

ISSN 2304-9081

Учредители:  
Уральское отделение РАН  
Оренбургский научный центр УрО РАН

**Бюллетень**  
**Оренбургского научного центра**  
**УрО РАН**  
(электронный журнал)



**2014 \* № 3**

On-line версия журнала на сайте  
<http://www.elmag.uran.ru>

© Коллектив авторов, 2014

УДК 579.2:519.2

А.Л. Турская<sup>1,2</sup>, Ю.С. Букин<sup>3</sup>, А.А. Ульданова<sup>1</sup>, Ю.А. Маркова<sup>1</sup>,  
В.В. Верхотуров<sup>2</sup>

## **ИСПОЛЬЗОВАНИЕ СРЕДЫ ПРОГРАММИРОВАНИЯ R ДЛЯ ОЦЕНКИ УСЛОВИЙ ФОРМИРОВАНИЯ МИКРОБНЫХ БИОПЛЕНОК**

<sup>1</sup> Сибирский институт физиологии и биохимии растений СО РАН, Иркутск, Россия

<sup>2</sup> Иркутский государственный технический университет, Иркутск, Россия

<sup>3</sup> Лимнологический институт СО РАН, Иркутск, Россия

*Цель.* Используя средства языка программирования R, статистически оценить зависимость формирования биопленки бактериальным фитопатогеном *Pectobacterium carotovorum* от источника углерода в среде культивирования.

*Материалы и методы.* Изучено формирование биопленок *Pectobacterium carotovorum* (штамм В-1247) в лунках полистироловых планшетов в среде с различными источниками углерода. Произведен статистический анализ полученных результатов методом однофакторного дисперсионного анализа с помощью непараметрического критерия Краскела-Уоллиса.

*Результаты.* С помощью однофакторного дисперсионного анализа показано достоверное влияние вносимого в среду источника углерода на процесс формирования биопленки *Pectobacterium carotovorum*, наиболее благоприятным субстратом для этого процесса оказался многоатомный спирт инозитол.

*Заключение.* Язык программирования R является удобным средством анализа данных по оценке условий формирования микробных биопленок.

*Ключевые слова:* микробные биопленки, *Pectobacterium carotovorum*, среда программирования R, однофакторный дисперсионный анализ, статистический критерий Краскела-Уоллиса.

---

---

A.L. Turskaya<sup>1,2</sup>, Yu.S. Bukin<sup>3</sup>, A.A. Uldanova<sup>2</sup>, Yu. A. Markova<sup>2</sup>,  
V.V. Verkhoturov<sup>2</sup>

## **USE R PROJECT TO ASSESS CONDITIONS OF FORMATION MICROBIAL BIOFILMS**

<sup>1</sup> Siberian Institute of Plant Physiology and Biochemistry SB RAS, Irkutsk, Russia.

<sup>2</sup> Irkutsk State Technical University, Irkutsk, Russia.

<sup>3</sup> Limnological Institute SB RAS, Irkutsk, Russia.

*Objective.* Using means of programming language R tools statistically to estimate dependence of biofilm formation with bacterial phytopathogen *Pectobacterium carotovorum* on a carbon source in the culture medium.

*Materials and methods.* Biofilm formation with *Pectobacterium carotovorum* (strain B-1247) in wells of polystyrene tablets in the culture medium with different carbon sources has been studied. The statistical analysis of the obtained results using a method of the One-Way ANOVA by means the nonparametric Kruskal–Wallis test has been made.

*Results.* By means of a method of the One-Way ANOVA reliable influence of a carbon source brought in the culture medium on process of biofilm formation with *Pectobacterium carotovorum* has been shown, polyatomic alcohol inositol appeared an optimum substrate for this process.

*Conclusions.* The programming language R is convenient means for the analysis of data connected with conditions of formation microbial biofilms.

*Key words:* microbial biofilms, *Pectobacterium carotovorum*, R project, One–Way ANOVA, Kruskal–Wallis test.

## **Введение**

По современным представлениям большинство микроорганизмов в естественной среде существует в виде биопленок. Формирование биопленочных сообществ оказалось одной из основных стратегий выживания бактерий в занимаемых ими экологических нишах [1]. В составе биопленки микроорганизмы способны обмениваться сигналами и проявлять координированную активность, свойственную многоклеточным организмам. Таким образом они лучше защищены от повреждающих факторов внешней среды и действия антибактериальных веществ в организме хозяина [2, 3].

