

УДК 635.63:631.527

03.00.00 Биологические науки

**ВИЗУАЛЬНОЕ ФЕНОТИПИРОВАНИЕ В СЕЛЕКЦИИ РАСТЕНИЙ**

Цаценко Людмила Владимировна  
д-р. биол. наук, профессор, кафедра генетики,  
селекции и семеноводства  
[lvt-lemna@yandex.ru](mailto:lvt-lemna@yandex.ru)  
ID 2120-6510

Савиченко Дмитрий Леонидович  
Магистрант  
[d\\_savichenko@mail.ru](mailto:d_savichenko@mail.ru)  
*Кубанский государственный аграрный  
Университет имени И.Т.Трубилина, Россия,  
Краснодар 350044, Калинина 13*

Статья посвящена визуальному анализу и его применению в селекционном процессе. Рассмотрены новые возможности применения визуального отображения информации в исследовательском процессе. Приведены примеры использования визуального фенотипирования в селекции, оценены его преимущества и недостатки в этой области. В частности описан алгоритм программы LHDetect для оценки опущения листовых пластин пшеницы, основанный на анализе цифровых изображений сгиба листа. Так же рассмотрена платформа Phenotiki, использующая цифровые снимки профиля растений для визуального фенотипирования с целью оценки роста растений. На основе анализа множества исследований было отмечено, что применение визуального анализа для одних целей крайне эффективно, а для других, напротив, непригодно. Освещены наиболее актуальные проблемы правильного представления и конвертации информации. Показаны различные причины потери и искажения данных, а так же возможные пути решения этих проблем. Рассмотрено применение АСК-анализа в селекционном процессе. В качестве примера описаны работы по оценке полиморфизма сортогрупп винограда. Представлены, найденные авторами этих работ, интересные решения и пути для минимизации искажения и потерь информации в процессе ее преобразования. Раскрыто понятие «компьютерной биологии» и описаны принципиальные отличия от других направлений. Определены задачи, проблемы, возможные области исследований, а так же перспективные направления развития компьютерной биологии

Ключевые слова: ВИЗУАЛЬНОЕ ФЕНОТИПИРОВАНИЕ, ЦИФРОВОЙ ОБРАЗ, ВИЗУАЛЬНЫЙ АНАЛИЗ, ТЕХНОЛОГИИ ОБРАБОТКИ ОБРАЗОВ, ГЕНЕТИКО-

UDC 635.63:631.527

Biological sciences

**VISUAL PHENOTYPING IN PLANT BREEDING**

Tsatsenko Luidmila Vladimirovna,  
Dr.Sci.Biol., professor,  
Chair of genetic, plant breeding and seeds  
[lvt-lemna@yandex.ru](mailto:lvt-lemna@yandex.ru)  
ID 2120-6510

Savichenko Dmitriy Leonidovich  
student  
[d\\_savichenko@mail.ru](mailto:d_savichenko@mail.ru)  
*Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, Russia*

The article is devoted to visual analysis and its application in the selection process. New possibilities of application of visual display of information in the research process are considered. We have presented the examples of the use of visual phenotyping in breeding, evaluated the advantages and disadvantages in this area. In particular, the algorithm of the program of LHDetect for estimating leaf pubescence of wheat wafers is described, based on the analysis of digital images of the sheet bend. We have also considered Phenotiki platform that uses digital images of the profile of plants for visual phenotyping to assess plant growth. Based on the analysis of various studies it was noted that the application of visual analysis for some purposes is very effective, but for others, on the contrary, not effective. We highlight the actual problems of proper representation and converting information. The article shows the various causes of loss and distortion of data, as well as possible solutions to these problems. The work describes the application of ASK-analysis in the selection process. As an example, it describes the evaluation of the polymorphism of variety groups of grapes. We have presented interesting solutions and ways to minimize distortion and loss of information in the process of its transformation found by the authors of these works. The article reveals the concept of "computational biology" and describes fundamental differences from other areas as well as tasks, problems, possible research areas and promising directions of development of computer biology

