МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В.ЛОМОНОСОВА

На правах рукописи

ЛИСИЦЫНА ОЛЬГА МИХАЙЛОВНА

ИНТЕГРАЦИЯ СИГНАЛОВ ЯДЕРНОЙ И ЯДРЫШКОВОЙ ЛОКАЛИЗАЦИИ В СОСТАВЕ ФУНКЦИОНАЛЬНЫХ ДОМЕНОВ БЕЛКОВ

03.01.08 – Биоинженерия

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата биологических наук

Москва – 2019

Работа выполнена в отделе электронной микроскопии Научно-исследовательского института физико-химической биологии имени А.Н. Белозерского Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова»

Научный руководитель – ***Шеваль Евгений Валерьевич***

доктор биологических наук

Официальные оппоненты – ***Северин Фёдор Фёдорович*** *– доктор биологических наук, заведующий лабораторией фотохимии биомембран отдела молекулярной энергетики Научно-исследовательского института физико-химической биологии имени А.Н.Белозерского Московского государственного университета имени М.В. Ломоносова*

Кантидзе Омар Леванович *- доктор биологических наук, заведующий лабораторией стабильности генома Федерального государственного бюджетного учреждения науки «Институт биологии гена Российской академии наук».*

Семашко Татьяна Александровна *– кандидат химических наук, научный сотрудник лаборатории протеомного анализа отдела молекулярной биологии и генетики Федерального государственного бюджетного учреждения «Федеральный научно-клинический центр физико-химической медицины Федерального медико-биологического агентства»*

Защита диссертации состоится 9 октября 2019 г. в 18 часов 00 минут на заседании диссертационного совета МГУ.03.04 Московского государственного университета имени М.В. Ломоносова по адресу: 119234, Москва, Ленинские горы, дом 1, стр. 73, Факультет биоинженерии и биоинформатики, ауд. 406.
E-mail: dissovet@belozersky.msu.ru
С диссертацией можно ознакомиться в отделе диссертаций научной библиотеки МГУ имени М.В. Ломоносова (Ломоносовский просп., д. 27) и на сайте ИАС «ИСТИНА»: http://istina.msu.ru/dissertations/230005070/.
Автореферат разослан «__» сентября 2019 года

Ученый секретарь диссертационного совета, кандидат химических наук И.В. Шаповалова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования.

Появление ограниченного двумембранной оболочкой ядра – одно из ключевых событий в эволюции жизни на Земле, приведшее к появлению эукариотической клетки. Ядерная оболочка является не только физическим барьером, но и системой, активно перемещающей макромолекулы между ядром и цитоплазмой. Изучение путей её возникновения и эволюции является одной из актуальных проблем биоинженерии.

Селективность транспорта биологических макромолекул между ядром и цитоплазмой обеспечивается сложными структурами ядерного импорта и экспорта, способными распознавать и переносить в нужном направлении определенные макромолекулы. В случае импорта белков в ядро специальные адаптерные молекулы – импортины – распознают короткие сигнальные последовательности белка – сигналы ядерной локализации (NLS). Все NLS делят на классические и неклассические. Классические NLS представляют собой одну или две коротких аминокислотных последовательности, обогащенные положительно заряженными аминокислотными остатками, которые взаимодействуют с импортином-α (Feldherr and Akin 1997). Остальные NLS, не удовлетворяющие известным консенсусным последовательностям классических NLS, и, как правило, прямо взаимодействующие с импортином-β, принято относить к неклассическим NLS. Наличие консенсусных последовательностей позволяет предсказывать классические NLS в белках.

NLS встречаются в аминокислотных последовательностях не только эукариотических белков. Так NLS были идентифицированы в белках VirD2 и VirE агробактерии *Agrobacterium tumefaciens*. Основной функцией данных белков является доставка ДНК агробактерии в ядра клеток растений (Ballas and Citovsky 1997). Аналогичными примерами могут являться прокариотические белки, ответственные за развитие патогенеза в эукариотических клетках-хозяинах: HsdM бактерии *Klebsiella pneumonia*, обеспечивающий метилирование ДНК в ядре зараженной клетки (Lee et al. 2009), и белок HP0425 *Helicobacter pylori*, обладающий активностью ДНКазы (Kim et al. 2016). Считается, что присутствие NLS в некоторых прокариотических белках может быть следствием их функции в ядре эукариотической клетки-хозяина. Однако NLS также обнаруживаются и в белках бактериофагов (Redrejo-Rodriguez et al. 2012). В вирусных и бактериальных белках NLS часто ассоциированы с нуклеотид-связывающими доменами (Cokol, Nair, and

Rost 2000; Lacasse and Lefebvre 1995; Lai et al. 2018; Redrejo-Rodriguez et al. 2012), что может предполагать их тесную эволюционную взаимосвязь.

Внутри ядра белки накапливаются в составе различных структур – ядерных телец. Накопление белков в составе не ограниченных мембранами ядерных телец может зависеть от наличия сигнальных последовательностей, таких как сигналы ядрышковой локализации (NoLS), которые способствуют аккумуляции белков в ядрышках. В составе эукариотических белков NLS часто перекрываются с NoLS (Boyne and Whitehouse 2006; Horke et al. 2004; Sheng, Lewis, and Chirico 2004), которые также весьма вариабельны по последовательности, но при этом обогащены положительно заряженными аминокислотными остатками лизина и аргинина. Как и NLS, NoLS могут располагаться внутри РНК-связывающих доменов (Schmidt-Zachmann and Nigg 1993), необходимых для взаимодействия с компонентами ядрышка.

Таким образом, в литературе имеются данные, свидетельствующие о тесной интеграции некоторых типов функциональных доменов белков, NLS и NoLS, что могло существенно повлиять на эволюцию этих двух типов сигнальных последовательностей.

Цель и задачи работы

Целью данной работы является выявление возможной эволюционной взаимосвязи сигнальных последовательностей (NLS и NoLS) с нуклеотид-/белок-связывающими доменами белков. Для достижения данной цели были поставлены следующие задачи:

1. Провести сравнительный анализ программ, предсказывающих NLS в белках.
2. Изучить возможность эволюционной интеграции NLS и функциональных доменов белков.
3. Изучить интеграцию NLS в составе нуклеотид-связывающих доменов белков прокариот.
4. Изучить интеграцию NLS в составе доменов цитоплазматических белков.
5. Выявить механизм интеграции NLS и NoLS.

Научная новизна работы

На основании анализа последовательностей NLS и белков, в состав которых они входят, предложена гипотеза происхождения этого типа сигналов, как функциональной части нуклеотид/белок-связывающего домена, распознаваемой

системой импортинов для транспорта белка в ядро. На примере прокариотических белков с нуклеотид-связывающими доменами приведены доказательства в пользу вышеописанной гипотезы: 1) идентифицированы прокариотические белки с нуклеотид-связывающими доменами, способные проникать через ядерный поровый комплекс посредством активного, импортин-зависимого транспорта; 2) в составе данных прокариотических белков картированны NLS. На примере тропонина I показана взаимосвязь сигнальных последовательностей и доменов, участвующих в белок-белковых взаимодействиях: в отсутствии факторов, индуцирующих образование фибрилл, аминокислотные последовательности актин- и тропонин С-связывающих доменов могут играть роль сигналов ядерной локализации, релокализуя цитоплазматический тропонин I в ядро. Среди программ, предсказывающих сигналы ядерной локализации, выявлены наиболее эффективные (обладающие бóльшей предсказательной силой). Для последовательностей сигналов ядрышковой локализации, которые довольно часто перекрываются с сигналами ядерной локализации, выявлен механизм накопления белков в ядрышке.

