

УДК: 575.852

## Эффективность элонгации генов дрожжей коррелирует с плотностью нуклеосомной упаковки в 5'-нетранслируемом районе

Матушкин Ю.Г.\* , Левицкий В.Г., Соколов В.С.,  
Лихошвай В.А., Орлов Ю.Л.

*Институт цитологии и генетики, Сибирское отделение Российской академии наук,  
Новосибирск, 630090, Россия*

**Аннотация.** Ранее был предложен индекс эффективности элонгации трансляции (индекс эффективности элонгации – ИЭЭ), позволяющий адекватно оценивать эффективность экспрессии генов в одноклеточных организмах на основе нуклеотидного контекста кодирующей части генов. Мы проанализировали связь между ИЭЭ и профилем формирования нуклеосом в промоторных районах на выборке генов двух видов дрожжей. Теоретические оценки потенциала формирования нуклеосом (ПФН) были получены с помощью метода Reson. Экспериментальная оценка расположения нуклеосом была получена из данных по высокопроизводительному секвенированию нуклеосомной ДНК в *S. cerevisiae*. Вычислялся коэффициент корреляции между двумя векторами: вектор значений ПФН для определенной позиции фазированных относительно старта трансляции (AUG) всех последовательностей и вектор значений ИЭЭ для всех последовательностей. Профили коэффициентов корреляции между ПФН и ИЭЭ были получены для (–600; +600) участков, относительно старта трансляции экстрагированных из GenBank последовательностей генов. Для *S. pombe*: для всех генов и для 15% высокоэкспрессирующихся генов (15% последовательностей с максимальным ИЭЭ) мы нашли участки негативной зависимости между ПФН и ИЭЭ. Для *S. cerevisiae*: для всех генов и для низкоэкспрессирующихся генов (15% последовательностей с минимальным ИЭЭ) мы нашли участки позитивной зависимости между ПФН и ИЭЭ. Выявленная зависимость между ПФН и ИЭЭ может быть объяснена отбором в процессе эволюции контекстных характеристик последовательностей направленным на оптимизацию экспрессии генов. В высокоэкспрессирующихся последовательностях (высокое значение ИЭЭ) инициация транскрипции должна быть облегчена, следовательно нуклеосомная упаковка в промоторной области не должна быть слишком плотной. Для низкоэкспрессирующихся последовательностей – наоборот.

**Ключевые слова:** эффективность трансляции, нуклеосомный потенциал, корреляции, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*.

### ВВЕДЕНИЕ

Анализ кодирующих районов генов показывает, что у различных организмов почти для любой аминокислоты синонимичные кодоны используются с разной частотой, даже если в результате формируется одинаковая структура белка. В работе [1] было

---

\*[mat@bionet.nsc.ru](mailto:mat@bionet.nsc.ru)

показано, что каждый геном относится к определенному типу и имеет свою стратегию кодирования, то есть выбор синонимичных кодонов среди геномов каждого типа сходный (т. н. «геномная гипотеза»). Также показано, что для *Escherichia coli* и *Saccharomyces cerevisiae* организм-специфичный выбор кодона связан с организм-специфичной популяцией изоакцепторных тРНК [2–5].

Для оценки неравномерности в использовании кодонов были предложены различные индексы, основанные на подсчете частоты кодонов, встречающихся в кодирующих частях генов. В статье [6] описывается простая и эффективная мера подсчета частоты использования синонимичных кодонов индекс адаптации кодонов. В *E. coli* и *S. cerevisiae* выбор синонимичного кодона в высокоэкспрессирующихся генах находится под строгим контролем естественного отбора, т.е., зная паттерн кодонов в этих генах, можно определить, какой из синонимичных кодонов наиболее эффективен для трансляции.

Однако, не только дифференциальное использованием кодонов является определяющим для эффективности экспрессии. Показано [7], что существует пять групп организмов в зависимости от факторов влияния на эффективность экспрессии генов на уровне трансляции – эти факторы: кодонный состав гена, наличие и распределение вторичных структур в мРНК, стабильность этих структур. Разработанный ИЭЭ позволяет достоверно классифицировать более 700 одноклеточных прокариот и ряд одноклеточных эукариотических организмов.

На уровне транскрипции эффективность экспрессии генов зависит, в том числе, от 5'- регуляторной области, в частности от локализации нуклеосом в ней. Контекстно-направленное позиционирование нуклеосомы, обеспечиваемое правильными взаимодействиями функциональных сайтов ДНК с негистоновыми белками, играет важную функциональную роль среди факторов, определяющих регулярность формирования нуклеосом. Наряду с компактизацией ДНК, нуклеосомы участвует в обеспечении доступа транскрипционных факторов к регуляторным районам гена, которые являются определяющими для активации экспрессии гена [8].

Для оценки распространенности механизмов оптимизации экспрессии генов в виде согласованности процессов на разных уровнях экспрессии мы исследовали корреляции плотности нуклеосомной упаковки в позициях 5'- нетранслируемых областей генов дрожжей видов *S. cerevisiae* и *S. pombe* со значением индекса эффективности элонгации соответствующих генов.

Проверяемая гипотеза заключается в следующем: для эффективной экспрессии генов необходимы согласованно оптимизированные процессы трансляции и транскрипции, в частности – инициации транскрипции и элонгации трансляции. Такая корреляция была найдена между ФНП в 5' нетранслируемых областях окрестности AUG кодона генов дрожжей видов *S. cerevisiae* и *S. pombe* со значением ИЭЭ соответствующих генов.