Многочисленными исследованиями показано, что бактерии могут образовывать биопленки на любых биотических и абиотических поверхностях, что создает большие проблемы в медицинской практике [4, 5] и в различных областях хозяйственной деятельности [6]. В частности, поражение растений бактериями, способными образовывать биопленки, наносит значительный урон сельскому хозяйству. Один из важных представителей фитопатогенов, *Pectobacterium carotovorum*, обладает набором пектолитических ферментов, образуя мягкие гнили у разных видов растений [7]. Понимание процесса образования бактериальных биопленок в различных условиях окружающей среды даст объективное представление о механизмах взаимодействия бактерий с растительными организмами, благодаря чему станет возможной разработка методов защиты растений для повышения их устойчивости. В связи с этим, данная работа посвящена исследованию динамики образования биопленок *P. carotovorum* в среде с различными источниками углерода в лабораторных условиях.

Для обработки, анализа и визуализации данных на сегодняшний день существует множество программ, таких как Statistica, Minitab, MatLab, GenStat и пр. Однако в последнее время особое внимание исследователей привлекли возможности среды программирования R [8, 9].

Среда R позволяет проводить как первичный анализ данных (таблицы, графики), так и математическую статистику со всеми приложениями. R является свободно распространяемым программным продуктом, который доста-

точно просто устанавливается под Windows, MacOS X, Linux, базовая комплектация R занимает немного места на жестком диске и включает в себя все функции, необходимые для статистического анализа; для более серьезной работы можно дополнительно устанавливать вспомогательные пакеты, представляющие собой библиотеки для работы специфических функций [10, 11].

Цель работы – с помощью средств языка программирования R статистически оценить зависимость формирования биопленки бактериальным фитопатогеном *Pectobacterium carotovorum* от источника углерода в среде культивирования.

### **Материалы и методы**

В качестве объекта исследования использовался бактериальный штамм *P. carotovorum* В-1247, полученный из Всероссийской коллекции микроорганизмов ИБФМ им. Г.В. Скрыбина (г. Пущино). Для изучения процесса образования биопленки *P. carotovorum* в качестве среды культивирования использовали забуференный физиологический раствор с добавлением 5 г/л следующих источников углерода: глюкоза, фруктоза, рибоза, мальтоза, лактоза, L-триптофан, инозитол, дульцит, маннитол, D-ксилоза. Образование биопленок в 96-луночных полистироловых планшетах (США, Германия) выявляли с помощью 1%-ного раствора кристаллического фиолетового при длине волны 495 нм планшетного ридера Infinite 200 (Tecan, Швейцария) [12]. Интенсивность окрашивания содержимого лунок соответствовала степени пленкообразования. Количественным выражением степени служили полученные с помощью спектрофотометра значения оптической плотности, которые определяли после разведения и внесения в лунки, на 3, 5 и 7-е сутки инкубации.

Для статистической обработки экспериментальных данных использовалась среда программирования R (версия 3.1.0) [13]; оценку выборок на наличие выбросов проводили с помощью теста Диксона; соответствие выборок закону нормального распределения проверяли с помощью критерия Шапиро-Уилка; значимость результатов определяли методом однофакторного дисперсионного анализа с помощью непараметрического критерия Краскела-Уоллиса; для визуализации распределения величин в выборках применялся графический метод боксплотов [14].

### **Результаты и обсуждение**

Для удобства статистической оценки данных были приняты условные

обозначения субстратов среды и вариантов эксперимента: Забуференный физиологический раствор – *PBS*, глюкоза – *glu*, фруктоза – *fru*, рибоза – *rib*, мальтоза – *malt*, лактоза – *lact*, L-триптофан – *Ltrp*, инозитол – *ins*, дульцит – *dul*, маннитол – *man*, D-ксилоза – *xyl*. Прирост биомассы (концентрация бактерий в среде культивирования) обозначали *OD*, а интенсивности образования биопленок, то есть измерения раствора после окрашивания – *Bio*, время (сутки) инкубации обозначали: 0 – начальный титр бактерий в момент постановки эксперимента, третьи сутки инкубации – 3, пятые – 5 и седьмые, соответственно, цифрой 7. Контрольные варианты (измерения без внесения бактерий) обозначали как *contr*, а опытные измерения – *exp*. Таким образом, для последующего статистического анализа были использованы следующие варианты выборок: *OD0contr*, *OD0exp*, *OD3contr*, *OD3exp*, *OD5contr*, *OD5exp*, *OD7contr*, *OD7exp*, *Bio3contr*, *Bio3exp*, *Bio5contr*, *Bio5exp*, *Bio7contr*, *Bio7exp*.

