

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Суханова Ксения Владимировна

**Эволюционный консерватизм амилоидных свойств нуклеопоринов с  
FG повторами.**

Выпускная квалификационная работа  
по направлению подготовки: «Биология»  
основная образовательная программа бакалавриата: «Биология»

Научный руководитель:  
н.с., к.б.н. Бондарев Станислав Александрович

Санкт-Петербург  
2021

# Оглавление

	Стр.
<b>Введение</b> . . . . .	4
<b>Цель и задачи</b> . . . . .	5
<b>1. Обзор литературы</b> . . . . .	6
1.1. История изучения амилоидов . . . . .	6
1.2. Понятие амилоидов и общая характеристика . . . . .	9
1.3. Классификация амилоидов и их биологическая роль . . . . .	10
1.3.1. Патологические амилоиды . . . . .	11
1.3.2. Функциональные амилоиды . . . . .	15
1.4. Белки нуклеопорины . . . . .	21
1.4.1. История открытия и описания ядерно-порового комплекса . . . . .	21
1.4.2. Структурные особенности ядерной поры и её функции . . . . .	21
1.4.3. Транспорт через ядерную пору . . . . .	26
1.4.4. Барьерная функция ядерно-порового комплекса и значение FG-повторов . . . . .	28
1.4.5. Структурные особенности белка Nsp1 и Nup100 . . . . .	30
1.4.6. Амилоидные свойства Nsp1 . . . . .	32
1.4.7. Амилоидные и прионоподобные свойства Nup100 . . . . .	34
1.5. Биоинформатические методы анализа амилоидогенных свойств . . . . .	36
1.5.1. Общие принципы . . . . .	36
1.5.2. Краткое описание современных методов биоинформатического предсказания потенциально амилоидогенных белков . . . . .	37
<b>2. Материалы и методы</b> . . . . .	42
2.1. Материалы . . . . .	42
2.1.1. Поиск ортологов нуклеопоринов дрожжей <i>S.cerevisiae</i> . . . . .	42
2.1.2. Таксономическое положение видов по базе данных NCBI . . . . .	44
2.2. Построение выравниваний для предварительной проверки последовательностей для всех ортологов . . . . .	44
2.3. Анализ амилоидогенности с помощью программы ArchCandy . . . . .	45
2.4. Список использованных пакетов R . . . . .	45
2.5. Оценка долей потенциальных амилоидов среди наборов ортологов . . . . .	46
2.6. Поиск потенциальных консервативных амилоидогенных участков, не входящих в структурные домены . . . . .	47

2.6.1. Получение подвыборок набора ортологов и их последующее выравнивание . . . . .	47
2.6.2. Анализ выравнивания белковых последовательностей . . . . .	47
2.6.3. Картирование участков с известной структурой в выравнивании . . . . .	48
<b>3. Результаты и обсуждение . . . . .</b>	<b>49</b>
3.1. Проверка набора ортологов нуклеопоринов . . . . .	49
3.2. Анализ амилоидогенности последовательностей набора ортологов нуклеопоринов . . . . .	49
3.3. Поиск участков с консервативными амилоидными свойствами . . . . .	53
3.3.1. Подсчет параметров амилоидогенности, консервативности и аминокислотного состава . . . . .	53
3.3.2. Поиск участков с консервативными амилоидными свойствами . . . . .	54
<b>4. Выводы . . . . .</b>	<b>62</b>
<b>Список литературы . . . . .</b>	<b>63</b>
<b>Приложение А. Приложения . . . . .</b>	<b>76</b>

