

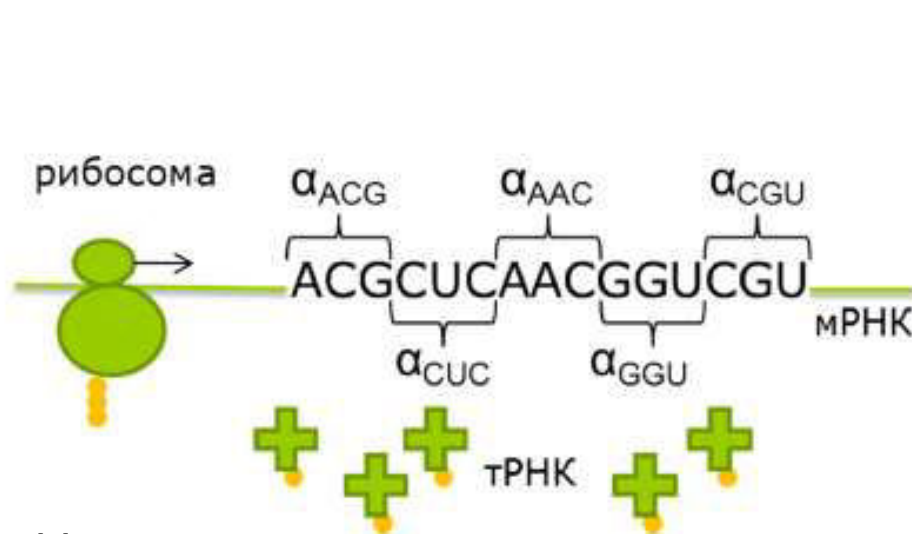
# Биоинформатический анализ характеристик элонгации трансляции у бактерий рода *Ralstonia*

Студентка ФЕН НГУ Коренская Александра Евгеньевна  
Научный руководитель к.б.н., н.с. Клименко Александра Игоревна  
Рецензент к.б.н., н.с. Алемасов Николай Александрович

Работа выполнена в лаборатории молекулярно-генетических систем ИЦиГ СО РАН

Новосибирск, 2021

# Оценка эффективности элонгации трансляции

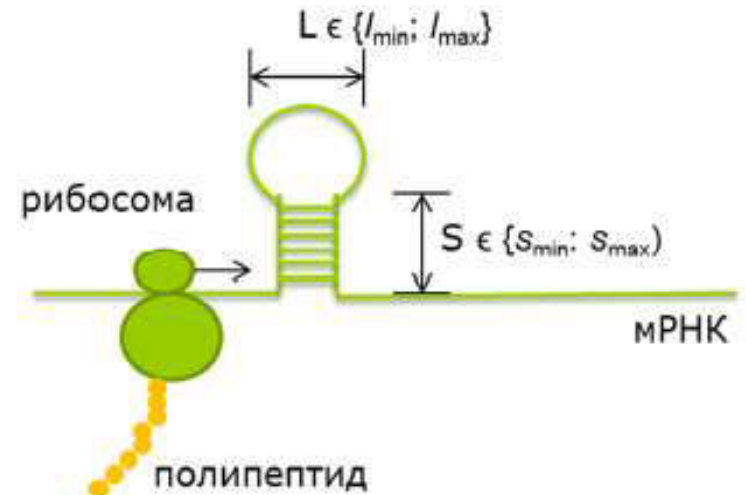


Учет кодонного состава

$\alpha$  – частота использования кодона

$L$  – длина петли шпильки

$S$  – длина стебля шпильки



Учет количества и энергии вторичных структур

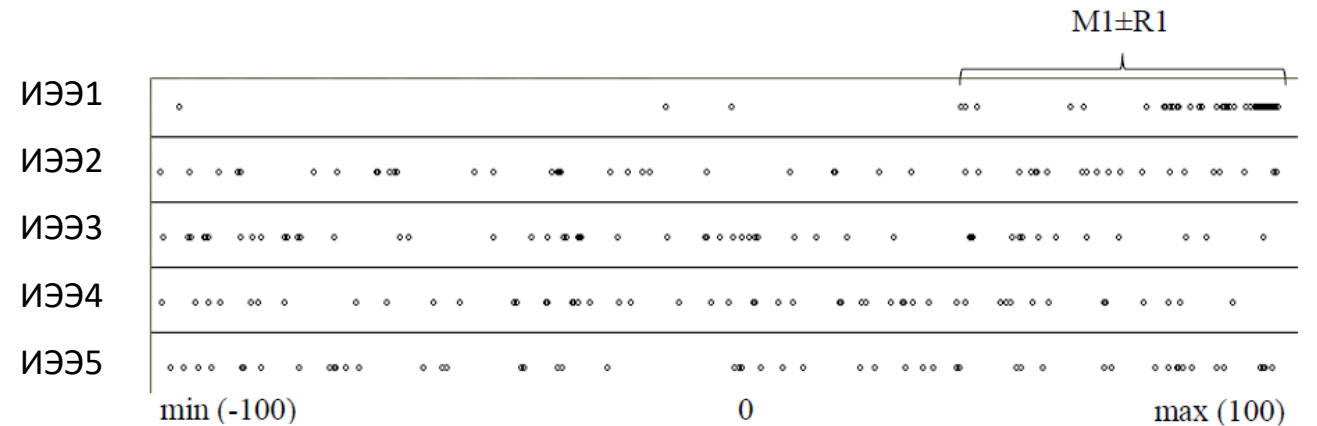
Взято из кандидатской диссертации Соколова В.С. (Mohammad et al., 2019; Xie & Chen, 2017)

# Программный комплекс для оценки эффективности элонгации трансляции (EloE)

| Тип эффективности элонгации трансляции | Кодонный состав | Количество вторичных структур | Энергия вторичных структур |
|--|-----------------|-------------------------------|----------------------------|
| ИЭЭ1                                   | +               | -                             | -                          |
| ИЭЭ2                                   | -               | +                             | -                          |
| ИЭЭ3                                   | -               | -                             | +                          |
| ИЭЭ4                                   | +               | +                             | -                          |
| ИЭЭ5                                   | +               | -                             | +                          |

Классификация типов индексов эффективности элонгации (ИЭЭ) трансляции в соответствии с учитываемыми факторами

M – среднее значение ранга ИЭЭ генов рибосомных белков (ГРБ)  
 R – стандартное отклонение ранга ИЭЭ ГРБ