Keywords: VISUAL PHENOTYPING, DIGITAL IMAGE, VISUAL ANALYSIS, TECHNOLOGY OF IMAGE PROCESSING, GENETIC AND BREEDING WORKS

## СЕЛЕКЦИОННЫЕ РАБОТЫ

**Doi: 10.21515/1990-4665-128-071**

Сегодня исследователи имеют большой арсенал вспомогательных средств: электронные микроскопы, цифровые камеры, компьютеры и сканеры, мощные программные продукты, а также робототехнику. Однако графические методы анализа – сложное дело, требующее определенных навыков и культуры в области подачи визуальной информации. Полезно сочетать огромные вычислительные ресурсы современных компьютеров с творческим и гибким человеческим мышлением. Визуальный анализ данных призван вовлечь человека в процесс поиска знаний в данных. Основная идея заключается в том, чтобы представить большие объёмы данных в такой форме, в которой человек смог бы увидеть то, что трудно выделить алгоритмически. Имел возможность погрузиться в данные, работать с их визуальным представлением, понять их суть, сделать выводы и напрямую взаимодействовать с данными. Из-за сложности информации, это не всегда возможно в простейших графических видах представления знаний, таких как деревья решений, дейтаграммы, двумерные графики и т.п.[2].

В связи с этим возникает необходимость в более сложных средствах отображения информации и результатов анализа. С помощью новых технологий исследователь способен оценивать: большие объекты и маленькие, далеко они находятся или близко. Пользователь в реальном времени может двигаться вокруг объектов или кластеров объектов и рассматривать их со всех сторон. Это позволяет использовать для анализа естественные человеческие перцепционные навыки в обнаружении неопределённых образцов в визуальном трёхмерном представлении данных.

Визуальный анализ данных особенно полезен, когда о самих данных мало что известно и цели исследования до конца не ясны. За счёт того, что пользователь напрямую работает с данными, представленными в виде ви-

зуальных образов, которые он может рассматривать с разных сторон и под любыми углами зрения, в прямом смысле этого слова, он может получить дополнительную информацию, которая поможет ему более четко сформулировать цели исследования [4,5].

Визуальное фенотипирование, в последнее время, интенсивно используется в селекционном процессе и включает в себя работу приборами, что, по сути, в основном автоматический процесс. Под визуальным понимается производимый невооруженным глазом или с помощью оптических приборов анализ или направленное наблюдение за объектом (в отличие от фотографирования или фотоэлектрических наблюдений).

Растения движутся по ленте в камеру с нормальным светом, одновременно на растения со всех сторон воздействуют инфракрасным светом. Если комбинировать с приемами для анализа изображения, то можно изучать на такой установке до 100000 растений риса в год, и при этом учитывать урожайность, число зерен и их размер, фотосинтетическую активность, биомассу листьев и корней, а также ряд других признаков. Основных преимуществ такого метода меньше, поскольку большое количество генотипов может быть охарактеризовано, по-прежнему возделыванием в поле продуктивными «высокопроизводительными» приемами. Например, при линейной селекции при самой маленькой селекционной программе каждые 500  $F_2$ -растений от 200 скрещиваний в поле, т.е. 100000 генотипов только в  $F_2$  поколении [1].

Преимущество автоматизированных систем фенотипирования при контролируемых условиях состоит в том, что многие признаки (площадь листьев, биомасса и другие) могут быть оценены значительно точнее, и, прежде всего, наблюдения могут проводиться непрерывно и без нарушений.

Эти наблюдения проводятся, конечно, при искусственных условиях и многие важные свойства, такие, как глубина проникновения корней или

полевая устойчивость против главных болезней, не исследуются. Обобщив имеющуюся информацию можно сделать заключение, что для некоторых селекционных целей эта система идеальна, для других непригодна [2,7].

Примером инновационного подхода к анализу опущения листа могут служить разработки на основе высокопроизводительного фенотипирования (рисунок 1).



Рисунок 1 – Растения в теплице. Конвейер визуального фенотипирования.

