

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В. Ломоносова**

---

**ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И КИБЕРНЕТИКИ**

**ISSN 2411-1473**

**Современные  
информационные технологии  
И  
ИТ-образование**

**Научный журнал**

**Том 2 (№ 11)**

**Москва  
2015**

УДК [004:377/378](063)  
ББК 74.5(0)я431+74.6(0)я431+32.81(0)я431  
С 56

**Современные информационные технологии и ИТ-образование. Т. 2 (№ 11),  
2015. - 614 с. (ISSN 2411-1473)**

В данном выпуске журнала представлены доклады X Юбилейной международной научно-практической конференции «Современные информационные технологии и ИТ-образование», прошедшей в Московском государственном университете имени М.В. Ломоносова 20-22 ноября 2015 года.

Журнал «Современные информационные технологии и ИТ-образование» включен в наукометрическую базу «Российский индекс научного цитирования» с размещением полнотекстовых версий в научной электронной библиотеке eLIBRARY.RU. URL: [http://elibrary.ru/title\\_about.asp?id=52785](http://elibrary.ru/title_about.asp?id=52785)



*Издание осуществлено при финансовой поддержке  
Российского фонда фундаментальных исследований  
(Грант РФФИ № 15-07-20760\_з)*

**Учредитель:**

Фонд содействия развитию интернет-медиа, ИТ-образования, человеческого потенциала «Лига интернет-медиа»

**Издатель:**

Фонд содействия развитию интернет-медиа, ИТ-образования, человеческого потенциала «Лига интернет-медиа»

**Адрес редакции:**

119991, г. Москва, ГСП-1, Ленинские горы, д. 1, стр. 52, факультет ВМК МГУ имени М.В. Ломоносова, каб. 375. E-mail: [sukhomlin@mail.ru](mailto:sukhomlin@mail.ru), тел./факс (495) 939-46-26.

Журнал зарегистрирован Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций (Роскомнадзор).

Свидетельство о регистрации средства массовой информации ПИ № ФС77-61433 от 10 апреля 2015 г.

Издается с 2005 года. Выходит 1 раз в год.

**Редакционная коллегия журнала:**

**Главный редактор:**

**Сухомлин В.А.** - доктор технических наук, профессор, заведующий лабораторией ОИТ факультета ВМК МГУ имени М.В. Ломоносова, Президент Фонда «Лига интернет-медиа»;

**Члены редакционной коллегии:**

Веремей Е.И. - доктор физ.-мат. наук, профессор, СПбГУ;

Гергель В.П. - доктор физ.-мат. наук, профессор, ННГУ им. Н.И. Лобачевского;

Самуйлов К.Е. - доктор физ.-мат. наук, профессор, РУДН;

Калиниченко Л.А. - доктор физ.-мат. наук, профессор, вед. н.с. ИПИ РАН ФИЦ ИУ РАН;

Лугачев М.И. - доктор экономических наук, профессор, МГУ имени М.В. Ломоносова;

Любецкий В.А. - доктор физ.-мат. наук, профессор, ИППИ РАН им. А.А. Харкевича;

Нечаев В. В. - доктор технических наук, профессор, МИРЭА;

Посыпкин М.А. - доктор физ.-мат. наук, вед. н. с. ИППИ РАН им. А.А. Харкевича;

Язенин А.В. - доктор физ.-мат. наук, декан факультета ПМиК, профессор, ТвГУ;

Намиот Д.Е. - кандидат физ.-мат. наук, с.н.с. факультета ВМК МГУ имени М.В. Ломоносова;

Зубарева Е.В. - кандидат пед. наук, доцент, н.с. факультета ВМК МГУ имени М.В. Ломоносова;

Сотникова М.В. - кандидат физ.-мат. наук, доцент СПбГУ.

Статьи, поступающие в редакцию, рецензируются. За достоверность сведений, изложенных в статьях, ответственность несут авторы публикаций. Мнение редакции может не совпадать с мнением авторов материалов. При перепечатке ссылка на журнал обязательна.

