

Транскрипция в пластидах кокцидий

К.В. Лопатовская,
ИППИ РАН
kristina@iitp.ru

А.В. Селиверстов,
ИППИ РАН
slvstv@iitp.ru

В.А. Любецкий
ИППИ РАН
lyubetsk@iitp.ru

Аннотация

В работе описаны кодируемые в ядре субъединицы РНК-полимеразы бактериального типа у кокцидий *Eimeria tenella*, *Neospora caninum* Liverpool и *Toxoplasma gondii*. В пластидах *Eimeria tenella* и *Toxoplasma gondii* предсказаны промоторы бактериального типа и регуляция транскрипции, предполагается существенная роль РНК-полимераз фагового типа в транскрипции пластома у *E. tenella*, но не у *T. gondii*. У этих видов обсуждается различие в ответе на воздействие антибиотиков.

1. Введение

Кокцидии (Coccidia) – класс простейших, принадлежащий типу споровиков (Apicomplexa). Они – широко распространенные внутриклеточные паразиты эпителиальной ткани позвоночных и беспозвоночных. Для большинства кокцидий характерно правильное чередование бесполого размножения (шизогонии), полового процесса и спорогонии. Благодаря наличию спор кокцидии остаются жизнеспособными вне тела хозяина в течение нескольких месяцев, пока ооциста не будет проглочена животным-хозяином.

У некоторых кокцидий обнаружены полуавтономные органеллы двух типов: митохондрии и пластыды (апикопласты). Вероятно, пластыды кокцидий ведут свое происхождение от пластид водорослей [1; 2]. Транскрипционный и трансляционный аппараты пластид близки к таковым у бактерий, что делает пластыды удобной мишенью для антибиотиков. Как и у растений, кокцидии ведут транскрипцию генов в пластидах с помощью ДНК-зависимых РНК-полимераз (РНКП) бактериального и фагового типов, последние гомологичны РНКП бактериофага T7. У кокцидий роль РНКП экспериментально не изучена.

В инициацию транскрипции РНКП бактериального типа вовлекаются σ -субъединицы, у которых выделяют несколько консервативных регионов. Второй регион важен для распознавания -10 бокса промотора, третий

– для связывания σ -субъединиц с другими субъединицами РНКП и для распознавания TG-расширения -10 бокса промотора, четвертый – для распознавания -35 бокса промотора [3; 4].

Известно, что в пластидах *Toxoplasma gondii* происходят многие биохимические процессы, включая метаболизм углеводов [5], синтез липоевой кислоты [6] и жирных кислот [7]. При этом большинство ферментов кодируется в ядре. Непосредственно в пластиде кодируются только рРНК, тРНК, несколько рибосомных белков, β -, β' - и β'' -субъединицы РНКП бактериального типа, фактор элонгации Tu, шаперон ClpC и белок SufB, участвующий в формировании железосероцентров. Также имеется несколько коротких рамок считывания, вероятно, не имеющих существенного значения [8]. Отметим гипотетический рибосомный белок S5, кодируемый в пластидах многих видов; у кокцидий он чрезвычайно отличается от одноименных белков водорослей и, возможно, несёт иную функцию.

2. Материалы

Мы рассматриваем два вида кокцидий, для которых полностью секвенированы пластома: *Eimeria tenella* (NC_004823) и *Toxoplasma gondii* (NC_001799). Ядерный геном *T. gondii*, фрагменты ядерного генома *Neospora caninum* Liverpool и обсуждаемых водорослей и бактерий доступны в базе данных NCBI. Некоторые фрагменты ядерных геномов *E. tenella* и *N. caninum* получены из базы данных Sanger Institute по адресу <http://www.sanger.ac.uk>.

3. Методы

При поиске промоторов сначала отбирались участки вида YTG-(20..22N)-TA-(7..13N)-R. Затем эти участки сортировались по интенсивности связывания РНКП с Sig1, предсказанной на основе данных о влиянии мутаций в промоторе перед геном *psbA* в хлоропластах горчицы [9]. В частности, нами учитывался нуклеотидный состав -35 и -10 боксов промотора с его TG-расширением и кривизна ДНК между боксами. Затем лучшие кандидаты перед ортологичными генами у *E.*

tenella и *T. gondii* выравнивались между собой. Аналогичный метод, без учета кривизны ДНК, был применен в [10].

