

ISSN 2518-1726 (Online),  
ISSN 1991-346X (Print)



«ҚАЗАҚСТАН РЕСПУБЛИКАСЫ  
ҰЛТТЫҚ ҒЫЛЫМ АКАДЕМИЯСЫ» РҚБ  
«ХАЛЫҚ» ЖҚ

# Х А Б А Р Л А Р Ы

---

---

## ИЗВЕСТИЯ

РОО «НАЦИОНАЛЬНОЙ  
АКАДЕМИИ НАУК РЕСПУБЛИКИ  
КАЗАХСТАН»  
ЧФ «Халық»

## N E W S

OF THE ACADEMY OF SCIENCES  
OF THE REPUBLIC OF  
KAZAKHSTAN  
«Halyk» Private Foundation

**SERIES**  
**PHYSICS AND INFORMATION TECHNOLOGY**

**4 (348)**

**OCTOBER – DECEMBER 2023**

PUBLISHED SINCE JANUARY 1963  
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK



## ЧФ «ХАЛЫҚ»

В 2016 году для развития и улучшения качества жизни казахстанцев был создан частный Благотворительный фонд «Халық». За годы своей деятельности на реализацию благотворительных проектов в областях образования и науки, социальной защиты, культуры, здравоохранения и спорта, Фонд выделил более 45 миллиардов тенге.

Особое внимание Благотворительный фонд «Халық» уделяет образовательным программам, считая это направление одним из ключевых в своей деятельности. Оказывая поддержку отечественному образованию, Фонд вносит свой посильный вклад в развитие качественного образования в Казахстане. Тем самым способствуя росту числа людей, способных менять жизнь в стране к лучшему – профессионалов в различных сферах, потенциальных лидеров и «великих умов». Одной из значимых инициатив фонда «Халық» в образовательной сфере стал проект *Ozgeris powered by Halyk Fund* – первый в стране бизнес-инкубатор для учащихся 9-11 классов, который помогает развивать необходимые в современном мире предпринимательские навыки. Так, на содействие малому бизнесу школьников было выделено более 200 грантов. Для поддержки талантливых и мотивированных детей Фонд неоднократно выделял гранты на обучение в Международной школе «Мирас» и в *Astana IT University*, а также помог казахстанским школьникам принять участие в престижном конкурсе «*USTEM Robotics*» в США. Авторские работы в рамках проекта «Тәлімгер», которому Фонд оказал поддержку, легли в основу учебной программы, учебников и учебно-методических книг по предмету «Основы предпринимательства и бизнеса», преподаваемого в 10-11 классах казахстанских школ и колледжей.

Помимо помощи школьникам, учащимся колледжей и студентам Фонд считает важным внести свой вклад в повышение квалификации педагогов, совершенствование их знаний и навыков, поскольку именно они являются проводниками знаний будущих поколений казахстанцев. При поддержке Фонда «Халық» в южной столице был организован ежегодный городской конкурс педагогов «*Almaty Digital Ustaz*».

Важной инициативой стал реализуемый проект по обучению основам финансовой грамотности преподавателей из восьми областей Казахстана, что должно оказать существенное влияние на воспитание финансовой грамотности и предпринимательского мышления у нового поколения граждан страны.

Необходимую помощь Фонд «Халык» оказывает и тем, кто особенно остро в ней нуждается. В рамках социальной защиты населения активно проводится работа по поддержке детей, оставшихся без родителей, детей и взрослых из социально уязвимых слоев населения, людей с ограниченными возможностями, а также обеспечению нуждающихся социальным жильем, строительству социально важных объектов, таких как детские сады, детские площадки и физкультурно-оздоровительные комплексы.

В копилку добрых дел Фонда «Халык» можно добавить оказание помощи детскому спорту, куда относится поддержка в развитии детского футбола и карате в нашей стране. Жизненно важную помощь Благотворительный фонд «Халык» оказал нашим соотечественникам во время недавней пандемии COVID-19. Тогда, в разгар тяжелой борьбы с коронавирусной инфекцией Фонд выделил свыше 11 миллиардов тенге на приобретение необходимого медицинского оборудования и дорогостоящих медицинских препаратов, автомобилей скорой медицинской помощи и средств защиты, адресную материальную помощь социально уязвимым слоям населения и денежные выплаты медицинским работникам.

В 2023 году наряду с другими проектами, нацеленными на повышение благосостояния казахстанских граждан Фонд решил уделить особое внимание науке, поскольку она является частью общественной культуры, а уровень ее развития определяет уровень развития государства.

Поддержка Фондом выпуска журналов Национальной Академии наук Республики Казахстан, которые входят в международные фонды Scopus и Wos и в которых публикуются статьи отечественных ученых, докторантов и магистрантов, а также научных сотрудников высших учебных заведений и научно-исследовательских институтов нашей страны является не менее значимым вкладом Фонда в развитие казахстанского общества.

**С уважением,  
Благотворительный Фонд «Халык»!**

### **БАС РЕДАКТОР:**

**МУТАНОВ Ғалымқайыр Мұтанұлы**, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР БҒМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының м.а. (Алматы, Қазақстан), **Н=5**

### **БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ:**

**МАМЫРБАЕВ Өркен Жұмажанұлы**, ақпараттық жүйелер мамандығы бойынша философия докторы (Ph.D), ҚР БҒМ Ғылым комитеті «Ақпараттық және есептеуші технологиялар институты» РМК жауапты хатшысы (Алматы, Қазақстан), **Н=5**

### **РЕДАКЦИЯ АЛҚАСЫ:**

**ҚАЛИМОЛДАЕВ Мақсат Нұрәділұлы**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі (Алматы, Қазақстан), **Н=7**

**БАЙГУНЧЕКОВ Жұмаділ Жанабайұлы**, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Кибернетика және ақпараттық технологиялар институты, Сатпаев университетінің Қолданбалы механика және инженерлік графика кафедрасы, (Алматы, Қазақстан), **Н=3**

**ВОЙЧИК Вальдемар**, техника ғылымдарының докторы (физика), Люблин технологиялық университетінің профессоры (Люблин, Польша), **Н=23**

**БОШКАЕВ Қуантай Авғазыұлы**, Ph.D. Теориялық және ядролық физика кафедрасының доценті, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=10**

**QUEVEDO Nemando**, профессор, Ядролық ғылымдар институты (Мехико, Мексика), **Н=28**

**ЖҮСПОВ Марат Абжанұлы**, физика-математика ғылымдарының докторы, теориялық және ядролық физика кафедрасының профессоры, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=7**

**КОВАЛЕВ Александр Михайлович**, физика-математика ғылымдарының докторы, Украина ҰҒА академигі, Қолданбалы математика және механика институты (Донецк, Украина), **Н=5**

**РАМАЗАНОВ Тілекқабұл Сәбитұлы**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, әл-Фараби атындағы Қазақ ұлттық университетінің ғылыми-инновациялық қызмет жөніндегі проректоры, (Алматы, Қазақстан), **Н=26**

**ТАКИБАЕВ Нұрғали Жабағаұлы**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=5**

**ТИГИНЯНУ Ион Михайлович**, физика-математика ғылымдарының докторы, академик, Молдова Ғылым Академиясының президенті, Молдова техникалық университеті (Кишинев, Молдова), **Н=42**

**ХАРИН Станислав Николаевич**, физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Қазақстан-Британ техникалық университеті (Алматы, Қазақстан), **Н=10**

**ДАВЛЕТОВ Асқар Ербуланович**, физика-математика ғылымдарының докторы, профессор, әл-Фараби атындағы Қазақ ұлттық университеті (Алматы, Қазақстан), **Н=12**

**КАЛАНДРА Пьетро**, Ph.D (физика), Нанокұрылымды материалдарды зерттеу институтының профессоры (Рим, Италия), **Н=26**

**«ҚР ҰҒА Хабарлары. Физика және информатика сериясы».**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Меншіктеуші: «Қазақстан Республикасының Ұлттық ғылым академиясы» РҚБ (Алматы қ.). Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігінің Ақпарат комитетінде 14.02.2018 ж. берілген **№ 16906-Ж** мерзімдік басылым тіркеуіне қойылу туралы куәлік.

