



КОНФЕРЕНЦИЯ МОЛОДЫХ УЧЕНЫХ И СПЕЦИАЛИСТОВ

**Информационные  
ТЕХНОЛОГИИ И СИСТЕМЫ**  
**ИТиС'09**

СБОРНИК ТРУДОВ КОНФЕРЕНЦИИ

**ISBN 978-5-901158-11-1**

пос. д/о Бекасово,  
15 – 18 декабря 2009 г.

Информационные технологии и системы (ИТиС'09):  
сборник трудов конференции. – М.: ИППИ РАН, 2009. – 463 с.

Издание содержит труды ежегодной конференции молодых ученых и специалистов «Информационные технологии и системы» (ИТиС'09), традиционно организуемой Советом молодых ученых и специалистов ИППИ РАН.

Основная цель Конференции ИТиС'09 – дать возможность молодым ученым и специалистам различных подразделений ИППИ РАН, а также студентам, аспирантам и молодым ученым других институтов РАН, отраслевых институтов, университетов и вузов, познакомиться с коллегами и обменяться научными достижениями по основным для ИППИ РАН направлениям научной деятельности: теория передачи и защиты информации; математическая теория информации и управления, многокомпонентные случайные системы; информационно-коммуникационные технологии и их применение в сложных системах и сетях; информационные процессы в живых системах и биоинформатика; компьютерная лингвистика и моделирование естественного языка; высшая математика.

Все включенные в данный сборник работы прошли рецензирование и опубликованы в том виде, в котором они были представлены авторами, среди которых молодые ученые и специалисты ИППИ РАН, ИПУ РАН, ИККИ РАН, ИПМ РАН, ИПТМУ РАН, ИМБ РАН, ИОГен РАН, МГУ им. М.В. Ломоносова, МФТИ (ГУ), МАИ, НГУ, СПбГУ ИТМО, ЛГТУ, МПГУ, МТУСИ, РХТУ и др.

Труды Конференции могут представлять интерес для ученых, студентов и аспирантов, специализирующихся в областях науки, связанных с перечисленными выше научными направлениями.

Конференция проведена при финансовой поддержке ИППИ РАН, Отделения нанотехнологий и информационных технологий РАН, Целевой программы Президиума РАН «Поддержка молодых ученых 2009 г.» и РФФИ (грант 09-07-06810моб\_г)



**ISBN 978-5-901158-11-1**

© Коллектив авторов, 2009

© Учреждение Российской академии наук Институт проблем передачи информации им. А.А. Харкевича РАН, 2009

## Моделирование конкуренции РНК-полимераз: влияние нокаута сигма-субъединицы и температуры на экспрессию генов

О.А. Зверков, А.В. Селиверстов, Л.И. Рубанов, В.А. Любецкий  
ИППИ РАН  
lyubetsk@iitp.ru

### Аннотация

Разработаны алгоритм и программа для многопроцессорной вычислительной системы, моделирующие конкуренцию РНК-полимераз, транскрибирующих гены на разных цепях ДНК. Показано, что модель воспроизводит два основных экспериментально установленных феномена у растений: изменение экспрессии генов после нокаута определенной сигма-субъединицы холофермента РНК-полимеразы пластидного кодирования и аналогично – при изменении температуры клетки. Например, у ячменя *Hordeum vulgare* locus, включающий ген *psbA*, другие гены и ген *rpl23* с экзонем гена *rpl2*, содержит опероны на комплементарных цепях ДНК, и на них происходит конкуренция РНК-полимераз. При повышении температуры в изолированных хлоропластах наблюдается почти двукратное уменьшение числа транскриптов *psbA* и более чем двукратное увеличение числа транскриптов *rpl23-rpl2*, как в эксперименте, так и в нашей модели.

### 1. Введение

Пластиды – полуавтономные органеллы растений, которые обладают собственной транскрипционной системой. В пластидах растений и водорослей транскрипцию осуществляют РНК-полимеразы разных типов: одна-две фагового типа (NEP) и одна эубактериального типа (PEP): NEP – моносубъединичные полимеразы ядерного кодирования – связываются с NEP-промоторами, а PEP – многосубъединичные РНК-полимеразы пластидного кодирования – связываются с PEP-промоторами, причем инициация транскрипции требует участия одной из нескольких сигма-субъединиц. Каждой из сигма-субъединиц соответствуют в той или иной степени предпочитаемые PEP-промоторы и, тем самым, регулируется активность инициации транскрипции, осуществляемой PEP. Последовательности ДНК, кодирующие сигма-субъединицы у растений, описаны недавно, и, например, у *Arabidopsis thaliana* их определено шесть: Sig1-Sig6. Одни сигма-субъ-