Материалы публикуются в авторской редакции. При перепечатке и цитировании материалов ссылка на журнал «Современные информационные технологии и ИТ-образование» обязательна.

**Зверков О.А.<sup>1</sup>, Селиверстов А.В.<sup>2</sup>, Любецкий В.А.<sup>3</sup>**

<sup>1</sup>ИППИ РАН, г. Москва, к.ф.-м.н., н.с., [zverkov@iitp.ru](mailto:zverkov@iitp.ru)

<sup>2</sup>ИППИ РАН, г. Москва, к.ф.-м.н., в.н.с., [slvstv@iitp.ru](mailto:slvstv@iitp.ru)

<sup>3</sup>ИППИ РАН, г. Москва, д.ф.-м.н., зав.лаб., [lyubetsk@iitp.ru](mailto:lyubetsk@iitp.ru)

## **О ТРАНСКРИПЦИОННЫХ ФАКТОРАХ, КОДИРУЕМЫХ В ПЛАСТИДАХ РОДОФИТНОЙ ВЕТВИ**

### **КЛЮЧЕВЫЕ СЛОВА**

*Пластида, белок, транскрипционный фактор, кластеризация.*

### **АННОТАЦИЯ**

*Расширена ранее полученная кластеризация белков, кодируемых в пластидах родофитной ветви. Результаты представлены в общедоступной базе данных по адресу <http://lab6.iitp.ru/ppc/redline67>. База данных позволяет проводить быстрый поиск кластера (семейства белков) как по фрагменту аминокислотной последовательности одного из белков, так и по филогенетическому профилю белка. На этой основе нами предсказаны регулоны транскрипционных факторов Ycf28, Ycf29 и Ycf30, кодируемых в пластидах.*

Быстрый рост числа секвенированных геномов пластид позволяет развивать предположения об эволюции и регуляции этих геномов не только у водорослей, но и у нефотосинтезирующих видов простейших, содержащих пластиды. Среди последних возбудители опасных протозойных инфекций, таких как малярия и токсоплазмоз. Именно пластиды являются одной из эффективных мишеней для терапевтического воздействия и для получения невирулентных штаммов с целью быстрого создания вакцин.

Все известные пластиды происходят от цианобактерий. Выделяют три ветви первичных пластид с независимым происхождением, которые представлены зелёными водорослями и растениями, *Synechococcus paradoxus* и багрянками. Однако многие далёкие от перечисленных виды имеют вторичные или третичные пластиды, происходящие от первичных пластид. В этой работе мы ограничимся исследованием пластид родофитной ветви, имеющих общее происхождение с пластидами багрянков.

Весной 2015 года были опубликованы четыре новых генома пластид из родофитной ветви: *Lepidodinium chlorophorum* (класс Dinophyceae, GenBank: NC\_027093, дата 14.05.2015), *Choreocolax polysiphoniae* (отдел Rhodophyta, GenBank: NC\_026522, дата 27.03.2015), *Vertebrata lanosa* (отдел Rhodophyta, GenBank: NC\_026523, дата 27.03.2015) и *Trachydiscus minutus* (отдел Eustigmatophyceae, GenBank: NC\_026851, дата 22.04.2015). В пластиде *Choreocolax polysiphoniae* кодируются, в частности, некоторые гены путей синтеза триптофана и разветвлённых аминокислот, а также гены метаболизма жирных кислот. Однако генов фотосистем нет. Поэтому он существенно отличается от других багрянков.

Для кластеризации белков применён метод, описанный в [1] и успешно апробированный в серии работ [2–4]. Для визуализации кластеров (семейств белков) использована программа *sfdp* из пакета *Graphviz* [5]. Поиск промоторов осуществлён оригинальным алгоритмом, основанном на консервативности известных промоторов [6–7] и экспериментальных данных о влиянии нуклеотидных замен на эффективность инициации транскрипции РНК-полимеразой бактериального типа [8].

Нами выполнена кластеризация белков, кодируемых в пластидах родофитной ветви. Результаты представлены в общедоступной базе данных по адресу <http://lab6.iitp.ru/ppc/redline67>. База данных позволяет проводить быстрый поиск кластера как по фрагменту аминокислотной последовательности одного из белков, так и по филогенетическому профилю. Четыре добавленных протеома содержат в сумме 462 белка, что увеличивает общее число белков до 8940. Число синглетонов равно 259; число кластеров равно 294.