Тақырыптық бағыты: *физика және ақпараттық коммуникациялық технологиялар сериясы*. Қазіргі уақытта: *«ақпараттық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*

Мерзімділігі: *жылына 4 рет.*

Тиражы: *300 дана.*

Редакцияның мекен-жайы: *050010, Алматы қ., Шевченко көш., 28, 219 бөл., тел.: 272-13-19*  
*<http://www.physico-mathematical.kz/index.php/en/>*

## ГЛАВНЫЙ РЕДАКТОР:

**МУТАНОВ Галимкаир Мутанович**, доктор технических наук, профессор, академик НАН РК, и.о. генерального директора «Института информационных и вычислительных технологий» КН МОН РК (Алматы, Казахстан), **Н=5**

## ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА:

**МАМЫРБАЕВ Оркен Жумажанович**, доктор философии (PhD) по специальности Информационные системы, ответственный секретарь РГП «Института информационных и вычислительных технологий» Комитета науки МОН РК (Алматы, Казахстан), **Н=5**

## РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

**КАЛИМОЛДАЕВ Максат Нурадилович**, доктор физико-математических наук, профессор, академик НАН РК (Алматы, Казахстан), **Н=7**

**БАЙГУНЧЕКОВ Жумадил Жанабаевич**, доктор технических наук, профессор, академик НАН РК, Институт кибернетики и информационных технологий, кафедра прикладной механики и инженерной графики, Университет Сагпаева (Алматы, Казахстан), **Н=3**

**ВОЙЧИК Вальдемар**, доктор технических наук (физ.-мат.), профессор Люблинского технологического университета (Люблин, Польша), **Н=23**

**БОШКАЕВ Куантай Авгазыевич**, доктор Ph.D, преподаватель, доцент кафедры теоретической и ядерной физики, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=10**

**QUEVEDO Hemando**, профессор, Национальный автономный университет Мексики (UNAM), Институт ядерных наук (Мехико, Мексика), **Н=28**

**ЖУСУПОВ Марат Абжанович**, доктор физико-математических наук, профессор кафедры теоретической и ядерной физики, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=7**

**КОВАЛЕВ Александр Михайлович**, доктор физико-математических наук, академик НАН Украины, Институт прикладной математики и механики (Донецк, Украина), **Н=5**

**РАМАЗАНОВ Тлексабул Сабитович**, доктор физико-математических наук, профессор, академик НАН РК, проректор по научно-инновационной деятельности, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=26**

**ТАКИБАЕВ Нургали Жабигаевич**, доктор физико-математических наук, профессор, академик НАН РК, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=5**

**ТИГИНЯНУ Ион Михайлович**, доктор физико-математических наук, академик, президент Академии наук Молдовы, Технический университет Молдовы (Кишинев, Молдова), **Н=42**

**ХАРИН Станислав Николаевич**, доктор физико-математических наук, профессор, академик НАН РК, Казахстанско-Британский технический университет (Алматы, Казахстан), **Н=10**

**ДАВЛЕТОВ Аскар Ербуланович**, доктор физико-математических наук, профессор, Казахский национальный университет им. аль-Фараби (Алматы, Казахстан), **Н=12**

**КАЛАНДРА Пьетро**, доктор философии (Ph.D, физика), профессор Института по изучению наноструктурированных материалов (Рим, Италия), **Н=26**

## «Известия НАН РК. Серия физика и информатики».

ISSN 2518-1726 (Online),

ISSN 1991-346X (Print)

Собственник: *Республиканское общественное объединение «Национальная академия наук Республики Казахстан» (г. Алматы).*

Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и общественного развития Республики Казахстан **№ 16906-Ж** выданное 14.02.2018 г.

Тематическая направленность: *серия физика и информационные коммуникационные технологии.* В настоящее время: *вошел в список журналов, рекомендованных ККСОН МОН РК по направлению «информационные коммуникационные технологии».*

Периодичность: *4 раз в год.*

Тираж: *300 экземпляров.*

Адрес редакции: *050010, г. Алматы, ул. Шевченко, 28, оф. 219, тел.: 272-13-19*

*<http://www.physico-mathematical.kz/index.php/en/>*

#### **EDITOR IN CHIEF:**

**MUTANOV Galimkair Mutanovich**, doctor of technical Sciences, Professor, Academician of NAS RK, acting director of the Institute of Information and Computing Technologies of SC MES RK (Almaty, Kazakhstan), **H=5**

#### **DEPUTY EDITOR-IN-CHIEF**

**MAMYRBAYEV Orken Zhumazhanovich**, Ph.D. in the specialty "Information systems, executive secretary of the RSE "Institute of Information and Computational Technologies", Committee of Science MES RK (Almaty, Kazakhstan) **H=5**

#### **EDITORIAL BOARD:**

**KALIMOLDAYEV Maksat Nuradilovich**, doctor in Physics and Mathematics, Professor, Academician of NAS RK (Almaty, Kazakhstan), **H=7**

**BAYGUNCHEKOV Zhumadil Zhanabayevich**, doctor of Technical Sciences, Professor, Academician of NAS RK, Institute of Cybernetics and Information Technologies, Department of Applied Mechanics and Engineering Graphics, Satbayev University (Almaty, Kazakhstan), **H=3**

**WOICIK Waldemar**, Doctor of Phys.-Math. Sciences, Professor, Lublin University of Technology (Lublin, Poland), **H=23**

**BOSHKAYEV Kuantai Avgazievich**, PhD, Lecturer, Associate Professor of the Department of Theoretical and Nuclear Physics, Al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=10**

**QUEVEDO Hemando**, Professor, National Autonomous University of Mexico (UNAM), Institute of Nuclear Sciences (Mexico City, Mexico), **H=28**

**ZHUSSUPOV Marat Abzhanovich**, Doctor in Physics and Mathematics, Professor of the Department of Theoretical and Nuclear Physics, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=7**

**KOVALEV Alexander Mikhailovich**, Doctor in Physics and Mathematics, Academician of NAS of Ukraine, Director of the State Institution «Institute of Applied Mathematics and Mechanics» DPR (Donetsk, Ukraine), **H=5**

**RAMAZANOV Tlekkabul Sabitovich**, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, Vice-Rector for Scientific and Innovative Activity, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=26**

**TAKIBAYEV Nurgali Zhabagaevich**, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=5**

**TIGHINEANU Ion Mikhailovich**, Doctor in Physics and Mathematics, Academician, Full Member of the Academy of Sciences of Moldova, President of the AS of Moldova, Technical University of Moldova (Chisinau, Moldova), **H=42**

**KHARIN Stanislav Nikolayevich**, Doctor in Physics and Mathematics, Professor, Academician of NAS RK, Kazakh-British Technical University (Almaty, Kazakhstan), **H=10**

**DAVLETOV Askar Erbulanovich**, Doctor in Physics and Mathematics, Professor, al-Farabi Kazakh National University (Almaty, Kazakhstan), **H=12**

**CALANDRA Pietro**, PhD in Physics, Professor at the Institute of Nanostructured Materials (Monterotondo Station Rome, Italy), **H=26**

#### **News of the National Academy of Sciences of the Republic of Kazakhstan.**

**Series of physics and informatics.**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Owner: RPA «National Academy of Sciences of the Republic of Kazakhstan» (Almaty). The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Social Development of the Republic of Kazakhstan **No. 16906-ЖК**, issued 14.02.2018  
Thematic scope: *series physics and information technology.*

Currently: *included in the list of journals recommended by the CCSES MES RK in the direction of «information and communication technologies».*

Periodicity: *4 times a year.*

Circulation: *300 copies.*

Editorial address: *28, Shevchenko str., of. 219, Almaty, 050010, tel. 272-13-19*

*<http://www.physico-mathematical.kz/index.php/en/>*

UDC 681.3:578.5(083.94)

© V. Shevtsov\*, A. Ismailova, Zh. Beldeubayeva, A. Satybaldiyeva,  
A. Nurpeisova, 2023

S. Seifullin Kazakh Agrotechnical Research University, Astana, Kazakhstan.

E-mail: shevtsovvladislav111@gmail.com

### MLVA AS A METHOD OF GENOTYPING AND ALGORITHMS FOR ITS IMPLEMENTATION USING GENOME-WIDE DATA

**Shevtsov V.A.** — Master of Science, PhD student, NCJSC «S. Seifullin Kazakh AgroTechnical Research University», Astana, Kazakhstan

E-mail: shevtsovvladislav111@gmail.com, <https://orcid.org/0000-0001-6202-2123>;

**Ismailova A.A.** — PhD, associate professor, NCJSC «S. Seifullin Kazakh AgroTechnical Research University», Astana, Kazakhstan

E-mail: a.ismailova@mail.ru, <https://orcid.org/0000-0002-8958-1846>;

**Beldeubayeva Zh.** — PhD, senior lecturer, NCJSC «S. Seifullin Kazakh AgroTechnical Research University», Astana, Kazakhstan

E-mail: zh.beldeubayeva@gmail.com, <https://orcid.org/0000-0003-4056-6220>;

**Satybaldiyeva A.** — PhD, senior lecturer, NCJSC «S. Seifullin Kazakh AgroTechnical Research University», Astana, Kazakhstan

E-mail: satekbayeva@gmail.com, <https://orcid.org/0000-0001-5740-7934>;

**Nurpeisova A.** — PhD, senior lecturer, NCJSC «S. Seifullin Kazakh AgroTechnical Research University», Astana, Kazakhstan

E-mail: naa11317@mail.ru, <https://orcid.org/0000-0002-1245-8313>.

**Abstract.** In the realm of modern molecular biology and epidemiology, where extensive genetic data undergoes intricate processing and analysis, information technologies assume a pivotal role in both propelling and refining the implementation of Multi-Locus Variable Number Tandem Repeat Analysis (MLVA). Within this dynamic landscape, these technologies emerge as catalysts that empower and elevate the MLVA methodology. The comparison of MLVA profiles across diverse individuals or locales reveals the rhythmic patterns of transmission, identifies sources of infection, and illuminates the path of pathogen propagation. This knowledge serves as the bedrock for focused interventions, targeted control strategies, and the preemptive containment of potential outbreaks. In the context of outbreak detection and management, MLVA's use shines particularly bright. The comparison of MLVA profiles across diverse individuals or locales reveals the rhythmic patterns of transmission, identifies sources of infection, and illuminates

the path of pathogen propagation. This knowledge serves as the bedrock for focused interventions, targeted control strategies, and the preemptive containment of potential outbreaks. In the endeavor to combat infectious diseases, Multi-Locus Variable Number Tandem Repeat Analysis (MLVA), underpinned by the formidable capabilities of information technologies, emerges as an indispensable instrument. It not only elucidates the intricate dynamics governing pathogenic agents but also furnishes the essential knowledge requisite for safeguarding public health. This review article describes the significant importance of information technology's role in revealing outbreaks sources and epidemiological control as well as comparative analysis of algorithms is conducted.