## Введение

На сегодняшний день актуально исследование патологических и функциональных амилоидов. И список последних значительно увеличился за последнее десятилетие. Функциональные амилоиды показаны как для прокариотических организмов (курлины *Escherichia coli* (CsgA-CsgB); чаплины *Streptomyces coelicolor* (ChpA-H) и др.), так и для эукариот (Pmel17, RIP1 и RIP3, семеноглеин *Homo sapiens*; Orb2 *Drosophila melanogaster*; СРЕВ *Aplysia californica* и др.). Например, функциональный амилоид Pmel17 участвует в накоплении специального пигмента меланина с образованием меланосом в клетках меланоцитах, что обеспечивает как защитную функцию человека и других млекопитающих от солнечных лучей, так и формирование окраса. Также структурные белки шелка шелкопряда спидроин и фиброин, помимо своего функционального значения в природе, сейчас приобрели практическое применение во многих биомедицинских и промышленных направлениях. Любопытно и то, что среди функциональных амилоидов есть те, которые у разных организмов за счет своей амилоидной природы выполняют сходные функции. К последним относится, например, белок СРЕВ моллюска *A. californica* и ортологичный ему белок Orb2 *D. melanogaster*: данный белок участвует в процессе так называемой консолидации памяти, то есть трансляции её из кратковременной в долговременную. Последнее свидетельствует о консервативности амилоидных свойств среди далеких друг от друга таксономических групп. Примечательно, что для таких важных белков, как нуклеопоринов, являющихся структурным компонентом ядерных поровых комплексов, также были показаны амилоидные свойства. Продемонстрировано *in vitro* и *in vivo*, что дрожжевые белки Nsp1 и Nup100 способны образовывать амилоидоподобные структуры и агрегаты, и что в их формировании участвуют особые домены данных белков - FG-обогащенные участки. Кроме этого, как было замечено ранее, амилоидные свойства могут быть эволюционно консервативными для одних и тех же белков у разных организмов, поэтому мы заинтересовались, соблюдается ли аналогичное свойство для всех нуклеопоринов среди разных таксономических групп или нет.

## Цели и задачи

### **Цель:**

Оценка консервативности амилоидных свойств нуклеопоринов с FG повторами.

### **Задачи:**

1. Биоинформатический анализ амилоидных свойств у ортологов нуклеопоринов с FG-повторами;
2. Поиск участков, которые обладают амилоидными свойствами у большинства организмов.