Для анализа данных в среде R каждая выборка эксперимента была оформлена в виде таблицы в программе Microsoft Excel и сохранена в формате *.csv* с разделителями «запятая» между столбцами.

Поскольку помимо случайных факторов при проведении микробиологических и биохимических исследований часто действуют факторы неслучайного характера (погрешности в ходе выполнения экспериментов, измерений приборов), такие значения необходимо исключать из анализа общей выборки. Поэтому начальным этапом статистической обработки данных служит проверка исследуемых выборок на наличие отбрасываемых величин. В языке программирования R эта задача реализуется в виде функции: *>dixon.test(x)*, где *x* – вектор, содержащий массив данных. Для работы этой функции необходимо подключить дополнительный пакет «*outlier*». Таким образом, для расчета данных был создан скрипт вычисления в среде R (прил. 1), где первой строкой происходит считывание массива данных в переменную *>mydata* типа фрейм. Фреймы в языке программирования R представляют собой сложные структуры, содержащие в себе несколько векторов (массивов данных) одинаковой размерности. При выполнении теста Диксона на проверку наличия отбрасываемых величин в выборке формулируется гипотеза  $H_0$  о том, что выбросов в выборке нет. Если *p-value* больше вероятностного порога 0.05, принимается гипотеза  $H_0$ , в противном случае – отвергается гипотеза  $H_0$  и принимается альтернативная гипотеза  $H_1$  о наличии отбрасываемого значе-

ния в выборке.

### Приложение 1.

Содержание скрипта на наличие выбросов в выборке:

```
>mydata<-read.csv("C:\\R\\biofilm.csv")  
>dixon.test(mydata$data)  
>x<-mydata$data
```

Следующим этапом статистического анализа данных была оценка соответствия выборок закону нормального распределения, для этого применяли критерий Шапиро-Уилка. В языке программирования R критерий реализован в виде функции:  $>shapiro.test(x)$ , где  $x$  – вектор, содержащий значения проверяемой выборки. Данный критерий проверяет гипотезу  $H_0$  о том, что исследуемая выборка распределена по нормальному закону. При оценке выборок оказалось, что некоторые из них не соответствовали закону нормального распределения, следовательно, для дальнейшего анализа необходимо использовать непараметрические критерии оценки.

Дальнейшим шагом исследования стала необходимость определить, влияет ли вид питательного субстрата в среде культивирования на показатели оптической плотности. Для этого брались контрольные измерения (растворы без добавления бактерий) и производился расчет с помощью критерия Краскела-Уоллиса. Критерий Краскела-Уоллиса является непараметрической версией классического однофакторного дисперсионного анализа и предназначен для проверки равенства средних значений нескольких выборок [14]. Данный критерий проверяет гипотезу  $H_0$  – выбранные градации фактора не влияют на средние значения в сравниваемых выборках. В языке программирования R критерий реализован в виде функции:  $>kruskal.test(x\sim fac)$ , где  $x$  – вектор, содержащий массив данных,  $fac$  – переменная типа фактор, указывающая принадлежность измерения определенной градации фактора. В нашем случае, критерий проверяет гипотезу  $H_0$  о том, что зависимости вида субстрата в среде культивирования на показатели оптической плотности прибора нет.

В таблице 1 представлены значения вероятностей (p-value) принятия гипотезы  $H_0$  об отсутствии достоверных различий между средними значениями оптической плотности сред культивирования с добавлением различных источников углерода.

Поскольку значения p-value всех вариантов оказались меньше вероят-

ностного порога 0.05, то гипотеза  $H_0$  отвергается, и принимается альтернативная гипотеза  $H_1$  о зависимости субстрата питательной среды на показателе оптической плотности.