**Keywords:** Information systems, bioinformatics, MLVA, VNTR, algorithms

*This work was supported by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan in the framework of program funding for research 2023–2025, project number AP19678041 (“Development of software for the identification of tandem repeats in whole genome sequencing”).*

© В. Шевцов\*, А. Исмаилова, Ж. Белдеубаева, А. Сатыбалдиева,  
А. Нурпеисова, 2023

С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті,  
Астана, Қазақстан.

E-mail: shevtsovvladislav111@gmail.com

## МЛВА ГЕНОТИПТЕУДІҢ ӘДІСІ ЖӘНЕ ОНЫ ЖҮЗЕГЕ АСЫРУ АЛГОРИТМДЕРІ РЕТІНДЕГІ ГЕНОМДЫҚ ДЕРЕКТЕРДІ ПАЙДАЛАНУ

**Шевцов В.А.** — техника ғылымдарының магистрі, докторант, «С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті» КеАҚ, Астана, Қазақстан Республикасы  
E-mail: shevtsovvladislav111@gmail.com, <https://orcid.org/0000-0001-6202-2123>;

**Исмаилова А.А.** — PhD, қауымдастырылған профессор, «С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті» КеАҚ, Астана, Қазақстан Республикасы  
E-mail: a.ismailova@mail.ru, <https://orcid.org/0000-0002-8958-1846>;

**Белдеубаева Ж.** — PhD, аға оқытушы, «С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті» КеАҚ, Астана, Қазақстан Республикасы  
E-mail: zh.beldeubayeva@gmail.com, ORCID ID: <https://orcid.org/0000-0003-4056-6220>;

**Сатыбалдиева А.** — PhD, аға оқытушы, «С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті» КеАҚ, Астана, Қазақстан Республикасы  
E-mail: satekbaeva@gmail.com, <https://orcid.org/0000-0001-5740-7934>;

**Нурпеисова А.** — PhD, аға оқытушы, «С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті» КеАҚ, Астана, Қазақстан Республикасы  
E-mail: naa11317@mail.ru, <https://orcid.org/0000-0002-1245-8313>.

**Аннотация.** Кең ауқымды генетикалық деректер күрделі өндеуден және талдаудан өтетін заманауи молекулалық биология және эпидемиология саласында ақпараттық технологиялар көп локусты айнымалы сандар тандемді



қайталау талдауын (MLVA) енгізуді ынталандыруда және нақтылауда шешуші рөл атқарады. Осы динамикалық ландшафтта бұл технологиялар MLVA әдістемесін күшейтетін және көтеретін катализаторлар ретінде пайда болады. Өртүрлі адамдар немесе жергілікті жерлерде MLVA профильдерін салыстыру берілудің ырғақты үлгілерін көрсетеді, инфекция көздерін анықтайды және патогеннің таралу жолын жарықтандырады. Бұл білім бағдарланған араласулар, мақсатты бақылау стратегиялары және ықтимал індеттердің алдын алу үшін негіз болады. Ауруды анықтау және басқару контекстінде MLVA қолдануы ерекше жарқырайды. Зардап шеккен адамдардан алынған патогендік штаммдардың MLVA профильдерін бөлу арқылы әдіс жалпы инфекция көздерін және жағдайлар арасындағы байланыстарды тез арада анықтайды. Бұл ептілік тез жауап беруге мүмкіндік береді, бақылауды және бақылауды жеңілдетеді. Өртүрлі адамдар немесе жергілікті жерлерде MLVA профильдерін салыстыру берілудің ырғақты үлгілерін көрсетеді, инфекция көздерін анықтайды және патогеннің таралу жолын жарықтандырады. Бұл білім бағдарланған араласулар, мақсатты бақылау стратегиялары және ықтимал індеттердің алдын алу үшін негіз болады. Жұқпалы аурулармен күресуде ақпараттық технологиялардың зор мүмкіндіктеріне негізделген Multi-Locus Variable Number Tandem Repeat Analysis (MLVA) таптырмас құрал ретінде пайда болады. Ол патогендік агенттерді басқаратын күрделі динамикасын ашып қана қоймайды, сонымен қатар қоғамдық денсаулықты сақтау үшін қажетті біліммен қамтамасыз етеді. Бұл шолу мақаласында індет ошақтарын және эпидемиологиялық бақылауды ашудағы ақпараттық технологиялар рөлінің маңыздылығы сипатталған, сонымен қатар алгоритмдерге салыстырмалы талдау жүргізілген.

**Түйін сөздер:** Ақпараттық жүйелер, биоинформатика, MLVA, VNTR, алгоритмдер

*Бұл жұмысты Қазақстан Республикасы Білім және ғылым министрлігінің Ғылым комитеті 2023–2025 жылдарға арналған ғылыми зерттеулерді бағдарламалық қаржыландыру аясында, № AP19678041 («Тұтас геномдық секвенирлеудегі тандемдік қайталауларды идентификациялау үшін бағдарламалық қамтамасыз етуді әзірлеу») жобасының қолдауымен өткізілді.*

© В. Шевцов\*, А. Исмаилова, Ж. Бельдеубаева, А. Сатыбалдиева,  
А. Нурпеисова, 2023

Казахский агротехнический исследовательский университет  
им. С. Сейфуллина, Астана, Казахстан  
E-mail: shevtsovvladislav111@gmail.com

## MLVA КАК МЕТОД ГЕНОТИПИРОВАНИЯ И АЛГОРИТМЫ ЕГО РЕАЛИЗАЦИИ С ИСПОЛЬЗОВАНИЕМ ПОЛНОГЕНОМНЫХ ДАННЫХ

**Шевцов В.А.** — магистр технических наук, докторант, НАО «Казахский агротехнический исследовательский университет им. С. Сейфуллина», Астана, Казахстан

E-mail: shevtsovvladislav111@gmail.com, <https://orcid.org/0000-0001-6202-2123>;

**Исмаилова А.А.** — PhD, ассоциированный профессор, НАО «Казахский агротехнический исследовательский университет им. С. Сейфуллина», Астана, Казахстан

E-mail: a.ismailova@mail.ru, <https://orcid.org/0000-0002-8958-1846>;

**Бельдеубаева Ж.** — PhD, старший преподаватель, НАО «Казахский агротехнический исследовательский университет им. С. Сейфуллина», Астана, Казахстан

E-mail: zh.beldeubayeva@gmail.com, <https://orcid.org/0000-0003-4056-6220>;

**Сатыбалдиева А.** — PhD, старший преподаватель, НАО «Казахский агротехнический исследовательский университет им. С. Сейфуллина», Астана, Казахстан

E-mail: satekbayeva@gmail.com, <https://orcid.org/0000-0001-5740-7934>;

**Нурпеисова А.** — PhD, старший преподаватель, НАО «Казахский агротехнический исследовательский университет им. С. Сейфуллина», Астана, Казахстан

E-mail: naa11317@mail.ru, <https://orcid.org/0000-0002-1245-8313>.

**Аннотация.** В сфере современной молекулярной биологии и эпидемиологии, где обширные генетические данные подвергаются сложной обработке и анализу, информационные технологии играют ключевую роль как в продвижении, так и в совершенствовании внедрения многолокусного тандемного повторного анализа с переменным числом (MLVA). В этом динамичном ландшафте эти технологии становятся катализаторами, которые расширяют возможности и совершенствуют методологию MLVA. Сравнение профилей MLVA у разных людей или мест выявляет ритмические закономерности передачи, определяет источники инфекции и освещает пути распространения патогена. Эти знания служат основой для целенаправленных мер, стратегий прицельного контроля и упреждающего сдерживания потенциальных вспышек. В контексте выявления и борьбы со вспышками использование MLVA проявляется особенно ярко. В борьбе с инфекционными заболеваниями незаменимым инструментом становится мультилокусный анализ тандемных повторов с переменным числом (MLVA), подкрепленный огромными возможностями информационных технологий. Он не только разъясняет сложную динамику патогенных агентов, но также предоставляет важные знания, необходимые для защиты общественного здоровья. В данной обзорной статье описывается значимость роли информационных технологий

в выявлении источников вспышек и эпидемиологическом контроле, а также проводится сравнительный анализ алгоритмов.

**Ключевые слова:** информационные системы, биоинформатика, MLVA, VNTR, алгоритмы

*Работа выполнена при поддержке Комитета науки Министерства науки и высшего образования Республики Казахстан в рамках программы финансирования научных исследований на 2023–2025 годы, ИРН проекта AP19678041 («Разработка программного обеспечения для идентификации тандемных повторов при полногеномном секвенировании»).*

### **Introduction**

In recent years, the field of molecular biology has witnessed tremendous advancements in genotyping techniques, revolutionising our understanding of microbial diversity and population dynamics (Messina et.al., 2018). Among these techniques, Multi-Locus Variable Number Tandem Repeat Analysis (MLVA) has emerged as a powerful tool for characterising and discriminating between strains of various microorganisms. MLVA utilises the polymorphism observed in tandemly repeated DNA sequences, known as variable number tandem repeats (VNTR) to provide insights into the genetic variability and relatedness of microbial populations.