Для аннотации белков, кодируемых в ядре, использовались их выравнивания и данные из базы Pfam. Поиск σ -субъединиц РНКП выполнялся с помощью программы BLAST посредством выравнивания с белком XP_002367841.1 из *T. gondii* и белками, кодируемыми в нуклеоморфах криптофитовых водорослей, NP_113247.1 из *Guillardia theta* и XP_001712214.1 из *Hemiselmis andersenii*, и другими белками. Поиск РНКП фагового типа выполнялся той же программой BLAST посредством выравнивания белков NP_177063.1 и NP_179989.1 из *Arabidopsis thaliana*, EEE23737.1 из *T. gondii* GT1, XP_002367014.1 из *T. gondii* ME49, EEE31947.1 из *T. gondii* VEG, CBZ55882.1 из *N. caninum*.

Для моделирования конкуренции РНКП использовалась оригинальная параллельная программа, доступная по адресу <http://lab6.iitp.ru/ru/rivals/>, позволяющая быстро проводить большой объем вычислений с использованием суперкомпьютера. Соответствующий алгоритм описан в [11], вычисления проводились на суперкомпьютере Межведомственного Суперкомпьютерного Центра РАН.

4. Результаты

В отличие от большинства видов растений и водорослей, α -субъединица *T. gondii* ME49 кодируется в ядре, соответствующий белок XP_002367289.1 имеет 836 аминокислотных остатков. В этом белке имеется отличие в одной позиции между штаммами *T. gondii* ME49 и GT1. В ядерном геноме *E. tenella* обнаружена близкая ($E=1.1e-71$) α -субъединица, для которой определены фрагменты четырёх экзонов на контиге dev_EIMER_contig_00028796 с координатами соответственно 5283..5453, 5682..6167, 6576..6785 и 7273..7965. В ядерном геноме *N. caninum* обнаружена близкая ($E=9.9e-288$) α -субъединица, для которой определены два экзона на контиге Contig892 с координатами соответственно 45655..47412 и 47940..48611.

В ядерном геноме *T. gondii* обнаружен только один ген, кодирующий σ -субъединицу РНКП. Ее длина – 1002 аминокислотных остатков у штаммов ME49 и GT1, 1001 у штамма VEG. Ниже рассматривается белок XP_002367841.1 штамма ME49. В ядерном геноме *Neospora caninum* Liverpool ген CBZ51366.1 кодирует σ -субъединицу РНКП

длиной 1206 аминокислотных остатков. С-концы σ -субъединиц РНКП у *T. gondii* и *N. caninum* чрезвычайно близки друг к другу. Однако они не имеют существенного сходства с σ -субъединицами ни диатомовых водорослей *Phaeodactylum tricornutum* CCAP 1055/1 и *Thalassiosira pseudonana* CCMP1335, ни золотистой водоросли *Aureococcus anophagefferens*, ни криптофитовых водорослей *Guillardia theta* и *Hemiselmis andersenii*. У кокцидий σ -субъединицы, ближайшие к этим σ -субъединицам, найдены у цианобактерий *Cyanothece* sp. PCC 7822 (YP_003885480.1), *Microcoleus chthonoplastes* PCC 7420 (ZP_05024793.1), *Acaryochloris marina* MBIC11017 (YP_001519047.1) и у δ -протеобактерии *Desulfarculus baarsii* DSM 2075 (YP_003809216.1). Бактериальные ортологи имеют длины от 260 до 363 аминокислотных остатков. У всех видов хорошо выравниваются С-концы второго региона, весь третий регион и N-концы четвертого региона σ -субъединиц РНКП. По всей длине четвертый регион выравнивается у *T. gondii*, *N. caninum* и *D. baarsii*.

Определить σ -субъединицу РНКП в ядерном геноме *E. tenella* не удалось.

В пластомах *E. tenella* и *T. gondii* найдено несколько участков, близких к промотору перед геном *psbA* у горчицы, которые одинаково расположены относительно соседних генов этих пластома. В частности, предсказаны хорошие промоторы для транскрипции аспарагиновой тРНК. Ген этой тРНК пересекается с промотором на комплементарной цепи ДНК, который позволяет инициировать транскрипцию длинного оперона, включающего 16S рРНК и большое количество тРНК. У *E. tenella* лучшим кандидатом в промоторы перед этим опероном является последовательность TTGGCT-TTATATAAATATATTAT-TATATG. У *T. gondii* соответствующий промотор TTGACT-6N-TATATA-7N-TATACT перед этим опероном также хорошо согласуется с консенсусом.