Распределение генов рибосомных белков по индексам

Взято из кандидатской диссертации Соколова В.С. (Лихошвай & Матушкин, 2000; Соколов et al., 2014)

# Классификация

Филум: Proteobacteria

Класс: Betaproteobacteria

Порядок: Burkholderiales

Семейство: Burkholderiaceae

Род: *Ralstonia*

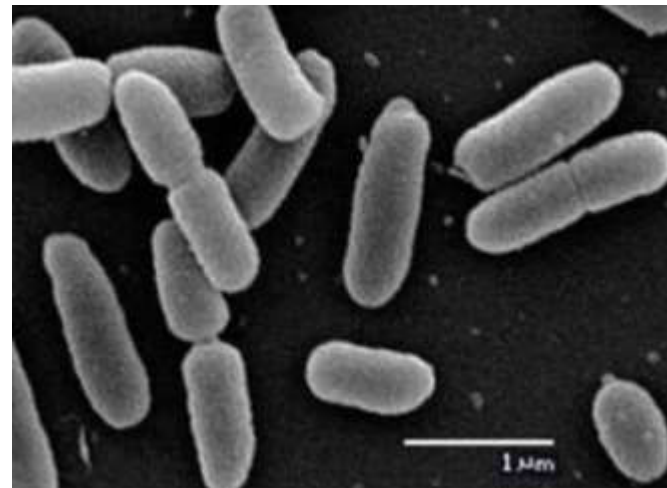


(Photo courtesy of T. Momol and S.M. Olson, University of Florida)

## *Ralstonia*

Почвенные, условные патогены  
человека

- *R. mannitolilytica*
- *R. pickettii*
- *R. insidiosa*



(Guarisch-Sousa et al., 2016)

Фитопатогены

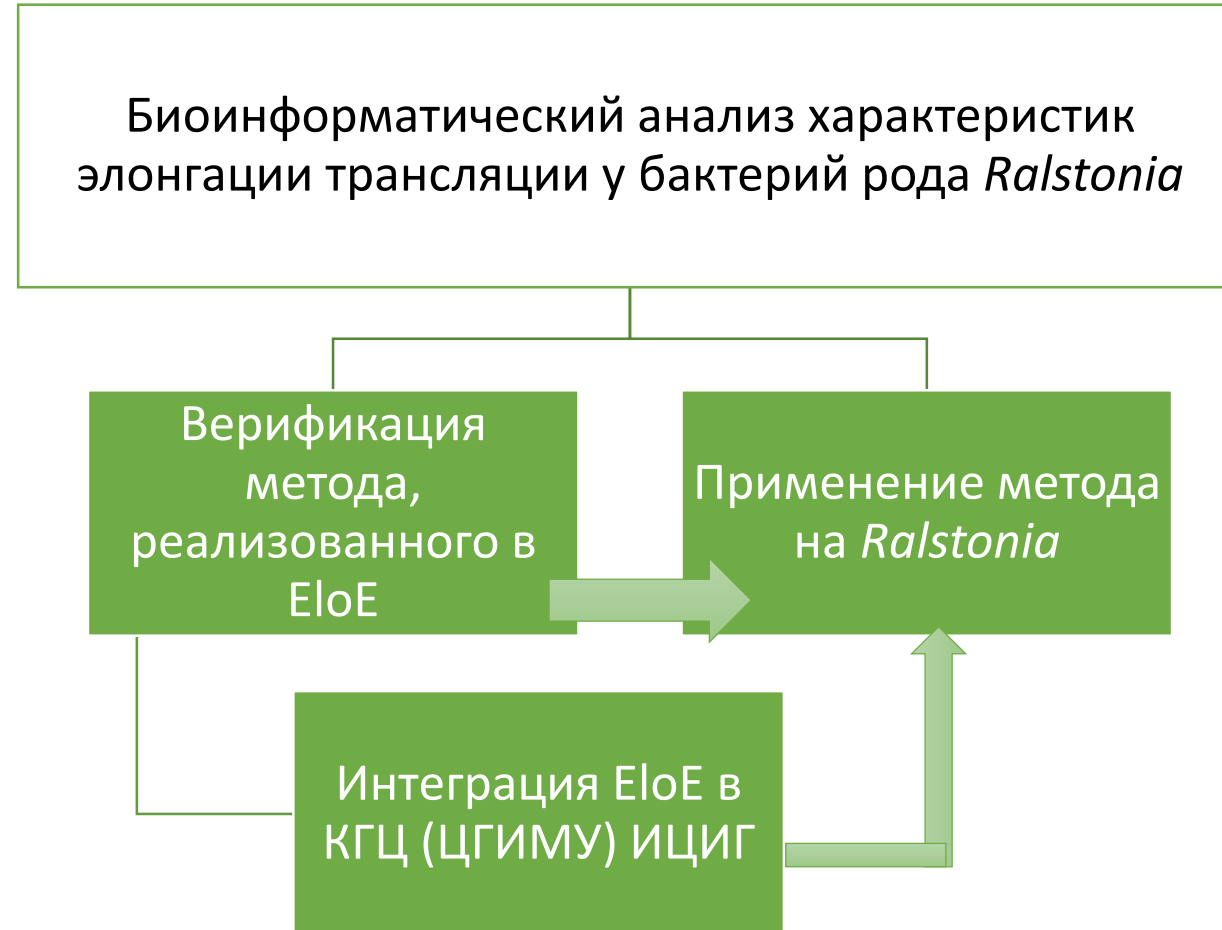
- *R. solanacearum*,
- *R. pseudosolanacearum*
- *R. syzygii*

Цель: биоинформатический анализ характеристик элонгации трансляции у бактерий рода *Ralstonia*

Задачи:

1. Анализ корреляции между значениями индексов эффективности элонгации трансляции и экспрессией генов на уровне белка у различных прокариот.
2. Интеграция программного комплекса для автоматической оценки эффективности элонгации трансляции генов различных организмов (EloE) в качестве модуля в систему обработки геномных данных КГЦ (ЦГИМУ) ИЦиГ СО РАН.
3. Проведение с помощью программного комплекса EloE биоинформатического анализа характеристик элонгации трансляции у бактерий видов *R. solanacearum*, *R. pseudosolanacearum*, *R. syzygii*, *R. mannitolilytica*, *R. pickettii* и *R. insidiosa* и сопоставление полученных результатов с филогенетическими взаимоотношениями в рамках данной группы прокариот.
4. Сравнительный анализ потенциально высоко экспрессирующихся групп генов у различных видов бактерий рода *Ralstonia* на основе биоинформатического анализа характеристик элонгации трансляции.