До недавнего времени описание морфологии опущения было основано на визуальной или тактильной оценке. Такая оценка являлась ненадежной

и требовала, чтобы вся работа была выполнена одним специалистом. Для определения морфологических характеристик опушения листьев пшеницы разработана инновационная технология высокопроизводительного фенотипирования. Она основана на анализе цифровых изображений сгиба листа, полученных при помощи микроскопа. Разработаны алгоритм и компьютерная программа LHDetect (<http://wheatdb.org/lhdetect>), которая по цифровому изображению позволяет оценить число трихом и их распределение по длине. Время обработки одного образца при помощи инновационной технологии занимает около 1 мин. Это позволяет проводить массовое фенотипирование растений. Алгоритм оценки степени опушенности может обрабатывать цветные изображения листа. Лист располагается с одной стороны микрофотографии, а его граница проходит приблизительно посередине изображения вертикально. Изображения подаются на вход алгоритма в формате битовых карт. Работа алгоритма состоит из нескольких этапов:

- определение цветовых характеристик области фона и листа;
- определение границы листа;
- подсчет числа трихом на различной удаленности от границы листа;
- пошаговое смещение и формирование вектора чисел пересечений с трихомами на разных удалениях от листа;
- пересчет чисел пересечений с трихомами в вектор распределения трихом по длине.

Иным инновационным подходом к анализу роста и развития растений является платформа Phenotiki для изображения растений и визуального фенотипирования. С помощью визуального сканирования обрабатывается визуальный профиль растения и подается на дальнейшую обработку [2,8]. Это доступная и простая в использовании система для сбора и анализа цифровых изображений, ориентированных на оценку роста растений (рисунки 2).

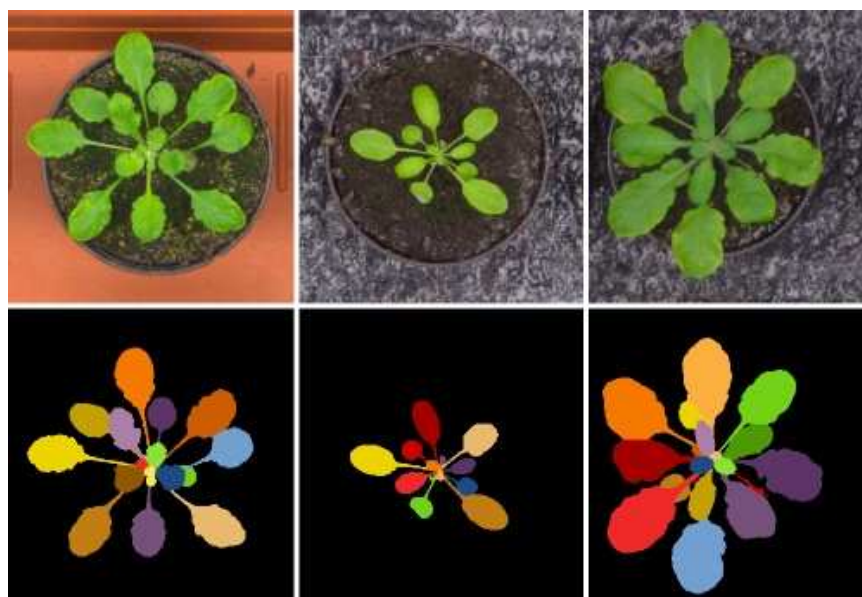


Рисунок 2 – Визуальный анализ роста растений. Учитывается количество листьев, степень разрастания розетки, окраска листьев, т. е. интенсивность фотосинтеза

Визуальное фенотипирование или современные методы высокопроизводительного фенотипирования позволяют получать более быстрое точное описание морфологических признаков, чем ранее используемые методы. Высокопроизводительное фенотипирование успешно применяется в процессе анализа опушения растений картофеля при подсчете количества трихом, набора поверхностных образований на листе, определяющие степень опушения. Плотность и степень опушения является селекционным признаком, и используется для предсказания устойчивости генотипов картофеля к насекомым. Постановка метода позволяющего быстро проводить количественную оценку листа картофеля на основе полученного изображения требует разработки протокола. В данном случае, представленная технология А. Дорошковым, М. Генаевым и Д. Афониным (2016) основывается на получении цифровых изображений сгибов листа с помощью оптического

микроскопа в проходящем свете и последующей автоматической обработке этих изображений на компьютере с помощью программы LHDetect2.