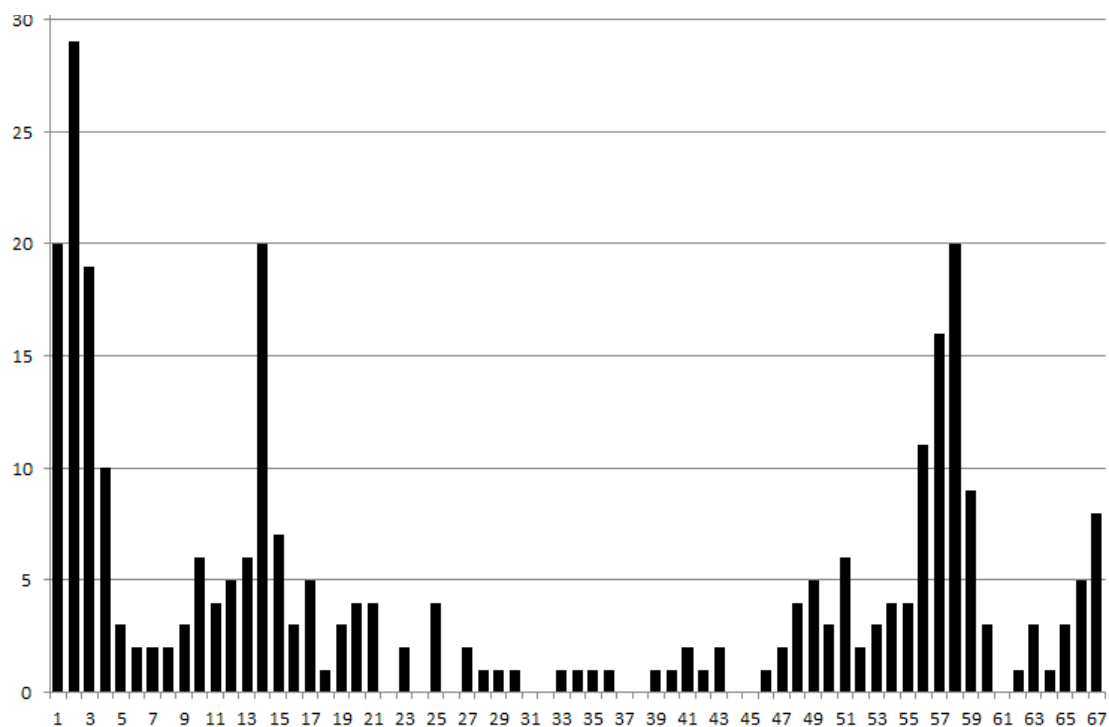


Рис.1. Зависимость числа кластеров белков от числа представленных видов водорослей

Зависимость числа кластеров белков от числа представленных видов показано на рисунке 1. Специфичными для всех пластид рассмотренных видов из отдела Rhodophyta (багрянки) являются гены *odpA/pdhA*, *odpB/pdhB*, *trpA*, *trpG*, *tilS/ycf62* и *infC*.

На рисунке 2 показан граф, вершины которого соответствуют белкам, а рёбра соединяют белки, связанные взаимными хитами (BBH) BLAST'a ( $E < 10^{-5}$ ). За редким исключением кластеры отлично разделяются выделением компонент связности. При этом многие из них удивительно плотные. Выделяются следующие пары и одна тройка близких по последовательности белков: PsaA + PsaB, PsbA + PsbD, PsbB + PsbC + синглетон Orf157, AtpA + AtpB, PetJ + PetV, ArcA + ArcD, ArcB + ArcF, CpeA + CpcA, CpcB + CpeB, CpcG + ArcE, TrpG + CarA, TufA + InfB. Здесь каждая сумма означает связный подграф, большинство вершин которого соответствуют одному из указанных типов белков. Параметры кластеризации выбираются так, чтобы указанные подграфы разделились в соответствии с аннотацией белков.