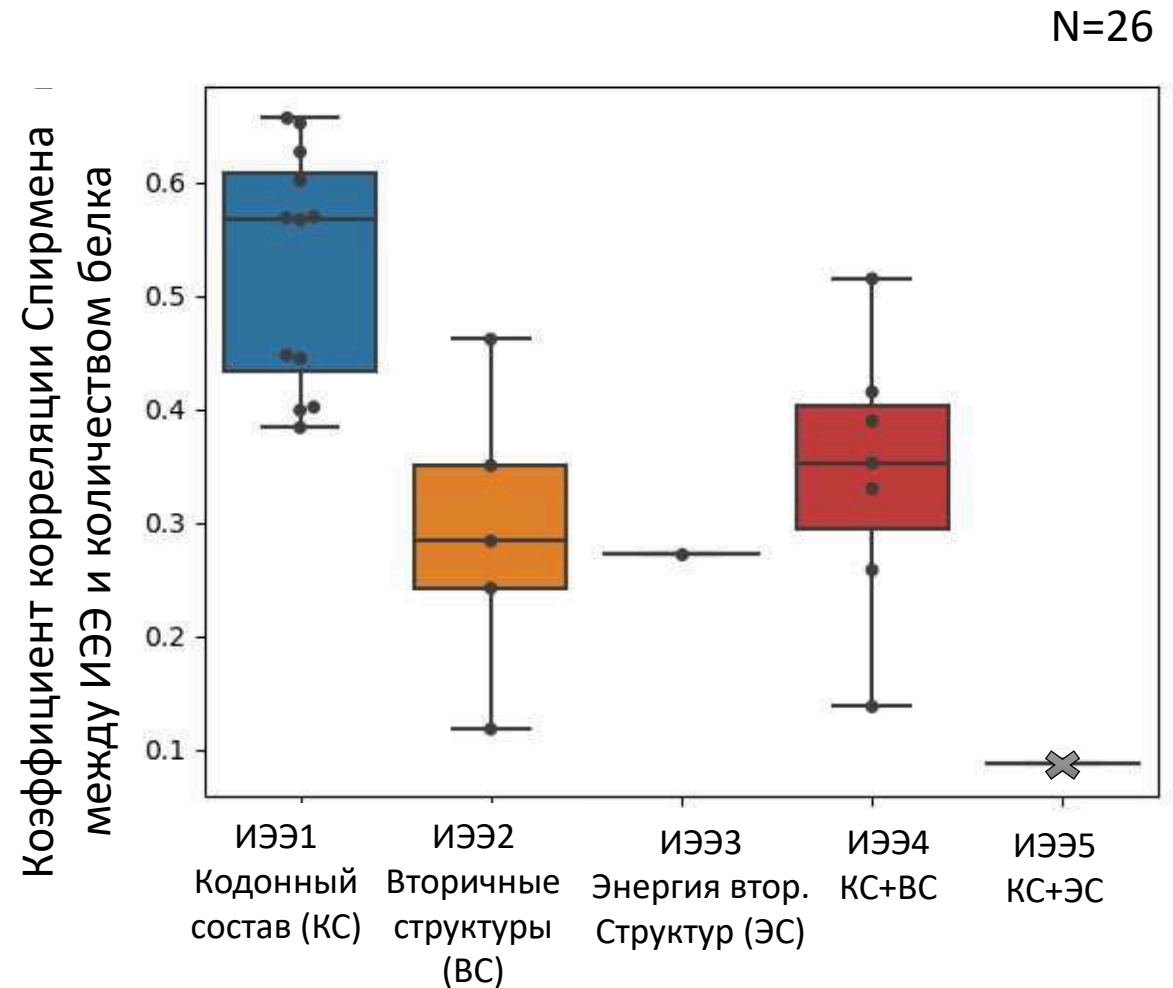
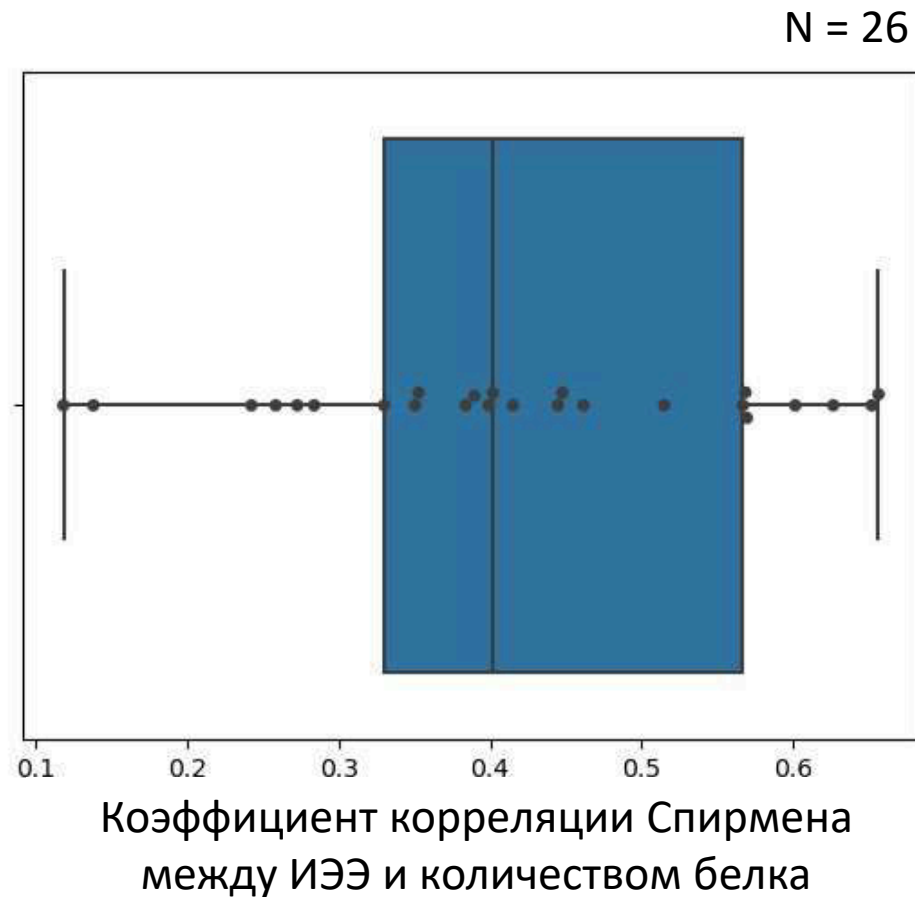
# Структура работы