единицы достаточно универсальные, например, Sig1-2, а другие – специфичные, например, Sig4-5 соответственно для генов *ndhF* и *psbD* [1]. Для определения специфичности Sig3 и Sig4 к промоторам проведены следующие эксперименты. Дикый тип *Arabidopsis thaliana* сравнивался с мутантом, в котором выполнен нокаут по Sig3 [2]. Для каждого гена подсчитаны усредненные частные sig3/WT уровня транскрипции мутантного типа к дикому типу и стандартные отклонения этих отношений. В мутантных растениях значительно изменяется содержание мРНК многих генов. Аналогично рассматривались нокауты по Sig4 [3] и влияние на уровни транскрипции изменения температуры клетки [4] и изменения уровня фитогормонов [5].

Исследование ответа на тепловой шок показало значительное различие в поведении хлоропластов в составе эукариотических клеток и в изолированном состоянии. Таким образом, изменение температуры запускает как внутренний механизм регуляции транскрипции (сразу после изменения температуры), так и внешний, связанный с поступлением в хлоропласты из цитоплазмы белков, кодируемых в ядре. Для моделирования поведения изолированных хлоропластов мы, в частности, используем данные о зависимости скорости элонгации РНК-полимеразы от температуры у кишечной палочки [6-7].

Консенсусы для NEP-промоторов разных типов известны (и в целом более изучены, чем PEP-промоторы для минорных сигма-субъединиц), но они, как и PEP-промоторы, определялись нами вместе с регуляциями соответствующих генов. Использовались данные о зависимости (от нуклеотидных замен) уровня связывания с промотором бактериального типа РНК-полимераз с субъединицами Sig1-3 [8]. Известны данные по Sig5-зависимому промотору перед геном *psbD* [9], по зависимости (от нуклеотидных замен) уровня связывания с промотором фагового типа RpoTr РНК-полимераз [10], по PEP- и NEP-зависимым промоторам пластид [11-13].

У зеленых водорослей из таксономической группы Chlorophyta положение генов на хромо-

сое быстро меняется. Это позволяет предполагать, что транскрипционная регуляция и конкуренция РНК-полимераз не играют важной роли для хлоропластов этих видов, а регуляция экспрессии идет в основном на уровне трансляции посредством кодируемых в ядре белков. В частности, это справедливо для хорошо изученной водоросли *Chlamydomonas reinhardtii* [14-15]. Например, у этой водоросли для *psbA* экспериментально показана именно такая регуляция [16]. С другой стороны, для водорослей и растений группы Streptophyta многие гены сохраняют позиционную связь на хромосоме, причем часто такие гены располагаются на комплементарных цепях ДНК, т.е. не образуют оперон. Это делает актуальным исследование роли конкуренции РНК-полимераз именно у Streptophyta. Некоторые гены хлоропластов у Streptophyta регулируются на уровне трансляции, что делает актуальным поиск и моделирование одновременно нескольких разных типов регуляции и определение вклада каждой из них.

## 2. Методы

В процессе транскрипции локуса ДНК могут одновременно участвовать много РНК-полимераз, которые связываются со своими промоторами и затем двигаются вдоль своей цели, вообще говоря, навстречу друг другу. При их столкновении на последовательности некоторого гена его транскрипция, вообще говоря, прерывается. Для каждой полимеразы этот процесс протекает асинхронно по отношению к другим и представляет собой композицию случайного (в первом приближении пуассоновского) и детерминированного процессов. Такую систему взаимодействующих процессов трудно исследовать аналитически (это – новый класс случайных процессов), поэтому мы построили (пока предварительную) модель монте-карловского типа.

Чтобы верно воспроизвести динамику реального процесса, важно было сохранить правильную очередность событий, из которых складывается конкуренция полимераз (т.е. очередность коллизий разного рода). Это не позволяет применить равномерное квантование времени из-за асинхронности процессов, описывающих поведение полимераз.

Вместо этого модель реализована как система с многими агентами, которые независимо действуют в пределах изучаемого локуса.

