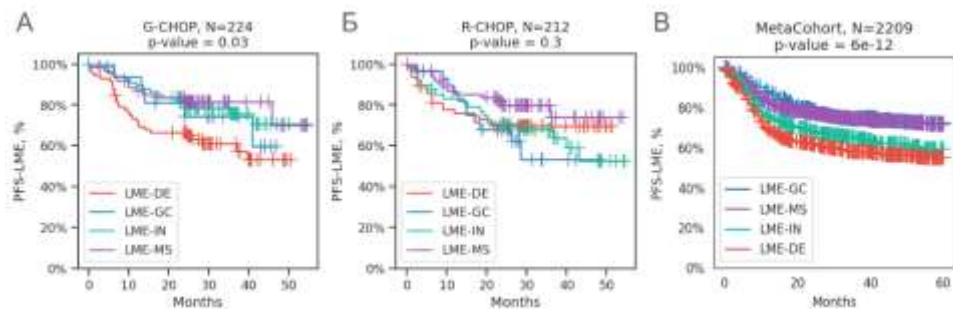


Рисунок 5. Сравнение выживаемости без прогрессии между группами LME. А – когорта G-CHOP GOYA, Б – когорта R-CHOP GOYA, В – Мета когорта. В заголовках представлены p-значения четырехстороннего лог-ранк теста.

## Заключение

Мы проанализировали прогностическую силу подтипов лимфоного микроокружения LME на данных рандомизированного клинического исследования GOYA. В результате мы подтвердили гипотезу, что при лечении обинутузумабом сохраняется прогностическая сила LME подтипов. Также мы показали прогностическую силу подтипов СОО. В когорте пациентов, леченных ритуксимабом, не удалось найти значимых отличий по выживаемости не только в подтипах LME, но даже между СОО, что может говорить об особенностях этой подкогорты.

## ЛИТЕРАТУРА

1. Alizadeh A.A., Eisen M.B., Davis R.E., Ma C., Lossos I.S., Rosenwald A., Boldrick J.C., Sabet H., Tran T., Yu X., Powell J.I., Yang L., Marti G.E., Moore T., Hudson J., Lu L., Lewis D.B., Tibshirani R., Sherlock G., ... Staudt L.M. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, 403, 6769, 503–511, 2000.
2. Atallah-Yunes S.A., Robertson M.J., Davé U.P., Ghione P., Perna F. Novel Immune-Based treatments for Diffuse Large B-Cell Lymphoma: The Post-CAR T Cell Era. *Frontiers in Immunology*, 13, 901365, 1–10, 2022.
3. Cheson B.D., Nowakowski G., Salles G. Diffuse large B-cell lymphoma: new targets and novel therapies. *Blood Cancer Journal*, 11, 4, 68, 2021.
4. Davies A., Kater A.P., Sharman J.P., Stilgenbauer S., Vitolo U., Klein C., Parreira J., Salles G. (2022). Obinutuzumab in the treatment of B-cell malignancies: a comprehensive review. *Future Oncology (London, England)*, 18, 26, 2943–2966, 2022.
5. Kline J., Godfrey J., Ansell S.M. The immune landscape and response to immune checkpoint blockade therapy in lymphoma. *Blood*, 135, 8, 523–533, 2020.
6. Kotlov N., Bagaev A., Revuelta M.V., Phillip J.M., Cacciapuoti M.T., Antysheva Z., Svelolkin V., Tikhonova E., Mihecheva N., Kuzkina N., Nos G., Tabbo F., Frenkel F., Ghione P., Tsiper M., Almog N., Fowler N., Melnick A.M., Leonard J.P., ... Cerchetti L. Clinical and biological subtypes of B-cell lymphoma revealed by microenvironmental signatures. *Cancer Discovery*, 11, 6, 1468–1489, 2021.
7. Le Gouill S., Ghesquières H., Oberic L., Morschhauser F., Tilly H., Ribrag V., Lamy T., Thieblemont C., Maisonneuve H., Gressin R., Bouhabdallah K., Haioun C., Damaj G., Fornecker L., Bouhabdallah R., Feugier P., Sibon D., Cartron G., Bonnet C., ... Casasnovas R.O. Obinutuzumab vs rituximab for advanced DLBCL: a PET-guided and randomized phase 3 study by LYSA. *Blood*, 137, 17, 2307–2320, 2021.
8. Lu T., Zhang J., Xu-Monette Z.Y., Young K.H. The progress of novel strategies on immune-based therapy in relapsed or refractory diffuse large B-cell lymphoma. *Experimental Hematology & Oncology*, 12, 1, 72, 1–38, 2023.
9. Paillassa J., Safa F. Novel biologic therapies in relapsed or refractory diffuse large B cell lymphoma: CAR-T is not the only answer. *Leukemia Research Reports*, 17, 100282, 1–6, 2022.
10. Sarkozy C., Sehn L.H. New drugs for the management of relapsed or refractory diffuse large B-cell lymphoma. *Annals of Lymphoma*, 3, 10, 1–19, 2019.
11. Scott D.W., Gascoyne R.D. The tumour microenvironment in B cell lymphomas. *Nature Reviews. Cancer*, 14(8), 517–534, 2014.
12. SEER. Cancer Stat Facts: NHL – Diffuse Large B-Cell Lymphoma (DLBCL), 2023: <https://seer.cancer.gov/statfacts/html/dlbcl.html>
13. Sehn L.H., Martelli M., Trněný M., Liu W., Bolen C.R., Knapp A., Sahin D., Sellam G., Vitolo U. A randomized, open-label, Phase III study of obinutuzumab or rituximab plus CHOP in patients with previously untreated diffuse large B-Cell lymphoma: final analysis of GOYA. *Journal of Hematology & Oncology*, 13, 71, 1–9, 2020.
14. Sehn L.H., Martelli M., Trněný M., Liu W., Bolen C.R., Knapp A., Sahin D., Sellam G., Vitolo U. Final Analysis of GOYA: A Randomized, Open-Label, Phase III Study of Obinutuzumab