## МАТЕРИАЛЫ И МЕТОДЫ

Для *S. cerevisiae* и *S. pombe* последовательности 5'- нетранслируемых областей окрестности старт-кодона трансляции генов от –600 до +600 нуклеотида относительно AUG кодона экстрагировались из генных карт этих организмов, взятых из базы данных GenBank (<http://www.ncbi.nlm.nih.gov/>). Экстракция проводилась на основе первичной разметки, имеющейся в этих генных картах с вырезанием интронов. Мы будем называть генами все экстрагированные последовательности, хотя для некоторых функциональность не доказана. Не рассматривались гены с неопределенными нуклеотидами (N) или длиной менее 30 кодонов. Также не рассматривались последовательности, маркированные как псевдогены Анализ проводился для 5698

генов *S. cerevisiae* и 5088 генов *S. pombe*, а также для 15% генов этих выборок с максимальным и минимальным значениями индекса эффективности элонгации.

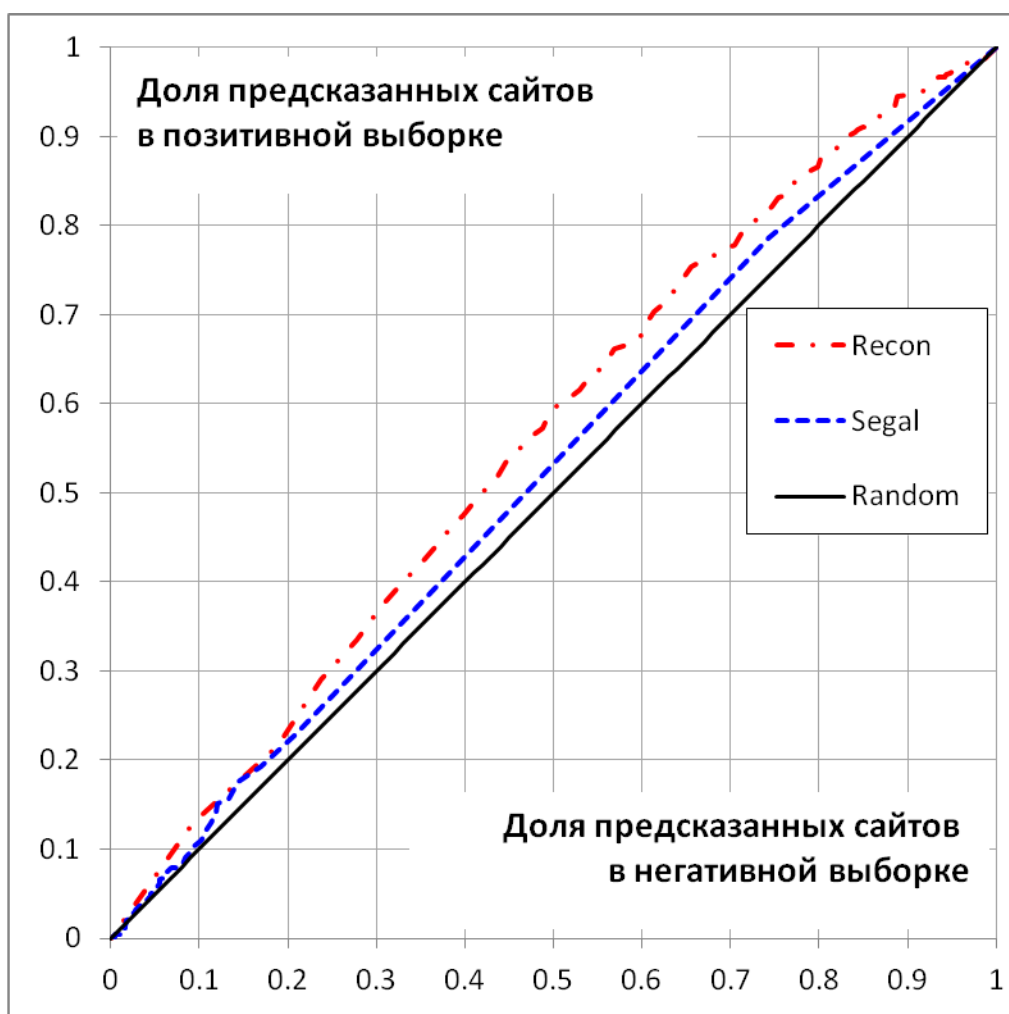
Для *S. cerevisiae* использовались экспериментальные данные по полногеномному секвенированию нуклеосомной ДНК [9] в различных условиях. Использовались данные только для 5418 генов. Профиль плотности расположения нуклеосом в референсном геноме дрожжей был построен из коротких секвенированных фрагментов («ридов» - от англ. read: прочтение). Данные секвенирования представлены в GEO: GSM351492. Только однозначно картируемые на геном дрожжей фрагменты, полученные с помощью технологии Иллюмина (Illumina Genome Analyzer II) были использованы при построении профиля плотности расположения нуклеосом. Для корректного компьютерного построения профиля каждый короткий фрагмент («рид») при картировании на хромосомы был увеличен до средней длины фрагмента, полученного при дроблении геномной ДНК. Чтобы построить профиль плотности расположения нуклеосом мы использовали исходные данные секвенирования для картирования нуклеосомной ДНК полученной *in vivo* при оптимальных условиях роста дрожжей (YPD). В итоге каждой нуклеотидной позиции в каждой хромосоме получила в соответствие значение плотности расположения нуклеосом – целое число соответствующее числу перекрывающихся фрагментов секвенирования, перекрывающих данную позицию.

