

Регуляция транскрипции гена *chlL* у *Viridiplantae*¹

К. В. Лопатовская, О. А. Зверков, А. В. Селиверстов, В. А. Любецкий

Институт проблем передачи информации им. А. А. Харкевича, Российская академия наук,
Москва, Россия

Поступила в редколлегию 17.07.2012

Аннотация—У хары и многих наземных растений для оперона, кодирующего ферменты синтеза хлорофилла в темноте, предсказана транскрипционная регуляция. С другой стороны, для некоторых видов мы предсказали регуляцию на уровне процессинга.

КЛЮЧЕВЫЕ СЛОВА: транскрипция, регуляция, множественное выравнивание.

1. ПОСТАНОВКА ЗАДАЧИ

Известно, что синтез хлорофилла у растений осуществляется двумя путями, из которых только один не зависит от света и может проходить в темноте [1]. Известно также, что экспрессия соответствующих генов регулируется в зависимости от освещения [2]. Нами предприняты попытки исследовать регуляцию генов синтеза хлорофилла, кодируемых в пластидах. Для этого проведён поиск консервативных сайтов связывания белков с ДНК вблизи промотора. Повторы сайтов бывают как прямые, так и обратные. Поскольку ДНК представляет собой две комплементарные цепи, то при обратном повторе те же сайты лежат на разных цепях ДНК и одинаково могут служить для связывания белка с ДНК.

Нами рассмотрена регуляция транскрипции оперона *chlLN*. Синтез хлорофилла *a* связан с ферментом протохлорофиллидредуктазой, состоящим из трёх субъединиц, кодируемых генами *chlL*, *chlN* и *chlB*. В пластидах некоторых зелёных водорослей ген *chlL* имеет интроны. Если транскрипция генов осуществляется РНК-полимеразой бактериального типа (то есть скорости транскрипции и трансляции сопоставимы), то трансляция этих генов находится под регуляцией, задерживающей трансляцию до завершения сплайсинга. Напротив, отсутствие хорошего кандидата для промотора бактериального типа и отсутствие консервативных структур в нетранслируемых областях мРНК свидетельствует о транскрипции РНК-полимеразой фагового типа.

Согласно имеющимся данным [2] экспрессия гена *chlL* у *Chlamydomonas reinhardtii* регулируется на уровне трансляции, причём в регуляцию вовлечены ядерные гены *y-1*, *y-5*, *y-6*, *y-8* и *y-10*, мутации в любом из которых препятствуют трансляции. С другой стороны, у близкого вида зелёных водорослей *Volvox carteri* ген *chlL* содержит интрон.

2. МАТЕРИАЛЫ И МЕТОДЫ

Рассмотрены сиквенсы из базы данных GenBank, содержащие гены или псевдогены *chlL* из пластид следующих видов: *Chara vulgaris* NC008097, *Aneura mirabilis* NC010359, *Marchantia polymorpha* NC001319, *Physcomitrella patens* NC005087, *Syntrichia ruralis* NC012052, *Anthoceros formosae* NC004543, *Huperzia lucidula* NC006861, *Selaginella uncinata* AB197035, *Polystichum*

¹ Работа выполнена при частичной финансовой поддержке госконтрактами 14.740.11.0624 и 14.740.11.1053 Министерства образования и науки РФ.

acrostichoides PAU00732, *Angiopteris evecta* NC008829, *Adiantum capillus-veneris* NC004766, *Cycas taitungensis* NC009618, *Keteleeria davidiana* NC011930, *Picea abies* AJ001025, *Pinus koraiensis* NC004677, *P. thunbergii* NC001631, *Cunninghamia lanceolata* AB232499, *Metasequoia glyptostroboides* AB232463, *Sequoia sempervirens* AB232465, *Sequoiadendron giganteum* AB232467, *Cryptomeria fortunei* AB232469, *C. japonica* NC010548, *Glyptostrobos pensilis* AB232473, *Taxodium distichum* AB232475, *Thuja occidentalis* AB232497, *T. plicata* AB232495, *T. standishii* AB232493, *Thujopsis dolobrata* AB232491, *Platycladus orientalis* AB232489, *Juniperus chinensis* AB232487, *J. rigida* AB232485, *Cupressus sempervirens* AB232483, *Chamaecyparis obtusa* AB232479, *C. lawsoniana* AB232481, *C. pisifera* AB232477, *Ephedra equisetina* NC011954, *Gnetum gnemon* AJ007508.