Практическая значимость исследования

Знания об эволюционной взаимосвязи NLS/NoLS с функциональными доменами белка могут быть использованы при разработке методов направленного транспорта макромолекул. В частности, это может иметь значение для модификации уже существующих противоопухолевых препаратов с целью улучшения их проникновения в ядро, увеличения их аффинности к активируемым участкам ДНК или уменьшения их взаимодействиях с «нежелательными партнерами» в ходе терапии. Данные проведенного исследования используются при чтении курса «Клеточная биология» для студентов факультета биоинженерии и биоинформатики МГУ.

Положения, выносимые на защиту

1. Функционально активные NLS могут быть эволюционно интегрированы в составе различных участков белков, обогащенных положительно заряженными аминокислотными остатками. При этом участок белка может функционировать в качестве NLS вне зависимости от необходимости и/или возможности для белка накапливаться в ядре.
2. Обогащенность положительно заряженными аминокислотными остатками некоторых NLS может быть причиной функционирования этих NLS в качестве

NoLS – сигнальных последовательностей, обеспечивающих накопление белков в ядрышке за счет электростатического взаимодействия белков с компонентами ядрышка.

Апробация работы

Работа докладывалась на семинаре отдела электронной микроскопии НИИ физико-химической биологии имени А.Н. Белозерского МГУ. Материалы работы были представлены на 22-ом, 25-ом и 26-ой симпозиумах Вильгельма Бернхарда по исследованию клеточного ядра, на VII международной конференции «From molecular to cellular events in human pathologies», на Международном молодежном научном форуме «Ломоносов-2018», на Всероссийском симпозиуме и школе-конференции для молодых ученых по биологии клетки в культуре, на I всероссийской конференции «Внутриклеточная сигнализация, транспорт, цитоскелет», на XVII российском симпозиуме по растровой электронной микроскопии и аналитическим методам исследования твердых тел, на 50-ой международной ежегодной конференции «American Society for Cell Biology», а также на XXIII российской конференции по электронной микроскопии.

Публикации

По материалам диссертации опубликовано 14 печатных работ (4 статьи в рецензируемых журналах и 10 тезисов докладов).

Личный вклад автора в проведении исследования

Основные результаты работы были получены автором. Личный вклад заключается в анализе данных литературы, планировании и проведении экспериментов, а также в обработке и анализе полученных данных, подготовке публикаций.

Структура и объем диссертации

Диссертация изложена на 337 страницах и состоит из следующих разделов: список сокращений, введение, обзор литературы, материалы и методы, результаты, заключение, выводы, список цитируемой литературы и приложения. Список литературы включает 204 источника. Диссертация содержит 20 таблиц и 36 рисунков, из них 12 таблиц входят в состав раздела «Приложения».

СОДЕРЖАНИЕ РАБОТЫ

Материалы и методы исследования

Геномная прокариотическая ДНК. Для исследований были взяты образцы геномной ДНК *Anabaena sp* и *Synechococcus sp* (любезно предоставленные О.А. Кокшаровой), *Vibrio harveyi* (любезно предоставленная Ю.В. Берцовой и А.В. Богачевым).

Культуры эукариотических клеток. В исследовании использовались следующие культуры эукариотических клеток HeLa, C2C12, U2OS, SK-N-SH, L959, клетки *Drosophila* S2 и фибробласты XL-2 *Xenopus laevis*.

Культуры прокариотических клеток. В качестве матрицы для получения целевых генов прокариотических белков использовали культуру клеток *Thermococcus sibiricus* (любезно предоставленную Е.А. Бонч-Осмоловской) и штамм *Escherichia coli* DH5α. Наработку плазмидной ДНК проводили в штамме *E. col*i DH5α. Выделение плазмидной ДНК из *E. coli* проводилось согласно протоколу с использованием коммерческого набора ThermoScientific (для трансфекции) либо Qiagen (для секвенирования).

Трансфекция. Трансфекцию производили непосредственно перед каждым экспериментом с использованием препарата Lipofectamin 2000 или Turbofect в соответствии с инструкцией производителя.

Иммуногистохимия. Кусочки ткани сердечной мышцы мыши фиксировали в 3,7% формальдегиде и замораживали. Из замороженного материала изготавливали криосрезы. Окрашивание проводили с использованием специфических антител MF4 (HyTest) к тропонину I с последующей детекцией посредством $F(ab)_2$-фрагментов с флуоресцентной меткой Alexa 555. Полученные препараты анализировали с использованием конфокального микроскопа Nikon C2, объектив 60x Plan Apo.

Получение химерных конструктов с EGFP. Для получения химерных конструктов с EGFP праймеры конструировались с помощью программы SerialCloner. ПЦР проводили с использованием ДНК полимеразы Phusion High-Fidelity (ThermoFisher Scientific). После проведения ПЦР продукты реакции очищали согласно протоколу PCR Purification (QIAquick, QIAGEN). Лигирование продуктов ПЦР осуществляли в вектор pEGFP-C1/N1, для этого вектор и продукты ПЦР обрабатывали эндонуклеазами рестрикции, после чего проводили их очистку выделением из 1% агарозного геля по протоколу Gene JET Gel Extraction Kit (ThermoScientific).

Лигирование осуществляли посредством T4 ДНК-лигазы (ThermoScientific) согласно инструкции производителя.

Сайт-направленный мутагенез. Получение мутантов прокариотических белков, в которых все положительно заряженные аминокислотные остатки были заменены на аланины, проводили в два этапа. На первом этапе на матрице гена целевого белка дикого типа проводили две параллельные ПЦР с использованием двух пар праймеров фланкирующего (реверс-праймер или форвард-праймер без внесения мутации) и внутреннего, содержащего мутации в нужном регионе. Затем ПЦР фрагменты, содержащие мутации, использовали в качестве матрицы для проведения ПЦР с использованием фланкирующих ген праймеров. Проведение ПЦР, очистка, рестрикция и лигирования продуктов ПЦР осуществлялась аналогично получению химерных конструктов с EGFP.

Конфокальная микроскопия. После трансфекции плазмидами клетки HeLa, экспрессирующие химерные конструкты EGFP с целевыми белками, анализировались посредствам конфокального микроскопа Nikon C2. Оценка интенсивности флуоресценции проводилась с использованием программного обеспечения ImageJ, с последующим анализом данных посредством GraphPad Prism 6.

Базы данных. Для биоинформатического анализа NLS и NoLS использовались базы данных Uniprot и NLSdb.