Для двух исследуемых видов дрожжей был рассчитан индекс эффективности элонгации (ИЭЭ) с помощью программы EEI-Calculator (<http://wwwmgs.bionet.nsc.ru/cgi-bin/mgs/eei-calculator/index.pl>) [7]. Рассчитывается ИЭЭ по трем факторам: кодонному составу кодирующей последовательности, ее насыщенности инвертированными повторами, свободной энергии потенциальных шпилек. Для каждого организма находится оптимальная индивидуальная комбинация этих факторов. Соответственно, определяется пять эволюционных стратегий оптимизации трансляции во всех исследованных организмах: 1) Учитывается только кодонный состав; 2) Учитывается только количество и длина вторичных структур, но не их энергия; 3) Учитывается количество и длина локальных комплементарных структур и их энергия; 4) Учитываются и кодонный состав, и длина локальных комплементарных структур; 5) Учитываются и кодонный состав длина локальных комплементарных структур и их энергия.

Ранее были получены численные характеристики элонгации трансляции и распределение 759 одноклеточных организмов (702 бактерий, 52 архей, 5 эукариот) по этим 5-ти классам (неопубликованные данные). Показаны достоверные корреляции между значениями ИЭЭ и уровнем экспрессии по данным микрочиповых экспериментов в *S. cerevisiae* и *H. pylori*. Выходной файл этой программы содержит информацию о значениях индекса эффективности элонгации всех генов исследуемого организма, расположенных в порядке убывания. Для нашей работы использовалась та форма ИЭЭ, (из 5 возможных) которая наиболее адекватно распознает рибосомные гены как высокоэкспрессирующиеся [7].

Для 5'-района каждого гена был рассчитан нуклеосомный потенциал – функция, которая характеризует вероятность расположения нуклеосомы в заданном сайте последовательности. Значения этой функции вычисляются на основе частот динуклеотидов. Для расчета нуклеосомного потенциала использовалась программа RECON (<http://wwwmgs.bionet.nsc.ru/mgs/programs/recon>) [10]. Эта программа рассчитывает нуклеосомный потенциал (ПФН) в окне  $W = 140$  нт по всей последовательности длиной  $L$ . Значение потенциала приписывается позиции центра окна. При анализе одной последовательности получается профиль ПФН длины  $L - W + 1$ . Значения ПФН всегда меньше 1, а значение +1 соответствует наибольшей вероятности сайта формирования нуклеосомы.

Мы оценили точность распознавания метода Recon и сравнили ее с известным аналогом [11, 9] [http://genie.weizmann.ac.il/software/nucleo\\_exe.html](http://genie.weizmann.ac.il/software/nucleo_exe.html). Этот веб-интерфейс позволяет оценить ПФН как только по нуклеосомной ДНК [11], так и по ее сочетанию с линкерной ДНК [9]. Так как метод Recon был построен только по нуклеосомной ДНК, мы сравнили с методом [11]. Точность была оценена на позитивной и негативной контрольных выборках, содержащих по 1000 случайно отобранных последовательностей нуклеосомной и линкерной ДНК человека [12, 13], <http://www.hgc.jp/~ytanaka/assess2009/index.html>. Потенциальный сайт считался распознанным, если значение функции распознавания превосходило заданный порог. Для оценки точности была построена кривая ROC (Receiver Operating Characteristic): зависимость числа верных предсказаний в позитивной контрольной выборке от таковых в негативной при изменении порога функции распознавания. Значение площади под кривой (AUC, Area under curve) использовано для количественной оценки точности. Для произвольного распознавания это значение равно 0.5, для совершенного оно равно 1. Значения AUC для методов Recon и Segal равны 0.560 и 0.526, соответственно (рис. 1). Эти оценки находятся с хорошим соответствием с полученными ранее [13]. Следовательно, мы можем заключить, что использованный нами метод превосходит популярный аналог [11] по точности.



**Рис. 1.** Сравнение точностей распознавания методов Recon и Segal. Кривая зависимости доли предсказанных сайтов в позитивной контрольной выборке от соответствующей доли в негативной, кривая ROC (Receiver operating characteristic).

Для каждого гена был получен профиль из 1061 значения ПФН для 1061 позиции. Далее вычислялся коэффициент корреляции между двумя векторами: вектор значений ПФН для определенной позиции фазированных относительно старта трансляции (AUG) всех последовательностей и вектор значений ИЭЭ для всех последовательностей. Для полученных результатов были рассчитаны критические значения коэффициента корреляции при уровне значимости 0.95 с использованием критерия Фишера. Мы усредняли эти профили: для всех последовательностей, упорядоченных по значению ИЭЭ, для 15% имеющих максимальные значения ИЭЭ и для 15% имеющих минимальные значения.

