

2. S2 SARS rep UniProt (<https://www.uniprot.org/uniprot/P0C6X7> – SARS rep UniProt, P0C6X7. Дата обращения 17 апреля 2020).

3. Altschul S.F. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs / S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman // *Nucleic Acids Res.* 1997 – Sep 1;25(17):3389-402.

4. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004 Mar 19;32(5):1792-7.

5. Tamura K. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0 / K. Tamura, G. Stecher, D. Peterson, A. Filipinski, S. Kumar // *Mol. Biol. Evol.* 2013. V. 30. P. 2725-2729.

6. Jones D.T., Taylor W.R., and Thornton J.M. The rapid generation of mutation data matrices from protein sequences / Jones D.T., Taylor W.R., and Thornton J.M. // *Computer Applications in the Biosciences.* 1992. 8: 275-282.

7. Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap / J. Felsenstein // *Evolution.* 1985. 39:783-791.

DOI:10.34617/05jd-vj11  
УДК 574.24:591

### **СПЕЦИФИЧНАЯ КЛАСТЕРИЗАЦИЯ ОКЕАНИЧЕСКИХ БЕЛКОВ DENV С БЕЛКАМИ БАКТЕРИЙ НАЗЕМНЫХ ЖИВОТНЫХ**

**Карманова Александра Николаевна<sup>1,2</sup>**

**Зимин Андрей Антонович<sup>1</sup>**, канд. биол. наук

<sup>1</sup>*Институт биохимии и физиологии микроорганизмов им. Г.К. Скрыбина РАН – обособленное подразделение ФИЦ*

*«Пушкинский научный центр биологических исследований РАН», г. Пушкино, Российская Федерация*

<sup>2</sup>*Вятский государственный университет, г. Киров, Российская Федерация*

В данной работе исследовано эволюционное сходство гомологов DenV из микробиоты пелагики океана и микробиоты наземных животных, в том числе и той, которая является обычным компонентом сточных вод животноводческих предприятий. Филогенетика данного набора аминокислотных последовательностей гомологов DenV позволяет сделать предположения о различиях и сходствах в резистентности к УФ-излучению у бактерий этих двух экологических ниш. Это открывает фундаментальные возможности прогнозирования в области очистки сточных вод ферм УФ-светом и возможности самоочистки морских вод за счет природной инсоляции от бактериальной составляющей сельскохозяйственной контаминации.

**Ключевые слова:** устойчивость к ультрафиолету; DenV; очистка сточных вод ферм; микробиота пелагики океана; микробиота наземных животных

### **SPECIFIC CLUSTERING OF THE OCEAN DENV PROTEINS WITH PROTEINS OF TERRESTRIAL ANIMAL BACTERIA**

**Karmanova Aleksandra Nikolaevna<sup>1,2</sup>**

**Zimin Andrei Antonovich**<sup>1</sup>, PhD Biol. Sci.

<sup>1</sup>*Institute of Biochemistry and Physiology of Microorganisms named after G. K. Scriabin RAS – a separate subdivision of the Federal Research Center «Pushchino Scientific Centre for Biological Research of the Russian Academy of Sciences», Pushchino, Russian Federation*

<sup>2</sup>*Vyatka State University, Kirov, Russian Federation*

In this paper, we studied the evolutionary similarity of DenV homologs from the microbiota of the ocean pelagic and microbiota of terrestrial animals, including one that is a common component of sewage of livestock farms. Phylogeny of this set of amino acid sequences DenV homologues allow us to make assumptions about differences and similarities in resistance to UV radiation in bacteria of these two ecological niches. As a consequence, it opens the fundamental forecasting possibilities in the field of farm wastewater treatment with UV light and self-cleaning capabilities seawater due to natural isolation from the bacterial component of agricultural contamination.