Предсказание NLS. Для предсказания NLS в составе аминокислотных последовательностей белков использовали следующие программы PSORT II, NucPred, NLSMapper, NLStradamus, NucImport и seqNLS.

Результаты и обсуждения

Сравнительный анализ программ для предсказания NLS

В нашей работе мы использовали не только NLS, последовательности которых были выявлены экспериментально, но и предсказанные с помощью различных программ. Это потребовало проведения на первом этапе сравнительного анализа существующих программ для предсказания NLS.

Для выбора наиболее подходящего алгоритма, предсказывающего NLS, мы проанализировали шесть программ: PSORT II, NucPred, NLSMapper, NLStradamus, NucImport и seqNLS. Оценку данных программ проводили с использованием выборки из 155 экспериментально определенных NLS из 128 человеческих белков, а также выборки из 155 неядерных белков (трансмембранных) в качестве отрицательного контроля.

Используя вышеописанные выборки белков с NLS и трансмембранных белков, мы оценили предсказательную способность каждой из программ по двум основным критериям: доле истинно положительных результатов (True Positive Rate, TPR) и доле ложноположительных результатов (False Positive Rate, FPR). В случае программ seqNLS, NLSMapper и NLStradamus, позволяющих варьировать минимально-допустимый уровень предсказания вероятности NLS, были получены ROC-кривые (Receiver Operating Characteristic) – графики, позволяющие оценить качество алгоритма предсказания, отображающие соотношение между количеством TPR и количеством FPR при варьировании значений минимально-допустимого уровня вероятности предсказать NLS. Для остальных было получено только по одному значению TPR и FPR.

Оценка предсказательной способности показала, что наиболее достоверно предсказывают NLS в последовательностях белков программы NucPred, NLStradamus NLSMapper и seqNLS. Программа и PSORTII согласно полученным значениям TPR и FPR предсказывает NLS чуть хуже.

Для оценки взаимосвязи между предсказанием и реальным результатом для каждой программы также был рассчитан коэффициент корреляции Мэтьюса (ККМ). Наибольший ККМ (~0,3) был получен для программ NucPred, seqNLS, NLStradamus и NLSMapper.

Важно отметить, что даже в наилучших случаях количество истинно положительных результатов, полученных программами NucPred, NLStradamus, NLSMapper и seqNLS, составляет только около 45% от числа экспериментальных NLS в выборке. Следовательно, любые предсказания NLS нуждаются в строгой экспериментальной проверке.

Коэволюция NLS и окружающего региона

Для изучения эволюционного происхождения NLS была составлена выборка из 496 белков человека, в составе которых было аннотировано 592 NLS согласно базам данных UniProt и NLSdb. 234 NLS из данной выборки были определены экспериментально, тогда как 358 последовательностей NLS были аннотированы *in silico.*

Каждый NLS из данной выборки был проанализирован на перекрытие с аннотированными доменами белков: 26% NLS из выборки перекрывались с доменами, участвующими в белок-белковых взаимодействиях, 19% NLS с нуклеотид-

связывающими доменами (77% из них - ДНК-связывающие домены, 20% - РНК-связывающие, и 3% - связывающие моно-/динуклеотиды) (рис. 1А).

Высокий процент NLS, входящих в состав какого-либо домена, может быть следствием функциональной и/или эволюционной взаимосвязи между NLS и самим доменом. Для проверки данной гипотезы мы проанализировали степень консервативности NLS при сравнении аминокислотных последовательностей белков человека с ортологами из протеомов хордовых: *Branchiostoma floridae, Danio rerio, Xenopus laevis, Pelodiscus sinensis* и *Gallus gallus.* Сравнительный анализ рассчитанных значений информационного содержания (Information content, IC) множественных выравниваний ортологичных белков, полученных для трех групп NLS, показал, что NLS, входящие в состав аннотированных доменов более консервативны (имеют бóльшие значения IC) по сравнению с вне-доменными NLS (дисперсионный анализ с поправкой Бонферрони для множественного сравнения, $p < 0,05$) (рис. 1Б).

Высокий уровень консервативности NLS, интегрированных в состав нуклеотид-/белок-связывающих доменов, по сравнению с вне-доменными NLS может предполагать эволюционную взаимосвязь NLS и окружающего его региона, т.е. консервативность NLS может зависеть от консервативности его ближайшего окружения (домена/вне-доменного региона). Для проверки данной гипотезы при выравнивании аминокислотных последовательностей человеческих белков с NLS с ортологами из пяти протеомов хордовых была оценена величина информационного содержания NLS относительно величины информационного содержания окружающего его региона (относительное IC NLS).

На рисунке 1В видно, что более половины NLS, входящих в состав нуклеотид-связывающих доменов, имеют ту же величину IC, что и окружающий их домен (для 51% NLS отношение IC сигнала к IC региона находится в интервале 0,9-1,1). Таким образом, эволюция NLS, интегрированных в нуклеотид-связывающие домены, может зависеть от эволюции самих доменов. В случае NLS, аннотированных в составе доменов, участвующих в белок-белковых взаимодействиях, для 39% NLS характерна та же степень консервативности, что и окружающего их региона. Для вне-доменных NLS относительное IC практически равномерно распределено на всех выбранных интервалах, что свидетельствует о низкой коэволюции вне-доменного NLS и окружающего его региона.