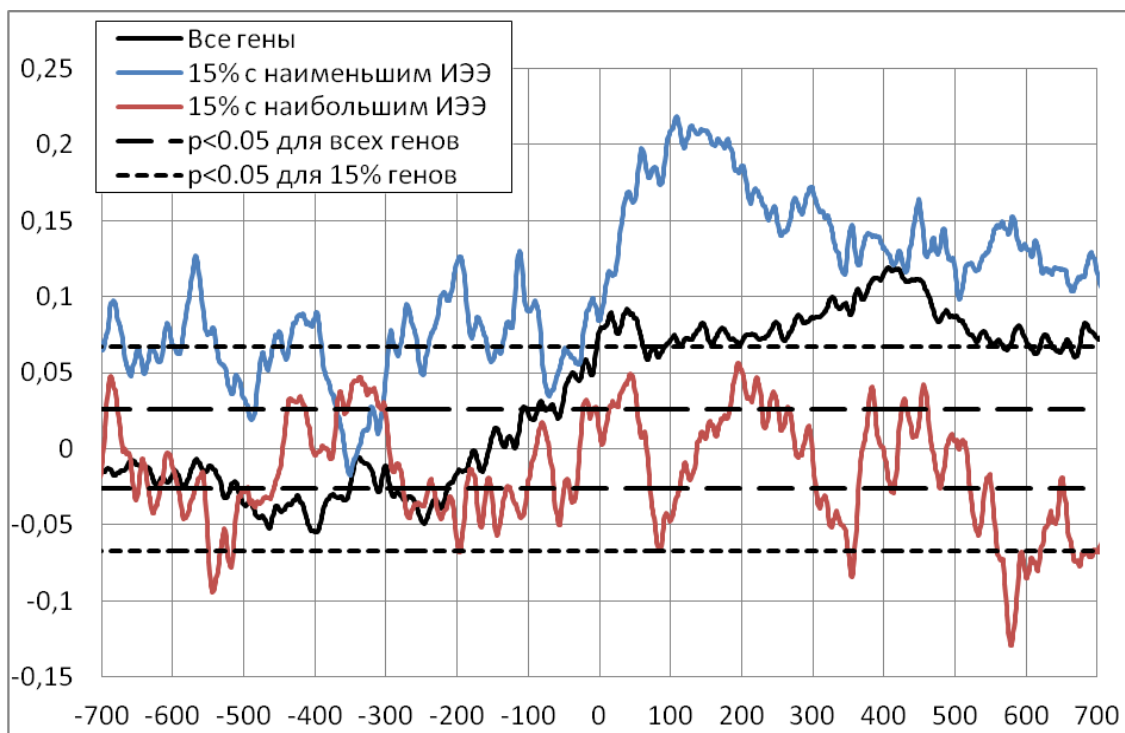
## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Корреляция между ПФН и ИЭЭ для *S. cerevisiae*

Как мы показали ранее [7] ИЭЭ хорошо работает для генов *S. cerevisiae* и *S. pombe*. Для *S. cerevisiae* коэффициент корреляции между ИЭЭ и экспериментально определенной эффективностью экспрессии составляет 0.8 при  $p < 0.00001$ .

Для *S. cerevisiae* и *S. pombe* были рассчитаны коэффициенты корреляции между двумя векторами: вектор значений ПФН для определенной позиции фазированных относительно старта трансляции (AUG) всех последовательностей и вектор значений ИЭЭ для всех последовательностей. Профили коэффициентов корреляции между ПФН и ИЭЭ были получены для (-600; +600) участков, относительно старта трансляции экстрагированных из GenBank последовательностей генов.

Будем далее для краткости называть последовательности с наименьшим ИЭЭ низкоэкспрессирующимися, а последовательности с наибольшим ИЭЭ – высокоэкспрессирующимися.



**Рис. 2.** Коэффициенты корреляции (ось ординат) между ИЭЭ и ПФН для *S. cerevisiae*: для всех последовательностей, для 15% высокоэкспрессирующихся по ИЭЭ и 15% низкоэкспрессирующихся по ИЭЭ. Позиция 0 соответствует положению нуклеотида А в старт-кодоне трансляции АТГ. Горизонтальные линии – зоны достоверности ( $p < 0.05$ ) по критерию Фишера. Каждая точка графика – усреднение по 10 точкам расчета.

Как видно из рис. 2, налицо достоверная положительная корреляция между ИЭЭ и НФП для низкоэкспрессирующихся последовательностей в интервале  $(-250, -120)$ , а также в  $(-590, -550)$  и в  $(-440, -390)$ . И достоверная и отрицательная – для высокоэкспрессирующихся последовательностей в интервале  $(-550, -520)$ . Т. е., чем ниже эффективность элонгации кодирующей части, тем крепче нуклеосомная упаковка в 5'-области. Высоко достоверная положительная корреляция между ИЭЭ и НФП для низкоэкспрессирующихся последовательностей в начале кодирующей части  $(0, +700)$ , видимо, объясняется тем, что не только инициация, но и элонгация транскрипции для низкоэкспрессирующихся дрожжей тормозится с помощью прочной нуклеосомной упаковки. Хотя, возможно, это следствие отбора кодонов в кодирующей части.