Использование компьютерных технологий позволяет сократить рутинные операции по получению, обработке и представлению данных, повысить точность и существенно повысить скорость данных, сделать анализ и представление данных более эффективными. Постановку любой задачи в рамках селекции и семеноводства, как и вообще любой научной задачи, можно представить как передачу информации между исследователем и исследуемым объектом, с одной стороны и сообществом исследователей (обществом) с другой стороны. Компьютер способен обрабатывать только данные, представленные в виде чисел. Поэтому первой задачей, которую должен решить экспериментатор, желающей применить в своей работе компьютерную технику, визуальный анализ и инновационные технологии, станет выбор способа представления и измеряемого параметра в численной форме. К счастью, большая часть таких способов уже реализована, и доступны периферийные устройства для перевода данных в числовую форму.

Перевод данных в компьютерную форму можно рассматривать как стартовый этап инновационных технологий по визуальному анализу. Требуется заранее определить, какой диапазон значений может принимать та или иная величина. Следовательно, необходимо решать, например, сколько оттенков требуется различать при изучении окрашивания клеток флуоресцентными красителями или какова требуемая точность определения концентрации интересующего нас вещества. Таким образом, любой перевод данных в компьютерную форму неизбежно сопряжен с потерей части информации, и на исследователя возложена ответственность решения, какой частью информации можно пожертвовать.

Примеры преобразования данных можно рассматривать как следующий этап. Представление физических параметров, таких как изменение оптической плотности, может быть осуществлено различными способами. Наименее точным и эффективным является способ визуальной регистрации. Исследователь сам читает компьютерную программу, показания прибора и вручную вводит их в ту или иную компьютерную программу. В некоторых случаях применение этого способа неизбежно, например, и проведении единичных измерений, когда затраты на разработку и внедрение компьютерной системы неоправданно.

Основным подходом к преобразованию изображений в цифровую форму является растеризация. Изображение при этом разбивают на участки (пиксели, точки), образующие прямоугольную матрицу (растр). Всё изображение, таким образом, представляет собой последовательность чисел, которые характеризуют цвет точек. Оцифровку изображений обычно осуществляют с помощью камеры или сканера.

При переводе изображения в цифровую форму происходит потеря части информации (точности). С одной стороны, теряются детали изображения при растеризации, с другой – даже при использовании совершенных алгоритмов цветопередачи бесконечное множество цветов будет описано конечным набором возможных значений. Таким образом, подбор параметров оцифровки изображения требует к себе ответственного отношения (рисунок 3).





Рисунок 3 – Анализ архитектуры растения кукурузы с помощью сканалайзера

При хранении изображений в цифровой форме часто прибегают к специальным способам сжатия. Сжатые данные занимают существенно меньшее пространство в памяти компьютера. Одни алгоритмы позволяют достигать хорошего сжатия изображений без потери информации за счет поиска и удаления повторов, другие дают возможность существенно уменьшать размер хранимого изображения за счет исключения из него несущественной информации.

Описанные способы преобразования данных позволяют представить некоторые данные в форме, доступной для обработки компьютером. Не следует, однако, забывать, что в основе преобразования любых данных лежит условность. Так, изображение текстовой страницы, переведенное в цифровую форму при помощи сканера, является не текстом, а изображением, и чтобы подойти к его обработке как к обработке текста, требуется предпринять дополнительные действия (произвести распознавание при помощи специальных программных средств).

На сегодняшний день сделаны интересные работы по биометрической оценке полиморфизма сортогрупп винограда на основе АСК-анализа [ 5,6 ].