Сначала перед генами *chlL* предсказано по несколько участков, сходных с промотором для РНК-полимеразы бактериального типа перед геном *psbA* у *Sinapis alba*, где известно влияние нуклеотидных замен на качество связывания РНК-полимеразы с промотором [4]. Для контроля такой же поиск проведён перед другими генами [5].

После определения потенциальных промоторов вблизи этих участков проведён поиск консервативных сайтов на основе построения множественного выравнивания с учётом дерева видов, т. е. посредством сравнения участков сначала у близких, а потом шаг за шагом у всё более далёких видов. Наконец, на основе так полученных консенсусов поиск сайтов по образцу распространён на весь пластом. Два первых этапа этого метода проводились с помощью оригинальных программ, подробно описанных и представленных для свободного пользования по адресам <http://lab6.iitp.ru/ru/treeal> [6] и <http://lab6.iitp.ru/ru/twobox> [7]. Эти выравнивания незначительно редактировались вручную. Для вычисления энергии РНК использовалась программа RNAstructure, являющаяся обновлением ранее описанной программы [8].

Гены *chlL*, *chlN* и *chlB* присутствуют в пластидах большинства зелёных водорослей и растений, кроме цветковых растений, *Welwitschia mirabilis* из отдела Gymnospermae и *Psilotum nudum* из класса псилотовых. Они присутствуют в пластомах у жёлто-зелёной водоросли *Vaucheria litorea*, у багрянок *Porphyra purpurea*, *P. yezoensis* и у *Cyanophora paradoxa*. Первые два гена обычно составляют оперон *chlLN*, исключением являются зелёные водоросли *Scenedesmus obliquus* и *Stigeoclonium helveticum*. *Welwitschia mirabilis*, *Psilotum nudum* и все цветковые растения не имеют гена *chlL*. У *Gnetum gnemon* найден псевдоген *chlL*.

3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1. Исследование области промоторов у *Streptophyta*

Промоторы перед опероном *chlLN* менее консервативны, чем промоторы перед генами *psbA*, *psbB*, *psaA* и *rbcL* [5]. Предсказана регуляция инициации транскрипции оперона *chlLN* посредством связывания транскрипционного фактора с сайтом ДНК вблизи промотора. Сайт перекрывает промотор и является тандемным повтором с консенсусом GATCTAT-11п. н. -GATCTAT. Сайт найден у единственной водоросли *Chara vulgaris*, мохообразного *Anthoceros formosae*, плауна *Huperzia lucidula*, папоротника *Adiantum capillus-veneris* и голосеменных растений *Cycas taitungensis*, *Keteleeria davidiana*, *Picea abies*, *Pinus koraiensis* и *P. thunbergii*. (Рис. 1). Перед другими генами сайт не удалось найти. Искажённый сайт с промотором низкого качества найден перед геном *chlN* у некоторых видов Coniferopsida и перед псевдогеном *chlL* у *Gnetum gnemon*. Сайт не найден у других водорослей, мхов *Aneura mirabilis*, *Marchantia polymorpha*, *Physcomitrella patens*, *Syntrichia ruralis*, папоротника *Angiopteris evecta*, *Polystichum acrostichoides* и голосеменных *Ephedra equisetina*, *Cunninghamia lanceolata*, *Metasequoia glyptostroboides*, *Sequoia sempervirens*, *Sequoiadendron giganteum*, *Cryptomeria japonica*, *C. fortunei*, *Glyptostrobos pensilis*, *Taxodium distichum*, *Thuja standishii*, *T. occidentalis*, *T. plicata*, *Thujopsis dolobrata*, *Platycladus orientalis*, *Juniperus rigida*, *J. chinensis*, *Cupressus sempervirens*, *Chamaecyparis pisifera*, *C. obtusa*, *C. lawsoniana*.