MLVA involves the amplification and subsequent analysis of multiple VNTR loci, each of which comprises a variable number of repeated DNA motifs (Lindstedt, 2005). The number of repeats at each locus varies among strains, leading to distinct patterns that can be used for strain typing and epidemiological investigations. By comparing the MLVA profiles obtained from different isolates, researchers can discern the relatedness between strains, identify clonal outbreaks, trace the source of infections, and monitor the spread of pathogens within a population or geographical region.

MLVA's significance spans a spectrum of vital contributions, including its role as a discriminating force among closely linked microorganism strains. By delving into the intricacies of variable number tandem repeats (VNTRs) within strain genomes, MLVA crafts distinct profiles that segregate strains with precision.

This micro-level resolution unveils the genetic mosaic within populations and meticulously traces the footprints of infectious agents' transmissions. Beyond mere differentiation, MLVA significantly bolsters epidemiological investigations (Li et.al., 2009). The comparison of MLVA profiles across diverse individuals or locales reveals the rhythmic patterns of transmission, identifies sources of infection, and illuminates the path of pathogen propagation. This knowledge serves as the bedrock for focused interventions, targeted control strategies, and the preemptive containment of potential outbreaks (Cheng et.al, 2015). In the context of outbreak detection and management, MLVA's use shines particularly bright. By dissecting MLVA profiles of pathogenic strains sampled from affected individuals, the technique

promptly pinpoints common infection sources and links between cases. This agility enables swift response, facilitating containment and control. Remarkably, MLVA has played an instrumental role in deciphering outbreaks spanning a gamut of infectious diseases, from foodborne ailments to nosocomial infections. Thus, in the summer of 2018, a small dairy farm located in the eastern region of Austria experienced an abnormal and advanced increase in the number of somatic cells in milk (MSCC). As a result of the bacteriological examination, it became clear that nine out of twenty lactating cows had at least a quarter infected with *P. aeruginosa*. Results obtained in this study suggested that one single strain was responsible for the whole outbreak. Complementing its dynamic roles, MLVA serves as a vigilant sentinel through constant surveillance and monitoring of microbial populations. Regular analyses of isolates from clinical samples, food sources, and environmental reservoirs reveal the evolution of pathogens over time. This vigilance aids in uncovering nascent strains, discerning evolutionary trajectories, and gauging the effectiveness of control measures. Furthermore, MLVA's utility extends to zoonotic diseases, elucidating transmission dynamics by identifying akin strains in animal reservoirs and tracing links to human infections (Cumbassa, 2013). It's also a cornerstone in foodborne outbreak investigations, identifying contamination sources and catalysing effective interventions. MLVA's contributions encompass strain discrimination, epidemiological insight, outbreak vigilance, surveillance, and the scrutiny of zoonotic and culinary pathogens. A fusion of high-resolution discrimination and transmission tracking, MLVA plays an indispensable role in molecular biology and public health, bolstering our understanding of microbial realms, advancing control strategies, and safeguarding public welfare.

### Methods

As the main method, comparative analysis of various algorithms for VNTR search was used when comparing types of algorithms that focused on finding already known tandem repeats and discovering new tandem repeats using machine learning methods.

MLVA analysis was implemented using the MLVA\_finder script in which RAW reads of *Brucella abortus* were used as input data.

### Results

#### 1.1 Process of MLVA analysis implementation

The process of MLVA typically involves three main steps: DNA extraction, PCR amplification of VNTR loci (Figure 1), and fragment analysis. Initially, DNA should be extracted from the target microorganism using established protocols, after that the stage of assessing concentration and purity of the DNA. Subsequently, PCR primers specifically need to be chosen, after performing PCR amplification, the resulting amplicons are then subjected to fragment analysis, typically utilizing capillary electrophoresis, to determine the size of the VNTR alleles present in each sample. These sizes are subsequently used to construct MLVA profiles for each isolate (Li et.al., 2009; Pourcel et.al., 2020).

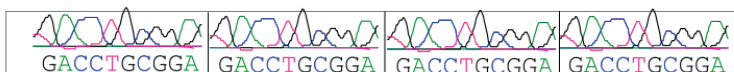
**Tandem repeats**



Repeat sequences types



Variable Number of Tandem Repeats (VNTR)



Repeat sequence length:  
10-100 base pairs/repeat

Short Tandem Repeats (STR)



Repeat sequence length:  
2-9 base pairs/repeat

Fig. 1. The representation of tandem repeats

1.2 The MLVA analysis input and output data

The input data for MLVA analysis include such files and/or information as:

Sequences, including “fasta” sequences or raw data converted to fasta as well as reference sequence in “fasta” format. Primers located near to repeat region together with additional information including <locus\_name>\_<pattern\_size>bp\_<insert\_size\_in\_reference\_genome>bp\_<corresponding allele coding convention>U forward\_primer, reverse\_primer (Table 1)

Table 1 – primers located near VNTR region

n/a	PrimerName	Forward primer	Reverse primer
1	Bruce23_9bp_51bp_1u	CGCCCTTGAGGGTGAGCGTGT	GGATATGGAATCAGAACACGC
2	Bruce24_8bp_108bp_8u	CATGGCCGCAGCGACCATATT	TGCTCCAGCGCGCGGGAGGT
3	Bruce35_10bp_98bp_4u	GCGCGGTGTTGGCATCAGC	ACGGCAGCCATGCTGATGGGG
4	Bruce36_15bp_95bp_4u	GCCCGTCGCCAGTGTGAAGT	CAATAGGGACAACGCATCGAA

Once the MLVA analysis has finished, it will produce a “.csv” file which contain information on numbers of VNTR repeats (Table 2)

Table 2 – output of the MLVA analysis using *MLVA\_finder* [10] and custom scripts

n/a		Bruce23_ 9bp_51bp_1u	Bruce24 8bp_108bp_8u	Bruce35_ 10bp_98bp_4u	Bruce36_15bp_ 95bp_4u
1	B-abortus-B0502_ S44	3	7	4	19
2	B-abortus-B0558_ S38	4	6	4	19
3	B-abortus- Kaz-030_S2	4	6	4	19
4	B-abortus- Kaz-050_S6	3	6	5	19

As can be seen from the example of artificial data, it can be seen that primers whose locus name is Bruce23 has highest discriminatory power, while Bruce36 loci has the lowest discriminatory power which means such data have no scientific interest (Delgrange et.al., 2004; Benson, 1999).

### 1.3 Algorithms for identification of VNTR repeats

VNTRs are sequences where a short motif is repeated in tandem, and the number of repeats varies among individuals. For the last 2 decades, various algorithms have been developed to detect these repeats in DNA sequences. Table 3 illustrates an overview of some common algorithms used for VNTR identification (Table 3)

Table 3 – Algorithms for searching VNTR repeats

n/a	Algorithm name	Description
1	STAR	Identifies all segments of the sequence that correspond to significant approximate tandem repetitions of the motif. In our model, an Exact Tandem Repeat (ETR) comes from the tandem duplication of the motif and an ATR derives from an ETR by a series of point mutations.
2	TRF	TRF is a widely used algorithm for detecting tandem repeats in DNA sequences. It employs a heuristic approach to identify repeating patterns and can handle a variety of repeat motifs. TRF has been extensively used in the field of genomics and bioinformatics for detecting and characterising VNTRs.
3	T-REKS	T-REKS is an algorithm for de novo detection and alignment of repeats in sequences based on K-mers algorithm. Minimal length of repeat arrays is 9 for true homorepeats and 14 for other repeats with potential biological meaning.
4	RPWM	The novel RPWM algorithm, which uses a novel algorithm for constructing multiple alignments based on the generation of random position weight matrices, is developed and applied to detect TRs of 2 to 50 nucleotides long in the rice genome, revealing that TRs occupied 5% of the genome and that most of them were 2 and 3 bases long
5	TRhist	An algorithm specifically developed to detect and correctly annotate TR expansions. Their identification of expanded TTTCA repeats within long TTTTA repeats causing BAFME1 was also facilitated by long-read sequencing technologies.

6	PTRStalker	PTRStalker is a new algorithm for ab-initio detection of fuzzy tandem repeats in protein amino acid sequences.
7	RepeatSeq	Determines genotypes for microsatellite repeats in high-throughput sequencing data.
8	STR-FM	

Various algorithms have a number of advantages and drawbacks in relation to each other.

For instance, STAR is a straightforward algorithm used for identifying simple tandem repeats (STRs) in DNA sequences. It works by sliding a fixed-size window across the sequence and searching for repeated patterns within that window. When a repeated pattern is found, it is marked as a tandem repeat, however, the algorithm is designed for basic repeat patterns and may not handle complex or compound repeats efficiently. To overcome the issue of searching a limited number of patterns, the other tools such RepeatSeq and STR-FM were developed. These tools implement search of VNTR repeats using machine learning algorithms such as linear regression or classification. Despite machine learning solutions for identifying VNTR repeats becoming more popular for identifying new repeat patterns, most VNTR analysis algorithms based on already known repeats remain the main direction when it comes to outbreak tracking, that is why the algorithms like RPWM play such an important role (Jorda et.al., 2009; Korotkov et.al., 2021; Gall-Duncan et.al., 2022).