# 1. Обзор литературы

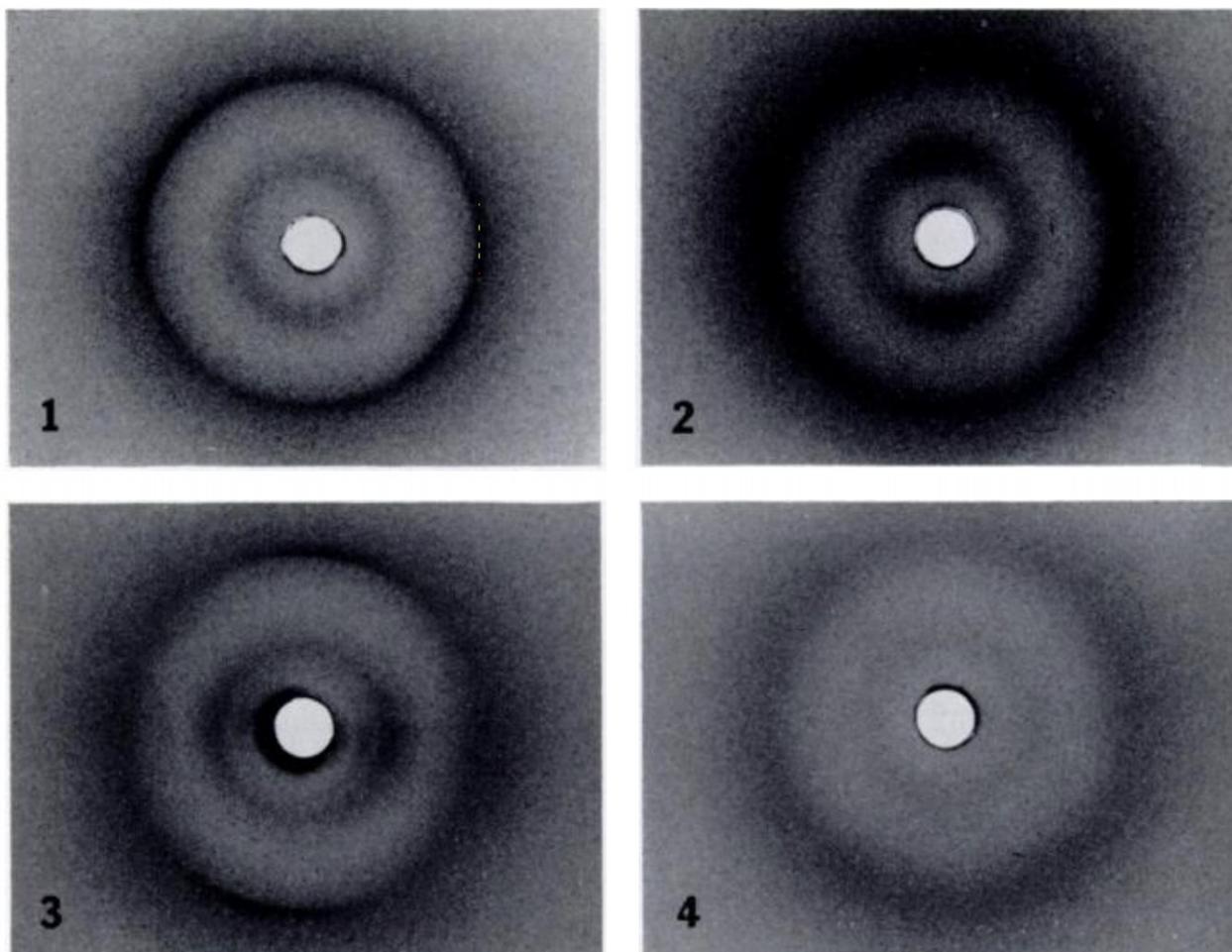
## 1.1. История изучения амилоидов

История изучения амилоидов началась около 200 лет назад. Сам термин «амилоид» происходит от латинского и греческого слов «amylon» и «amylum», которые в переводе означают «крахмал», и впервые стал употребляться М. Шлейденом в 1838 году для обозначения этого вещества. Однако применительно к человеку впервые амилоидами были названы крахмалоподобные белые отложения в тканях селезенки, которые позднее были обнаружены и в печени Р. Вирховым (Kyle, 2001). Такие отложения наблюдались у людей с различными заболеваниями, потому что именно отложения определяли как причину развития заболевания (Kyle, 2001) (подробнее в разделе «Патологические амилоиды»). Тогда предполагали, что они по биохимической природе являются углеводами ввиду синего окрашивания йодом, аналогично крахмалу. Но в 1859 году Ч. Шмидт показал высокое преобладание в них азота (Schmidt, 1859), и в этом же году Н. Фридрих и А. Кекуле полностью исключили, что основным компонентом амилоидных отложений являются крахмал и целлюлоза (Friedreich, Kekulé, 1859).

Развитие методов гистохимического окрашивания с начала 20 века позволило иначе посмотреть на природу амилоидов. На смену методике окрашивания йодом и серной кислотой пришло использование анилиновых красителей. Первыми были введены метакроматические красители (например, метил фиолетовый синий), которые позволили лучше определять наличие внеклеточных амилоидных образований ввиду более сильной специфичности связывания красителя с амилоидом. С 1922 года стали применять специфический краситель Конго красный (Bennhold, 1922) и с 1959 - Тиофлавин (Vassar, Culling, 1959, Elghetany, Saleem, 1988) для более точной идентификации амилоидных агрегатов в тканях. Характерные для окрашенных Конго красным амилоидов зелено-яблочное двулучепреломление и желто-оранжевое свечение в поляризованном свете были впервые показаны П. Диври и М. Флоркином в 1927. Однако позднее иные цвета данного окрашивания были показаны другими исследователями (Taylor *et al.*, 1974, Alexander, Douglas, 2008, Yakupova *et al.*, 2019).

С параллельным развитием методов электронной микроскопии удалось исследовать структуру амилоидов. Исследуя процессы деградации и денатурации глобулярных белков, У. Атсбери и С. Диккинсон первыми показали необычные диффракционные свойства агрегатов этих белков, которые позднее были показаны для всех амилоидов (Astbury *et al.*, 1935). Общее же свойство амилоидов - наличие кросс- $\beta$  структуры (Maji *et al.*, 2009a) - было продемонстрировано И. Йенсом и Д. Гленнером Б. с помощью анализа диффракции рентгеновских лучей выделенных амилоидных филаментов человека и животных из