В целом распределение по кластерам белков, кодируемых в пластидах багрянки, показывает значительное отличие *Porphyridium purpureum* от остальных видов, что сопровождается многочисленными перестройками ДНК пластид этих водорослей, а также значительное обособление клады из трёх видов *Galdieria sulphuraria*, *Cyanidium caldarium* и *Cyanidioschyzon merolae*.

По сравнению с нашими предыдущими результатами кластеры белков MoeB и Ycf28 одновременно пополнились белками, кодируемыми в пластидах *Vertebrata lanosa*, у *Choreocolax polysiphoniae* и всех рассмотренных видов вне отдела Rhodophyta ни один из этих белков не кодируется в пластидах. Так подтверждено ранее отмеченное нами совпадение филогенетических профилей этих белков [4].

Белок Ycf28 имеет значительное сходство с транскрипционным фактором NtcA цианобактерий. Поэтому мы предполагаем, что именно Ycf28 регулирует в пластидах транскрипцию гена *moeB*, связывая ДНК вблизи промотора, где нами предсказан консервативный мотив. При этом нет оснований считать, что Ycf28 связан с метаболизмом азота, то есть по сравнению с цианобактериями произошла смена специфичности транскрипционного фактора к субстрату. Отсутствие типичного -35 бокса промотора перед геном *moeB*, говорит о том, что Ycf28 является активатором транскрипции.