**Key words:** resistance to UV radiation; DenV; farm wastewater treatment; microbiota of the ocean pelagic; microbiota of terrestrial animals

Ультрафиолетовое излучение является одним из методов инактивации для широкого спектра бактерий, вирусов и паразитов. Инактивация микроорганизмов с помощью УФ-излучения в сточных водах животноводческих предприятий и ферм обусловлена появлением модификаций их геномной ДНК, из которых наиболее распространенным повреждением является появление тиминовых димеров. Принимая во внимание, что эти модификации ингибируют полимеризацию ДНК *in vivo*, а в их репарации ведущую роль играет ДНК-гликозилаза пириимидиновых димеров, продукт гена *denV* [6, 7], который очень широко распространен в природе, мы предприняли широкий сравнительный филогенетический анализ гомологов этого белка из метагеномных океанических проб с родственными белками различных бактерий, в том числе с теми, которые могут быть встречены в сточных водах животноводческих предприятий. Дополнительной актуальностью этой работы является наличие таких ферм по разведению сельскохозяйственных животных в различных прибрежных районах Российской Федерации в том числе и в Краснодарском и Приморском краях. Горизонтальный перенос генетической информации между бактериями весьма эффективен. В случае попадания бактерий, устойчивых к ультрафиолету из сточных вод животноводства в море мо-

жет происходить перенос генов этих бактерий в геномы морских бактерий и наоборот локализация мест животноводства в прибрежных морских районах может приводить к переносу генов устойчивости к ультрафиолету от морских бактерий к различным бактериальным возбудителям инфекций у сельскохозяйственных животных, таких как коровы, свиньи, производственная птица. Для понимания потенциала генов, устойчивости к УФ-излучению в двух этих бактериальных мирах - океане и животноводстве мы предприняли филогенетический анализ набора характерных бактериальных маркеров DenV- белков наземных бактерий и белков, последовательности которых мы извлекли ранее из метагенома поверхностных вод мирового океана из БД экспедиции Крега Вентера «Сорсер» и «Сорсер2».

Мы использовали для анализа эволюционных связей этих белков два метода: максимальной экономии (Maximum Parsimony) и наибольшего правдоподобия (Maximum Likelihood). Оба рассматриваемых нами метода относятся к дискретным. В данном типе методов сравниваются не расстояния между выровненными последовательностями, а отдельные сайты в них [3]. Также они позволяют проводить анализ всех возможных топологий дерева и выбора наиболее подходящего.

Maximum Parsimony – это критерий оптимальности, при котором строится филогенетическое дерево, минимизирующее общее количество изменений – мутаций того или иного белка. В этом методе такое дерево является предпочтительным. Согласно критерию максимальной экономии, оптимальное дерево минимизирует количество гомоплазии (то есть минимизирует конвергентную эволюцию, параллельную эволюцию или реверсии гена). Другими словами, согласно этому критерию, кратчайшее дерево, которое объясняет сходство между последовательностями в имеющемся наборе данных, считается наилучшим. Из-за того, что метод использует все возможные топологии дерева – его сложно использовать при достаточно большом количестве таксонов, ввиду слишком большого времени вычислений. В Maximum parsimony эта проблема решается путем рассмотрения только информативных сайтов в сравниваемых последовательностях, т.е. имеющих не менее двух разных типов нуклеотидов и каждый присутствует как минимум в двух последовательностях. Плюсы данного метода заключаются в том, что он не уменьшает объем исходной информации, исследует все возможные топологии дерева, пытается предсказать последовательности предков. К недостаткам можно отнести низкую скорость работы, использование только информативных сайтов, а не полных сиквенсов, отсутствие информации о длине ветвей, обладание высокой чувствительности к смещению в исходных данных, невозможность коррекции множественных мутаций [1].

В методе Maximum likelihood анализируются все сайты в сравниваемых последовательностях. В основе метода лежит максимизация критерия выбора правдоподобия на основе выбранной модели – все полученные топологии деревьев оцениваются по правдоподобию и выбирается именно то, которое имеет наибольшее значение критерия, аналогичного Maximum parsimony – варианта с

минимальным количеством мутаций. Проблема с большим количеством вычислений может быть решена путем построения деревьев только по некоторым частям сайтов. Плюсы метода – одновременно с построением дерева дается оценка его правдоподобия, возможность задавать различные скорости мутаций, а также использовать различные модели эволюции. Данный метод позволяет обнаруживать сходство даже для эволюционно удаленных последовательностей. Основной недостаток – это большое количество вычислений, и как следствие низкая скорость работы [1].