тканей с развившимися патологиями и соответствующей диффракционной картины (Eanes, Glenner, 1968). Кросс- $\beta$  структура характеризуется образованием белковыми молекулами межмолекулярных  $\beta$ -листов, в которых мономеры белка находятся на расстоянии 0,47 нм между собой, а соседние  $\beta$ -листы располагаются в 0,6-1 нм друг от друга. Такой принцип укладки дает определенную диффракционную картину (Рис.1). Позднее, в 1969



**Рисунок 1 — Диффракционные картины полученных амилоидных филаментов человека и животных.** На схеме представлены диффракционные картины амилоидных филаментов в разной ориентации: (1) — неориентированные очищенные филаменты человека; (2) — механически ориентированные филаменты; (3) — полученные путем дегидратации жидкой суспензии филаментов на плоской поверхности мембраны филаментов; (4) — фибриллоподобные агрегаты глобулярных единиц. Внешнее кольцо соответствует расстоянию между  $\beta$ -слоями внутри одной аминокислотной цепи белка и имеет размер в 0,47 нм, а внутреннее - расстоянию размером в 0,6-1 нм между соседними  $\beta$ -слоями двух параллельных тяжей фибриллы белка. Данные взяты из статьи И. Йенсена и Д. Гленнера (Eanes, Glenner, 1968).

году, с помощью электронной микроскопии было показано, что все амилоидные белки образуют неветвящиеся фибриллярные структуры (Bonar *et al.*, 1969, Iadanza *et al.*, 2018). И благодаря дальнейшему развитию методов микроскопии (электронная микроскопия, криоэлектронная микроскопия, атомно-силовая микроскопия) и спектроскопии (кругового

дихроизма, ядерно-магнитного резонанса) удалось описать структурные особенности фибрилл (Iadanza *et al.*, 2018) и явление их полиморфизма (Goldsbury *et al.*, 2000).

История изучения амилоидов включает в себя и исследования прионов, которые были открыты в конце 1980-х годов С. Прузинером (Prusiner, 1982). Прионы являются уникальным классом инфекционных агентов. Большинство прионов — амилоиды с инфекционными свойствами. Под «инфекционностью» подразумевается способность белка в амилоидной форме вызывать переход других белков в соответствующую структуру и последующую их агрегацию. Первым описанным прионом, давшим начало самому термину, стала охарактеризованная Прузинером особая форма белка PrP, который в норме является трансмембранным белком, функции которого до конца не изучены (Prusiner, 1982, Wulf *et al.*, 2017). Большой вклад в исследование прионов внесли описания Р. Викнером первых дрожжевых прионов: [URE3], [PSI<sup>+</sup>] — соответствующих таким дрожжевым белкам, как Ure2p, который в норме репрессирует катаболизм бедных азотом субстратов, и Sup35p, являющийся фактором терминации трансляции (Wickner, 1994). Со временем обнаруживали все больше дрожжевых прионов: [PIN<sup>+</sup>] (Derkatch *et al.*, 2001), [Het-s] (Coustou *et al.*, 1997), [SWI<sup>+</sup>] (Du *et al.*, 2008) и т.д. И в последнее время все больше стали обнаруживать прионы, не обладающие амилоидной структурой (Chakravarty, Jarosz, 2018). Впервые такие свойства были показаны Б. Робертсом и Р. Викнером для  $\beta$ -протеазы дрожжей, которая вовлечена в процессы споруляции в условиях голодания (Roberts, Wickner, 2003). Впоследствии были также обнаружены и другие не амилоидные прионы у разных организмов: например, прион [GAR<sup>+</sup>], позволяющий клеткам дрожжей обходить недостаток глюкозы в среде; ответственный за процесс консолидации памяти прион [CPEB] моллюска *Aplysia californica*; ответственный за реализацию противовирусного сигнального каскада прион млекопитающих [MAVS] (Chakravarty, Jarosz, 2018) и недавно описанная прионная форма белка Vts1 [SMAUG<sup>+</sup>], способствующая усилению роста клеток дрожжей при бедности питательной среды (Chakravarty *et al.*, 2020).