Таблица 1. Значения вероятностей принятия гипотезы  $H_0$  об отсутствии достоверных различий между средними значениями показателей оптической плотности исследуемых выборок

	Концентрация бактерий в среде культивирования, нм				Интенсивность образования биопленок, нм		
	0 сут	3 сут	5 сут	7 сут	3 сут	5 сут	7 сут
p-value	0.0033	$1.1 \cdot 10^{-7}$	$1.7 \cdot 10^{-5}$	$2.5 \cdot 10^{-5}$	0.0015	0.0006	0.0334

Для того чтобы оценить во сколько раз опытные показатели превышают контрольные, рассчитаны средние значения контрольных показателей по каждому субстрату. Результаты вычислений представлены в таблице 2, для чего был создан скрипт в среде R (прил. 2).

Таблица 2. Средние значения контрольных показателей оптической плотности по каждому субстрату питательной среды

Субстрат питательной среды	Оптическая плотность бактериальной суспензии, нм				Интенсивность образования биопленок, нм		
	0 сут	3 сут	5 сут	7 сут	3 сут	5 сут	7 сут
PBS	0.0539	0.0464	0.0498	0.0467	0.0743	0.0797	0.0800
glu	0.0517	0.0489	0.0530	0.0496	0.0744	0.0764	0.0755
fru	0.0594	0.0573	0.0616	0.0580	0.0771	0.0864	0.0816
rib	0.0695	0.0616	0.0653	0.0627	0.0854	0.0866	0.0835
malt	0.0576	0.0523	0.0558	0.0551	0.0772	0.0782	0.0796
lact	0.0509	0.0500	0.0534	0.0504	0.0789	0.0804	0.0786
Ltrp	0.0647	0.0700	0.0727	0.0720	0.0856	0.0842	0.0827
ins	0.0541	0.0536	0.0524	0.0496	0.0796	0.0812	0.0744
dul	0.0475	0.0459	0.0490	0.0469	0.0779	0.0779	0.0737
xyl	0.0590	0.0578	0.0629	0.0603	0.0802	0.0809	0.0799
man	0.0549	0.0504	0.0561	0.0524	0.0772	0.0837	0.0774

## Приложение 2

Содержание скрипта: дисперсионный анализ выборок

```
>data_k<-read.csv(file="C:\\R\\Bio3contr.csv")
>kruskal.test(data_k$ms~data_k$sbt)
>boxplot(data_k$ms~data_k$sbt)
>x<-c(1:11)
>sred<-c("PBS","glu","fru","rib","malt","lact","Ltrp","ins","dul","xyl","man")
>for(i in 1:11) x[i]<-mean(data_k[data_k$sbt==sred[i],2])
>s_k<-x
>data_s<-data.frame(sred,s_k)
>data_3<-read.csv(file="C:\\R\\Bio3exp.csv")
>for(i in 1:11)
{
data_3[data_3$sbt==data_s$sred[i],2]<-
```

```
(data_3[data_3$sbt==data_s$red[i],2])/data_s$s_k[i]
}
>boxplot(data_3$ms~data_3$sbt,range=6)
>kruskal.test(data_3$ms~data_3$sbt)
```

Для визуализации распределения величин в выборках использовался метод боксплотов или «ящичков с усами» – график, использующийся в описательной статистике, компактно изображающий одномерное распределение вероятностей (рис. 1).