Часть таких агентов неподвижны – это гены, промоторы, сайты посадки белка-репрессора, терминаторы, внешние источники полимераз, служащие для компенсации искусственного выделения локуса из всей геномной последовательности. Их поведение выражается изменением внутрен-

них параметров (например, счетчика числа транскрипций гена) или возникновением/исчезновением самого этого агента (скажем, репрессора), либо порождением других агентов (например, для РЕР-промотора это появление на цепи ДНК нового холофермента в результате успешного связывания). Другие агенты, в дополнение к перечисленному, еще и меняют свое положение на ДНК – это холоферменты и полимеразы зубактериального (кор-ферменты) и фагового типов. Моделируя поведение каждого агента, начиная с момента запуска модели, мы определяем момент времени наступления следующего события с этим агентом. Например, таким событием может быть попытка связывания холофермента с промотором, сдвиг полимеразы на следующий нуклеотид, появление новой полимеразы, исчезновение полимеразы в результате коллизии или выхода за пределы локуса и т.д. Моменты наступления всех событий в модели организованы в единую очередь, которая обслуживается в хронологическом порядке.

Таким образом, в модели реализована схема неравномерного дискретного времени, отсчеты которого задаются возникающими событиями. В этом отношении модель построена по типу многозадачной операционной системы, управляемой событиями, а не системы квантования времени.

Благодаря отказу от равномерной временной сетки, удалось достичь высокой скорости моделирования (в среднем на 1-2 порядка быстрее физического процесса в клетке). Это важно для получения сопоставимых данных, поскольку биологические эксперименты длятся часами или сутками на культуре из многих сотен тысяч клеток, так что величины экспрессии генов усредняются естественным образом, а на модели приходится усреднять результаты по многочисленным траекториям моделирования, что зачастую заставляет применять высокопроизводительные кластеры.

Результатом каждой траектории моделирования являются значения числа транскрипций каждого гена из данного локуса; тем самым после усреднения и учета физического времени моделирования мы получаем оценку частоты транскрипции каждого гена. После варьирования неизвестных параметров (интенсивностей связывания промоторов и т.п.) в исследуемой области выбираются пары точек – кандидатов на нокаутированный и дикий тип, следующим образом.

Пусть локус образован генами  $g_1, \dots, g_n$ . Обозначим  $R_M(g_i)$  полученное в модели отношение частоты транскрипции гена  $g_i$  при нокауте к соответствующей частоте в диком типе. То же отношение, известное из лабораторного эксперимента (точнее, его среднее значение), обозначим

$R_E(g_i)$ , а наблюдавшийся в эксперименте разброс значений –  $R_D(g_i)$ . Пусть  $D(g_i) = |R_M(g_i) - R_E(g_i)|$ ,  $R_N(g_i) = D(g_i) / R_D(g_i)$ . Определим меру качества  $Q$  данного набора констант связывания как максимум  $Q = \max_{i=1, \dots, n} [R_N(g_i)]$ . Лучшим для данного локуса набором констант связывания считался набор с наименьшим значением  $Q$ .

### 3. Результаты

На основе прототипа компьютерной модели нами рассмотрены несколько локусов, в которых имеет место конкуренция РНК-полимераз. Из них ниже приведены результаты для трех локусов и для них получено хорошее согласие с экспериментами по нокаутам сигма-субъединиц и по изменению температуры клетки. Расчеты проводились на суперкомпьютере МВС-100К Межведомственного Суперкомпьютерного Центра РАН.

Первый локус (рис. 1а): **P-psbB-psbT-(T1-psbN-P\*)-psbH-petB-petD-(T2-rpoA-rps11-P^-rpl36-trnI-N)**, где P, P<sup>^</sup>, P\* – РЕР-промоторы, N – NЕР-промотор, в скобках пишутся объекты на комплементарной цепи. T1, T2 – возможные терминаторы, влияющие на процесс транскрипции. Здесь после нокаута Sig3 происходят сложные изменения интенсивностей транскрипции генов оперонов, которые были экспериментально определены у *Arabidopsis thaliana* [2]. Они хорошо воспроизводятся в нашей предварительной модели.

Второй локус (рис. 1б): **N1-N2-P-ycf1-(ndhF-P\*)-rpl32**, где некоторые РЕР- и NЕР-промоторы известны, а другие были определены нами. Здесь после нокаута по Sig4 подсчитывалось [2] отношение sig4/WT, которое также хорошо воспроизводится в нашей предварительной модели.

Третий локус (рис. 1в): *Hordeum vulgare* содержит две одинаковые копии локуса: **trnI-rpl23-rpl2-(trnH)-rps19**. В работе [4] экспериментально определено, как меняется суммарная интенсивность транскрипции этих генов и связанного с ними гена *psbA* при изменении температуры клетки.