# Анализ корреляции типов индексов эффективности элонгации (ИЭЭ) трансляции с экспрессией белка у различных прокариот.

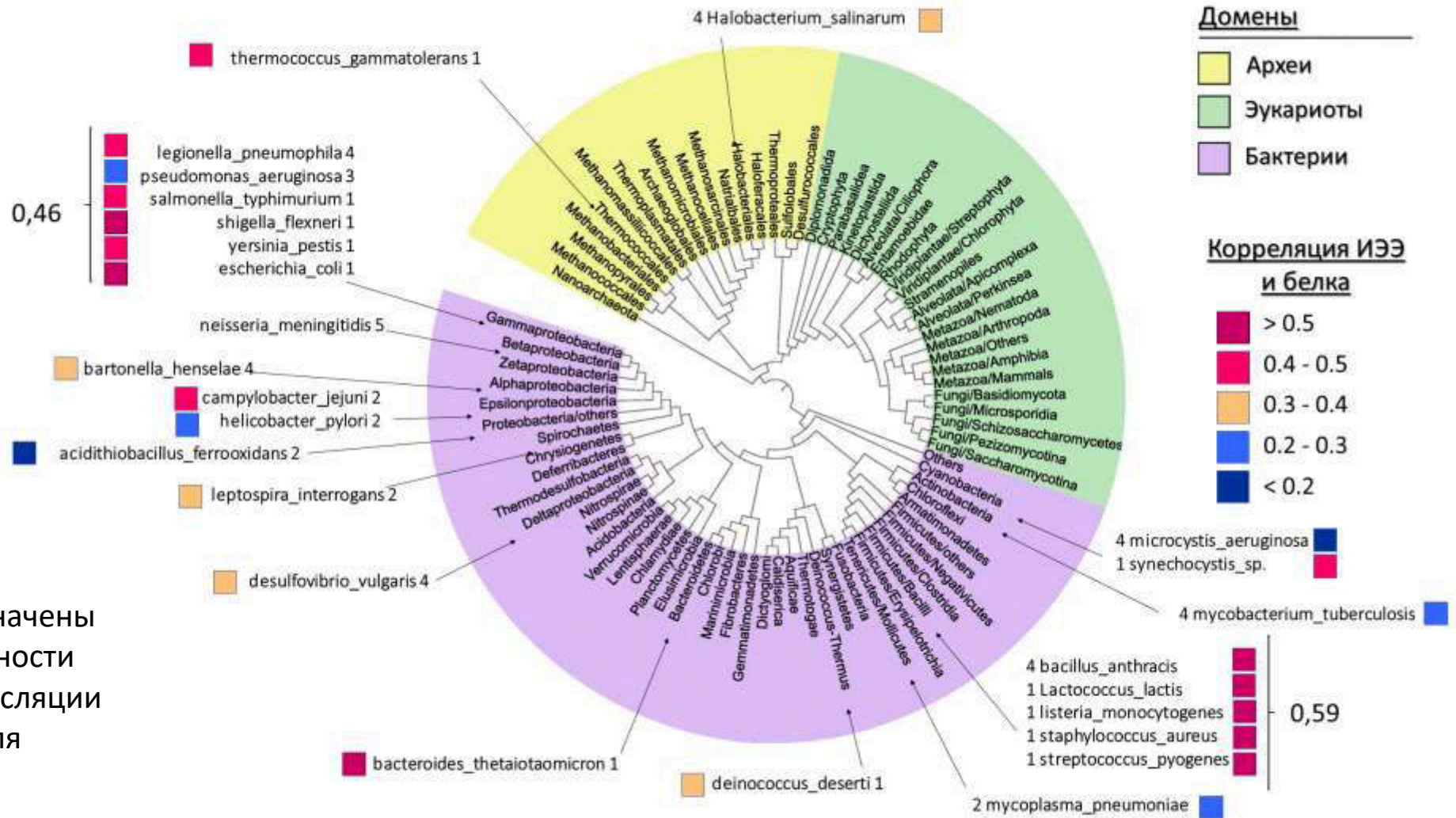


# Анализ корреляции типов индексов эффективности элонгации (ИЭЭ) трансляции с экспрессией генов на уровне белка у различных прокариот.