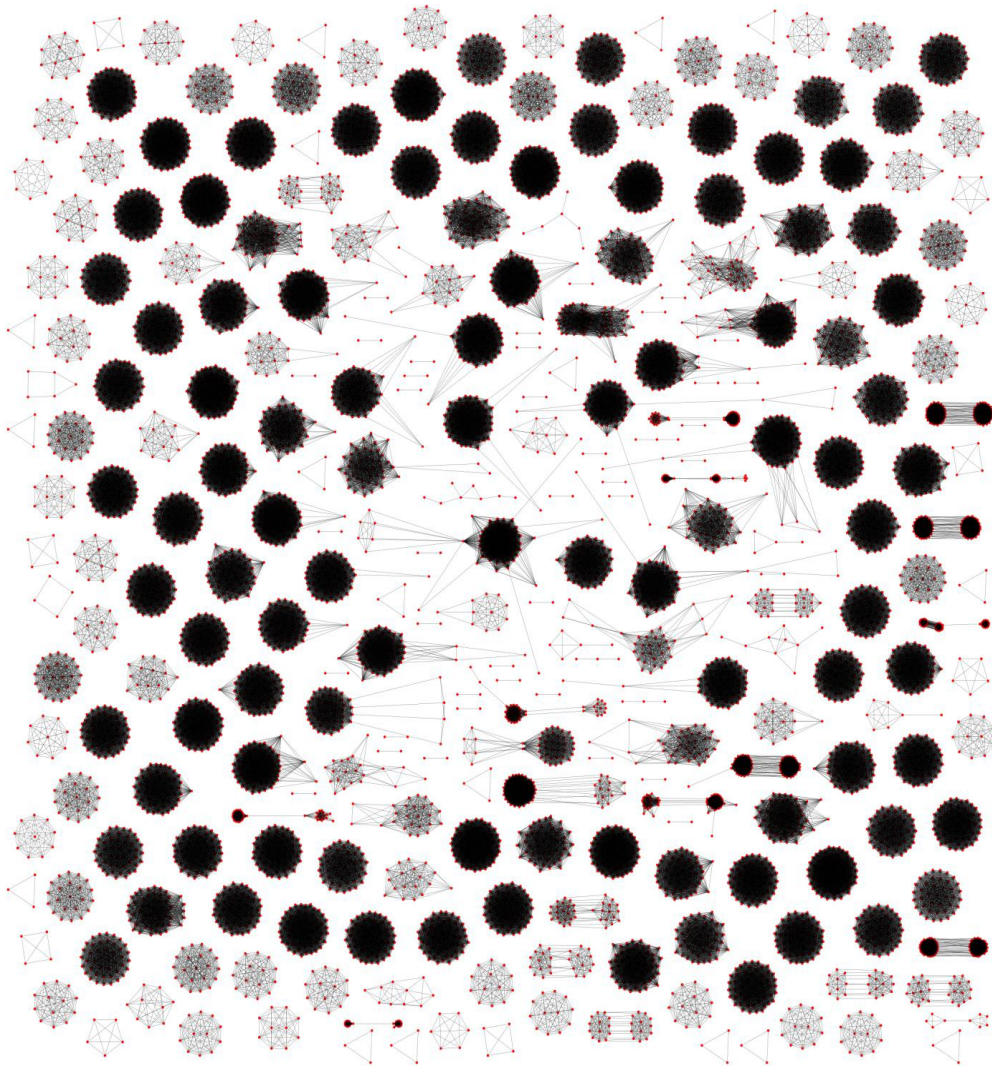


Рис.2. Компоненты связности графа белков в пластидах родофитной ветви

Транскрипционный регулятор Ycf29 кодируется в пластидах криптофитовых водорослей и багрянок кроме *Porphyridium purpureum*. Дерево Ycf29, построенное по выравниванию белков, показано на рисунке 3. Визуализация скобочной структуры выполнена программой TreeView 1.6.6 [9]. Другого белка с таким филогенетическим профилем не найдено. Близкий профиль имеет белок SemA, который есть у *Porphyridium purpureum*, но отсутствует у *Choreocolax polysiphoniae*. Белок SemA содержит домен PF03040 и найден на внутренней стороне наружной мембраны хлоропластов, но не в тилакоидной мембране. Ортологичные SemA белки у цианобактерий вовлечены в транспорт углекислого газа, но сами не являются транспортёрами [10]. Другой белок с близким филогенетическим профилем – это мембранный белок Ycf19. Перед геном *ycf19* предсказан консервативный промотор бактериального типа, близкий к консенсусу. С другой стороны, присутствие Ycf29 у нефотосинтезирующих видов *Cryptomonas paramecium* и *Choreocolax polysiphoniae* свидетельствует о том, что он регулирует процессы, не связанные с фотосинтезом. Поскольку Ycf29 входит в двухкомпонентную систему передачи сигнала, его регулон связан с реакцией на изменения внешних условий, а не внутри пластиды.

Известно, что в пластидах многих видов водорослей кодируется транскрипционный фактор Ycf30, регулирующий экспрессию генов *rbcLS*, кодирующих субъединицы рибулозобисфосфаткарбоксилазы (КФ 4.1.1.39), и гена *cbbX*. Экспериментально показана индуцируемая светом активация транскрипции этих генов в изолированных *Cyanidioschyzon merolae* и предсказан сайт связывания фактора Ycf30 в пластидах багрянок [11]. Построенные нами филогенетические профили этих белков согласуются с этими предсказаниями. Однако низкая консервативность сайта связывания Ycf30 не позволила точно предсказать его положение на ДНК.