Все это позволяет ставить и решать на практике ряд важных задач ампелографии:

- 1) оцифровка сканированных изображений листьев и создание их математических моделей;
- 2) формирование математических моделей конкретных листьев с применением теории информации;
- 3) формирование моделей обобщенных образов листьев различных сортов;
- 4) сравнение образа конкретного листа с обобщенным образом листа разных сортов и определение количественной степени сходства-различия между ними, т.е. идентификация сорта по листу;
- 5) количественное определение сходства-различия сортов, т.е. кластерно-конструктивный анализ обобщенных образов листьев различных сортов.

В результате оцифровки и АСК-анализа изображений получают иконография листа растений винограда без зашумления, т.е. получение информации об истинной форме листа данного сорта (чистый сигнал), без шума, искажающего эту истинную форму, обусловленного случайным воздействием окружающей среды.

Компьютерная биология. Ключевой элемент исследования в компьютерной биологии – возможность прижизненного анализа изображений с помощью специализированного программного обеспечения. Именно оригинальные разработки в этой области науки определяют прогресс или отставание в данном направлении. Практика исследования включает обязательный элемент – верификацию полученных данных. В компьютерной биологии одно и то же изображение должно быть проанализировано с помощью различного программного обеспечения и полученные результаты должны однозначно соответствовать друг другу. Основными объекта-

ми компьютерной биологии сегодня являются так называемые «плоские» объекты – те объекты, которые можно точно охарактеризовать с помощью двумерного изображения. К таким объектам относятся листья растений, цветы, лепестки.

Работа с цветом. Со времен К. Линнея, биологи не использовали цвет объекта, поскольку не было технических средств для работы с цветом. Электронные изображения живых биологических объектов могут быть охарактеризованы, в том числе, по цвету. Это открывает перспективы применения точных цветовых характеристик для описания объектов в целях оценки физиологического состояния растений, а также для задач биологической систематики.

Компьютерная биология, как наука, оперирующая электронными изображениями – образами (имидж) биологического объекта. Работа с образами биологических объектов – фундаментальное отличие компьютерной биологии от биоинформатики и математического моделирования биологических процессов. Принципиальным новшеством является возможность создания имиджей живых биологических объектов, что открывает перспективу неинвазивной работы с биологическим объектом. Это выводит биологию на качественно новый уровень исследований, за счет анализа и обработки изображений и распознавания живых объектов. Важнейшим компонентом компьютерной биологии становится анализ изображений. Разработка и применение различного софта для анализа одного и того же изображения – наиболее надежный путь верификации полученных данных.

Одним из перспективных подходов в решении этой задачи является методология создания и распознавания виртуальных образов, которая плодотворно разрабатывается во многих российских математических школах и находит реализацию в селекционно-генетических работах.

*Авторы выражают признательность профессору Луценко Е.В. за вдохновение и полезные комментарии во время написания статьи.*

### СПИСОК ЛИТЕРАТУРЫ

1. Беккер Х. Селекция растений / Х. Беккер. – М. : Товарищество научных изданий КМК, 2015. – 425 с.
2. Биологический контроль окружающей среды: генетический мониторинг : учеб. пособие / С. А. Гераськин, Е. И. Сарапульцева, Л. В. Цаценко [и др.]; под ред. С. А. Гераськина и Е. И. Сарапульцевой. – М. : Академия, 2010. – 208 с.
3. Дорошков А. В., Генаев М. А., Афонников Д. А. Протокол анализа количественных характеристик опушения листа картофеля //Zhurnal Genetiki i Seleksii=Vavilov Journal of Genetics and Breeding. – 2016. – Т. 20. – №. 6. – С. 863-868.
4. Малюга Н.Г., Цаценко Л.В. Перспективы растениеводства в будущем веке// Аграрная наука. 1998.- № 4.- С. 14-19.
5. Луценко Е.В. Биометрическая оценка полиморфизма сортогрупп винограда Пино и Рислинг по морфологическим признакам листьев среднего яруса кроны / Л.П. Трошин, Е.В. Луценко, П.П. Подваленко, А.С. Звягин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2009. – №08(052). С. 1 – 14. – Шифр Информрегистра: 0420900012\0097, IDA [article ID]: 0520908001. – Режим доступа: <http://ej.kubagro.ru/2009/08/pdf/01.pdf>, 0,875 у.п.л.
6. Луценко Е.В. Решение задач ампелографии с применением АСК-анализа изображений листьев по их внешним контурам (обобщение, абстрагирование, классификация и идентификация) / Е.В. Луценко, Д.К. Бандык, Л.П. Трошин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №08(112). С. 862 – 910. – IDA [article ID]: 1121508064. – Режим доступа: <http://ej.kubagro.ru/2015/08/pdf/64.pdf>, 3,062 у.п.л.
7. Цаценко Л.В. Метод скетчей в археогенетике и селекции сельскохозяйственных растений / Л.В. Цаценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2015. – №02(106). С. 1083 – 1097. – IDA [article ID]: 1061502071. – Режим доступа: <http://ej.kubagro.ru/2015/02/pdf/71.pdf>, 0,938 у.п.л.
8. Цаценко Л.В. Образовательные технологии, развивающие визуальное мышление, в преподавании биологических дисциплин / Л.В. Цаценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №04(118). С. 937 – 948. – IDA [article ID]: 1181604056. – Режим доступа: <http://ej.kubagro.ru/2016/04/pdf/56.pdf>, 0,75 у.п.л.