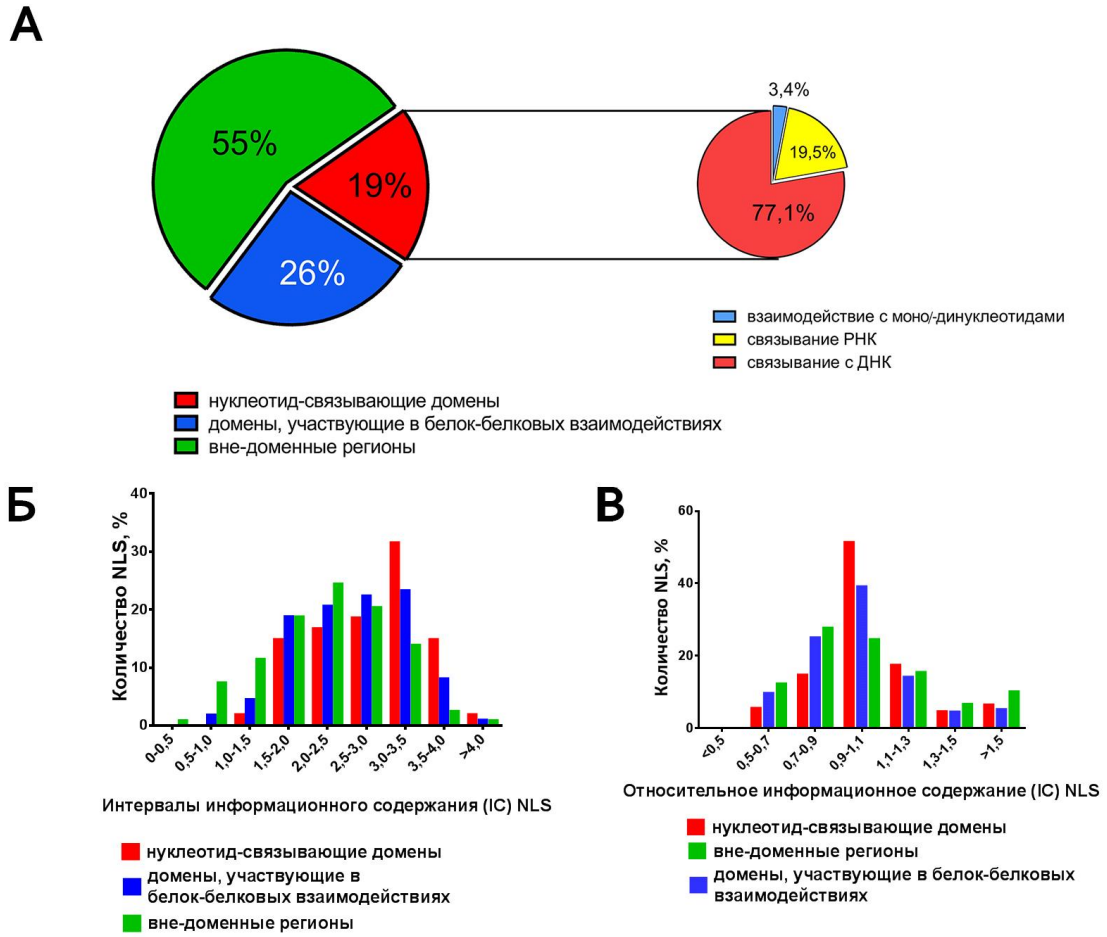

Рис. 1. Биоинформатический анализ NLS. А - Принадлежность NLS какому-либо типу доменов. **Б** - IC NLS, определенное при множественном выравнивании аминокислотных последовательностей ортологов белков человека, *B. floridae, D. rerio, X. laevis, P. sinensis* и *G. Gallus.* **В** - Относительное IC NLS, определенное при множественном выравнивании аминокислотных последовательностей ортологов белков человека, *B. floridae, D. rerio, X. laevis, P. sinensis* и *G. gallus*.

Полученные данные свидетельствуют о том, что NLS, интегрированные в состав функциональных доменов, в отличие от вне-доменных NLS, часто обладают той же степенью консервативности, что и окружающий их участок белка. Так как функциональные домены белка находятся под меньшим эволюционным давлением, чем вариабельные вне-доменные регионы (Takahata and Kimura 1994), то и NLS, входящие в их состав, имеют бóльшую степень консервативности по сравнению с вне-доменными NLS (рис. 1Б). Таким образом, наблюдается ко-эволюция функционального домена и интегрированного в него NLS. Следовательно, можно предположить, что исходно NLS представляли собой часть домена, которая стала выполнять дополнительную функцию транслокации молекулы белка в ядро. Вполне вероятно, что NLS могли предсуществовать в последовательностях белков еще до

формирования ядра, например, в составе нуклеотид-связывающих доменов, положительно заряженные остатки которых, как правило, неспецифично взаимодействуют с сахаро-фосфатным остовом ДНК/РНК. После формирования ядерной оболочки наличие обогащенных положительно заряженными аминокислотными остатками доменов могло стать удобным признаком для отбора ядерных белков при прохождении через ядерный поровый комплекс.

NLS интегрированы внутри функциональных доменов прокариотических белков

Интеграция NLS в составе функциональных доменов белков была нами обнаружена в ходе проведения биоинформатического анализа. Прямым образом верифицировать эту гипотезу экспериментально не представляется возможным, однако некоторые следствия этой гипотезы могут быть проверены экспериментально.

Нуклеотид-связывающие домены присутствуют не только в составе белков эукариот, но и в белках прокариот. В последнем случае отбор принципиально не мог способствовать появлению и стабилизации NLS, так как в отсутствии ядра такая функция существовать не может. Однако такие домены обогащены положительно заряженными аминокислотными остатками ровно так же, как и их эукариотические аналоги. При этом совершенно не важно, относится ли прокариотический организм к предкам эукариот (скажем, к археям) или нет. Любой домен может потенциально содержать участок, который способен работать в качестве NLS. Если это предположение верно, то после экспрессии такого прокариотического белка в клетках эукариот, будет наблюдаться накопление белка в ядре (т.е. в белке есть NLS).

Для проверки описанного выше предположения из 5 различных прокариотических организмов (*A. sp, S. sp, E. coli, V. harveyi* и *T. sibiricus*), были выбраны 12 белков, взаимодействующих с ДНК или РНК (согласно базе данных UniProt). Мы отбирали белки с молекулярной массой более 40 кДа, т.е. не способные эффективно перемещаться в ядро за счет диффузии через ядерные поровые комплексы. Посредством программы NLSMapper (Kosugi et al. 2009) в каждом белке были предсказаны NLS, большинство из которых входило в состав РНК- или ДНК-связывающего домена. Для контрольной выборки было выбрано 15 белков с молекулярным весом более 40 кДа, в составе которых не предсказывались NLS.

Для анализа локализации прокариотических белков внутри эукариотической клетки были сделаны химерные конструкции данных белков с зеленым флуоресцирующим белком EGFP. На рисунке 2А приведены репрезентативные

фотографии, иллюстрирующие локализацию выбранных 12-ти прокариотических белков, слитых с EGFP, в клетках HeLa. Степень накопления белков в ядре различна, но в случае шести белков визуально концентрация белка выше в ядре, чем в цитоплазме. Следует отметить, что белок SigA1 из *S. sp* не только накапливался в ядре, но и в ядрышке.

Определение степени накопления химерного белка в ядре проводилось с использованием конфокальной микроскопии живых клеток. На фотографиях измеряли среднюю флуоресценцию нуклеоплазмы (F_{nuc}) и цитоплазмы (F_{cyt}). Отношение F_{nuc}/F_{cyt} позволяло оценить накопление белка в ядре – чем оно выше, тем эффективнее белок накапливается в ядре.

Для установки порога мы использовали отношение F_{nuc}/F_{cyt} для EGFP, которое составляет 1,16. В основной выборке 6 прокариотических белков из 12 демонстрировали способностью накапливаться в ядре эукариотической клетки (рис. 2Б). При этом не было выявлено какой-либо корреляции между степенью накопления белка в ядре и его молекулярным весом (коэффициент корреляции Пирсона 0,13), что косвенно свидетельствует об отсутствии вклада диффузии в это накопление (диффузия пропорциональная размеру). В контрольной выборке только один белок (CasA) накапливался в ядре (F_{nuc}/F_{cyt} 1,71±0,25) (рис. 2В).

Сложно объяснить отсутствие накопления в ядре половины белков из основной экспериментальной выборки. Одним из возможных объяснений может быть ошибочность предсказания NLS или инактивация NLS, например, в результате фосфорилирования, метилирования или ацетилирования (Cao et al. 2017; Dormann et al. 2012; Hou et al. 2018; Li et al. 2018; von Morgen et al. 2018; Nakada, Hirano, and Matsuura 2017). Также нельзя исключать вероятность наличия в этих белках сигналов экспорта из ядра (NES), которые могли бы приводить к экспорту импортированных белков. Таким образом невозможно точно установить причину без проведения отдельного исследования, что выходило за рамки настоящей работы.