Здесь также некоторые промоторы были определены нами, другие пока не учитывались, это предварительное исследование дало хорошее согласие с данными из [4].

Итак, получены следующие результаты. Обозначение промотора ниже используется и для указания константы связывания с ним.

**По 1-му локусу** в модели получен результат:  $P=0.35+0.45$ ,  $P^*=0.15+0.35$ ,  $P^{\wedge}=0.1+0$ ,  $N=0.35$ ,  $T1=0.2$ ,  $T2=0.25$  с качеством  $Q=1.65$ . Здесь и далее первые слагаемые указывают значения константы связывания холофермента или полимеразы фагового типа со своим промотором в случае нокаута, а вторые слагаемые – добавки к ним для случая дикого типа. В таблице 1 приведено сравнение изменения при нокауте частот транскрипции генов в эксперименте и в модели.

**По 2-му локусу** в модели получен результат:  $N2=0.26$ ,  $P=0.05+0.15$ ,  $P^*=0.35+0.20$  с качеством  $Q=1.32$ . В таблице 2 приведено сравнение изменения при нокауте частот транскрипции генов в эксперименте и в модели.

**Для 3-го локуса** место мутанта в эксперименте и модели занимает клетка в условиях температурного шока. Один из вариантов полученного в модели результата:  $trnI-56=1.4$ ,  $trnH-36=0.7$ ,  $psbA-79=0.3$ ,  $R1=20$ ,  $R2=70$ , где R1, R2 – скорости элонгации РНК-полимеразы в условиях низких и высоких температур соответственно. В таблице 3 приведено сравнение изменения при увеличении температуры частот транскрипции генов в эксперименте и в модели.

Статистически расхождения везде укладываются в известное правило «трех сигма».

### 4. Обсуждение

Хорошо известно, что при тепловом шоке изолированные хлоропласты ведут себя не так, как в составе клетки. Моделирование конкуренции РНК-полимераз позволяет объяснить быстрые изменения уровней транскрипции при тепловом шоке в изолированных хлоропластах, которые не подвергаются воздействию белков, кодируемых в

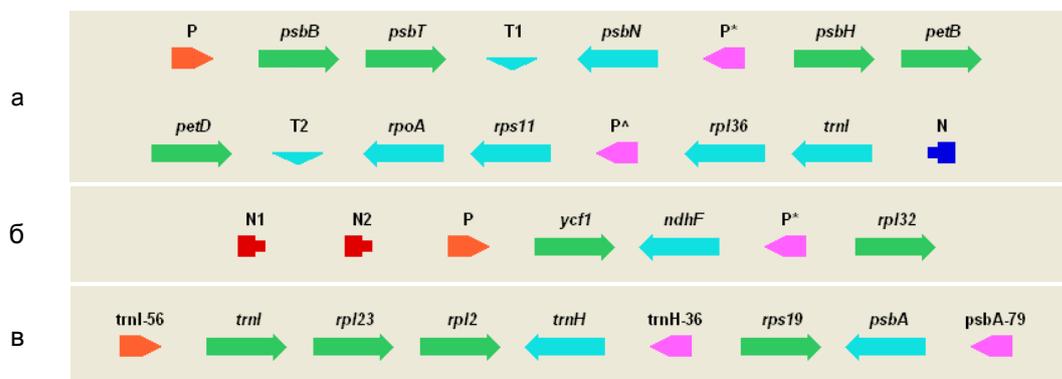


Рисунок 1. Схематическое изображение промоторов и генов трех локусов.

**Таблица 1. Сравнение изменения частот транскрипции генов в эксперименте и в модели по локусу 1.**

ген	эксперимент	модель
<i>psbB</i>	1.02 ± 0.36	0.60
<i>psbT</i>	0.98 ± 0.25	0.57
<i>psbN</i>	0.49 ± 0.46	1.12
<i>psbH</i>	1.31 ± 0.05	1.26
<i>pefB</i>	0.91 ± 0.15	1.02
<i>pefD</i>	0.92 ± 0.09	0.88
<i>rpoA</i>	0.94 ± 0.14	1.00
<i>rps11</i>	0.92 ± 0.33	1.00
<i>rp136</i>	0.88 ± 0.11	1.00
<i>trnI</i>	нет данных	1.00