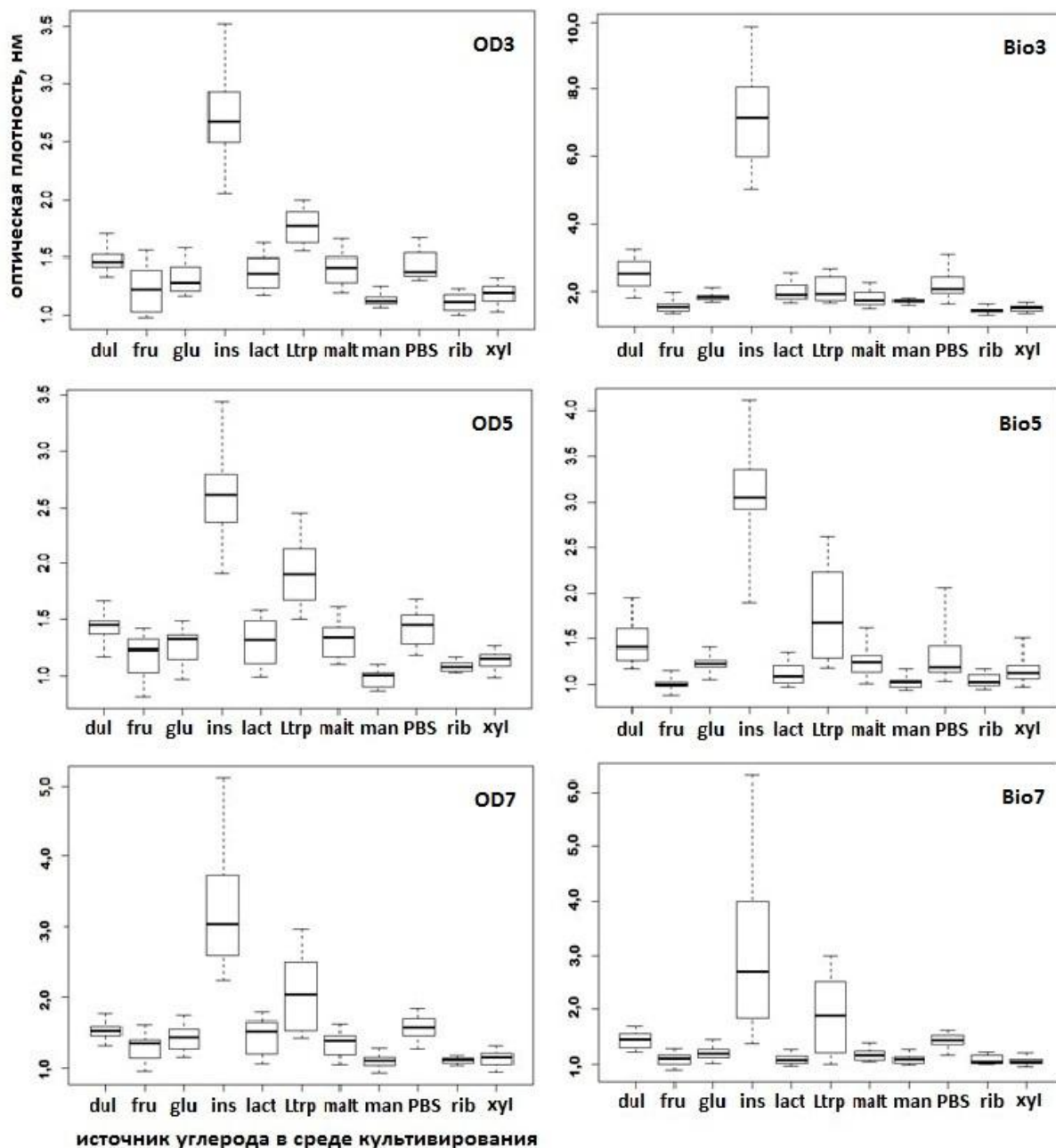


Рис. 1 Образование биопленок бактериальным фитопатогеном *P. carotovorum* в среде с различными источниками углерода. OD – оптическая плотность; Bio – оптическая плотность биопленки; цифрами обозначено время инкубации бактерий – 3, 5 и 7-е сутки.



Такой вид диаграммы в удобной форме показывает медиану, нижний и верхний квартили, минимальное и максимальное значения выборки, а также отбрасываемые величины [15].

В языке программирования R графическая функция боксплот имеет следующий вид:  $> \text{boxplot}(x \sim \text{fac})$ , где  $x$  – вектор, содержащий массив данных,  $\text{fac}$  – переменная типа фактор, указывающая принадлежность измерения определенной градации фактора.

На рисунке 1 можно увидеть, что боксплоты показателей оптической плотности, полученных при разном времени культивирования бактерий, демонстрируют сходную тенденцию. Наиболее благоприятной средой для формирования биопленок *P. carotovorum* оказалась среда, содержащая многоатомный спирт инозитол, где значения оптической плотности по сравнению с остальными вариантами эксперимента превышали аналогичные показатели в несколько раз.