<i>Chara vulgaris</i>	AG GATCTTTT GGAAAAAC==CCAAA ATCTAT GCSTAATATTATTA -180
<i>Anthoceros formosae</i>	AAGAAACATA GATAAAT TTTAACTCTTCCT GATCTAG TTTCACT -96
<i>Huperzia lucidula</i>	СТТАТТГАТТ GGATACCT ААТААСТА=GG GATCTAA TTAGAG -46
<i>Adiantum capillus-veneris</i>	GATTTCGTAT TTGATCAAT CACCAATGAGT= GGTTTCTAT TCCCC -156
<i>Cycas taitungensis</i>	GTGGAGAG TTAAACTAT TGGTATGACCG GATCTAT TATTCCSTA -137
<i>Keteleeria davidiana</i>	TAGAGTTAT TTGAACST TAAATACAATAG GATCTAT TATTATTA -130
<i>Picea abies</i>	TAGAGTTAT TTGAACST TAAATACAATAG GATCTAT TATTATTA -110
<i>Pinus koraiensis</i>	TAGAGTTAT TTGAACST TAAATACAATAG GATCTAT TATTATTG -120
<i>Pinus thunbergii</i>	TAGAGGTAT TTGACSTAT TAAATACAATAG GATCTATA AATTATTG -101

Рис. 1. Промоторы и сайты связывания транскрипционного фактора перед геном *chlL*. Участки выравнивались по боксам промоторов. Указаны координаты 3'-края каждого участка. У промоторов –35 и –10 боксы и TG-расширение подчёркнуты, а предполагаемые сайты выделены полужирным.

Также найдены сайты другого типа — инвертированные повторы с консенсусом GATCTAT–11п. н.–ATAGATC, которые перекрывают промоторы у голосеменных растений *Chamaecyparis lawsoniana*, *C. obtusa*, *Cunninghamia lanceolata*. (Рис. 2). Этот повтор у *C. lanceolata* не имеет аналогов у других исследованных видов. Здесь есть два варианта промотора (оба подчёркнуты): CATTTGCCATGCAGTAAAT**TTGGTTAT**TCCATTCTATGATTT**TAATAGAGAGA**ACTAAGTCTAA

<i>Chamaecyparis lawsoniana</i>	AAAA AAACAAT GTCAAACA ACTATAATT CAATTTATAAAATAAGGCAAAG
<i>Chamaecyparis obtusa</i>	AAAA AAACAAT GTCAAACA ACTACAATT CAATTTATAAAATAAGGCAAAG
<i>Chamaecyparis pisifera</i>	AAAAAAAACAATGTCAAACAATGTAACATGCTACAAT TTCAATTT TACAAAT
<i>Cupressus sempervirens</i>	AACACACACAATGTTACAAAGAATGTAATGTAACATACTAGAAATTTATTT
<i>Juniperus rigida</i>	AAAGCAAACAATGTTACAAAGAATGTAATGTAACATACTAGAAATTTATTT
<i>Juniperus chinensis</i>	AAAACAACAATGTTACAAAGAATGTAATGTAACATACTAGAAATTTATTT
<i>Platycladus orientalis</i>	AAAAAAAACAATGTTACAAAGAATGTAACATACTAGAAATTTATTTTCATAT
<i>Thujopsis dolabrata</i>	AAAAAAAACATTTGTCAAACAATGCAATGTAACATGCTACAATTTATAAAT

Рис. 2. Выравнивание участков из лидерных областей гена *chlL* у *Cupressoideae*. У *C. lawsoniana* и *C. obtusa* –35 и –10 боксы промоторов подчёркнуты, предполагаемые сайты выделены полужирным.