**Таблица 2. Сравнение изменения частот транскрипции генов в эксперименте и в модели по локусу 2.**

ген	эксперимент	модель
<i>psbB</i>	1.02 ± 0.36	0.57
<i>psbT</i>	0.98 ± 0.25	0.64
<i>Rpl32</i>	1.52 ± 0.06	1.48

**Таблица 3. Сравнение изменения частот транскрипции генов в эксперименте и в модели по локусу 3.**

ген	эксперимент	модель
<i>rp123</i>	2.42 ± 0.27	2.29
<i>rp12</i>	2.42 ± 0.27	2.28
<i>psbA</i>	0.54 ± 0.01	0.55

ядре. Кроме того, поиск оптимальных параметров модели позволяет уточнить некоторые трудно-измеримые в прямых экспериментах параметры РНК-полимеразы, например, интенсивность связывания холофермента с промотором в зависимости от его нуклеотидного состава и типа сигма-субъединицы.

## 5. Список литературы

- [1] Е.А. Лысенко "Plant Sigma Factors and Their Role in Plastid Transcription" *Plant Cell Reports*, 2007, 26, 845-859.
- [2] J.-J. Favory, M. Kobayashi, K. Tanaka, G. Peltier, M. Kreis, J.-G. Valay, S. Lerbs-Mache "Specific function of a plastid sigma factor for *ndhF* gene transcription" *Nucleic Acids Research*, 2005, 33 (18), 5991-5999.
- [3] W. Zghidi, L. Merendino, A. Cottet, R. Mache, S. Lerbs-Mache "Nucleus-encoded plastid sigma factor SIG3 transcribes specifically the *psbN* gene in plastids" *Nucleic Acids Research*, 2007, 35(2), 455-464.
- [4] Я.О. Зубо, Е.А. Лысенко, А.Ю. Алейникова, В.В. Кузнецов, Н.Л. Пшибыгто «Изменение листовых

параметров как показатель смены функциональных типов степных растений вдоль градиента аридности» *Физиология растений*, 2008, 55(3), 323-331.

[5] Ya.O. Zubo, M.V. Yamburenko, S.Yu. Selivankina, F.M. Shakirova, A.M. Avalbaev, N.V. Kudryakova, N.K. Zubkova, K. Liere, O.N. Kulaeva, V.V. Kusnetsov, Th. Börner "Cytokinin stimulates chloroplast transcription in detached barley leaves" *Plant Physiology*, 2008.

[6] E.A. Abbondanzieri, J.W. Shaevitz, S.M. Block "Picocalorimetry of Transcription by RNA Polymerase" *Biophysical Journal: Biophysical Letters*, 2005, L61-L63.

[7] J. Ryals, R. Little, H. Bremer "Temperature dependence of RNA synthesis parameters in *Escherichia coli*." *Journal of Bacteriology*, 1982, 151(2), 879-887.

[8] 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" *European Journal of Biochemistry* 2003, 270(6), 1288-1300.

[9] P.H. Hoffer, D.A. Christopher "Structure and Blue-Light-Responsive Transcription of a Chloroplast *psbD* Promoter from *Arabidopsis thaliana*" *Plant Physiol.* 1997. 115, 213-222.

[10] 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" *The EMBO Journal*, 1999, 18(1), 249-257.

[11] M. Swiatecka-Hagenbruch, K. Liere, Th. Börner "High diversity of plastidial promoters in *Arabidopsis thaliana*" *Molecular Genetics and Genomics*, 2007, 277, 725-734.

[12] M. Swiatecka-Hagenbruch, C. Emanuel, B. Hedtke, K. Liere, Th. Börner "Impaired function of the phage-type RNA polymerase RpoTp in transcription of chloroplast genes is compensated by a second phage-type RNA polymerase" *Nucleic Acids Research*, 2008, 36(3), 785-792.

[13] P. Westhoff, R.G. Herrmann "Complex RNA maturation in chloroplasts" *European Journal of Biochemistry*, 1988, 171, 551-564.

[14] W. Zerges "Translation in chloroplasts" *Biochimie*, 2000, 82(6-7), 583-601.

[15] E.J. Stauber, M. Hippler "Chlamydomonas reinhardtii proteomics" *Plant physiology and biochemistry : PPB / Société française de physiologie végétale* 2004, 42(12), 989-1001.

[16] C.R. Hauser, N.W. Gillham, J.E. Boynton "Translational regulation of chloroplast genes. Proteins binding to the 5'-untranslated regions of chloroplast mRNAs in *Chlamydomonas reinhardtii*" *The Journal of biological chemistry*, 1996, 271(3), 1486-97.