### Корреляция между ПФН и ИЭЭ для *S. pombe*

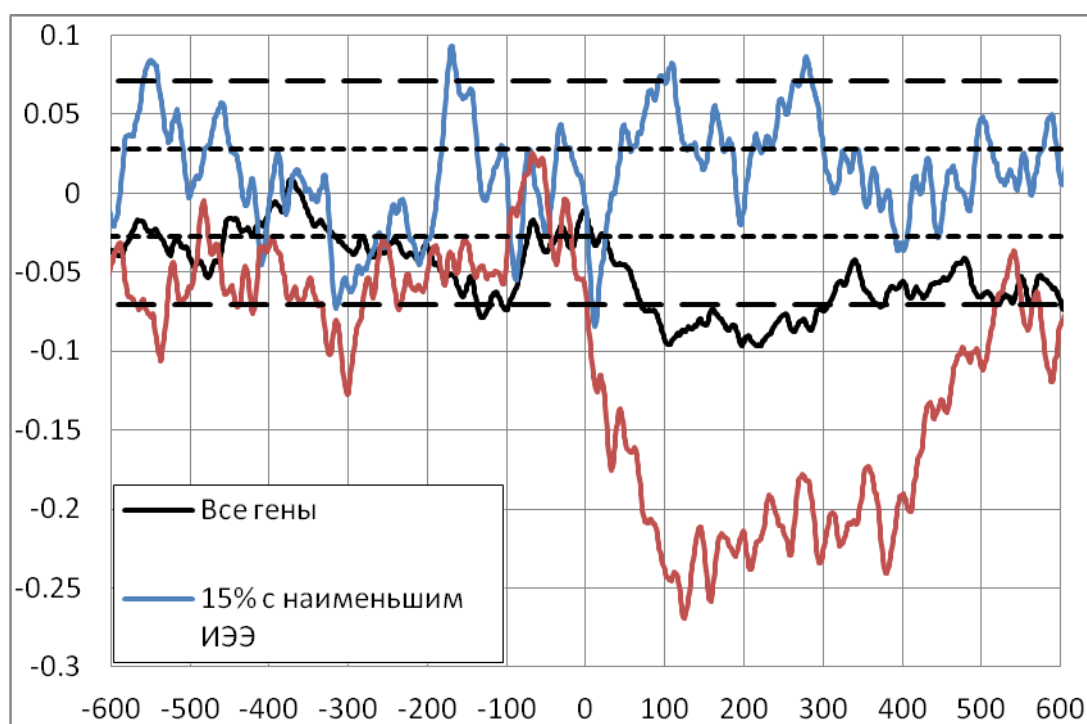


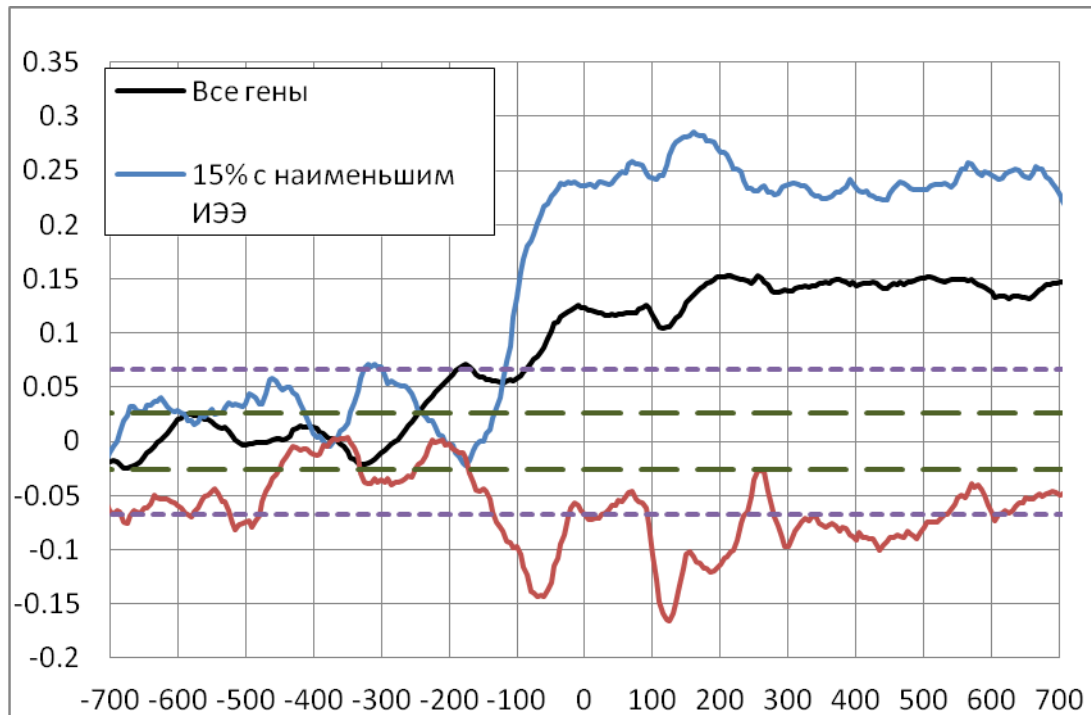
Рис. 3. Коэффициенты корреляции (ось ординат) между ИЭЭ и НФП для *S. pombe*. Обозначения такие же как на рис. 1.