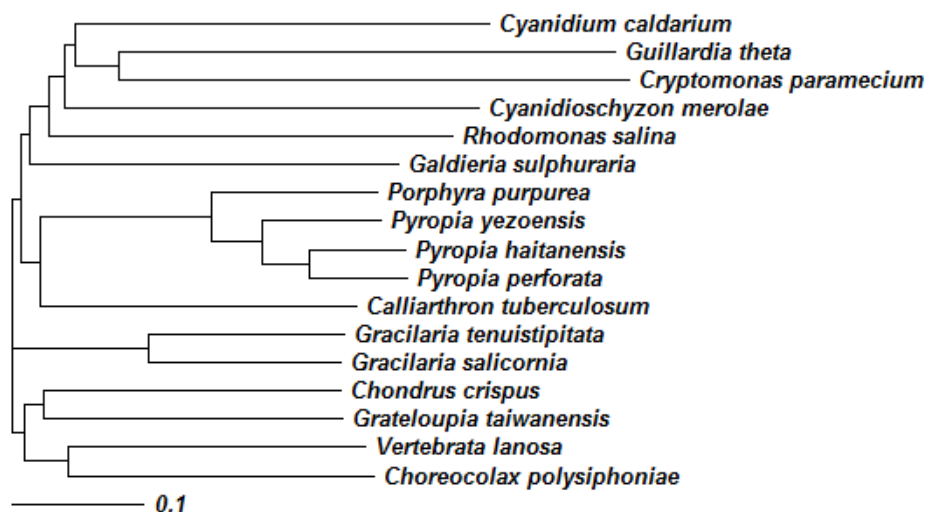


Рис.3. Дерево факторов Ycf29

Ген *ftsH* (синоним *ycf25*) кодирует АТФ-зависимую протеазу, содержащую цинк и предположительно участвующую в клеточном делении. Ортологичные гены широко распространены у прокариот [12]. У бактерии *Opencoccus oeni* белок FtsH важен для защиты от теплового или осмотического стрессов [13].

У видов из группы Stramenopiles: золотисто-бурая водоросль *Heterosigma akashiwo*, жёлто-зелёная *Vaucheria litorea* и бурые водоросли *Ectocarpus siliculosus*, *Fucus vesiculosus* и *Saccharina japonica* нами выполнено выравнивание 5'-лидерных областей генов *ftsH*. Предсказан консервативный промотор, типичный для РНК-полимераз бактериального типа. Это хорошо согласуется с близким филогенетическим положением пластид этих видов.

Дадим подробное описание предсказанных промоторов генов *ftsH*. -35 бокс промотора имеет нуклеотидный состав TTGTAT у *H. akashiwo* и TTGTTT у *V. litorea*, *E. siliculosus*, *F. vesiculosus* и *S. japonica*. -10 бокс во всех случаях имеет 5'-расширение TG. Консенсус этого бокса TGnTAnwwA, где w обозначает любой из двух нуклеотидов Т или А. Транскрипция предположительно начинается в позициях -41, -33, -29, -22 и -29 соответственно. Эти промоторы показаны на рисунке 4. Нуклеотидный состав наиболее консервативных позиций соответствует экспериментальным данным о влиянии мутаций на эффективность инициации транскрипции РНК-полимеразой бактериального типа [8].

<i>Heterosigma akashiwo</i>	AGTTGTATAAAAAGATTTT-TCATGTTATATATAATTTAA	-41
<i>Vaucheria litorea</i>	TATTGTTTTTGTAATTAATTTTATGGTACTTAAAAAATA	-33
<i>Ectocarpus siliculosus</i>	GGTTGTTTATTAAGATAAT-AAATGTTAATAACTTTAT	-29
<i>Fucus vesiculosus</i>	GTTTGTTTTTTAATACATT-ATATGTTAATAATAATATTA	-22
<i>Saccharina japonica</i>	AGTTGTTTTTACTCTGAT-ATGTGTTAATAACTATAT	-29

Рис.4. Промоторы перед генами *ftsH*. Выделены -35 и -10 боксы и 5'-расширение последнего. Справа указаны позиции потенциального сайта инициации транскрипции относительно иницирующего кодона

Менее консервативные потенциальные промоторы перед геном *ftsH* предсказаны и у других водорослей. Кроме перечисленных случаев промоторы бактериального типа с 5'-расширением -10 бокса предсказаны у *Porphyra purpurea*, *Pyropia yezoensis*, *Pyropia haitanensis*, *Cyanidium caldarium*, *Gracilaria tenuistipitata*, *Rhodomonas salina*, *Odontella sinensis*, *Thalassiosira pseudonana*, *Thalassiosira oceanica*, *Synedra acus*, *Durinskia baltica* и *Kryptoperidinium foliaceum*. Напомним, что хотя две последние водоросли принадлежат надтипу Alveolata, их пластиды имеют третичное происхождение от пластид диатомовых водорослей [14]. Промоторы без 5'-расширения -10 бокса предсказаны у *Cyanidioschyzon merolae*, *Fistulifera* sp. JPC DA0580 и *Phaeodactylum tricorutum*. У *Nannochloropsis* spp., *Porphyridium purpureum* и *Guillardia theta* определить промотор не удалось. Выделить консервативный сайт связывания какого-либо транскрипционного фактора вблизи промоторов также не удалось, что позволяет предполагать регуляцию экспрессии гена *ftsH* на каком-то другом уровне. Гены *ftsH* отсутствуют у представителей отделов Apicomplexa и Chromerida, классов Haptophyceae и Pelagophyceae, а также у нефотосинтезирующей

криптофитовой водоросли *Cryptomonas paramecium*.