При построении нами деревьев также использовался алгоритм бутстрэп, служащий для статистического подтверждения корректности результатов [1]. Метод основан на построении серии данных путем создания многочисленных последовательностей из случайных выборок исходной последовательности. Он позволяет проверить наличие статистических ошибок и смещения в исходных данных. Если данные не содержат смещения и других стохастических искажений, то все деревья, построенные на основе бутстрэп данных, будут иметь практически одинаковую топологию [1].

**Методика исследований.** Для сравнения а.к. последовательностей эндонуклеазы DenV бактериофага T4 с базами данных белковых последовательностей использовался алгоритм PSI-BLAST [2]. Уровень достоверности результатов E-value <  $3e^{-29}$ . Полученный файл с а.к. последовательностями гомологов DenV T4 в формате FASTA объединили в один файл и использовали для обработки в пакете программ MEGA6 [8]. Выравнивание осуществляли при помощи алгоритма MUSCLE [3]. В анализе приняли участие 65 аминокислотных последовательностей. Все позиции, содержащие пробелы и пропущенные данные, были исключены. Всего в итоговом наборе данных было 6 позиций. Эволюционный анализ был проведен в MEGA6 [8]. Эволюционная история

была выведена с использованием нескольких методов.

Первый используемый метод – Maximal likelihood, основанный на матричной модели JTT [5]. Начальное дерево консенсуса выведено из 1000 повторов с использованием бутстрепа. Исходные деревья для эвристического поиска были получены автоматически с использованием метода Maximum Parsimony к матрице попарных расстояний, оцененной с использованием модели JTT. В итоге получили дерево с наибольшим логарифмическим правдоподобием (-264,3223) – представлено на рисунке 1 А.

Второй используемый метод – Maximum Parsimony. Эволюционная история была выведена с использованием метода максимальной экономии. Начальное дерево консенсуса, выведенное из 1000 повторов, используется для представления эволюционной истории анализируемых таксонов [4]. Ветви, соответствующие разделам, воспроизведенным менее чем в 50% копий начальной загрузки, свернуты. В итоге получили дерево, представленное на рисунке 1 Б.

**Результаты исследований и их обсуждение.** В дереве, полученным методом Maximum likelihood (рисунок 1 А), можно выделить две ветви гомологов. В первую вошли эндонуклеазы всех используемых фагов, а также бактерии-возбудители острых кишечных инфекций и обитатели кишечной микрофлоры. Ветвь кажется не сильно однородной, разделение наблюдается скорее ступенчатое.

Большинство использованных нами эндонуклеаз морского метагенома и цианобактерий образовали вторую крупную ветвь гомологов. Помимо этого, в нее попали последовательности *Brucella suis* и *Verrucomicrobiaceae bacterium* – распространенной почвенной бактерии. Остальные гликозилазы метагеномов оказались объединены в небольшие подветви. Интересно расположение последовательности из морской бактерии *Alteromonas sp*,

попавшей в одну подветвь с патогенными бактериями и бактериями микрофлоры носоглотки.

Также интересен тот факт, что эндонуклеазы бруцелл схожи с различными последовательностями метагеномов, а не между собой и бактериями почвы. Это может говорить о наличие горизонтального переноса между бруцеллами и пикопланктоном, ввиду попадания сточных вод животноводства в океаны или из-за близкого нахождения животноводческих ферм к прибрежным морским водам, где и могут данные микроорганизмы пересекаться.

Гликозилаза *Escherichia virus T4* имеет больше всего сходств с последовательностью *Cellulomonas oligotrophica* – почвенной бактерией-целлюлозолитиком, образуя с ней отдельную ветвь. Ближайшая к ней ветвь объединила в себе белки *Clostridioides difficile* и *Escherichia phage vB EcoM-UFV13*.