Как видно на рис. 3, имеется достоверная отрицательная корреляция между ИЭЭ и НФП для высокоэкспрессирующихся последовательностей в интервале  $(-340, -280)$ , а также в  $(-550, -520)$ . Т. е., чем выше эффективность элонгации кодирующей части, тем слабее нуклеосомная упаковка в 5'-области. При этом есть достоверная положительная корреляция между ИЭЭ и НФП для низкоэкспрессирующихся последовательностей в интервалах  $(-560, -540)$  и  $(-180, -160)$ . Здесь мы видим в начале кодирующей части  $(0, +500)$  достоверную отрицательную корреляцию для высокоэкспрессирующихся последовательностей. Видимо, отбор шел не только на облегченную инициацию транскрипции, но и на облегченную элонгацию не только трансляции, но и транскрипции – для высокоэкспрессирующихся последовательностей. Хотя, возможно, это следствие отбора кодонов в кодирующей части.

Мы проверили корреляцию ИЭЭ генов *S. cerevisiae* с экспериментальными данными по расположению нуклеосом в геномной ДНК [9]. Короткие фрагменты ДНК, связанные с нуклеосомой, были наложены на геном. Мы отфильтровали полностью идентичные фрагменты и фрагменты, не имеющие однозначной локализации в геноме. Удлиненные до 146 нт фрагменты, что соответствует размеру нуклеосомы, формируют

наложенный друг на друга профиль в хромосомных координатах, который и является плотностью, или картой расположения нуклеосом. Таким образом, каждая позиция на хромосоме имеет в соответствии число фрагментов секвенированной нуклеосомной ДНК, перекрывающих данную позицию. Это число экспериментальной плотности расположения нуклеосом может быть сравнено с любым другим числом, характеризующим данную позицию, например на основе анализа окружающего контекста, либо также определенным в геноме экспериментально.

### Корреляция между ИЭЭ и экспериментальными данными по нуклеосомной упаковке для *S. cerevisiae*



**Рис. 4.** Коэффициенты корреляции (ось ординат) между ИЭЭ и экспериментальными данными по плотности нуклеосомной упаковки, полученной прямым секвенированием нуклеосомной ДНК для *S. cerevisiae*. Обозначения такие же, как на рис. 1.

Как видно из рис. 4, для низкоэкспрессирующихся последовательностей в интервале  $(-110, 0)$  корреляция (коэффициент корреляции до 0.25) достоверна и положительна. Для высокоэкспрессирующихся последовательностей корреляция (коэффициент корреляции до  $-0.15$ ) в этом же интервале достоверна и отрицательна.

Эти данные подтверждают проверяемую гипотезу, хотя по локализации областей значимых корреляций не вполне совпадают с найденными исходя из теоретических оценок. Отличие нуклеосомной локализации определенной по экспериментальным данным *in vivo* от теоретически предсказанных в промоторах, особенно в районе  $(-200; 0)$  было показано в статьях [9] и [14].

### Анализ нуклеотидного контекста промоторных районов для *S. cerevisiae* и *S. pombe*

Для объяснения обнаруженной связи между ИЭЭ и ПФН на уровне контекстных особенностей мы детально исследовали нуклеотидные последовательности 5'-районов генов *S. cerevisiae* и *S. pombe*.

Для *S. cerevisiae* обнаружено, что проксимальный промоторный район  $(-250; 0)$  генов обогащен polyA-трактами (длина тракта  $\geq 5$  и  $\leq 9$ ). В районе  $(-200; -100)$

обнаружено существенное повышение числа политрактов по сравнению с более дистальным районом промотора (-500;-250) (рис. 4, poly(A)<sub>5</sub>). Трактов подобного типа практически не наблюдается в районе ниже старта трансляции (0;+600).

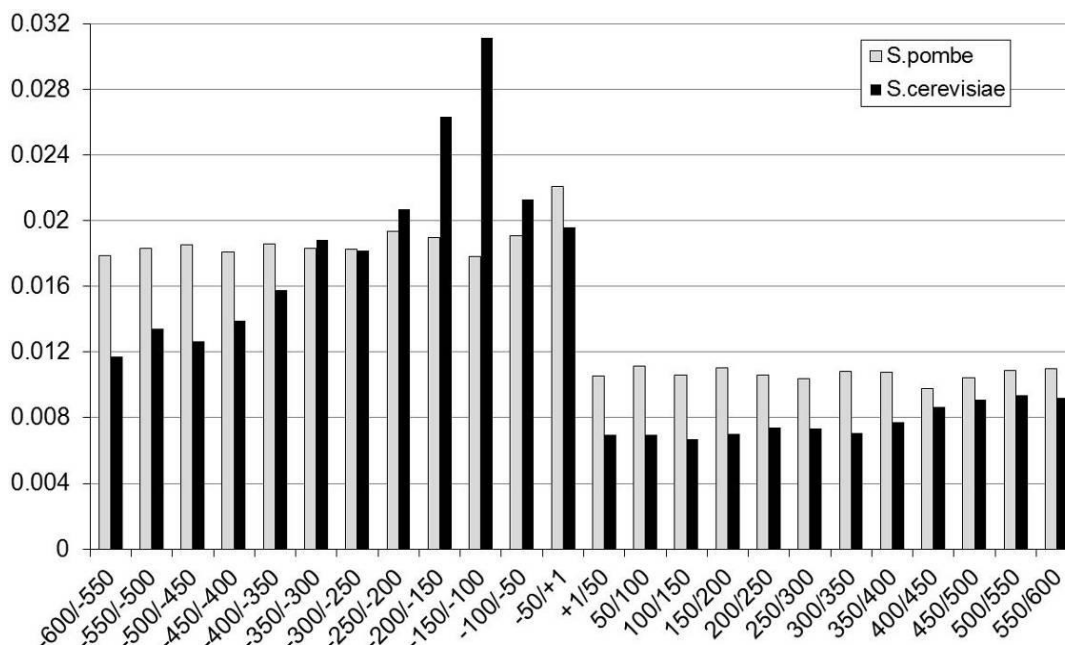
Мы также обнаружили, что плотность polyA-трактов в промоторном районе всех генов заметно превышает плотность подобных трактов в последовательностях сайтов формирования нуклеосом дрожжей [10] а также в случайных последовательностях (марковские цепи нулевого порядка). Эти случайные последовательности были сгенерированы нами на основе нуклеотидного состава как известных сайтов формирования нуклеосом дрожжей [15], так и исследованных нами проксимальных промоторных районов.