The comparative analysis between the RPWM (Reputation Weight Matrix) algorithm and alternative methods for tandem repeat (TR) identification yielded insightful revelations about their respective capabilities. Specifically, the study unveiled that RPWM displayed a unique proficiency in discerning tandem repeats characterised by elevated levels of nucleotide divergence (Pellegrini et.al., 2012; Guilmatre et.al., 2013; Korotkov et.al., 2021).

The parameter "average number of base substitutions per nucleotide (x)" denotes the extent of sequence variation or mutation present within a tandem repeat sequence. Within this context, a higher value of x signifies a greater departure from the original sequence due to various genetic changes. These alterations can encompass mutations, insertions, deletions, or evolutionary shifts that reshape the repeat motif while preserving the tandem arrangement (Bolognini et.al., 2020; Kumar et.al., 2013).

The research findings illuminated RPWM's exceptional acumen in identifying tandem repeats situated within the range of 1.5 to 3.2 for the average number of base substitutions per nucleotide (x). Essentially, RPWM excelled in capturing tandem repeats that underwent a moderate level of mutational transformation. This remarkable capability allowed for nuanced modifications in the repeat motif while preserving the essential tandem structure (Gao et.al., 2019).

Conversely, prevalent methods like T-REKS and TRF demonstrated limitations in detecting divergent tandem repeats characterised by higher x values (>1.5). This

implies that these methodologies might encounter challenges when accurately pinpointing tandem repeats subjected to significant mutational adjustments, consequently leading to a decreased sensitivity in recognizing such divergent repeats (Gelfand, 2014).

The significance of these findings resides in RPWM's potential to expand the horizons of tandem repeat discovery to sequences marked by evolutionary divergence. Such implications are far-reaching, particularly in evolutionary genomics, where the identification of more divergent tandem repeats enriches the comprehension of genetic variability and evolutionary dynamics.

#### 1.4 Tools for detecting VNTR repeats

These tools (Table 4) offer diverse algorithms and features tailored to cater to various aspects of tandem repeat analysis. Here is the list of software solutions dedicated to tandem repeat identification.

Table 4 – Tools for MLVA analysis

n/a	Tool name	Description
1	<i>ReviSTER</i> 0.1.7	<i>ReviSTER</i> (Revise Simple Tandem repeat Error Reads) is an automated pipeline using a local mapping reference reconstruction method to revise mismapped (mapped to incorrect position) or partially misaligned (mapped to correct position but one of ends misaligned) reads at STR (Simple Tandem Repeat) loci. It takes FASTQ-formatted files, a reference sequence file and a list file containing STR locations as inputs and utilises BWA as an initial mapping program.
2	<i>TRiCoLoR</i>	a freely available tool for tandem repeat profiling using error-prone long reads from third-generation sequencing technologies. The method can identify repetitive regions in sequencing data without a prior knowledge of their motifs or locations and resolve repeat multiplicity and period size in a haplotype-specific manner.
3	<i>Phobos</i>	<i>Phobos</i> is a tandem repeat search tool for complete genomes. <i>PHOBOS</i> can search for tandem repeats with a unit size of more than 5000 bp, which in the <i>STAMP</i> modules implies that primers can also be designed for minisatellites and tandem repeats with even longer units.
4	<i>Bionumerics</i>	The <i>BIONUMERICICS</i> software offers a fully automated workflow for multi-locus VNTR analysis, starting from raw capillary sequencer chromatogram files
5	<i>TideHunter</i>	Uses a fast seed-and-chain algorithm to efficiently recognize the underlying repeat pattern size, and then partition the original long-read into multiple repeat units.
6	<i>MLVA_finder</i>	<i>MLVA_finder.py</i> is a python script designed to do Multi loci VNTR analysis (VNTR stands for Variable Number of Tandem Repeats). <i>MLVA_finder.py</i> performs an in-silico PCR to extract sequences of tandem repeat from submitted fasta file(s) and call VNTR alleles.
7	<i>VNTRseek</i>	<i>VNTRseek</i> is a targeted, efficient VNTR (Variable Number of Tandem Repeats) detection software which can provide essential information on VNTR occurrence and characteristics of minisatellites
8	<i>RepWords</i>	<i>RepWords</i> detects tandem repeats in FASTA sequences.



These software solutions collectively cater to the various needs of researchers engaged in tandem repeat analysis. The choice of software depends on factors such as the complexity of the genome, the specific type of tandem repeats being investigated, the desired level of accuracy, and the ease of use. By utilising these tools, researchers can unravel the intricate patterns of tandem repeats, shedding light on the genomic tapestry and contributing to our understanding of genetic diversity and evolution.

*1.5 Sequences analysis equations (calculation and interpretation of the Discriminatory Power)*

Sequence analysis equations is an algorithm that is used for aligning the sequences. Although it is not used as the main tool in MLVA analysis, the algorithm is able to provide more information on gaps in sequences, sequence identity and similarity when it comes to missing data(nucleotides) in sequence or checking the direction of sequence.

The expression of Discriminatory Power (D), as initially demonstrated by Hunter, finds its formulation in the context of Simpson's index of diversity (1):

$$D = 1 - \frac{1}{N(N-1)} = \sum_{j=1}^s x_j(x_j - 1) \quad (1)$$

Within this equation, D represents the discriminatory power index, N signifies the count of unrelated strains subjected to testing, S denotes the variety of distinct types identified, and  $x_j$  corresponds to the number of strains belonging to the  $j$ th type. It's important to note that the assumption is made that the strains are classified into exclusive categories. In essence, a D value of 1.0 would denote a typing method capable of distinctly classifying every member within a strain population from all others. In contrast, an index of 0.0 would imply that all members of a strain population share an identical type. Notably, an index of 0.50 signifies that if one strain is randomly selected from a strain population, there exists a 50% likelihood that the subsequent randomly selected strain would be indistinguishable from the initial selection. Simpson's diversity index is applicable only to situations in which all strains can be placed into mutually exclusive groups (Hunter et.al., 1989).

To improve the algorithm, a generalised version of the first algorithm was proposed by Hunter (Doster, 2019):

$$D = 1 - \frac{1}{N(N-1)} = \sum_{j=1}^N a_j \quad (2)$$

In the context of this mathematical formulation (4), "aj" is indicative of the count of strains within the population that are non-distinguishable from the  $j$ th strain. Here, "N" stands for the total number of strains within the population. Put simply, this process involves sequentially comparing each individual strain with all other strains present within the population. Through this comparison, the goal is to ascertain the number of other strains that share an indistinguishable nature with the  $j$ th strain, thereby yielding the value of "aj".

The second equation is more effective when it comes to implementing statistical and epidemiological Interpretation. Based on the MLVA results, researchers can draw statistical and epidemiological conclusions. This might involve identifying potential sources of infection, characterising transmission dynamics, and making informed decisions for public health interventions.

### Conclusion

While new approaches like machine learning are involved in detecting new VNTR repeat patterns, it is important to improve already existing algorithms for identifying known repeats in assemblies. As it can be seen from the comparative analysis conducted in the article, each algorithm has its own range of application, some algorithms are strictly focused on identifying long repeats while missing the short ones.

Although there are a variety of tools and algorithms existing, still there is no gold standard protocol for detection of such repeats in various organisms. It has been revealed that most algorithms for detecting VNTR repeats are focused on searching already known tandem repeats in pathogens. T-REKS is one of the most successfully developed algorithms in terms of accurate detection of tandem repeats from several nucleotides long to hundreds of nucleotides long.

An important aspect in searching such repeats is that most algorithms and software created are working only with assemblies, meaning that some data could be missed from the final analysis as in most cases genome assembly contains gaps that may lack important information such repeats. Thus, creating an algorithm that could work directly with raw data (sequences) is another direction that is able to improve the quality of MLVA analysis.

### REFERENCES

- Benson G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic acids research*, — 27(2), — 573–580.
- Bolognini D., Magi A., Benes V., Korb J.O. & Rausch T. (2020). TRiCoLoR: tandem repeat profiling using whole-genome long-read sequencing data. *Gigascience*, — 9(10). — P. 101.
- Cheng J.M., Hiscoe L., Pollock S.L., Hasselback P., Gardy J.L. & Parker R. (2015). A clonal outbreak of tuberculosis in a homeless population in the interior of British Columbia, Canada, — 2008–2015. *Epidemiology & Infection*, — 143(15), — 3220–3226.
- Cumbassa A. (2013). Sources and routes of transmission of Q fever: detection, identification and molecular typing of *Coxiella burnetii* in domestic and wild animals (Doctoral dissertation).
- Delgrange O. & Rivals E. (2004). STAR: an algorithm to search for tandem approximate repeats. *Bioinformatics*, — 20(16), — 2812–2820.
- Doster E. (2019). *Epidemiological Investigation of Antimicrobial Resistance in Beef Production Using Metagenomic Sequencing* (Doctoral dissertation, Colorado State University).
- Gao Y., Liu B., Wang Y. & Xing Y. (2019). TideHunter: efficient and sensitive tandem repeat detection from noisy long-reads using seed-and-chain. *Bioinformatics*, — 35(14), — i200–i207.
- Gall-Duncan T., Sato N., Yuen R.K. & Pearson C.E. (2022). Advancing genomic technologies and clinical awareness accelerates discovery of disease-associated tandem repeat sequences. *Genome research*, — 32(1), — 1–27.
- Guilmatre A., Highnam G., Borel C., Mittelman D. & Sharp A.J. (2013). Rapid Multiplexed Genotyping of Simple Tandem Repeats using Capture and High-Throughput Sequencing. *Human*

mutation, — 34(9), — 1304–1311.