- or Rituximab Plus CHOP in Patients with Previously Untreated Diffuse Large B-Cell Lymphoma. *Blood*, 134, Supplement\_1, 4088–4088, 2019.
15. Wright G., Tan B., Rosenwald A., Hurt E.H., Wiestner A., Staudt L.M. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 17, 9991–9996, 2003.

**LYMPHOMA MICROENVIRONMENT SUBTYPES PROGNOSTIC POWER VALIDATION IN FRONTLINE OBINUTUZUMAB TREATED PATIENTS WITH DIFFUSE LARGE B-CELL LYMPHOMA**

*N.Yu. Kotlov*

*“Boston Gene Technologies” LLC*

**ABSTRACT**

The standard treatment for diffuse large B-cell lymphoma still remains the combination of immunochemotherapy agents R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone). However, up to 40% of patients do not achieve stable remission. Obinutuzumab – a monoclonal antibody against the B-cell receptor CD20, is being tested in clinical trials as an alternative to rituximab in the form of the combination G-CHOP. The microenvironment of lymphomas plays a significant prognostic role in the treatment with R-CHOP in the first line. In this study, the association of the LME-GC and LME-MS subtypes with a better prognosis was confirmed compared to the LME-DE subtype when treated with the combination of G-CHOP in the first line based on data from the clinical trial GOYA.

**Keywords:** Diffuse large B-cell lymphoma – lymphoma microenvironment – prognostic biomarkers.

## СВЕДЕНИЯ ОБ АВТОРАХ

**Алтунян В.Н.** – аспирант кафедры дискретной математики и теоретической информатики Ереванского государственного университета.

**Апинян Л.А.** – студент магистратуры второго курса по направлению подготовки «Прикладная математика и информатика» Российско-Армянского (Славянского) университета.

**Арамян Р.Г.** – д.ф.-м.н., профессор, заведующий кафедрой математической кибернетики Российско-Армянского (Славянского) университета.

**Багдасарян Т.** – студент первого курса магистратуры по специальности «Прикладная статистика и наука о данных» Ереванского государственного университета.