Тем не менее, 6 из 12 прокариотических белков могут в различной степени накапливаться в ядре эукариотической клетки. Это позволяет предполагать наличие NLS в этих прокариотических белках. Но как было показано выше, точность предсказания NLS невелика, поэтому мы провели более детальный анализ участков, потенциально вовлеченных в накопление белков в ядре.

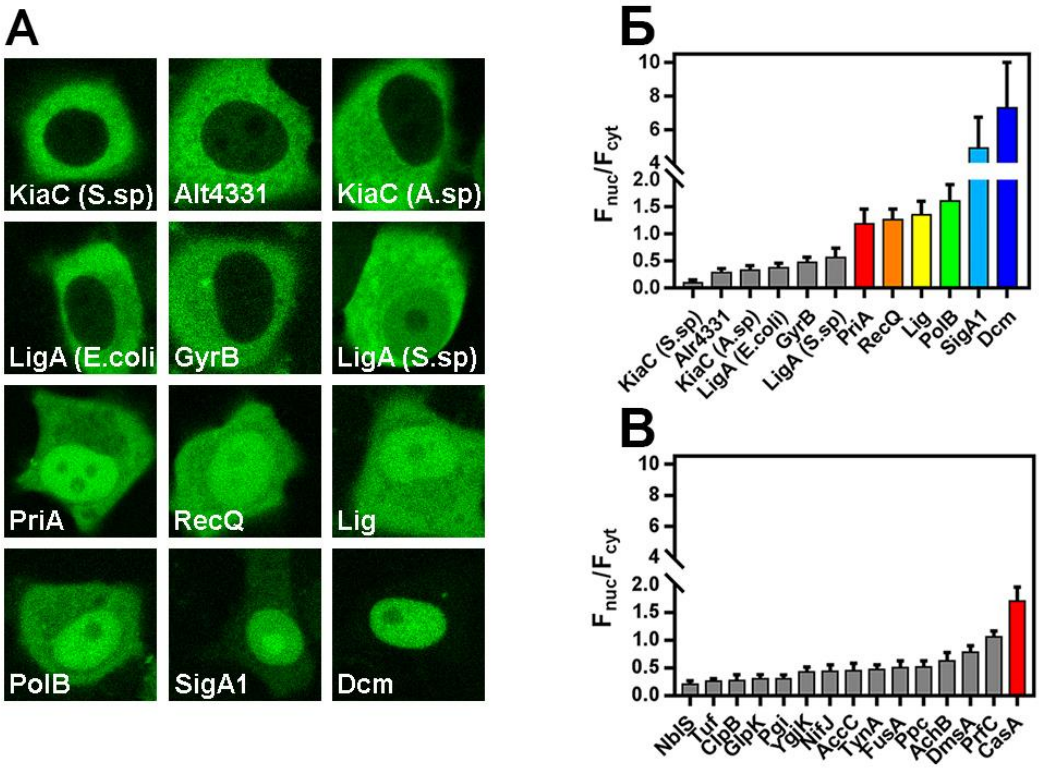

Рис. 2. NLS в составе прокариотических белков. А – Локализация прокариотических белков, слитых с EGFP, при экспрессии в клетках HeLa (прижизненные наблюдения). **Б** – Степень накопления химерных прокариотических белков, слитых с EGFP, с предсказанными NLS, в ядре эукариотической клетки. Цветом выделены белки, классифицированные как ядерные. **В** – Степень накопления контрольных химерных прокариотических белков, слитых с EGFP, в ядре эукариотической клетки. Цветом выделен белок, классифицированный как ядерный.

Картирование NLS в прокариотических белках с помощью сайт-направленного мутагенеза

Для идентификации NLS внутри прокариотических белков мы использовали сайт-направленный мутагенез. Положительно заряженные аминокислотные остатки в составе NLS необходимы для взаимодействия с импортинами, поэтому все они были заменены на аланин в предсказанных последовательностях NLS прокариотических белков. После чего анализировалась степень накопления мутантных белков (отношение F_{nuc}/F_{cyt}) при экспрессии в клетках HeLa.

В случае всех прокариотических белков с ядерной локализацией наблюдалось статистически достоверное снижение накопления белка при мутации хотя бы одного из предсказанных NLS (рис.3). При этом наиболее сильные эффекты давали мутации тех белков, которые в большей степени накапливались в ядре (SigA1 и Dcm). Таким

образом, в составе прокариотических белков было выявлено 9 NLS, 5 из которых входят в состав нуклеотид-связывающих доменов.

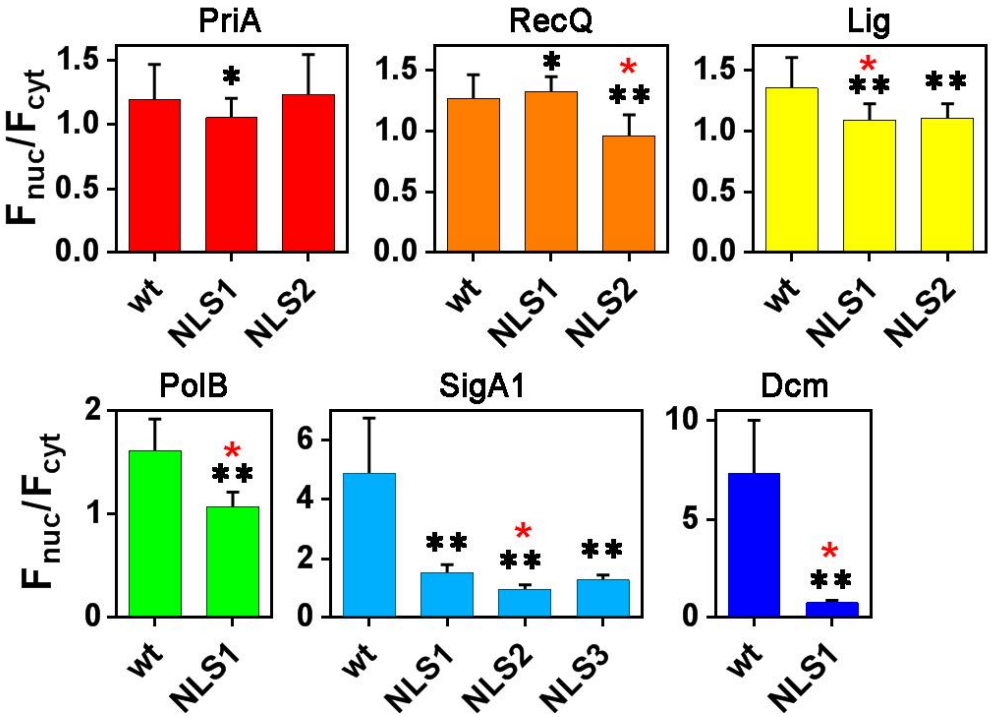

Рис. 3. Эффективность накопления белков в ядре с мутированными NLS по сравнению с диким типом. * - выделены NLS, входящие в состав нуклеотид-связывающих доменов. * – p<0.05 (критерий Манн-Уитни) ** – p<0.001 (критерий Манн-Уитни).