Судя по дереву, эти две ветви занимают промежуточное положение между ветвями фагов и крупной ветвью последовательностей из цианобактерий и метагенома. Можно предположить, что эндонуклеазы цианобактерий и метагенома достаточно схожи с таковой у фагов эшерихии и почвенными бактериями. Вполне возможно, что когда-то между ними произошел горизонтальный перенос, но нужно проводить дальнейшие исследования для окончательного вывода.

На дереве Maximum parsimony (рисунок 1 Б) получили интересное распределение выбранным нами объектов по небольшим группам, тенденции сходны с распределением на дереве, полученным методом Maximum likelihood, но прослеживаются не так четко. В целом, по всему дереву можно сказать, что гликозилазы большей части выбранных нами фагов схожи между собой и чаще объединены в мелкие клады по парам, ступенчатое распределение заметно в меньшей степени. Наиболее близкой большей части вирусов оказались эндонуклеазы *Salmonella*

*enterica*, серовары которой могут быть патогенными для человека и животных, и, ожидаемо, *Escherichia coli*.

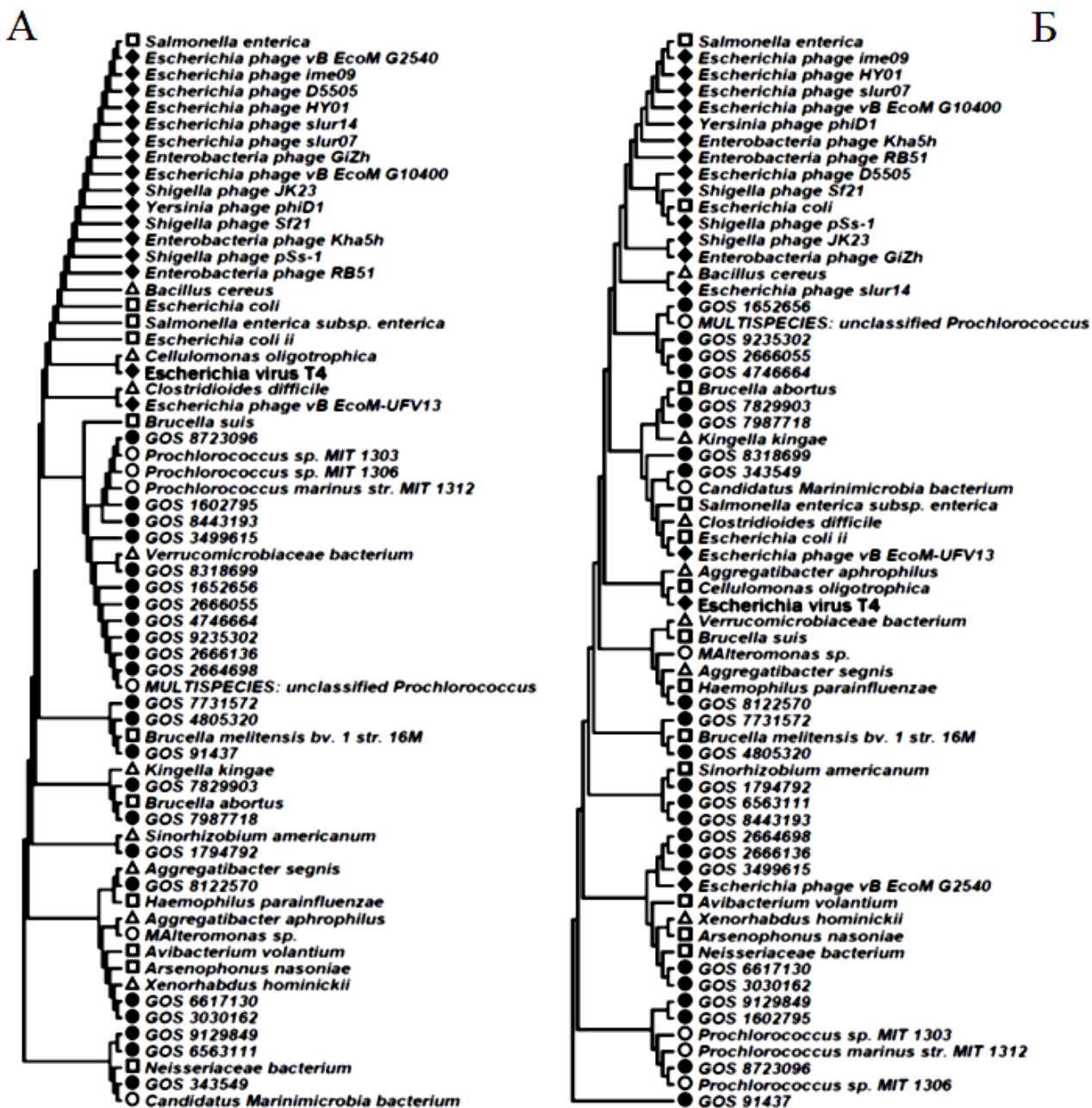


Рисунок 1 – Два филогенетических дерева эволюционного сходства белков гомологов DenV – эндонуклеазы фага *Escherichia virus T4* – отмечена черным ромбом и более крупным шрифтом. Под буквой А – филогенетическое дерево, построенное методом Maximum likelihood с учетом бустрепов, Б – методом Maximum parsimony с учетом бустрепов. Последовательности из метагенома океана обозначены черными кружочками, белыми – из морских бактерий родов *Prochlorococcus sp.*, *Malteromonas sp.*, *Candidatus Marinimicrobia bacterim sp.* Эндонуклеазы бактериофагов бактерий родов *Enterobacteria sp.*, *Shigella sp.*, *Yersinia sp.*, *Escherichia sp.*, обозначены черными ромбами, последовательности из наземных бактерий – белыми квадратами и треугольниками