С конца 20 века активно изучается вопрос о функциональных амилоидах, перечень которых значительно расширился за последние годы (Otzen, Riek, 2019). В таких случаях агрегаты становятся жизненно необходимыми, так как вовлечены в выполнение тех или иных биологических процессов. Среди функциональных групп амилоидные свойства были показаны и доказаны *in vitro* для многих структурных (Iconomidou *et al.*, 2000, Wösten, Vocht, 2000, Claessen *et al.*, 2003, Fowler *et al.*, 2006, Slotta *et al.*, 2007), сигнальных (Li *et al.*, 2012, Riek, Saure, 2016), запасующих (Maji *et al.*, 2009b) и регуляторных белков (Si *et al.*, 2003) в царствах как прокариот, так и эукариот. Кроме того, один и тот же белок обладает амилоидными свойствами в разных организмах (например, прион CPEB *Aplysia californica* и его ортолог Orb2 *Drosophila melanogaster*). Это говорит об эволюционной значимости амилоидного состояния белков для клеточных процессов. И, более того, некоторые из них обладают прионоподобными свойствами (Coustou *et al.*, 1997, Si *et al.*, 2003).

Однако, на сегодняшний день, еще больший интерес вызывает факт, что многие белки, не являющиеся амилоидами, в определенных условиях *in vitro* образуют амилоидоподобные агрегаты и фибриллы. На основании полученных данных зародилась концепция «амилоидного мира», согласно которой кросс- $\beta$ -структура является одной из предковых функциональных конформаций белков ввиду таких свойств, как высокая стабильность в жестких биохимических условиях, низкий энергетический потенциал, способность к самостоятельной укладке и проявление белковой активности (Greenwald, Riek, 2012, Greenwald *et al.*, 2018). Все вышеупомянутое говорит о том, что, несмотря на уже подробную изученность амилоидов, многое все еще предстоит открыть.

## 1.2. Понятие амилоидов и общая характеристика

Понятие «амилоиды», на сегодняшний день, имеет разные трактовки. Сейчас «амилоидом» без каких-либо уточнений считаются *in vivo* или *ex vivo* белковые отложения, обладающие кросс- $\beta$  структурой. С медицинской точки зрения амилоидами принято называть отложения белков, которые связаны с развитием патологий и идентифицируются по отличительным свойствам (Benson *et al.*, 2018). Обязательными характеристиками являются связывание внеклеточных агрегатов с красителем Конго красным и последующее зеленое, желтое или оранжевое свечение ввиду двулучепреломления при исследовании путем поляризационной микроскопии (Benson *et al.*, 2018). Однако с биологической точки зрения важна не столько локализация агрегатов, сколько их биофизические параметры и структура. Поэтому в биологии амилоидами называю белковые фибриллы с кросс- $\beta$  структурой (Iadanza *et al.*, 2018).

На сегодняшний день известно несколько вариантов положения кросс- $\beta$  листов, среди которых выделяют (Iadanza *et al.*, 2018):

- параллельные  $\beta$ -структуры;
- антипараллельные  $\beta$ -структуры;
- $\beta$ -спирали (в том числе,  $\beta$ -соленоиды);
- комбинации структур (например,  $\beta$ -листов и  $\beta$ -спиралей).

Амилоидные свойства и структура белка напрямую определяются его аминокислотным составом. С точки зрения аминокислотного состава амилоиды делят на обогащенные гидрофобными и алифатическими неполярными аминокислотами и на обогащенные полярными аминокислотами. К первым относятся амилоиды, в аминокислотной последовательности которых преобладают следующие аминокислоты: гидрофобные фенилаланин, тирозин, триптофан и неполярные алифатические лейцин, изолейцин, валин. Взаимодействие гидрофобных радикалов лежит в основе образования уникальной для амилоидов «стерической молнии» и её вариантов (Sawaya *et al.*, 2007, Iadanza *et al.*, 2018).