Взаимодействие с импортинами *in vivo*

Транспорт белков в ядро, опосредованный классическими NLS, основан на взаимодействии с комплексом адапторных белков, включающим импортины-*α* и -*β*. Неклассические NLS способны взаимодействовать с импортином-*β* напрямую. Наличие функционально активного NLS подразумевает способность белка взаимодействовать с импортинами, причем в отсутствие такого взаимодействия накопление белка в ядре должно снижаться.

Для выявления связывания с импортинами мы использовали пептидные ингибиторы Bimax2 (Kosugi et al. 2008), который высокоаффино и специфически связывается с молекулами импортина-*α*, и M9M (Cansizoglu et al. 2007), взаимодействующий с импортином-*β2*. Важно отметить, что в случае классических NLS на импорт должны влиять оба ингибитора, в случае неклассического пути импорта – только M9M.

Клетки HeLa одновременно трансфецировали плазмидой, несущей ген целевого белка слитого с EGFP, и плазмидой, кодирующей одну из ингибиторных молекул

(Bimax2 или M9M) слитую с TagRFP. После чего анализировали локализацию прокариотических белков в эукаритической клетке в сравнении с контрольными клетками, экспрессирующими только ген целевого белка. Анализ микроскопических изображений показал, что как коэкспрессия TagRFP-Bimax2, так и TagRFP-M9M снижает интенсивность флуоресценции ядра в случае белков PriA, Lig, PolB, SigA1 и Dcm. Эти данные говорят об использовании классического (т.е. импортин-*α/β*-зависимого) пути транспорта в ядро. Использованный подход не позволяет исключать параллельное использование и неклассического пути, но провести такое разделение не представляется на данный момент возможным. В случае белка RecQ снижение флуоресценции ядра наблюдалось только при использовании ингибитора TagRFP-M9M, т.е. данный белок накапливается по неклассическому пути.

NLS в составе цитоплазматических белков

Выявленная в ходе биоинформатического анализа интеграция NLS в составе функциональных доменов может быть характерна не только для ядерных белков, но и для любых доменов, обогащенных положительно заряженными аминокислотными остатками, причем не важно, принадлежат ли эти белки эукариотам или прокариотам. Проведенный анализ прокариотических белков подтвердил это предположение. Возникает вопрос о цитоплазматических белках – будут ли и они содержать NLS в случае наличия аналогичных по аминокислотному составу доменов?

Для проверки этого следствия гипотезы об интеграции NLS в составе функциональных доменов мы сформировали две выборки белков – цитозольных и нуклеоплазматических, согласно аннотации Uniprot. Выборка цитозольных белков состояла из 295 белков, выборка нуклеоплазматических – из 300. Посредством сервера PSORT II для каждого белка из выборок было предсказано наличие NLS. Оказалось, что в 52% случаев цитозольных белков программа PSORT II предсказывает в их последовательностях NLS, тогда как для нуклеоплазматических белков NLS предсказываются для 68% выборки. NLS предсказываются одинаково часто (34%) у цитозольных и нуклеоплазматических белков с молекулярным весом менее 40 кДа. Для белков с молекулярным весом более 40 кДа процент предсказания NLS выше в случае нуклеоплазматических белков (78%) по сравнению с белками цитозоля (61%).

Аналогичный анализ был также проведен на бóльшей по размеру выборки белков из опубликованного протеома ооцитов *X. laevis* (Wü et al. 2015). Посредством сервиса NLStradamus было проанализировано 746 ядерных белков, 1439

цитоплазматических белков и 560 белков, обладающих как ядерной, так и цитоплазматической локализацией. Оказалось, что, несмотря на то, что NLS предсказываются значительно чаще в ядерных белках с молекулярным весом более 40 кДа, более трети цитоплазматических белков с молекулярным весом выше 100 кДа могут также потенциально содержать NLS.

Таким образом, даже при отсутствии ядерной локализации и/или известной функции, связанной с клеточным ядром, вероятность наличия NLS в последовательностях цитоплазматических белков достаточно высока.

Сердечный тропонин I может накапливаться в клеточном ядре

Существующие методы предсказания NLS не очень надежны, и предсказания NLS в цитоплазматических белках могут быть артефактом. Поэтому мы проанализировали возможность накопления в ядре одного из цитоплазматических белков, сердечного тропонина I (cTnI). В аминокислотной последовательности человеческого cTnI (hcTnI) посредством программы PSORT II было предсказано несколько NLS, три из которых (NLS1, NLS2, NLS3) располагаются в тропонин C (TnC)-связывающем домене, один (NLS4) – в домене, взаимодействующем с актином и тропонином C, и два (NLS5 и NLS6) – во вне-доменном регионе (согласно базе данных UniProt).

cTnI является типичным цитоплазматическим белком, локализующимся практически исключительно внутри миофибрилл. В качестве контроля мы проанализировали мышиный cTnI в клетках сердца, где выявляется исключительно цитоплазматическая локализация этого белка (рис. 4А).

Для изучения локализации cTnI в недифференцированных мышечных, а также немышечных клетках мы использовали hcTnI, слитый с EGFP. При экспрессии в клетках культуры миобластов C2C12 EGFP-hcTnI имеет преимущественно ядерную локализацию (рис. 4Б). Ядерная локализация EGFP-hcTnI также была обнаружена в немышечных культурах клеток: человеческой карциномы (HeLa), человеческой остеосаркомы (U2OS), человеческой нейробластомы (SK-N-SH), мышиных фибробластов (L929), фибробластов *X. laevis* (XL-2) и даже в культуре клеток *D. melanogaster* (S2) (рис. 4В). Следует отметить, что hcTnI при экспрессии в немышечных клетках также накапливался и в ядрышках. Наблюдаемая ядерная локализация EGFP-hcTnI в различных типах клеток может действительно быть

обусловлена наличием в его аминокислотной последовательности NLS[1], которые в недифференцированных мышечных клетках, а также в культурах немышечных клеток могут обеспечивать импорт hcTnI в ядро клетки.