В обоих пластомах имеются хорошие потенциальные промоторы транскрипции генов рибосомных белков S5, L36, S11, S12, S7 и фактора элонгации Tu (ген *tufA*).

Предсказано несколько потенциальных промоторов перед участками, лишёнными важных генов на смысловой цепи ДНК. В частности, такой промотор имеется на комплементарной цепи к генам *rpo* и на 3'-конце гена *tufA* на его смысловой цепи. Потенциальные промоторы имеются внутри *tufA* вида TTGATT-17W-TATTAT у *E. tenella* и CTGATA-17W-TATTAT у *T. gondii*, рис. 1.

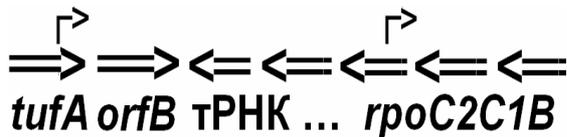


Рисунок 1. Потенциальные промотеры в локусе, содержащем гены *tufA* и *rpo* у *E. tenella* и *T. gondii*.

Двойными стрелками обозначены гены, тонкими угловыми стрелками - промотеры. Направление стрелки соответствует направлению транскрипции.

Несмотря на значительное сходство размеров и организации пластовов *E. tenella* и *T. gondii*, расположение в них некоторых предсказанных промотеров существенно различается. Например, в пластиде *T. gondii* в составе большого инвертированного повтора найден промотер TTGTAА-3N-АТАТА-9N-ТААGСТ-4N-А, с которого может транскрибироваться длинный оперон, кодирующий аргининовую и метиониновую тРНК, 26S рРНК, треониновую тРНК, SufB, субъединицы РНКП и рибосомный белок S2, рис. 2а. С другой копии этого промотера может происходить транскрипция длинного оперона, кодирующего вторые копии аргининовой и метиониновой тРНК, 26S рРНК, треониновой тРНК, рибосомный белок S4 и большое количество других тРНК. Хотя этот промотер перекрывает 5'-край валиновой тРНК на комплементарной цепи, расстояние между сайтами инициации транскрипции со встречных промотеров не превышает 34-х пар оснований.

Однако у *E. tenella* лучший кандидат в промотеры перед опероном, начинающимся с аргининовой тРНК, отделен от начала этого оперона генами трех тРНК на комплементарной цепи ДНК и имеет меньшее сходство с консенсусом, рис. 2b. Его боксы вида TTGAAC-19N-TAGAAC, то есть -10 бокс сильно отличается от ожидаемого, и отсутствует чередование нуклеотидов ТАТА между -10 и -35 боксами, существенное для изгиба ДНК.

Также ортологи σ -субъединиц РНКП найдены у простейших из отряда Naemosporida: *Plasmodium berghei* (XM_669238.1), *Pl. falciparum* 3D7 (XP_966194.1), *Pl. knowlesi* H (XM_002261430.1), *Pl. vivax* SaI-1 (XP_001616222.1), *Pl. yoelii yoelii* 17XNL (XP_724777.1), *Pl. chabaudi chabaudi* (XM_739944.1), рис. 3. В каждом из них отсутствуют другие σ -субъединицы. Не удалось определить σ -субъединицы РНКП у видов из отряда Piroplasmida: *Theileria parva*, *Th. annulata*, *Babesia bovis*.

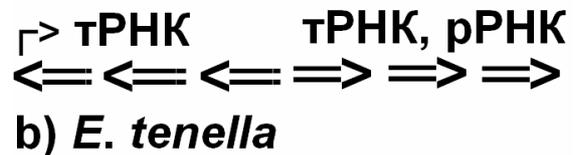
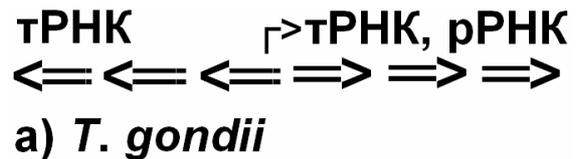


Рисунок 2. Различное положение потенциальных промотеров перед генами 26S рРНК у (a) *T. gondii* и (b) *E. tenella*.

Двойными стрелками обозначены гены, тонкими угловыми стрелками - промотеры. Направление стрелки соответствует направлению транскрипции.