«Стерическая молния» представляет собой пару комплементарных друг другу  $\beta$ -листов, образованных очень короткими пептидами. Стабильность такой структуры усиливается вовлечением в её образование одних и тех же аминокислотных последовательностей белка, обеспечивающих регулярность фибриллы (Kajava *et al.*, 2010, Lendel *et al.*, 2014). Гидрофобные и неполярные аминокислоты за счет своих биохимических и биофизических свойств радикалов формируют так называемый «гидрофобный кор». Этот кор является основополагающим в стабильности амилоидной структуры как внутри одного  $\beta$ -листа, так и в компактизации фибриллы вдоль и относительно двух входящих в неё  $\beta$ -тяжей пептидов вне зависимости от типа кросс- $\beta$  структуры (Iadanza *et al.*, 2018). Ко второй группе амилоидов относятся белки, богатые такими полярными аминокислотами, как аспарагин, глутамин, серин и треонин (Paz, Serrano, 2004). Особенность этой группы амилоидов заключается в том, что в стабильность их структуры больший вклад вносят взаимодействие друг с другом как между  $\beta$ -тяжами каждого пептида, так и внутри одного  $\beta$ -листа.

С момента начала изучения амилоидов и по сей день удалось описать их разные биофизические свойства, делающие эти белковые агрегаты уникальными на фоне обычных внутриклеточных макромолекулярных комплексов и отложений (Ваха, 2008, Iadanza *et al.*, 2018 Якурова *et al.*, 2019):

- окрашивание с помощью красителя Конго красным; и последующее характерное зелено-яблочное, оранжевое и другое свечение в поляризованном свете;
- окрашивание с помощью красителя Тиофлавин Т (ThT) и его аналогов;
- устойчивость агрегатов к действию детергентов и протеаз;
- неветвящаяся фибриллярная структура агрегатов.

Все вышеперечисленные свойства характерны для большинства амилоидных и амилоидоподобных агрегатов.

### 1.3. Классификация амилоидов и их биологическая роль

Исходно амилоиды рассматривали в контексте патогенных образований, связанных, в частности, с нейродегенеративными заболеваниями (Chiti, Dobson, 2006, Chiti, Dobson, 2017, Dobson *et al.*, 2019). Но в последние десятилетия более дискуссионным становится вопрос об их функциональности. Связь амилоидного состояния белка с выполнением определенных функций в организмах была показана как для эукариот, так и для прокариот (Otzen, Riek, 2019).

### 1.3.1. Патологические амилоиды

Заболевания, связанные с нарушением конформации белка, приводящего к потере функций, принято называть «протеинопатиями», их частным случаем являются «амилоидозы», связанные с отложением амилоидных агрегатов в тканях и органах. До начала XX века природа таких скоплений была неизвестна. Такие скопления проверялись специальным окрашиванием йодом в присутствии серной кислоты, приводящего к появлению красного или фиолетового оттенка на ткани в местах скопления отложений. Изначально данные агрегаты обнаруживали только в печени и почках, но Дж. Мекель показал, что аналогичное окрашивание возможно и в других тканях и структурах, таких как: аорте, артериях и кишечной стенке. После подтверждения белковой природы амилоидов заболевания, связанные с их накоплением в тканях, стали называть «амилоидозами». Тогда считали, что амилоидозы являются следствиями таких заболеваний, как сифилис, туберкулез или костная болезнь, однако в 1856 году С. Уилкис опубликовал в своей работе, что аналогичные патологии органов были характерны и для пациентов, не страдающих ни одним из выше упомянутых заболеваний. Позднее С. Уилкис также утверждал, что данные скопления могут откладываться в селезенке и надпочечниках, делая эти органы при этом более «крупными и примечательно твердыми» (Kyle, 2001).

Все описанные выше случаи амилоидозов согласно первой классификации относили к первичным (AL, amyloid light-chain (AL) amyloidosis) и вторичным амилоидозам (AA, serum amyloid A (SAA) amyloidosis). До момента развития биохимических и микроскопических методов исследования амилоидов и понимания их структуры классификация была основана лишь на медицинских данных об их локализации и ассоциированных болезнях. Так, к первичным амилоидозам относили не связанные с клиническим состоянием или развивающиеся на фоне множественной миеломы в таких тканях, как: сердце, язык и околосуставная ткань. В то время как ко вторичным амилоидозам относили ассоциированные с хроническими воспалительными состояниями. Среди последних выделяют наиболее консервативные случаи развития: в печени, селезенке и почках - и связанные с хроническими инфекциями: сифилис, туберкулез и остеомиелит. Обнаружение новых случаев возникновения амилоидозов позволило описать не только AL и AA амилоидозы, но и семейный, старческий системный и диализ-ассоциированный амилоидозы (Kyle, 2011).