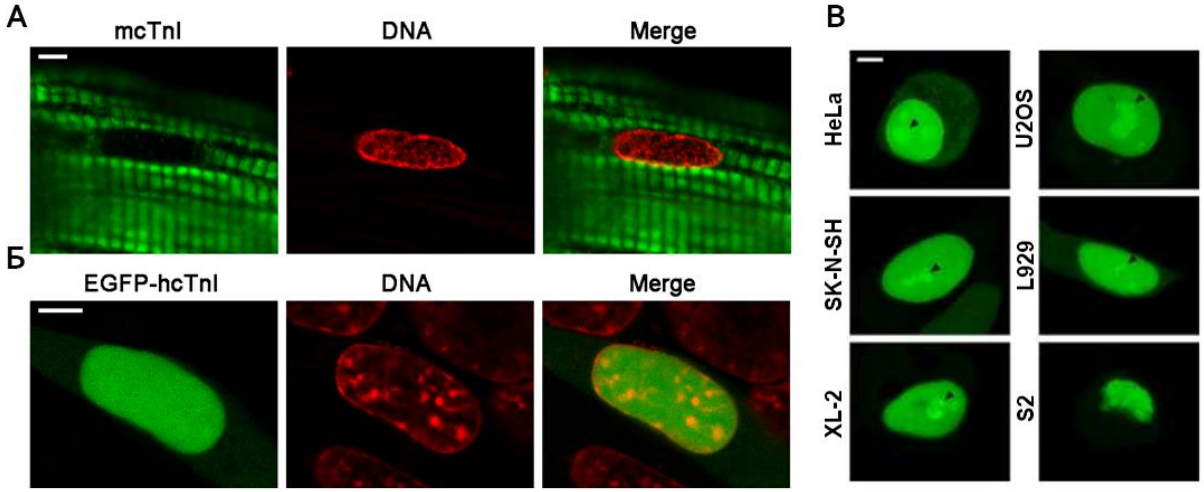

Рис. 4. Ядерная локализация hcTnI. А – Окрашивание специфическими антителами к мышиному cTnI кардиомиоцитов мыши (криосрез сердца мыши). **Б** – Ядерная локализация белка EGFP-hcTnI в культивируемых мышиных миобластах C2C12. **В** – Локализация EGFP-hcTnI в различных немышечных клетках. Ядрышко обозначено треугольными стрелками. Длина масштабной линейки 5 мкм.

Интеграция NLS и NoLS

Описанный выше феномен интеграции функциональных доменов и NLS является не единственным примером тесной интеграции последовательностей разного типа. Так, в литературе представлены данные о частом перекрывании двух типов сигнальных последовательностей – NLS и NoLS (Scott et al, 2010). С чем же связано такое совпадение этих сигналов? Для понимания этой интеграции необходимо было выявить механизм накопления белков в ядрышке за счет NoLS, который на момент начала работы не был известен.

Прежде всего, мы провели анализ аминокислотного состава NoLS. Для этого была использована выборка из 91 экспериментально определенных NoLS, принадлежащих белкам из различных царств живых организмов (растений, животных, вирусов и грибов). NoLS оказались достаточно гетерогенными по аминокислотному составу. Анализируя различные NoLS из выборки, мы не смогли

[1] Сделанные нами в этой работе предсказания о наличии функционально активных NLS в составе тропонина I было проверено с помощью сайт-направленного мутагенеза: все предсказанные монопартатные NLS оказались функционально активными, бипартатный NLS6 не проявлял активности в накоплении белка в ядре (данные М.Ю. Шубиной).

выявить какую-либо консенсусную последовательность. Единственным общим свойством всех NoLS является обогащенность положительно заряженными аминокислотными остатками лизина и аргинина. В среднем доля положительно заряженных аминокислотных остатков в последовательности белка составляет около 15%, тогда как в NoLS содержится около 50% аминокислотных остатков лизина и аргинина. При этом распределение частот встречаемости аминокислотных остатков в NoLS животных практически полностью совпадает с таковым для NLS человека (рис.5).

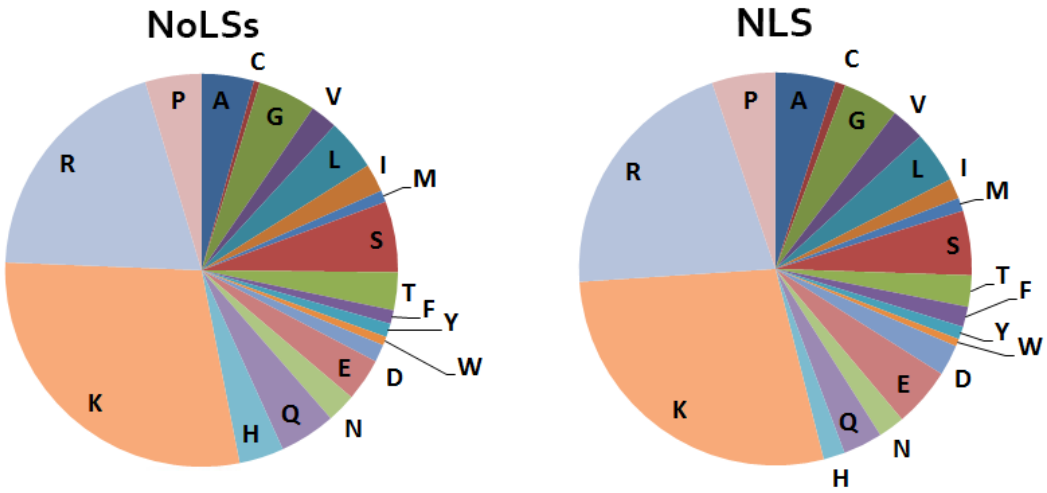

Рис.5. Распределения аминокислотного состава в NoLS и NLS

Эффективность накопления в ядрышке зависит от заряда участка, выступающего в качестве NoLS

Ранее в нашей лаборатории было показано, что последовательности из девяти аминокислотных остатков аргинин или лизина способны эффективно накапливать EGFP в ядрышке. Нами было проанализировано накопление EGFP в ядрышке в зависимости от количества добавленных аминокислотных остатков лизина и аргинина на его С-конце. Уже при визуальном анализе было видно, что чем больше аргинина или лизина содержится в добавленной последовательности, тем эффективнее накапливается EGFP в ядрышке. Для количественной оценки эффекта мы измеряли отношение средней яркости флуоресценции в ядрышке к средней яркости флуоресценции в нуклеоплазме (F_n/F_{nuc}). Мы считали, что последовательность является NoLS, если $F_n/F_{nuc} > 1$. Так было показано, что последовательность можно считать NoLS, если она состоит из более, чем 3 остатков аргинина или 5 остатков лизина.

Также мы предположили, что степень накопления в ядрышке может зависеть не только от заряда, но и от плотности распределения положительно заряженных аминокислотных остатков. Для проверки данной гипотезы были сконструированы последовательности из 6 или 9 аминокислотных остатков лизина или аргинина, расположенных подряд или перемежающихся незаряженными аминокислотными остатками глицина. Было показано, что в случае высокой плотности расположения положительно заряженных аминокислотных остатков, интенсивность накопления EGFP выше, по сравнению с конструктами, где положительно заряженные аминокислотные остатки NoLS разделены незаряженными аминокислотными остатками глицина.

Приведенные выше эксперименты свидетельствуют в пользу электростатического механизма накопления белков в ядрышке, зависящего не только от количества положительно заряженных аминокислотных остатков в NoLS, но и от плотности их распределения.

Электростатический механизм действия NoLS, совмещение функции NoLS и NLS в одной сигнальной последовательности, интегрированность NLS/NoLS в состав нуклеотид-/белок-связывающих доменов и обогащенность их по положительному заряду подтверждают гипотезу о предсуществовании NoLS/NLS в составе функциональных доменов. Таким образом, наличие положительно заряженного участка в аминокислотной последовательности белка в отсутствии каких-либо более аффинных взаимодействий в цитоплазме клетки может приводить к транслокации белка в ядро клетки, где также за счет электростатического взаимодействия белок может удерживаться в составе ядрышка.