Анализ показал, что для промоторов *S. cerevisiae* между ИЭЭ и наличием polyA-трактов в последовательностях есть достоверная положительная корреляция, в частности для тракта poly(A)<sub>5</sub> в районе (-250; -50) коэффициент корреляции  $r = 0.065$  ( $p < 10^{-7}$ ). Аналогичный коэффициент для *S. pombe* статистически недостоверен.

Подробный анализ (табл. 1) показал, что по сравнению с ожиданием на основе динуклеотидного состава последовательностей для вида *S. cerevisiae* наблюдается умеренное значимое обогащение трактов в дистальном районе (-600;-500), умеренное обеднение трактов poly(A)<sub>5</sub> в районе ниже старта трансляции (+100;+200) и значительное обогащение в проксимальном районе выше старта (-200;+100). В аналогичных районах генов *S. pombe* значимых обогащения и обеднения не наблюдается.

Ранее для различных модельных организмов эукариот была показана связь между короткими олигонуклеотидами и экспериментально определенным положением нуклеосом [16], в частности негативная корреляция нуклеосом с содержанием polyA-трактами [17].

Это подтверждает найденную нами независимо корреляционную зависимость между ИЭЭ и polyA-трактами.



**Рис. 5.** Плотность политрактов A<sub>5</sub> во всех генах двух видов дрожжей. Ось X – (район относительно старта трансляции) – число трактов, деленное на число возможных положений пентануклеотида. Ось Y – отношение числа обнаруженных политрактов в пределах заданного района к общему числу возможных положений пентануклеотида в пределах района.



Как видно из рис. 5, распределение polyA-трактов у двух изученных видов дрожжей разное. Мы полагаем, что этой разницей и объясняется различие в корреляциях между ПФН и ИЭЭ для *S. cerevisiae* и *S. pombe*. Резкое различие в поведении корреляций от старт-кодона трансляции к 3'-концу мы объясняем тем, что *S. cerevisiae* принадлежит к первому классам оптимизации индекса эффективности элонгации – отбор шел только по кодонному составу, а *S. pombe* – к четвертому – отбор шел и по кодонному составу, и длине и числу локальных комплементарных структур [7]. Возможно, именно этим объясняется также отмеченное в [18] различие в механизмах позиционирования нуклеосом для этих двух родственных организмов.

**Таблица 1.** Обогащение/обеднение A<sub>5</sub>-трактов в различных районах генов *S. cerevisiae* и *S. pombe* в сравнении ожидаемых для выборок с таким же динуклеотидным составом. Обогащение/обеднение представлено как отношение пропорций [Obs<sub>poly</sub>/Obs<sub>tot</sub>] / [Exp<sub>poly</sub>/Exp<sub>tot</sub>], где Obs<sub>poly</sub> – число реальных последовательностей содержащих тракты в заданном районе, Obs<sub>tot</sub> – объем выборки реальных последовательностей, Exp<sub>poly</sub> и Exp<sub>tot</sub> – аналогичные числа для выборки, сгенерированной по реальной, с сохранением динуклеотидного состава. Объемы случайной и реальной выборок приняты равными, Exp<sub>tot</sub> = Obs<sub>tot</sub>. Так как для каждого вида сделано по три сравнения, то может быть введена поправка Бонферони, т. е. принят критический порог значимости  $p < 0.05 / 3 = 0.0167$

Локализация <sup>1</sup>	<i>S. cerevisiae</i> <sup>2</sup>	<i>S. pombe</i> <sup>2</sup>
(-600; -500)	0.455 / 0.432 ( $p < 7 \cdot 10^{-3}$ )	0.594 / 0.593 ( $p > 0.05$ )
(-200; -100)	0.678 / 0.592 ( $p < 9 \cdot 10^{-22}$ )	0.607 / 0.622 ( $p > 0.05$ )
(+100; +200)	0.359 / 0.376 ( $p < 0.02$ )	0.471 / 0.474 ( $p > 0.05$ )

<sup>1</sup> – локализация относительно сайта старта трансляции;

<sup>2</sup> – значимость согласно критерию Фишера  $\phi^*$ .