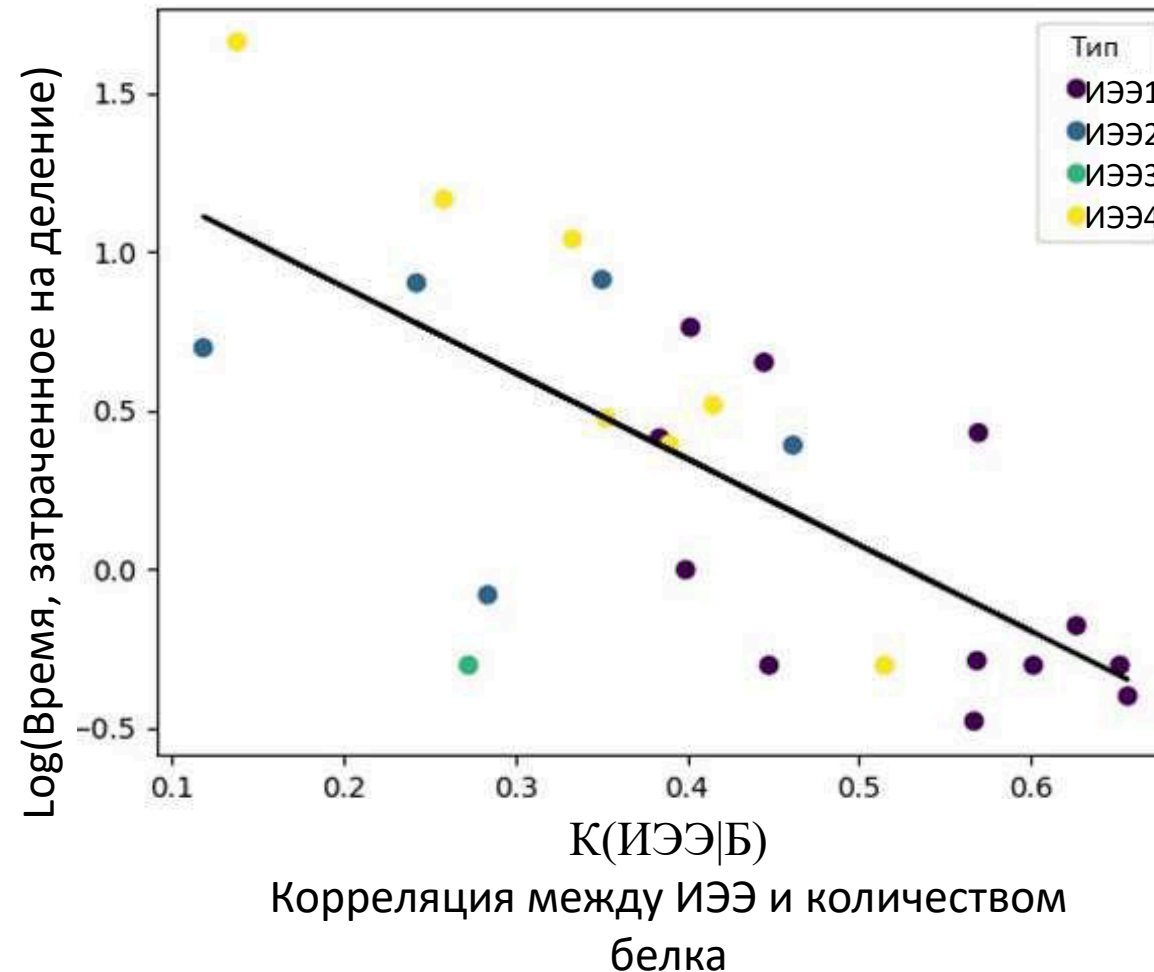


✕ - Статистически недостоверная ( $p > 0.05$ ) корреляция

# Сопоставление филогении и корреляции ИЭЭ с экспрессией генов на уровне белка



# Корреляция между ИЭЭ и количеством белка зависит от скорости деления организмов

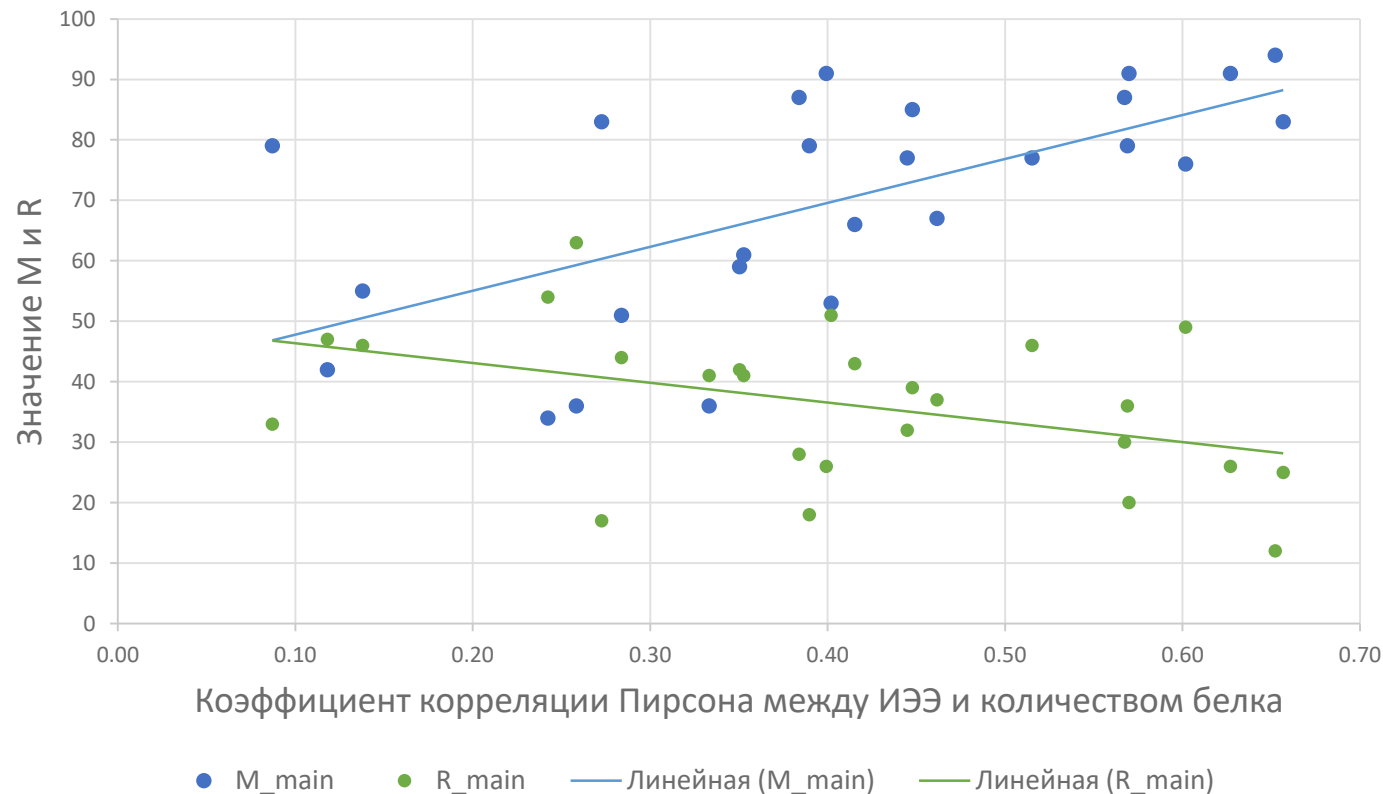


Точками на графике обозначены исследуемые организмы (N=25)

Корреляция Пирсона между временем на деление и  $K(\text{ИЭЭ}|\text{Б})$  - 0.56 ( $p < 0.05$ )