3.2. Исследование области промоторов у *Chlamydomonadales*

Вероятный промотор у *Chlamydomonas reinhardtii* перед *chlL* соответствует 5'-UTR длиной 530н. У двух близких видов *C. reinhardtii* и *Volvox carteri* из *Chlamydomonadales* в 5'-нетранслируемой области мРНК *chlL* найден повтор с консенсусом UCCCCUHHGGGNA. Здесь H={A, C, U}. Наиболее консервативен мотив с консенсусом UCCCCUNGGG. Он весьма часто встречается в пластиде у *C. reinhardtii*, хотя отсутствует или редок у других видов, чьи пластомеры полностью секвенированы и собраны в единый контиг. Участки встречаются неравномерно: иногда повторы разделены десятками нуклеотидов, а иногда — тысячами.

Указанные повторы могли образоваться в результате многократных хромосомных перестроек, которые привели к многочисленным и различным у близких видов перестановкам генов, кодируемых в пластидах у видов Chlamydomonadales. При этом происходил распад многих оперонов, обычно присутствующих в пластидах других зелёных водорослей.

3.3. Корреляция светозависимой регуляции на уровне процессинга и вставки внутри кодирующей области гена chlL у зелёных водорослей

У трёх водорослей из таксономической группы Chlorophyta обнаружены длинные вставки, не нарушающие рамку считывания *chlL*. Такая вставка есть у *C. reinhardtii*, *Volvox carteri* (из Chlamydomonadales) и у *Rychnococcus provasolii* (из Pseudosourfieldiales), но отсутствует у других исследованных видов. Эта вставка отсутствует у близких видов *Chlamydomonas moewusii* и *Dunaliella salina* из Chlamydomonadales и *Nephroselmis olivacea* из Pseudosourfieldiales. Поскольку аминокислотная последовательность белка ChlL чрезвычайно консервативна за исключением указанной вставки вблизи С-конца у перечисленных видов, можно думать, что эта вставка удаляется в ходе процессинга.

СПИСОК ЛИТЕРАТУРЫ

1. Fujita Y., Bauer C.E. Reconstitution of light-independent protochlorophyllide reductase from purified Bchl and BchN-BchB subunits. *In vitro* confirmation of nitrogenase-like features of a bacteriochlorophyll biosynthesis enzyme. *The journal of biological chemistry*, 2000, vol. 275, no. 31, pp. 23583–23588.
2. Cahoon A.B., Timko M.P. Yellow-in-the-dark mutants of chlamydomonas lack the CHLL subunit of light-independent protochlorophyllide reductase, *The Plant Cell*, 2000, vol. 12, pp. 559–568.
3. Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F., Higgins D.G.. The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 1997, vol. 25, pp. 4876–4882.
4. 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. *European Journal of Biochemistry*, 2003, vol. 270, no. 6, pp. 1288–300.
5. Lyubetsky V.A., Rubanov L.I., Seliverstov A.V. Lack of conservation of bacterial type promoters in plastids of Streptophyta. *Biology Direct*, 2010, vol. 5, no. 34.
6. Rubanov L., Seliverstov A., Lyubetsky V. Multiple alignment based on species tree. *Proceedings of the Sixth International Conference on Bioinformatics of Genome Regulation and Structure (BGRS 2008)*. Novosibirsk, 2008, p. 212.
7. Селиверстов А.В., Любецкий В.А. Алгоритмы для поиска многобоксовых сигналов и их применение. *Труды 51-й научной конференции МФТИ Современные проблемы фундаментальных и прикладных наук*. М.: МФТИ, 2008, часть I, с. 147–150.
8. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 2003, vol. 31, pp. 3406–3415.

Transcription regulation of the chlL gene in Viridiplantae

K. V. Lopatovskaya, O. A. Zverkov, A. V. Seliverstov, V. A. Lyubetsky

In Chara as well as in many land plants the transcription regulation is proposed for an operon, which encodes chlorophyll synthesis enzymes. On the other hand, we proposed a regulation on processing level in some species.

KEYWORDS: transcription, regulation, multiple alignment.