Интенсивное образование биопленок *P. carotovorum*, которые начали формироваться на пятые сутки инкубации, показала среда с содержанием аминокислоты L-триптофан. Триптофан является предшественником образования бактериями индола, одного из важнейших вторичных метаболитов ароматической природы, который, как было показано ранее, играет немаловажную роль в процессе образования биопленок [16]. Очевидно, многоатомный спирт инозитол также опосредованно участвует в формировании прокариотами многоклеточных сообществ, однако механизм этого участия на сегодняшний день абсолютно не изучен.

### **Заключение**

Таким образом, на примере серии экспериментов по изучению условий формирования микробных биопленок показано, что язык программирования R является удобным средством анализа данных. В системе R имеются широкие возможности для статистической обработки данных, в том числе и для работы с графикой. К числу ее основных достоинств нужно отнести возможность удобного обмена данными с электронными таблицами, а отличительным преимуществом является то, что однократно написав необходимые программы-скрипты, а затем, изменяя только имя входного файла с данными, можно использовать их в качестве шаблонов для последующих расчетов и анализа экспериментальных данных.

## ЛИТЕРАТУРА

1. Ильина Т.С., Романова Ю.М., Гинцбург А.Л. Биопленки как способ существования бактерий в окружающей среде и организме хозяина: феномен, генетический контроль и системы регуляции их развития. *Генетика*. 2004. 40 (11): 1445-1456.
2. Романова Ю.М., Смирнова Т.А., Андреев А.Л. и др. Образование биопленок – пример «социального» поведения бактерий. *Микробиология*. 2006. 75 (4): 556-561.
3. Николаев Ю.А., Плакунов В.К. Биопленка – «город микробов» или аналог многоклеточного организма? *Микробиология*. 2007. 76 (2): 149-163.
4. Сидоренко С.В. Роль бактериальных биопленок в патологии человека. *Инфекции в хирургии*. 2004. 2 (3): 16-20.
5. Коробов В.П., Лемкина Л.М., Монахов В.И. Анализ чувствительности процессов формирования биопленок *Staphylococcus epidermidis* 33 к некоторым факторам внешней среды. *Вестник Пермского университета. Биология*. 2010. 1 (1): 59-63.
6. Борецкая М.А., Сулова О.С., Горчев В.Ф. и др. Изучение структуры экзополимерного комплекса биопленки коррозионно-активных бактерий. *Биотехнология*. 2012. 5: 78-84.
7. Pollumaa L., Alamae T., Mae A. Quorum Sensing and Expression of Virulence in *Pectobacteria*. *Sensors*. 2012. 12: 3327-3349.
8. Waayen R.H. *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. UK: Cambridge University Press, 2008. 369 p.
9. Crawley M.J. *The R Book*. England: John Wiley and Sons Ltd, 2007. 951 p.
10. Зарядов И.С. Введение в статистический пакет R: типы переменных, структуры данных, чтение и запись информации, графика. М.: Издательство Российского университета дружбы народов, 2010. 207 с.
11. Zoonekynd V. *Statistics with R*. [http://zoonek2.free.fr/UNIX/48\\_R/all.html](http://zoonek2.free.fr/UNIX/48_R/all.html) (дата обращения 12.07.14).
12. Шагинян И.А., Алексеева Г.В., Чернуха М.Ю. и др. Формирование биопленок клиническими штаммами бактерий комплекса *Burkholderia cepacia* в зависимости от их фенотипических и генотипических характеристик. *Журнал микробиологии, эпидемиологии и иммунобиологии*. 2007. 1: 3-9.
13. R Foundation for Statistical Computing. Vienna, Austria. <http://www.R-project.org> (дата обращения 12.07.14).
14. Ивантер Э.В., Коросов А.В. *Элементарная биометрия: учеб. пособие*. Петрозаводск: Издательство ПетрГУ, 2010. 104 с.
15. Орлов А.И. *Прикладная статистика*. М.: Издательство «Экзамен», 2004. 656 с.
16. Pandey R., Swamy K.W., Khetmalas M.B. Indole: A novel signaling molecule and its applications. *Indian Journal of Biotechnology*. 2013. 12: 297-310.

Поступила 30.07.2014

(Контактная информация: **Турская Анна Леонидовна** – кандидат биологических наук, научный сотрудник лаборатории фитоиммунологии Сибирского института физиологии и биохимии растений СО РАН (г. Иркутск); e-mail: [turskaya-anna@mail.ru](mailto:turskaya-anna@mail.ru))