У штаммов *T. gondii* ME49 (XP_002367014.1), *T. gondii* VEG (EEE31947.1), *T. gondii* GT1 (EEE23737.1) и у *N. caninum* (CBZ55882.1) найдено по одному экземпляру РНКП фагового типа с номерами, указанными в скобках. При этом белки у штаммов *T. gondii* ME49 и VEG совпадают, а у штамма GT1 имеются замены аминокислотных остатков в нескольких позициях и вставка в позициях от 347 до 354. У *E. tenella* определить РНКП фагового типа не удалось.

Также гомологи РНКП фагового типа найдены у многих споровиков, не являющихся кокцидиями: *Plasmodium berghei* (XP_676913.1), *Pl. falciparum* 3D7 (XP_001347935.1), *Pl. knowlesi* H (XP_002259256.1), *Pl. vivax* SaI-1 (XP_001615369.1), *Pl. yoelii yoelii* 17XNL (XP_727223.1), *Pl. chabaudi chabaudi* (XP_739650), *Babesia bovis* T2Bo (XP_001611431.1), *Theileria annulata* (XP_953797.1), *Theileria parva* strain Muguga (XP_766496.1), рис. 4. Однако ортологичный белок не найден у кокцидии *Cryptosporidium parvum*, которая в отличие от многих споровиков не имеет пластид.

5. Обсуждение

Наличие единственной σ -субъединицы РНКП у *T. gondii* позволяет думать, что все промотеры бактериального типа у *T. gondii* близки к одному консенсусу. Наличие консервативного четвертого региона у σ -субъединиц РНКП подтверждает существенную роль -35 бокса промотера, а высокая консервативность третьего бокса подтверждает важность TG-расширения -10 бокса промотера.

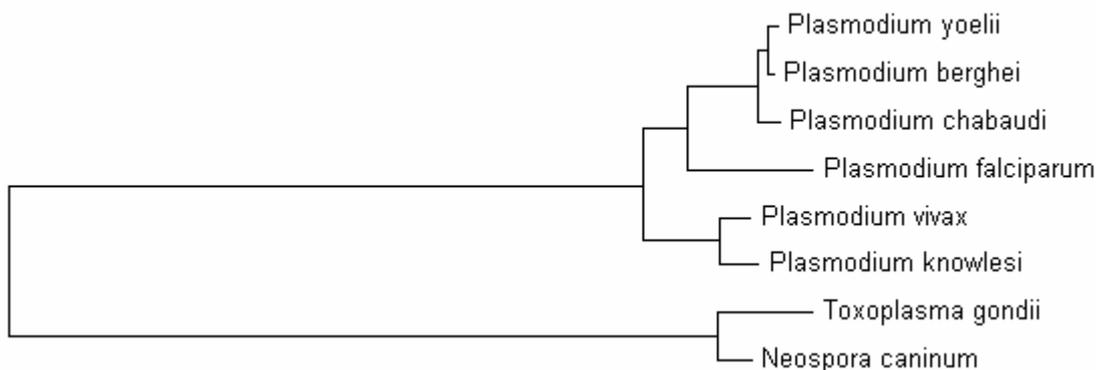


Рисунок 3. Дерево σ -субъединиц РНКП у споровиков.