Интересно, что наиболее близкими к последней оказались фаги шигелл, а не ее собственные. Как и в предыдущем дереве, гликозилаза бактериофага Т4 оказалась сильно схожей с *Cellulomonas Oligotrophica*. Отличие состоит в том, что к данной ветви прибавилась последовательность из *Aggregatibacter aphrophilus* – бактерии, которая в норме может быть представителем микрофлоры носоглотки человека.

Последовательности из морских бактерий распределились в небольшие группы вместе с эндонуклеазами метагена. Самой разнородной получилась группа, которая наиболее близка с вышеописанной ветвью с последовательностью из Т4. В нее вошли как и последовательности метагена и морской *Alteromonas sp.*, так и белки из почвенных бактерий, бактерий микрофлоры человека, патогенных *Haemophilus parainfluenzae* и *Brucella suis*. Последняя образует отдельную ветвь с *Verrucomicrobiaceae bacterium* – широко распространенной почвенной бактерии. Расположение остальных эндонуклеаз представителей рода бруцелл схоже с расположением на первом дереве, они также оказались близки к последовательностям метагена. Также, в отдельную ветвь, как и в дереве Maximum likelihood выделилась большая часть белков из цианобактерий, в нее же вошли несколько эндонуклеаз метагена. Последовательность из *Multispecies: unclassified Prochlorococcus* вошла в отдельную ветвь, но также имеет больше всего сходств с гликозилазами метагена.

В целом, по этому дереву можно сказать, что есть выраженное сходство между последовательностями метагена и морскими бактериями, а также между эндонуклеазами фагов. Общее сходство между двумя этими группами выделить сложнее.

**Выводы.** 1. В ходе исследования получили два филогенетических дерева построенных разными методами, но относящихся к общему типу и построенных на относительно общих признаках.

2. На каждом из филогенетических деревьев наблюдается тенденция объединения эндонуклеаз фагов и последовательностей из океанического метагена, цианобактерий и морских бактерий в собственные ветви.