## ЗАКЛЮЧЕНИЕ

Таким образом, мы подтвердили предложенную гипотезу, впервые показав взаимосвязь между локализацией нуклеосом в промоторных районах и индексом эффективности элонгации. Индекс эффективности элонгации отрицательно коррелирует с потенциалом формирования нуклеосом. Отбор может идти как на уменьшение ПФН для высокоэкспрессирующихся последовательностей, так и на увеличение ПФН для низкоэкспрессирующихся. Вероятно, это связано с разными распределениями polyA-трактов в 5'-нетранслируемых районах разных видов дрожжей.

Для генов дрожжей обоих видов наблюдаются районы с достоверной отрицательной корреляцией индекса эффективности элонгации и нуклеосомного потенциала в 5'-нетранслируемых областях для генов с умеренным и высоким уровнем экспрессии (по ИЭЭ), с более выраженной корреляцией для вида *S. pombe*. Тогда как для генов с низким уровнем экспрессии наблюдаются районы с достоверной положительной корреляцией, с более выраженными значениями для генов вида *S. cerevisiae*. Кроме того, для дрожжей вида *S. cerevisiae* в кодирующей части генов с низким и средним уровнем экспрессии наблюдается сильная достоверная положительная корреляция между индексом эффективности элонгации и нуклеосомным потенциалом. Тогда как для дрожжей вида *S. pombe* в кодирующей части генов с высоким и средним уровнем экспрессии в этом же районе наблюдается сильная отрицательная корреляция нуклеосомного потенциала и индекса эффективности элонгации, что может быть связано с принадлежностью данных организмов к различным классам оптимизации индекса эффективности элонгации – первому и четвертому соответственно. К первому классу принадлежат организмы, в которых в процессе эволюции отбор шел по пути оптимизации кодонного состава кодирующей части генов (быстрые кодоны для высокоэкспрессирующихся последовательностей). Для организмов четвертого класса

отбор шел как по кодонному составу, так и по длине и числу локальных комплементарных структур в кодирующей части мРНК.

Работа поддержана грантами РФФИ №№ 13-04-00620, 12-07-00671; НШ 5278.2012, Программами фундаментальных исследований Президиума РАН № 6 и 28.

### СПИСОК ЛИТЕРАТУРЫ

1. Grantham R., Gautier C., Gouy M., Jacobzone M., Mercier R. *Nucl. Acids Res.* 1981. V. 9. P. 43–74.
2. Post L.E., Strycharz G.D., Nomura M., Lewis H., Dennis P.P. *Proc. Natl. Acad. Sci. USA.* 1979. V. 76. P. 1697–1701.
3. Post L.E., Nomura M. *J. Biol. Chem.* 1980. V. 255. P. 4660–4666.
4. Ikemura T. In: *Genetics and evolution of RNA polymerase, tRNA and ribosomes.* Eds. Osawa S., Ozeki H., Uchida H. & Yura T. Tokyo: University of Tokyo Press and Amsterdam: Elsevier/North Holland, 1980. 519–523 p.
5. Bennetzen J.L., Hall B.D. *J. Biol. Chem.* 1982. V. 257. P. 3026–3031.
6. Sharp P.M., Li W.H. *Nucleic Acid Res.* 1987. V. 15. P. 1281–1295.
7. Владимиров Н.В., Лихошвай В.А., Матушкин Ю.Г. Корреляция частот кодонов и потенциальных вторичных структур с эффективностью трансляции мРНК в одноклеточных организмах. *Молекулярная биология.* 2007. Т. 41. № 5. 843–850.
8. Richmond T.J., Widom J. In: *Chromatin Structure and Gene Expression.* Eds. Workman J.L. and Elgin S.C. 2nd edn. Oxford, UK: Oxford University Press, 2000. P. 1–23.
9. Kaplan N., Moore I.K., Fondufe-Mittendorf Y., Gossett A.J., Tillo D., Field Y., LeProust E.M., Hughes T.R., Lieb J.D., Widom J., Segal E. *Nature.* 2009. V. 458. № 7236. P. 362–366.
10. Levitsky V.G. *Nucl. Acids. Res.* 2004. V. 32. P. 346–349.
11. Segal E., Fondufe-Mittendorf Y., Chen L., Thastrom A., Field Y., Moore I.K., Wang J.P., Widom J. *Nature.* 2006. V. 442. № 7104. P. 772–778.
12. Schones D.E., Cui K., Cuddapah S., Roh T.Y., Barski A., Wang Z., Wei G., Zhao K. *Cell.* 2008. V. 132. P. 887–898.
13. Tanaka Y., Nakai K. *Genome Inform.* 2009. V. 23. № 1. P. 169–178.
14. Goh W.S., Orlov Y., Li J., Clarke N.D. *PLoS Comput Biol.* 2010. V. 6. № 1. P. e1000649.
15. Levitsky V.G., Katokhin A.V., Podkolodnaya O.A., Furman D.P., Kolchanov N.A. *Nucl. Acids. Res.* 2005. V. 33. P. 67–70.
16. Reynolds S.M., Bilmes J.A., Noble W.S. *PLoS Comput Biol.* 2010. V. 6. № 7. P. e1000834.
17. Segal E., Widom J. *Trends Genet.* 2009. V. 25. № 8. P. 335–343.
18. Lantermann A.B., Straub T., Strålfors A., Yuan G.-C., Ekwall K., Korber P. *Nat. Struct. Mol. Biol.* 2010. V. 17. № 2. P. 251–257.

Материал поступил в редакцию 28.05.2013, опубликован 17.06.2013.