Gelfand Y., Hernandez Y., Loving J. & Benson G. (2014). VNTRseek—a computational tool to detect tandem repeat variants in high-throughput sequencing data. *Nucleic acids research*, — 42(14), — 8884–8894.

Hunter P.R. & Fraser C.A. (1989). Application of a numerical index of discriminatory power to a comparison of four physicochemical typing methods for *Candida albicans*. *Journal of clinical microbiology*, — 27(10), — 2156–2160.

Jorda J. & Kajava A.V. (2009). T-REKS: identification of Tandem REpeats in sequences with a K-meanS based algorithm. *Bioinformatics*, — 25(20), — 2632–2638.

Korotkov E.V., Kamionskya A.M. & Korotkova M.A. (2021). Detection of highly divergent tandem repeats in the rice genome. *Genes*, — 12(4), — 473.

Korotkov E.V., Kamionskya A.M. & Korotkova M.A. (2021). Detection of highly divergent tandem repeats in the rice genome. *Genes*, — 12(4), — 473.

Kumar S., Kumar D. & Chaudhury A. (2013). Mining and Analysis of Tandem Repeated Patterns in Oncogenic Sequences involved in Cancer progression. *International Journal of Computer Science Issues (IJCSI)*, — 10(6), — 250.

Lindstedt B.A. (2005). Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis*, — 26(13), — 2567–2582.

Li W., Raoult D. & Fournier P.E. (2009). Bacterial strain typing in the genomic era. *FEMS microbiology reviews*, — 33(5), — 892–916.

Messina F., Di Corcia T., Ragazzo M., Sanchez Mellado C., Contini I., Malaspina P. & Jodice C. (2018). Signs of continental ancestry in urban populations of Peru through autosomal STR loci and mitochondrial DNA typing. *PLoS One*, — 13(7), — e0200796.

Pourcel C., Midoux C., Vergnaud G. & Latino L. (2020). The basis for natural multiresistance to phage in *Pseudomonas aeruginosa*. *Antibiotics*, — 9(6), — 339.

Pellegrini M., Renda M.E. & Vecchio A. (2012, December). Ab initio detection of fuzzy amino acid tandem repeats in protein sequences. In *Bmc Bioinformatics*. — Vol. 13. — No. 3. — Pp. 1–13. BioMed Central.

## МАЗМҰНЫ

<b>Г.Б. Абдикеримова, Р.М. Аманов, Г.Т. Азиева, А.М. Заманбекова, Қ. Жеңсқанқызы</b> <i>МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНА ОТЫРЫП, АККОРДТЫ ТАҢУ ТАПСЫРМАСЫНДАҒЫ ДЫБЫСТЫ ӨҢДЕУ ӘДІСТЕРІН САЛЫСТЫРМАЛЫ ТАЛДАУ</i> .....	7
<b>Л.А. Абдыкеримова, Г.Е. Мырзабекова, Г.С. Омарова, Л. Ақзуллақызы, Г.Ш. Мусагулова</b> ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ АРҚЫЛЫ ЖҮРЕК ПАТОЛОГИЯСЫН АНЫҚТАУ.....	21
<b>А.Е. Әбжанова, Е.Ә. Әбжанов, А.А. Мырзамуратова, А.Г. Батырханов, А.Б. Бексейтова</b> ҚАШЫҚТАН ЗОНДТАУ АРҚЫЛЫ АЛЫНҒАН ТОПЫРАҚ ЫЛҒАЛДЫЛЫҒЫ.....	35
<b>У.Ж. Айтимова, М.Ж. Айтимов, Э.Н. Тулегенова, А.У. Есиркепова, Ж.Т. Абилдаева</b> СУРЕТТЕН ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІ АРҚЫЛЫ ӨРТ ОШАҒЫН АНЫҚТАУ.....	50
<b>К.М. Алдабергенова, М.Ж. Жасұзақова, М.Ж. Айтимов, Н.Т. Мұстафаева, К.К. Дауренбеков</b> АУЫЛ ШАРУАШЫЛЫҒЫН ЦИФРЛАНДЫРУ: ДАМУ МҮМКІНДІКТЕРІ МЕН ПЕРСПЕКТИВАЛАРЫ.....	64
<b>А.С. Баегизова, Г.И. Мухамедрахимова, Ж.Б. Ламашева, А.З. Абдрахманова, Т.Т. Оспанова</b> ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІМЕН КЕСКІННІҢ САПАСЫН ЖАҚСARTУ.....	78
<b>Г.Т. Бекманова, А.С. Омарбекова, М.А. Кантуреева, Н.О. Байгабылов, М.М. Құдабеков</b> ӘЛЕУМЕТТАНУЛЫҚ САУАЛНАМАЛЫҚ ЗЕРТТЕУЛЕРДЕГІ АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР.....	91
<b>М.Ә. Берсүгір, Г.У. Маматова, А.А. Нурпейсова, М.Б. Онгарбаева, Ж.Т. Алтынбекова</b> ТЕКСТУРАЛЫҚ ТИПТЕГІ СУРЕТТЕРДІ ЖАҚСARTУ ҮШІН МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ.....	104
<b>М.А. Болатбек, К.Д. Байсылбаева, М. Сағынай, Ш.Ж. Мусиралиева, А.Н. Жумаханова</b> ИНТЕРНЕТ КЕҢІСТІГІНДЕГІ ЖАСТАРҒА БАҒЫТТАЛҒАН ДЕСТРУКТИВТІ МӘТІНДЕРДІ ЖИНАҚТАУҒА ҚАЖЕТТІ ПАРСЕР БАҒДАРЛАМАСЫН ӨЗІРЛЕУ.....	117
<b>М.Қ. Болсынбек, Г.Б. Абдикеримова, Г.С. Омарова, А.Б. Остаева, А.Г. Батырханов</b> ТОПЫРАҚ ДАЙЫНДАУДЫ БОЛЖАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУДЫ ПАЙДАЛАНУ.....	132
<b>Ш.К. Ележанова, А.Г. Батырханов, А.Е. Чукуров, Б.С. Хайржанова, Д.А. Тагиев</b> АҚПАРАТТЫҚ БЕЛГІСІЗДІК ТИПОЛОГИЯСЫ ЖӘНЕ АҚПАРАТТЫ ІЗДЕУ ТҮРЛЕРІ.....	151
<b>М.М. Есмагамбетова, Т.Т. Оспанова, Л.К. Бобров, Т.Л. Тен, Т.У. Есмагамбетов</b> ҒАРЫШТЫҚ ТӨТЕНШЕ ЖАҒДАЙЛАРДЫ БАҚЫЛАУ КЕСКІНДЕРІН ӨҢДЕУДЕ ТҮСТЕРДІ ӨЛШЕУ БАҒДАРЛАМАЛЫҚ ҚҰРАЛЫН ТАҢДАУ.....	161