$K(\text{ИЭЭ}|\text{Б})$  – корреляция между ИЭЭ и количеством белка

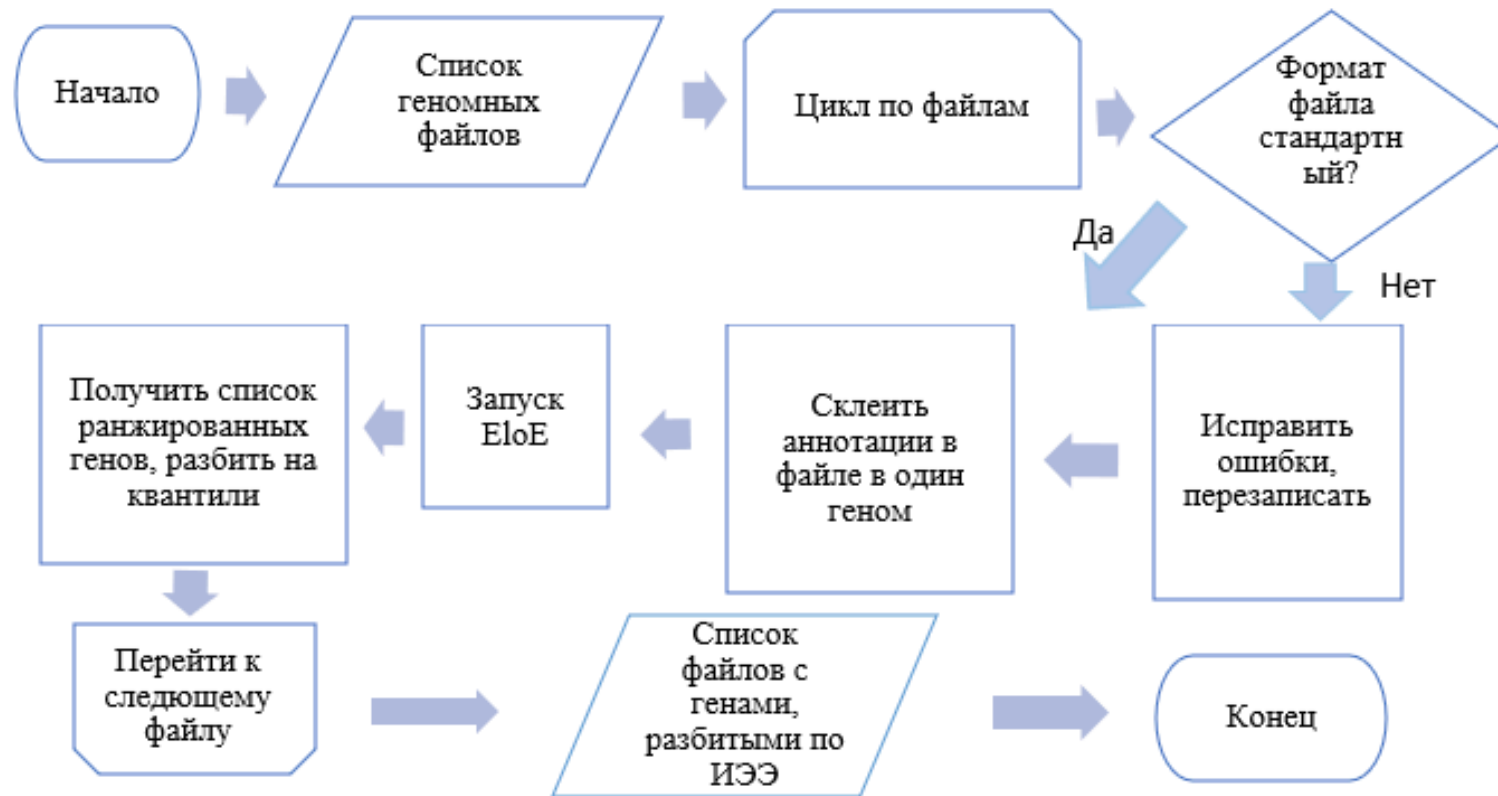
У организмов с высокими  $M$  (среднее значение ИЭЭ генов рибосомных белков) и низкими  $R$  (стандартное отклонение ГРБ) корреляция выше



Кoeffициент корреляции Спирмана для  $M$  – среднего значения ИЭЭ генов рибосомных белков (ГРБ) 0.73 ( $p < 0.05$ )

Кoeffициент корреляции Спирмана для  $R$  – стандартного отклонения ГРБ -0.47 ( $p < 0.05$ )

# Интеграция программного комплекса EюE в качестве модуля в систему обработки геномных данных КГЦ (ЦГИМУ) ИЦИГ.

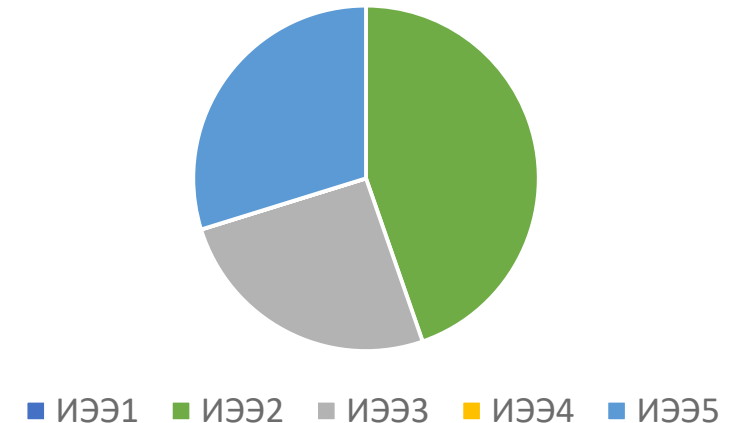


- Разработанный скрипт был интегрирован в систему обработки геномных данных КГЦ (ЦГИМУ) ИЦИГ аспирантом Мухиным Алексеем Максимовичем

# Характеристика фитопатогенов группы *Ralstonia*

- Систематика фитопатогенов недавно была пересмотрена, вид *R. solanacearum* разделен на *R. solanacearum*, *R. pseudosolanacearum*, *R. syzygii*
- Геном 5,5 - 6 Мб ( $\approx 3.5$  Мб – хромосома, 2 Мб – мегаплазмида)
- Хромосома – гены домашнего хозяйства
- Мегаплазмида – большая часть генов, связанная с патогенностью (компоненты жгутика, продукция экзополисахаридов) и штаммоспецифичностью
- Обе части генома имеют схожий паттерн повторов, что говорит о длительной совместной эволюции

Типы ИЭЭ *R. solanacearum* по 1 хромосоме



(Genin & Denny, 2012; Coenye & Vandamme, 2003)