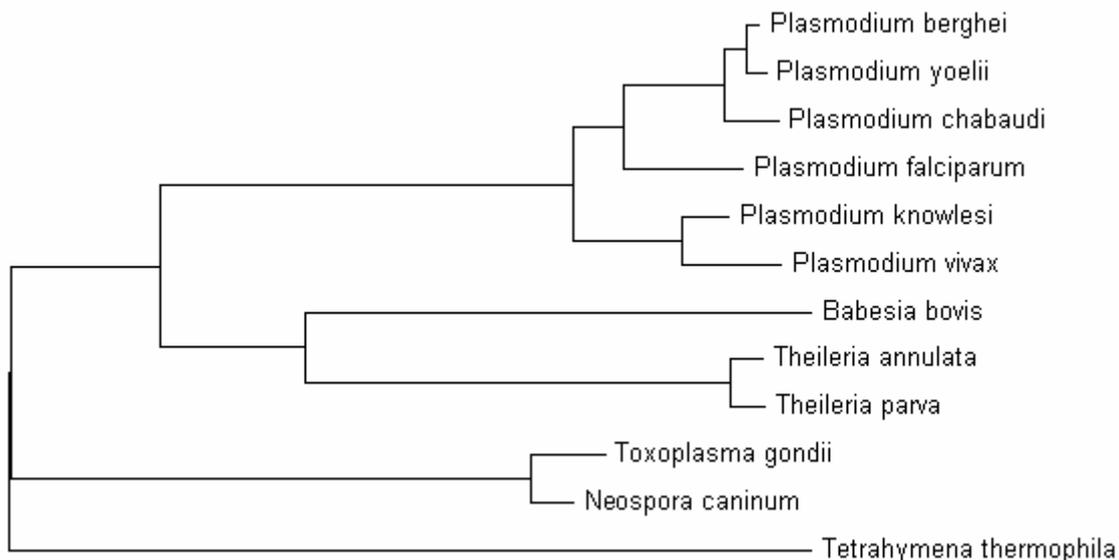


Рисунок 4. Дерево РНКП фагового типа у простейших надтипа Alveolata.

Вместе с предсказанием большого числа промоторов перед интенсивно транскрибируемыми генами это позволяет считать, что оптимальный промотор близок к таковому перед геном *psbA* у горчицы. Существенное различие между σ -субъединицами РНКП у кокцидий, с одной стороны, и диатомовыми, золотистыми и криптофитовыми водорослями, с другой стороны, позволяет предполагать независимое происхождение пластид этих групп от багрянков. Аналогичные выводы можно сделать и на основе сопоставления механизмов регуляции трансляции в пластидах [12]. Пластиды *Durinskia baltica* и *Kryptoperidinium foliaceum* близки к пластидам диатомовой водоросли *Phaeodactylum tricoratum* [13]. Таким образом, у разных видов группы Alveolata пластиды существенно отличаются.

Быстрая эволюция набора σ -субъединиц РНКП и состава соответствующих промоторов хорошо известна у цветковых растений [14].

Наши результаты позволяют предполагать качественные различия между *E. tenella* и *T. gondii* в наборе РНКП и составе соответствующих промоторов у пластид. В отличие от *T. gondii*, у *E. tenella* не найдено хороших промоторов бактериального типа, обеспечивающих интенсивную транскрипцию 26S рРНК и некоторых тРНК, рис. 2. Последнее позволяет думать, что в транскрипции этих генов основную роль играют необычные σ -субъединицы [4] или РНКП фагового типа, как у гена *rpoB* в пластидах табака [15]. В случае участия РНКП фагового типа непродолжительное воздействие антибиотиков, блокирующих РНКП бактериального типа, *E. tenella* может сохранять пластиды. Напомним, что в пластидах нефотосинтезирующей лианы *Epifagus virginiana* транскрипция осуществляется исключительно РНКП фагового типа.

Промоторы нередко расположены внутри кодирующих областей других генов, что можно

объяснить высокой плотностью генов и относительно низкой консервативностью многих белков. Заметим, что консервативные промоторы внутри кодирующих областей известны и хорошо изучены, например, перед геном *psbC* в пластидах растений.

Напомним, что β , β' и β'' -субъединицы РНКП бактериального типа кодируются в пластоме генами оперона *rpoBC12*. Интересно, что предсказана серия близких к консенсусу промоторов, с которых иницируется транскрипция локуса, содержащего оперон *rpoBC12*, в противоположном направлении, рис. 1. Это может свидетельствовать об авторегуляции транскрипции этого оперона в результате конкуренции РНКП. Действительно, повышение количества РНКП бактериального типа увеличивает интенсивность транскрипции с промоторов бактериального типа и при превышении некоторого порога транскрипция оперона *rpoBC12* от начала до конца станет невозможной из-за столкновений РНКП друг с другом. Этот эффект усиливает образование дуплексов РНК, транскрибированных с комплементарных цепей ДНК. Конкуренция приводит здесь к тому, что концентрация РНКП бактериального типа остается низкой, что однако естественно для сильно редуцированных пластов. Этот механизм дополняет ранее описанную регуляцию транскрипции *rpoB* [12].