<b>Т.К. Жукабаева, А. Адамова, Б.А. Ху Вен-Цен, Е.М. Марденов, Л.З. Жолшиева</b> СЫМСЫЗ СЕНСОР ЖЕЛІСІНДЕГІ SYBIL ЖӘНЕ WORMHOLE ШАБУЫЛДАРЫН АНЫҚТАУ.....	171
<b>А.А. Исмаилова, Ж.Т. Бельдеубаева, А.А. Нурпейсова, Г.О. Исакова, Ж.З. Жантасова</b> ӨСІМДІК АУРУЛАРЫН ТЕРЕҢ ОҚЫТУ ӘДІСТЕРІ АРҚЫЛЫ АНЫҚТАУ.....	184
<b>А.Х. Касымова, М.Б. Есенова, М.У. Худойберганов, А.Б. Остаева, М.Г. Байбулова</b> ДАҚЫЛДАРДЫҢ АУРУЛАРЫН ЖІКТЕУ ҮШІН ТЕРЕҢ ОҚЫТУ АЛГОРИТМДЕРІН ҚОЛДАНУ.....	198
<b>А.Ұ. Мұхиядин, М.У. Мукашева, У.Т. Махажанова, А.А. Муханова, Ж.Б. Ламашева</b> ПРОГРАММАЛЫҚ ҚҰРАЛДАР КӨМЕГІМЕН ЭКСТРЕМАЛДЫ ҚАШЫҚТЫҚТАН ОҚЫТУДЫҢ ОҚУШЫЛАРҒА ӘСЕРІН ЗЕРТТЕУ.....	209
<b>Б.Б. Оразбаев, Л.Т. Салыбек, К.Н. Оразбаева, Ш.К. Коданова, С.Ш. Исакова</b> МҰНАЙДЫ АЛҒАШҚЫ ӨНДЕУДЕ ЭЛЕКТРОТҰЗСЫЗДАНДЫРЫРУ ЖӘНЕ СУСЫЗДАНДЫРУ ПРОЦЕССТЕРІН ОПТИМИЗАЦИЯЛАУ ҮШІН МОДЕЛЬДЕР ҚҰРУ ТӘСІЛІ.....	224
<b>С.К. Серикбаева, М.Қ. Болсынбек, А.Д. Абдувалова, А.Т. Абдыхалық, Д.Е. Ануарбек</b> ТОПЫРАҚ САПАСЫН БОЛЖАУ ҮШІН МАШИНАЛЫҚ ОҚЫТУДЫ ҚОЛДАНУ: АЛГОРИТМДЕР МЕН ӘДІСТЕР.....	237
<b>А.Ж. Танирбергенов, Ж.К. Тасжурекова, С.К. Серикбаева, А.А. Шораев, А.Д. Абдувалова</b> ТОЛЫҚ МӘТІНДІ ҚҰЖАТТАРДЫ ІЗДЕУДІҢ МОДЕЛІ МЕН АҚПАРАТТЫҚ ЖҮЙЕСІН ҚҰРУ ӘДІСТЕРІ.....	253
<b>А.Ә. Таурбекова, Ө.Ж. Мамырбаев, К.Ж. Тұрғанбай</b> СЕЙСМИКАЛЫҚ ҚЫЗМЕТТІ БАҒАЛАУ ҮШІН ГИДРОДИНАМИКАЛЫҚ ТҰРАҚСЫЗДЫҚ ПРОЦЕСІН ЗЕРТТЕУ.....	268
<b>Н. Т. Тұржанов, Ш. К. Ележанова, С. Н. Идрисов, Ж. К. Дюсембина</b> АҚПАРАТТЫҚ ҮДЕРІСТЕРДІҢ РЕИНЖИНИРИНГІНІҢ ИННОВАЦИЯЛЫҚ КУРСЫН ӘЗІРЛЕУ.....	290
<b>В. Шевцов, А. Исмаилова, Ж. Белдеубаева, А. Сатыбалдиева, А. Нурпейсова</b> МЛВА ГЕНОТИПТЕУДІҢ ӘДІСІ ЖӘНЕ ОНЫ ЖҮЗЕГЕ АСЫРУ АЛГОРИТМДЕРІ РЕТІНДЕГІ ГЕНОМДЫҚ ДЕРЕКТЕРДІ ПАЙДАЛАНУ.....	300
<b>А.Ә. Шекербек, А.А. Некесова, Ж.Ж. Молдашева, А.И. Онгарбаева, А.О. Тохаева</b> ФРАКТАЛДЫҚ ӘДІСПЕН ӨКПЕНІҢ ПАТОЛОГИЯЛЫҚ ЖАҒДАЙЫН ТАЛДАУ.....	313

## СОДЕРЖАНИЕ

<b>Г.Б. Абдикеримова, Р.М. Аманов, Г.Т. Азиева, А.М. Заманбекова, К. Женсканкызы</b> СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ОБРАБОТКИ ЗВУКА В ЗАДАЧЕ РАСПОЗНАВАНИЯ АККОРДОВ С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ.....	7
<b>Л.А. Абдыкеримова, Г.Е. Мурзабекова, Г.С. Омарова, Л. Акзуллакызы, Г.Ш. Мусагулова</b> ОБНАРУЖЕНИЕ СЕРДЕЧНОЙ ПАТОЛОГИИ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ.....	21
<b>А.Е. Абжанова, Е.А. Абжанов, А.А. Мырзамуратова, А.Г. Батырханов, А.Б. Бексейтова</b> ВЛАЖНОСТЬ ПОЧВЫ, ПОЛУЧЕННАЯ ДИСТАНЦИОННЫМ ЗОНДИРОВАНИЕМ.....	35
<b>У.Ж. Айтимова, М.Ж. Айтимов, Э.Н. Тулегенова, А.У. Есиркепова, Ж.Т. Абилдаева</b> ОБНАРУЖЕНИЕ ОЧАГОВ ПОЖАРА С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ ПО ИЗОБРАЖЕНИЮ.....	50
<b>К.М. Алдабергенова, М.Ж. Жасузакова, М.Ж. Айтимов, Н.Т. Мустафаева, К.К. Дауренбеков</b> ЦИФРОВИЗАЦИЯ СЕЛЬСКОГО ХОЗЯЙСТВА: ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ РАЗВИТИЯ.....	64
<b>А.С. Баегизова, Г.И. Мухамедрахимова, Ж.Б. Ламашева, А.З. Абдрахманова, Т.Т. Оспанова</b> УЛУЧШЕНИЕ КАЧЕСТВА ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ МЕТОДОВ ГЛУБОКОГО ОБУЧЕНИЯ.....	78
<b>Г.Т. Бекманова, А.С. Омарбекова, М.А. Кантуреева, Н.О. Байгабылов, М.М. Кудабеков</b> ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В СОЦИОЛОГИЧЕСКИХ ОПРОСНЫХ ИССЛЕДОВАНИЯХ.....	91
<b>М.А. Берсугир, Г.У. Маматова, А.А. Нурпейсова, М.Б. Онгарбаева, Ж.Т. Алтынбекова</b> ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ ТЕКСТУРНОГО ТИПА.....	104
<b>М.А. Болатбек, К.Д. Байсылбаева, М. Сагынай, Ш.Ж. Мусиралиева, А.Н. Жумаханова</b> РАЗРАБОТКА ПРОГРАММЫ ПАРСЕРА ДЛЯ СБОРА ДЕСТРУКТИВНЫХ ТЕКСТОВ, ОРИЕНТИРОВАННЫХ НА МОЛОДЕЖЬ В ИНТЕРНЕТ-ПРОСТРАНСТВЕ.....	117
<b>М.К. Болсынбек, Г.Б. Абдикеримова, Г.С. Омарова, А.Б. Остаева, А.Г. Батырханов</b> ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ПОДГОТОВКИ ПОЧВЫ.....	132
<b>Ш.К. Ележанова, А.Г. Батырханов, А.Е. Чукуров, Б.С. Хайржанова, Д.А. Тагиев</b> ТИПОЛОГИЯ ИНФОРМАЦИОННОЙ НЕОПРЕДЕЛЕННОСТИ И ТИПЫ ПОИСКА ИНФОРМАЦИИ.....	151

<b>М.М. Есмагамбетова, Т.Т. Оспанова, Л.К. Бобров, Т.Л. Тен, Т.У. Есмагамбетов</b> ВЫБОР ПРОГРАММНЫХ СРЕДСТВ ЦВЕТОМЕТРИИ В ОБРАБОТКЕ ИЗОБРАЖЕНИЙ КОСМИЧЕСКОГО МОНИТОРИНГА ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЙ.....	161
<b>Т.К. Жукабаева, А. Адамова, В.А. Ху Вен-Цен, Е.М. Марденов, Л.З. Жолшиева</b> ОБНАРУЖЕНИЕ SYBIL И WORMHOLE АТАК В БЕСПРОВОДНОЙ СЕНСОРНОЙ СЕТИ.....	171
<b>А.А. Исмаилова, Ж.Т. Бельдеубаева, А.А. Нурпейсова, Г.О. Исакова, Ж.З. Жантасова</b> ОБНАРУЖЕНИЕ БОЛЕЗНЕЙ РАСТЕНИЙ МЕТОДОМ ГЛУБОКОГО ОБУЧЕНИЯ.....	184
<b>А.Х. Касымова, М.Б. Есенова, М.У. Худойбергенов, А.Б. Остаева, М.Г. Байбулова</b> ПРИМЕНЕНИЕ АЛГОРИТМОВ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ КЛАССИФИКАЦИИ БОЛЕЗНЕЙ СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР.....	198
<b>А.У. Мухиядин, М.У. Мукашева, У.Т. Махажанова, А.А. Муханова, Ж.Б. Ламашева</b> ИЗУЧЕНИЕ ВЛИЯНИЯ ЭКСТРЕМАЛЬНОГО ДИСТАНЦИОННОГО ОБРАЗОВАНИЯ НА УЧАЩИХСЯ С ИСПОЛЬЗОВАНИЕМ ПРОГРАММНЫХ СРЕДСТВ.....	209
<b>Б.Б. Оразбаев, Л.Т. Салыбек, К.Н. Оразбаева, Ш.К. Коданова, С.Ш. Исакова</b> МЕТОД РАЗРАБОТКИ МОДЕЛЕЙ ДЛЯ ОПТИМИЗАЦИИ ПРОЦЕССОВ ЭЛЕКТРООБЕССОЛИВАНИЯ И ОБЕЗВОЖИВАНИЯ ПРИ ПЕРВИЧНОЙ ПЕРЕРАБОТКЕ НЕФТИ.....	224
<b>С.К. Серикбаева, М.К. Болсынбек, А.Д. Абдувалова, А.Т. Абдыхалык, Д.Е. Ануарбек</b> ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ КАЧЕСТВА ПОЧВЫ: АЛГОРИТМЫ И МЕТОДИКИ.....	237
<b>А.Ж. Танирбергенов, Ж.К. Тасжурекова, С.К. Серикбаева, А.А. Шораев, А.Д. Абдувалова</b> МЕТОДЫ ПОСТРОЕНИЯ МОДЕЛИ И ИНФОРМАЦИОННОЙ СИСТЕМЫ ПОИСКА ПОЛНОТЕКСТОВЫХ ДОКУМЕНТОВ.....	253
<b>А.Ә. Taurbekova, O.Zh. Mamyrbayev, K.Zh. Doshtaev, T.K. Eginbaykyzy</b> HYDRODYNAMIC INSTABILITY MECHANISM PROCESS FOR ASSESSMENT SEISMIC ACTIVITY.....	268
<b>Н.Т. Туржанов, Ш.К. Ележанова, С.Н. Идрисов, Ж.К. Дюсембина</b> РАЗРАБОТКА ИННОВАЦИОННОГО КУРСА ПО РЕИНЖИНИРИНГУ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ.....	290
<b>В. Шевцов, А. Исмаилова, Ж. Бельдеубаева, А. Сатыбалдиева, А. Нурпейсова</b> MLVA КАК МЕТОД ГЕНОТИПИРОВАНИЯ И АЛГОРИТМЫ ЕГО РЕАЛИЗАЦИИ С ИСПОЛЬЗОВАНИЕМ ПОЛНОГЕНОМНЫХ ДАННЫХ.....	300
<b>А.А. Шекербек, А.А. Некесова, Ж.Ж. Молдашева, А.И. Онгарбаева, А.О. Тохаева</b> АНАЛИЗ ПАТОЛОГИЧЕСКИХ СОСТОЯНИЙ ЛЕГКИХ С ИСПОЛЬЗОВАНИЕМ ФРАКТАЛЬНОГО МЕТОДА.....	313