Работа выполнена за счёт гранта Российского научного фонда (проект № 14–50–00150).

## Литература

1. Любецкий В.А., Селиверстов А.В., Зверков О.А. Построение разделяющих паралоги семейств гомологичных белков, кодируемых в пластидах цветковых растений // *Математическая биология и биоинформатика*. 2013. Т. 8, № 1. С. 225–233.
2. Зверков О.А., Селиверстов А.В., Любецкий В.А. Белковые семейства, специфичные для пластомов небольших таксономических групп водорослей и простейших // *Молекулярная биология*. 2012. Т. 46, № 5. С. 799–809.
3. Lyubetsky V., Seliverstov A., Zverkov O. Transcription regulation of plastid genes involved in sulfate transport in Viridiplantae // *BioMed Research International*. 2013. V. 2013. Article ID 413450. 6 p.
4. Zverkov O.A., Seliverstov A.V., Lyubetsky V.A. A database of plastid protein families from red algae and Apicomplexa and expression regulation of the *moeB* gene // *BioMed Research International*. 2015. V. 2015. Article ID 510598, 5 p.
5. Yifan Hu. Efficient high-quality force-directed graph drawing // *The Mathematica Journal*. 2006. V. 10, № 1. P. 37–71.
6. Селиверстов А.В., Лысенко Е.А., Любецкий В.А. Быстрая эволюция промоторов пластомных генов *ndhF* у цветковых растений // *Физиология растений*. 2009. Т. 56, № 6. С. 926–934.
7. Lyubetsky V., Rubanov L., Seliverstov A. Lack of conservation of bacterial type promoters in plastids of Streptophyta // *Biology Direct*. 2010. V. 5, № 34. 11 p.
8. Homann A., Link G. DNA-binding and transcription characteristics of three cloned sigma factors from mustard (*Sinapis alba* L.) suggest overlapping and distinct roles in plastid gene expression // *Eur. J. Biochem.* 2003. V. 270, № 6. P. 1288–1300.
9. Page R.D.M. TREEVIEW: An application to display phylogenetic trees on personal computers // *Computer Applications in the Biosciences*. 1996. V. 12. P. 357–358.
10. Katoh A., Lee K.S., Fukuzawa H., Ohya K., Ogawa T. *cemA* homologue essential to CO<sub>2</sub> transport in the cyanobacterium *Synechocystis* PCC6803 // *Proc Natl Acad Sci USA*. 1996. V. 93. P. 4006–4010.
11. Minoda A., Weber A.P., Tanaka K., Miyagishima S.Y. Nucleus-independent control of the rubisco operon by the plastid-encoded transcription factor Ycf30 in the red alga *Cyanidioschyzon merolae* // *Plant Physiol*. 2010. V. 154, № 3. P. 1532–1540.
12. Herman C., Prakash S., Lu C.Z., Matouschek A., Gross C.A. Lack of a robust unfoldase activity confers a unique level of substrate specificity to the universal AAA protease FtsH // *Mol Cell*. 2003. V. 11, P. 659–669.
13. Bourdineaud J.P., Nehme B., Tesse S., Lonvaud-Funel A. The *ftsH* gene of the wine bacterium *Oenococcus oeni* is involved in protection against environmental stress // *Appl Environ Microbiol*. 2003. V. 69. P. 2512–2520.
14. Imanian B., Pombert J.F., Keeling P.J. The complete plastid genomes of the two 'dinotoms' *Durinskia baltica* and *Kryptoperidinium foliaceum* // *PLoS ONE*. 2010. V. 5, № 5. Article E10711.