ЗАКЛЮЧЕНИЕ

В ходе выполнения работы нами были получены данные, которые позволяют высказать предположение о том, что NLS могут быть эволюционно интегрированы в состав некоторых типов доменов белков. NLS – короткие и довольно гетерогенные последовательности, как правило, обогащенные положительно заряженными аминокислотными остатками. Если в составе домена белка присутствует участок, обогащенный положительно заряженными аминокислотными остатками, то наравне со своей типичной функцией, он может играть роль NLS, вне зависимости от того, необходима белку транслокация в ядро, или нет. Именно эту особенность мы и обозначили как интеграция NLS. Можно предполагать, что наличие таких интегрированных NLS в составе белков прокариот облегчало эволюцию системы импорта белков в ядро, а также диктовало, до некоторой степени, состав формирующегося ядерного протеома.

Другим следствием идеи об интеграции NLS было предположение о широком распространении NLS в цитоплазматических белках. В отсутствие более аффинных взаимодействий в цитоплазме положительно заряженные участки функциональных доменов, могут неспецифично взаимодействовать с системой импортинов, приводя к транслокации белка в ядро.

Еще одним примером интеграции, может быть описанное в литературе совпадение NLS и NoLS. Наши данные свидетельствуют в пользу того, что NoLS накапливаются в ядре за счет электростатического взаимодействия с компонентами ядрышка. Т.е. любой участок белка, обогащенный положительно заряженными аминокислотными остатками, включая значительную часть NLS, может играть роль NoLS.

ВЫВОДЫ

1. Наиболее точные результаты предсказания NLS показывают программы NucPred, seqNLS, NLStradamus и NLSMapper.
2. NLS могут быть интегрированы в составе некоторых – прежде всего, нуклеотид-связывающих – доменов, эволюционируя как часть этих доменов.
3. Прокариотические белки с интегрированными NLS могут проникать в клеточное ядро, взаимодействуя с системой ядерного импорта эукариотической клетки.
4. В аминокислотные последовательности цитоплазматических белков могут быть интегрированы NLS, посредством которых данные белки способны импортироваться в ядро.
5. Перекрывание NLS и NoLS может быть связано с обогащенностью NLS положительно заряженными аминокислотными остатками, что ведет к возможности накопления белка в ядрышке.

НАУЧНЫЕ СТАТЬИ ПО ТЕМЕ ДИССЕРТАЦИИ, ОПУБЛИКОВАННЫЕ В ЖУРНАЛАХ SCOPUS, WOS, RSCI:

1. **Lisitsyna O.M**., Seplyarskiy V.B., Sheval E.V. Comparative analysis of nuclear localization signal (NLS) prediction methods // *Biopolymers and Cell* – 2017 – 33 – p. 147–154;
2. **Лисицына О.М.,** Шеваль Е.В. Происхождение и ранние этапы эволюции ядерной оболочки // *Биологические мембраны* – 2016 – 33 – с. 243-251; IF = 0,598
3. Musinova Y.R., Kananykhina E.Y., Potashnikova D.M., **Lisitsyna O.M**., Sheval E.V. A charge-dependent mechanism is responsible for the dynamic accumulation of proteins inside nucleoli // *Biochimica et Biophysica Acta – Molecular Cell Research* – 2015 – 1853 – p. 101-110; IF = 4,739
4. Musinova Y.R., **Lisitsyna O.M.**, Golyshev S.A., Tuzhikov A.I., Polyakov V.Y., Sheval E.V. Nucleolar localization/retention signal is responsible for transient accumulation of histone H2B in the nucleolus through electrostatic interactions // *Biochimica et Biophysica Acta – Molecular Cell Research* 2011 – 1813 – p. 27–38; IF = 4,739

ДРУГИЕ НАУЧНЫЕ РАБОТЫ, ОПУБЛИКОВАННЫЕ ПО ТЕМЕ ДИССЕРТАЦИИ:

1. **Lisitsyna O.M**., Kurnaeva M.A., Shubina M.Y., Arifulin E.A., Musinova Y.R., Mironov A.A., Sheval E.V. Origin and evolution of nuclear localization signals // 26th Wilhelm Bernhard Workshop on the Cell Nucleus – 2019 – p. 190.
2. **Лисицына О.М.** Эволюционное происхождение сигналов ядерной локализации // Материалы Международного молодежного научного форума «Ломоносов-2018» М.: МАКС Пресс, 2018.
3. Kurnaeva M**.**, **Lisitsyna O**., Sheval E.V., Vassetzky Y., Musinova Y. Mechanism of nuclear and nucleolar localization signal overlapping // 25th Wilhelm Bernhard Workshop on the Cell Nucleus – 2017 – p. 93.
4. Musinova Y.R., Kananykhina E.Y., Potashnikova D.M., **Lysitsina O.M**., Sheval E.V. Charge-dependent mechanism for retention of proteins in the nucleoli via dynamic interaction with nucleolar RNA // VII International meeting “From molecular to cellular events in human pathologies”. Biopolymers and Cell – 2014 - 30 (Supplementary 5) – p. 29.

5. Мусинова Я.Р., Свистунова Д.М., Кананыхина Е.Ю., **Лисицына О.М.**, Шеваль Е.В. Анализ механизмов взаимодействия сигналов ядрышковой локализации с компонентами ядрышка // Материалы Всероссийского симпозиума и школы-конференции для молодых ученых по биологии клетки в культуре. Цитология – 2013 - Т. 55. № 9 – с. 650-651.
6. Мусинова Я.Р., **Лисицына О.М.**, Кананыхина Е.Ю., Поляков В.Ю., Шеваль Е.В. Механизмы накопления белков в ядрышке за счет неспецифических сигналов ядрышковой локализации // I Всероссийская конференция “Внутриклеточная сигнализация, транспорт, цитоскелет” Цитология – 2011 - Т. 53 - с.709.
7. Musinova Y.R., **Lisitsyna O.M**., Kananykhina E.Y., Polyakov V.U., Sheval E.V. The region of histone H2B with potential nucleolar localization (NoLS) activity // 22nd Wilhelm Bernhard Workshop – 2011 – p. 96.
8. Мусинова Я.Р., Кананыхина Е.Ю., **Лисицына О.М.**, Шеваль Е.В. Количественный анализ выраженности действия естественных и искусственных сигналов ядрышковой локализации // XVII Российский симпозиум по растровой электронной микроскопии и аналитическим методам исследования твердых тел – 2011 - с. 254.
9. Musinova Y.R., **Lisitsyna O.M.**, Golyshev S.A., Polyakov V.Y., Sheval E.V. Non-specific retention as a possible mechanism for histone H2B accumulation in the nucleolus // American Society for Cell Biology 50th Annual Meeting – 2010 - p. 1133.
10. Мусинова Я.Р., **Лисицына О.М.**, Шеваль Е.В. Несопряженное с функционированием удержание белков в ядрышке: к вопросу о механизме действия сигналов ядрышковой локализации // XXIII Российская конференция по электронной микроскопии. Черноголовка – 2010 – с. 394.