## CONTENTS

<b>G.B. Abdikerimova, R.M. Amanov, G.T. Azieva, A.M. Zamanbekova, K. Zhengskankyzy</b> COMPARATIVE ANALYSIS OF SOUND PROCESSING METHODS IN THE CHORD RECOGNITION PROBLEM USING MACHINE LEARNING.....	7
<b>L. Abdykerimova, G. Murzabekova, G. Omarova, L. Akzullakyyzy, G. Mussagulova</b> DETECTION OF CARDIAC PATHOLOGY USING DEEP LEARNING METHODS.....	21
<b>A.E. Abzhanova, E.A. Abzhanov, A.A. Myrzamuratova, A.G. Batyrkhanov, A.B. Bekseitova</b> SOIL MOISTURE OBTAINED BY REMOTE SENSING.....	35
<b>U. Zh Aitimova, M.Zh. Aitimov, E.N. Tulegenova, A.U. Yessirkepova, Zh.T. Abildaeva</b> FIRE FOCUS DETECTION USING DEEP LEARNING METHODS FROM IMAGE.....	50
<b>K.M. Aldabergenova, M.ZH. Zhasuzakova, M.Zh. Aitimov, N.T. Mustafaeva, K.K. Daurenbekov</b> DIGITALIZATION OF AGRICULTURE: OPPORTUNITIES AND PROSPECTS FOR DEVELOPMENT.....	64
<b>A.S. Baegizova, G.I. Mukhamedrakhimova, Zh.B. Lamasheva, A.Z. Abdrakhmanova, T.T. Ospanova</b> IMPROVE IMAGE QUALITY WITH DEEP LEARNING TECHNIQUES.....	78
<b>G. Bekmanova, A. Omarbekova, M. Kantureyeva, N. Baigabylov, M. Kudabekov</b> INFORMATION TECHNOLOGIES IN SOCIOLOGICAL SURVEY RESEARCH.....	91
<b>M.A. Bersugir, G.U. Mamatova, A.A. Nurpeisova, M.B. Ongarbayeva, Zh.T. Altynbekova</b> USING MACHINE LEARNING METHODS TO IMPROVE TEXTURE-TYPE IMAGES.....	104
<b>M. Bolatbek, K. Baisylbaeva, M. Sagynay, Sh. Mussiraliyeva, A. Zhumakhanova</b> DEVELOPMENT OF A PARSER PROGRAM FOR THE ACCUMULATION OF DESTRUCTIVE TEXTS AIMED AT YOUNG PEOPLE IN THE INTERNET SPACE.....	117
<b>M. Bolsynbek, G. Abdikerimova, G. Omarova, A. Ostayeva, A. Batyrkhanov</b> APPLICATION OF MACHINE LEARNING TO PREDICT SOIL PREPARATION....	132
<b>Sh.K. Yelezhanova, A.G. Batyrkhanov, A.Y. Chukurov, B.S. Khairzhanova, J.A. Taghiyev</b> TYPOLOGY OF INFORMATION UNCERTAINTY AND TYPES OF INFORMATION RETRIEVAL.....	151
<b>M. Yesmagambetova, T. Ospanova, L. Bobrov, T. Ten, T. Yesmagambetov</b> SELECTION OF COLORIMETRY SOFTWARE TOOLS IN IMAGE PROCESSING OF SPACE MONITORING OF EMERGENCY SITUATIONS.....	161
<b>T. Zhukabayeva, A. Adamova, B. Khu Ven-Tsen, Y. Mardenov, L. Zholshiyeva</b> DETECTION OF SYBIL AND WORMHOLE ATTACKS IN A WIRELESS SENSOR NETWORK.....	171
<b>A.A. Ismailova, Zh.T. Beldeubayeva, A.A. Nurpeisova, G.O. Issakova, Zh.Z. Zhantassova</b>	



DETECTION OF PLANT DISEASES USING DEEP LEARNING METHODS.....	184
<b>A.K. Kassymova, M.B. Yessenova, M.U. Khudoyberganov, A.B. Ostayeva, M.G. Baibulova</b>	
APPLICATION OF DEEP LEARNING ALGORITHMS FOR CLASSIFICATION OF DISEASES OF AGRICULTURAL CROPS.....	198
<b>A. Mukhiyadin, M. Mukasheva, U. Makhazhanova, A. Mukhanova, Zh. Lamasheva</b>	
STUDYING THE EFFECTS OF EXTREME DISTANCE EDUCATION ON STUDENTS USING SOFTWARE TOOLS.....	209
<b>B. Orazbayev, L. Salybek, K. Orazbayeva, Sn. Kodanova, S. Iskakova</b>	
METHOD FOR DEVELOPING MODELS FOR OPTIMIZING PROCESSES OF ELECTRICAL DESALTING AND DEHYDRATION DURING PRIMARY OIL PROCESSING.....	224
<b>S.Serikbayeva, M.Bolsynbek, A. Abduvalova, A. Abdykhalyk, D. Anuarbek</b>	
APPLICATION OF MACHINE LEARNING TO PREDICT SOIL QUALITY: ALGORITHMS AND TECHNIQUES.....	237
<b>A. Tanirbergenov, Zh. Tashhurekova, S. Serikbayeva, A. Shorayev, A. Abduvalova</b>	
METHODS OF CONSTRUCTING A MODEL AND AN INFORMATION SYSTEM FOR SEARCHING FULL-TEXT DOCUMENTS.....	253
<b>A.Ə. Taurbekova, O.Zh. Mamyrbayev, K.Zh. Doshtaev, T.K. Eginbaykyzy</b>	
HYDRODYNAMIC INSTABILITY MECHANISM PROCESS FOR ASSESSMENT SEISMIC ACTIVITY.....	268
<b>N.T. Turzhanov, Sh.K. Yelezhanova, S.N. Idrissov, Zh.K. Dyusseminina</b>	
DEVELOPMENT OF AN INNOVATIVE COURSE REENGINEERING OF INFORMATION PROCESSES.....	290
<b>V. Shevtsov, A. Ismailova, Zh. Beldeubayeva, A. Satybaldiyeva, A. Nurpeisova</b>	
MLVA AS A METHOD OF GENOTYPING AND ALGORITHMS FOR ITS IMPLEMENTATION USING GENOME-WIDE DATA.....	300
<b>A.A. Shekerbek, A.A. Nekesova, Zh.Zh. Moldasheva, A.I. Ongarbayeva, A. Tokhaeva</b>	
ANALYSIS OF PATHOLOGICAL CONDITIONS OF THE LUNG USING THE FRACTAL METHOD.....	313

**Publication Ethics and Publication Malpractice  
the journals of the National Academy of Sciences of the Republic of Kazakhstan**

For information on Ethics in publishing and Ethical guidelines for journal publication see <http://www.elsevier.com/publishingethics> and <http://www.elsevier.com/journal-authors/ethics>.

Submission of an article to the National Academy of Sciences of the Republic of Kazakhstan implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see <http://www.elsevier.com/postingpolicy>), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The National Academy of Sciences of the Republic of Kazakhstan follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct ([http://publicationethics.org/files/u2/New\\_Code.pdf](http://publicationethics.org/files/u2/New_Code.pdf)). To verify originality, your article may be checked by the Cross Check originality detection service <http://www.elsevier.com/editors/plagdetect>.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the National Academy of Sciences of the Republic of Kazakhstan.

The Editorial Board of the National Academy of Sciences of the Republic of Kazakhstan will monitor and safeguard publishing ethics.

Правила оформления статьи для публикации в журнале смотреть на сайтах:

**[www.nauka-nanrk.kz](http://www.nauka-nanrk.kz)**

**<http://physics-mathematics.kz/index.php/en/archive>**

**ISSN 2518-1726 (Online),**

**ISSN 1991-346X (Print)**

Подписано в печать 28.12.2023.

Формат 60x881/8. Бумага офсетная. Печать – ризограф.

21,0 п.л. Тираж 300. Заказ 4.