# Тип оптимизации эффективности трансляции и филогения рода *Ralstonia*

231 геном

Количество образцов по виду и индексу

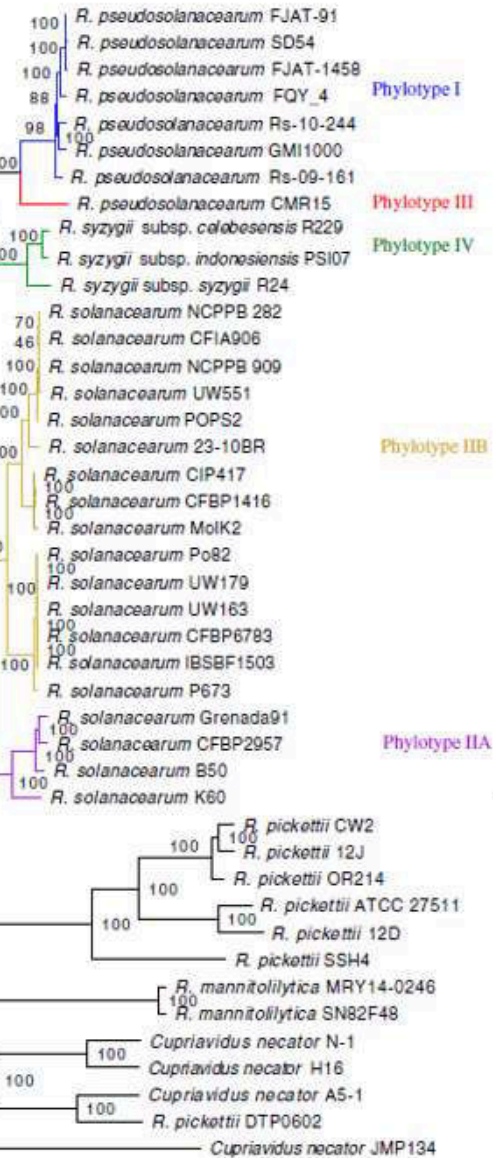
| Вид                       | ИЭЭ1 | ИЭЭ2 | ИЭЭ3 | ИЭЭ4 | ИЭЭ5 | Категория    |
|---------------------------|------|------|------|------|------|--------------|
| <i>solanacearum</i>       | 0    | 0    | 2    | 0    | 29   | фитопатогены |
| <i>pseudosolanacearum</i> | 0    | 3    | 0    | 0    | 101  |              |
| <i>syzygii</i>            | 0    | 0    | 1    | 0    | 12   | почвенные    |
| <i>insidiosa</i>          | 0    | 0    | 0    | 6    | 0    |              |
| <i>mannitolilytica</i>    | 0    | 0    | 0    | 7    | 3    |              |
| <i>pickettii</i>          | 0    | 0    | 0    | 18   | 0    |              |
| ?                         | 0    | 2    | 0    | 16   | 40   |              |

Результаты EIoE по полному геному (хромосома 1 + мегаплазида)

фитопатогены

почвенные

0.03



ИЭЭ5  
Кодонный состав + энергия вторичных структур

ИЭЭ4  
Кодонный состав + количество вторичных структур

ИЭЭ5

(Zhang & Qiu, 2016)

# Сравнительный анализ генов с высокой теоретической эффективностью элонгации трансляции



# Гены с высокой теоретической эффективностью элонгации трансляции



# Выводы

1. Проведённый анализ показал, что наибольший коэффициент корреляции между базовым индексом эффективности элонгации трансляции и экспериментально измеренным уровнем белка наблюдается у прокариот:
  - a. повышающих эффективность этого этапа за счет оптимизации кодонного состава;
  - b. относящихся к филуму Firmicutes класса Bacilli и филуму Proteobacteria класса Gammaproteobacteria;
  - c. имеющих более высокую максимальную скорость клеточного деления;
  - d. для которых оптимальный индекс эффективности элонгации трансляции имеет высокие значения параметра  $M$  (средний ранг) и низкие значения параметра  $R$  (стандартное отклонение) для генов рибосомных белков.

# Выводы

2. Программный комплекс для автоматической оценки эффективности элонгации трансляции генов различных организмов (EloE) был интегрирован в качестве модуля в систему обработки геномных данных КГЦ (ЦГИМУ) ИЦиГ СО РАН.

3. Типы оптимизации элонгации трансляции бактерий рода *Ralstonia* различаются в соответствии с крупными филогенетическими группами внутри рода: в то время как для таких фитопатогенов, как *R. solanacearum*, *R. pseudosolanacearum*, *R. syzygii*, а также почвенной бактерии *C. necator* характерна оптимизация с учётом кодонного состава и стабильности вторичных структур в мРНК (ИЭЭ5), для ветви, содержащей штаммы видов почвенных бактерий *R. mannitolilytica*, *R. pickettii* и *R. insidiosa*, характерна оптимизация с учётом кодонного состава и количества вторичных структур в мРНК (ИЭЭ4). Т. е. при дивергенции данной группы вектор отбора был направлен против наличия каких-либо вторичных структур в мРНК высоко экспрессирующихся генов.

# Выводы

4. Сравнительный анализ потенциально высоко экспрессирующихся групп генов у различных видов бактерий рода *Ralstonia* выявил, что помимо общих для различных видов генов домашнего хозяйства, для фитопатогенов в класс потенциально высоко экспрессирующихся генов попали гены, ассоциированные с патогенностью, а для почвенных бактерий в него попали гены, связанные с синтезом клеточной мембраны, ответом на тепловой и осмотический шок.