### References

1. Bekker H. Selekcija rastenij / H. Bekker. – M. : Tovarishhestvo nauchnyh izdaniy KMK, 2015. – 425 s.

2. Biologicheskij kontrol' okruzhajushhej sredy: geneticheskij monitoring : ucheb. posobie / S. A. Geras'kin, E. I. Sarapul'ceva, L. V. Cacenko [i dr.]; pod red. S. A. Geras'kina i E. I. Sarapul'cevoj. – M. : Akademija, 2010. – 208 s.

3. Doroshkov A. V., Genaev M. A., Afonnikov D. A. Protokol analiza ko-lichestvennyh harakteristik opushenija lista kartofelja //Zhurnal Genetiki i Seleksii= Vavilov Journal of Genetics and Breeding. – 2016. – T. 20. – №. 6. – S. 863-868.

4. Maljuga N.G., Tsatsenko L.V. Perspektivy rastenievodstva v budushhem veke// Agrarnaja nauka. 1998.- № 4.- S. 14-19.

5. Lucenko E.V. Biometricheskaja ocenka polimorfizma sortograpp vino-grada Pino i Risling po morfologicheskim priznakam list'ev srednego jarusa krony / L.P. Troshin, E.V. Lucenko, P.P. Podvalenko, A.S. Zvjagin // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2009. – №08(052). S. 1 – 14. – Shifr Informregistra: 0420900012\0097, IDA [article ID]: 0520908001. – Rezhim dostupa: <http://ej.kubagro.ru/2009/08/pdf/01.pdf>, 0,875 u.p.l.

6. Lucenko E.V. Reshenie zadach ampelografii s primeneniem ASK-analiza izobrazhenij list'ev po ih vneshnim konturam (obobshhenie, abstragirova-nie, klassifikacija i identifikacija) / E.V. Lucenko, D.K. Bandyk, L.P. Troshin // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2015. – №08(112). S. 862 – 910. – IDA [article ID]: 1121508064. – Rezhim dostupa: <http://ej.kubagro.ru/2015/08/pdf/64.pdf>, 3,062 u.p.l.

7. Tsatsenko L.V. Metod sketchej v arheogenetike i seleksii sel'skohozjaj-stvennyh rastenij / L.V. Tsatsenko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2015. – №02(106). S. 1083 – 1097. – IDA [article ID]: 1061502071. – Rezhim dostupa: <http://ej.kubagro.ru/2015/02/pdf/71.pdf>, 0,938 u.p.l.

8. Tsatsenko L.V. Obrazovatel'nye tehnologii, razvivajushhie vizual'noe myshlenie, v prepodavanii biologicheskikh disciplin / L.V. Tsatsenko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2016. – №04(118). S. 937 – 948. – IDA [article ID]: 1181604056. – Rezhim dostupa: <http://ej.kubagro.ru/2016/04/pdf/56.pdf>, 0,75 u.p.l.