3. Более логичные и наглядные результаты дало филогенетическое дерево, построенное методом Maximum likelihood. Вышеописанные тенденции там проявляются лучше. Данный метод подходит для выявления филогенетических связей между последовательностями, имеющими мало общего между собой, поэтому именно на данные результаты стоит обращать больше внимания.

4. Гликозилаза Escherichia virus Т4 и самая близкая к ней гликозилаза *Cellulomonas Oligotrophica* в обоих случаях занимала промежуточное положение между этими ветвями, но все же они имели больше сходств с фаговыми последовательностями. Согласно этому, можно предположить, что когда-то между фагами эшерихии, почвенными бактериями и морскими бактериями произошел горизонтальный гена данной эндонуклеазы, но нужно проводить дальнейшие исследования для окончательного вывода и прослеживания данной цепочки событий.

5. На обоих эволюционных деревьях мы получили интересное распределение последовательностей из бактерий, патогенных для человека и животных, в частности, они часто оказывались близки с эндонуклеазами океанического метагена. Это говорит в пользу наличия горизонтального переноса гена, который влияет на устойчивость к УФ-излучению, между группами морских бактерий и бактерий наземных, в том числе и сельскохозяйственных животных.

6. Для подтверждения данных предположений стоит проводить больше подобных исследований.

7. Открытые нами зависимости между наличием белка, участвующего в репарации после повреждения ДНК УФ-излучением у разных групп микроорга-

низмов стоит учитывать при выборе способов обеззараживания сточных вод и прочих жидких отходов животноводства.

**Благодарности.** Исследование было частично поддержано грантом РФФИ №20-54-53018 ГФЕН\_а для Зимина А.А. и выполнено им в рамках этого проекта.

### Список литературы

1. Волков Ю.П. Анализ эффективности некоторых методов построения филогенетических деревьев, используемых при оценке эволюционного родства микроорганизмов / Ю.П. Волков, Г.А. Ерошенко // «Проблемы особо опасных инфекций». 2009. Вып.99. С.35-41.
2. Altschul S.F. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs / S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman // Nucleic Acids Res. 1997. 25: 3389-3402.
3. Edgar R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput / R.C. Edgar // Nucleic Acids Res. 2004. 32: 1792-1797.
4. Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap / J. Felsenstein // Evolution. 1985. 39: 783-791.
5. Jones D.T. The rapid generation of mutation data matrices from protein sequences / D.T. Jones, W.R. Taylor, J.M. Thornton // Comput Appl Biosci. 1992. 8: 275-282.
6. McMillan S. Den V gene of bacteriophage T4 codes for both pyrimidine dimer-DNA glycosylase and apyrimidinic endonuclease activities / S. McMillan, J. Edenberg, E.H. Radany, R.C. Friedberg, E.C. Friedberg // J Virol. 1981. 40: 211-223.
7. Rubin J.S. The molecular genetics of the incision step in the DNA excision repair process / J.S. Rubin // Int J Radiat Biol. 1988. 54:309-365.
8. Tamura K. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. / K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar // Molecular Biology and Evolution. 2013. 30: 2725-2729.

DOI:

УДК 636.237.21.082

## ЭФФЕКТИВНОСТЬ РАЗВЕДЕНИЯ МОЛОЧНОГО СКОТА ЧЕРНО-ПЕСТРОЙ ПОРОДЫ ПО ЛИНИЯМ

**Ковалева Галина Петровна**, канд. с-х. наук, доцент  
ФГБНУ «Северо-Кавказский федеральный научный аграрный центр»,  
г. Михайловск, Российская Федерация

В статье приведены исследования по влиянию линейной принадлежности на продуктивность животного и качественные показатели молока и промеров животных. Работа проводилась в СПК колхозе-племзаводе «Казьминский» Кочубеевского района Ставропольского края на поголовье скота черно-пестрой породы. Установлено влияние линии на хозяйственно-полезные качества коров.

**Ключевые слова:** крова; линия; молочная продуктивность; содержание жира и белка в молоке