**Давитавян С.С.** – аспирант по специальности «Математическая биология, биоинформатика» Российско-Армянского (Славянского) университета

**Дарбинян А.А.** – к.ф.-м.н., доцент, директор Института математики и информатики Российско-Армянского (Славянского) университета.

**Зернов Н.И.** – аспирант по направленности «Молекулярная биология» Санкт-Петербургского политехнического университета имени Петра Великого, инженер-исследователь лаборатории молекулярной нейродегенерации.

**Камарян В.С.** – аспирант кафедры биоинженерии, биоинформатики и молекулярной биологии, научный сотрудник лаборатории структурной биоинформатики Российско-Армянского (Славянского) университета.

**Киракосян Л.А.** – аспирант(-ка) первого года обучения по специальности «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей» Российско-Армянского (Славянского) университета.

**Котлов Н.Ю.** – сотрудник ООО “BostonGene Technologies”.

**Манукян А.** – аспирант кафедры численного анализа и математического моделирования Ереванского государственного университета.

**Мелентьева Д.М.** – магистрант по направлению подготовки «Медицинская биоинженерия» Санкт-Петербургского политехнического университета имени Петра Великого.

---

**Мелконян В.Г.** – аспирант первого года обучения по специальности «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей» Российско-Армянского (Славянского) университета.

**Оганесян О.** – аспирантка первого года обучения по специальности «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей» Российско-Армянского (Славянского) университета.

**Паронян А.К.** – аспирант по специальности «Математическая биология, биоинформатика» Российско-Армянского (Славянского) университета.

**Попугаева Е.А.** – д.б.н., старший научный сотрудник лаборатории молекулярной нейродегенерации Санкт-Петербургского политехнического университета имени Петра Великого.

**Саакян В.Р.** – аспирант первого года обучения по специальности «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей» Российско-Армянского (Славянского) университета.

**Саргсян С.С.** – к.ф.-м.н., заведующий кафедрой системного программирования Российско-Армянского (Славянского) университета.

**Унанян Л.С.** – к.б.н., руководитель лаборатории структурной биоинформатики Российско-Армянского (Славянского) университета.

**Фаградян А.Р.** – студент четвертого курса по направлению подготовки «Прикладная математика и информатика» Российско-Армянского (Славянского) университета.

**ВЕСТНИК РОССИЙСКО-АРМЯНСКОГО УНИВЕРСИТЕТА.  
СЕРИЯ: ФИЗИКО-МАТЕМАТИЧЕСКИЕ  
И ЕСТЕСТВЕННЫЕ НАУКИ,  
№1 2024г.**

**VESTNIK OF RUSSIAN-ARMENIAN UNIVERSITY  
SERIES: PHYSICAL-MATHEMATICAL  
AND NATURAL SCIENCES,  
No. 1/2024**

*Главный редактор РНИ – М.Э. Авакян  
Корректор – А.С. Есяян  
Компьютерная верстка – А.С. Бжикян*

*Editor-in-Chief of the EOSP – M.E. Avakyan  
Proofreader – A.S. Yesayan  
Computer layout – A.S. Bzhikyan*

Адрес Редакции научных изданий  
Российско-Армянского университета:  
0051, г. Ереван, ул. Овсепя Эмина, 123  
тел./факс: (+374 12) 77-57-75 (внутр. 392)  
e-mail maria.avakian@rau.am

Address of the Editorial Board of Scientific  
Publications of the Russian-Armenian University:  
0051, Yerevan, st. Hovsep Emin, 123  
tel./fax: (+374 12) 77-57-75 (ext. 392)  
e-mail: maria.avakian@rau.am

Заказ № 14  
Подписано к печати 16.04.2024г.  
Формат 70x100<sup>1</sup>/<sub>16</sub>. Бумага офсетная № 1.  
Объем 11.9 усл. п.л. Тираж 100 экз.

Order No. 14  
Signed for publication on April 16, 2024  
Format 70x100<sup>1</sup>/<sub>16</sub>. Offset paper No. 1.  
Volume 11.9 conv. p.l. Circulation 100 copies