Другим неожиданным примером является предсказание промотора бактериального типа на 3'-конце гена *tufA*, рис. 1. Ниже него лежит лишь слабо консервативная рамка считывания ORF-B, значение которой не известно и далее на комплементарной цепи – ген фенилаланиновой тРНК. Можно думать, что этот промотор внутри гена *tufA* служит исключительно для сбрасывания РНКП, идущих по ДНК навстречу гену *tufA*, на основе конкуренции. Таким образом, инициация транскрипции с этого промотора выступает в роли своеобразного терминатора транскрипции. Терминаторы другой природы между сходящимися навстречу друг другу интенсивно транскрибируемыми генами наблюдаются у бактерий [16].

Работа выполнена при частичной финансовой поддержке Министерства образования и науки РФ (П2370 и 14.740.11.0624).

Список литературы

[1] A.O.T. Lau, T.F. McElwain, K.A. Brayton, D.P. Knowles, E.H. Roalson, *Babesia bovis: A comprehensive phylogenetic analysis of plastid-encoded genes supports green algal origin of apicoplasts*, Experimental Parasitology (2009), 123, 236–243.

[2] J. Janousek, A. Horak, M. Obornik, J. Lukes, P.J. Keeling, *A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids*, PNAS (2010), 107(24), 10949–10954.

[3] E.A. Campbell, O. Muzzin, M. Chlenov, J.L. Sun, C.A. Olson, O. Weinman, M.L. Trester-Zedlitz, S.A. Darst, *Structure of the bacterial RNA polymerase promoter specificity sigma subunit*, Mol. Cell (2002), 9, 527–539.

[4] E.A. Lysenko, *Plant sigma factors and their role in plastid transcription*, Plant Cell Rep. (2007), 26, 845–859.

[5] T. Fleige, K. Fischer, D.J.P. Ferguson, U. Gross, W. Bohn, *Carbohydrate metabolism in the Toxoplasma gondii apicoplast: localization of three glycolytic isoenzymes, the single pyruvate dehydrogenase complex, and a plastid phosphate translocator*, Eucaryotic Cell (2007), 6, 984–996.

[6] N. Thomsen-Zieger, J. Schachtner, F. Seeber, *Api-complexan parasites contain a single lipoic acid synthase located in the plastid*, FEBS Letters (2003), 547, 80–86.

[7] J. Mazumdar, E.H. Wilson, K. Masek, C.A. Hunter, B. Striepen, *Apicoplast fatty acid synthesis is essential for organelle biogenesis and parasite survival in Toxoplasma gondii*, PNAS (2006), 103(35), 13192–13197.

[8] R.J.M. (Iain) Wilson, K. Rangachari, J.W. Saldanha, L. Rickman, R.S. Buxton, J.F. Eccleston, *Parasite plastids: maintenance and functions*, Phil. Trans. R. Soc. Lond. B. (2003), 358, 155–164.

[9] A. Homann, G. Link, *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), 270(6), 1288–300.

[10] V.A. Lyubetsky, L.I. Rubanov, A.V. Seliverstov, *Lack of conservation of bacterial type promoters in plastids of Streptophyta*, Biology Direct (2010), 5, 34.

[11] V.A. Lyubetsky, O.A. Zverkov, L.I. Rubanov, A.V. Seliverstov, *Modeling RNA polymerase competition: the effect of σ -subunit knockout and heat shock on gene transcription level*, Biology Direct (2011), 6, 3.

[12] Т.А. Садовская, А.В. Селиверстов, *Анализ 5'-лидерных областей в пластидах у простейших типа Apicomplexa и у красных водорослей*, Молекулярная биология (2009), 43(4), 599–604.

[13] B. Imanian, J-F. Pombert, P.J. Keeling, *The complete plastid genomes of the two 'dinotoms' Durinskia baltica and Kryptoperidinium foliaceum*, PLoS ONE (2010), 5(5), E10711.

[14] А.В. Селиверстов, Е.А. Лысенко, В.А. Любецкий, *Быстрая эволюция промоторов пластомных генов ndhF у цветковых растений*, Физиология растений (2009), 56(6), 926–934.

[15] K. Liere, P. Maliga, *In vitro characterization of the tobacco rpoB promoter reveals a core sequence motif conserved between phage-type plastid and plant mitochondrial promoters*, EMBO J. (1999), 18, 249–257.

[16] Е.В. Любецкая, А.В. Селиверстов, В.А. Любецкий, *У актинобактерий число длинных шпилек в межгенных трейлерных областях велико по сравнению с другими областями*, Молекулярная биология (2007), 41(4), 739–742.