# Тезисы конференций

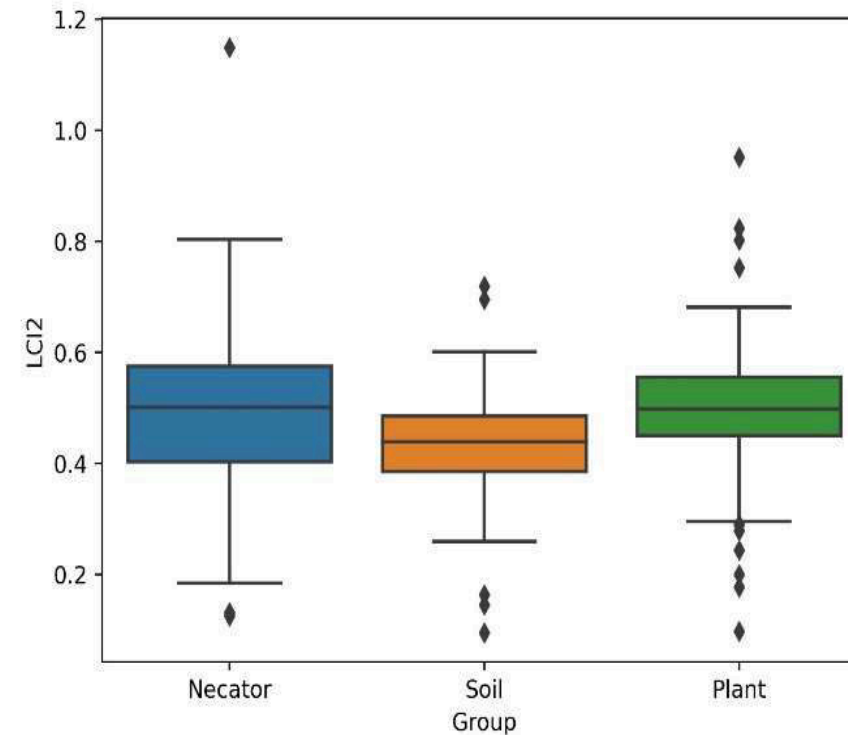
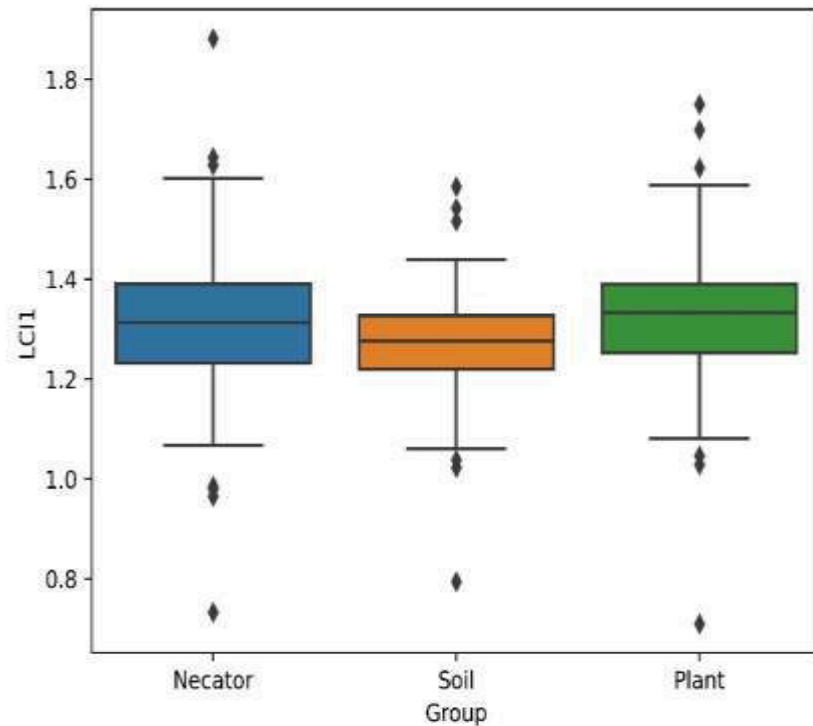
- Коренская А. Е., Биоинформатический анализ характеристик элонгации трансляции у бактерий рода *Ralstonia* //Биология: Материалы 59 Междунар. Студ. Конф. 12-23 апреля 2021. – Новосибирск, стр. 15
- Коренская А. Е., Корреляция теоретической оценки эффективности элонгации трансляции с экспрессией генов на уровне белка у различных прокариот //Материалы Международного молодежного научного форума «ЛОМОНОСОВ-2021»



Спасибо за внимание!

By Janice Carr

# Относительное количество и энергия вторичных структур у потенциально высоко экспрессирующихся генов



# Формула расчета индекса эффективности элонгации трансляции

$$EE_i(i) = K / (w_1 T_a(i) + w_2 T_e(i))$$

$i$  – номер гена,

$T_a(i)$  – параметр, имеющий смысл среднего времени размещения в А-сайте рибосомы изоакцепторной аминоксил-тРНК,

$T_e(i)$  – параметр, имеющий смысл среднего времени, затрачиваемого рибосомой на стадию транслокации,

$K$  – нормирующий множитель, обеспечивающий границы индекса от 0 до 10,

$w_1 = \{0, 1\}$  и  $w_2 = \{0, 1\}$  – индикаторные коэффициенты, определяющие учет слагаемых в значении индекса.

# Учет кодонного состава гена при расчете индекса EEI

$$T_a(i) = \sum_{j=1}^{n_i} \beta_{\delta(i,j)} / n_i$$

$$\beta_{\delta(i,j)} = \frac{\sum_{m=1}^C \sqrt{a_m}}{\sqrt{a_{\delta(i,j)}}}$$

$1/\beta_{\delta(i,j)}$  - оптимальная относительная концентрация аминоксил-тРНК, комплементарной j-ому учитываемому кодону,  
 $a_{\delta(i,j)}$  и  $a_m$  имеют смысл частот использования кодонов  $\delta(i,j)$  и  $m$  в выделенной подвыборке генов,  
 $n_i$  – количество кодонов в гене  $i$ ,  
 $C$  – общее число кодонов

# Учет потенциальных вторичных структур

$$Te(i) = t_{\min}(1 - p(i)) + t_{\max}p(i),$$

$$p(i) = \int_0^{LCI(i)} \frac{k^{n+1}x^n}{G(n+1)} e^{-kx} dx$$

$$k = \frac{m}{\sigma^2} \quad n = \left(\frac{m}{\sigma}\right)^2$$

$t_{\min}$  – минимальное условное время транслокации,  
 $t_{\max}$  – максимальное условное время транслокации,  
 $p(i)$  – вероятность реализации максимального условного времени транслокации.

$m$  и  $\delta^2$ , соответственно, математическое ожидание и дисперсия положительной случайной величины, имеющей плотность распределения  $\frac{k^{n+1}x^n}{G(n+1)} e^{-kx}$ , где  $G(n+1)$  – Гамма-функция

$$LCI1(i) = \frac{\sum_{m=1}^{m_i - s_{\max} - l_{\max}} \left\{ \sum_{s=s_{\min}}^{s_{\max}} \left[ \sum_{l=l_{\min}}^{l_{\max}} \zeta(\text{con}(m, m+s-1), \overline{\text{con}(m+s+l, m+2s+l-2)}) \right] \right\}}{m_i - s_{\max} - l_{\max}}$$

$$LCI2(i) = \frac{\sum_{m=1}^{m_i - s_{\max} - l_{\max}} \left\{ \sum_{s=s_{\min}}^{s_{\max}} \left[ \sum_{l=l_{\min}}^{l_{\max}} \psi(\text{con}(m, m+s), \overline{\text{con}(m+s+l-1, m+2s+l-2)}) \right] \right\}}{m_i - s_{\max} - l_{\max}}$$