Образование таких агрегатов может быть как тканеспецифично, так и повсеместно; наследуемым и нет. Однако агрегаты, образующиеся одновременно как в тканях и клетках ЦНС, так и других системах органов, показаны лишь для некоторых пептидов (например, пептид  $A\beta$ , транстирретин (Chiti, Dobson, 2017)). Таким образом, по современной классификации амилоидозы принято делить на системные, то есть затрагивающие разные ткани и системы органов, и на локализованные, являющиеся органо-специфичными. А для более детального уточнения природы отложений используется номенклатура с указанием типа предшественника амилоидных фибрилл (Табл.1, Benson *et al.*, 2018).

**Таблица 1** — Список известных амилоидозов по современной классификации и номенклатуре; «А» в белковом компоненте - от **А**милоид, \* — продукт того же гена, что и АВг1, \*\* — также известен как амилин, \*\*\* — не доказано по анализу аминокислотной последовательности, \*\*\*\* — требуется определение полной аминокислотной последовательности; *Benson et al., 2018*

<b>Белковый компонент фибрилл</b>	<b>Предшественник</b>	<b>Системный и/или локализованный</b>	<b>Приобретенный и/или наследственный</b>	<b>Место локализации и/или происхождение</b>
AL	Легкие цепи иммуноглобулина	С, Л	П, Н	Все органы, обычно кроме ЦНС
AH	Тяжелые цепи иммуноглобулина	С, Л	П	Все органы, кроме ЦНС
AA	(Аро) Сывороточный амилоид А	С	П	Все органы, кроме ЦНС
ATTR	Транстиретин	С	П, Н	Сердце (преимущественно у мужчин), железы, ПНС и АНС, глаза
A $\beta$ 2M	$\beta$ 2-микроглобулин	С	П, Н	Опорно-двигательная система, АНС
AApoAI	Апополипротеин А I	С	Н	Сердце, печень, почки, ПНС, яички, гортань (С-терминальные варианты), кожа (С-терминальные варианты)
AApoAII	Апополипротеин А II	с	Н, П	Почки
AApoAIV	Апополипротеин А IV	С	П	Почки
AApoCII	Апополипротеин С II	С	Н	Почки
AApoCIII	Апополипротеин С III	С	Н	Почки
Agel	Гельзолин	С	Н	ПНС, роговица
ALys	Лизозим	С	Н	Почки
ALEC2	Хемотаксический фактор-2 лейкоцитов	С	П	Почки
AFib	Фибриноген $\alpha$	С	Н	Почки
ACys	Цистатин С	С	Н	ПНС, кожа
Продолжение на следующей странице				

Белковый компонент фибрилл	Предшественник	Системный и/или локализованный	Приобретенный (П) или наследственный (Н)	Место локализации и/или происхождение
ABri	ABriPP	С	Н	ЦНС
ADan*	ADanPP	Л	Н	ЦНС
A $\beta$	A $\beta$ -белковый предшественник	Л	Н (wt), П (варианты)	ЦНС
A $\alpha$ Syn	$\alpha$ -синуклеин	Л	П	ЦНС
ATau	Tau	Л	П	ЦНС
APrP	PrP	Л, С(варианты)	П (wt), Н (варианты)	Болезнь Крейтцфельдта-Якоба, ПНС
ACal	(Про)кальцитонин	Л	П	С-клеточная опухоль щитовидной железы
AIAPP	Амилоидный полипептид** островков Лангерганса (амилин)	Л	П	Островки Лангерганса
AANF	Предсердный натрийуретический фактор	Л	П	Сердечное предсердие
APro	Пролактин	Л	П	Пролактиномы гипофиза, стареющий гипофиз
AIns	Инсулин	Л	П	Ятрогенные, локальные инъекции
ASPC***	Поверхностный белок легких	Л	П	Легкие
AGal7	Галектин 7	Л	П	Кожа
ACor	Корнеодесмозин	Л	П	Ороговевший эпителий, волосные фолликулы
AMed	Лактадгерин	Л	П	Старческая среда аорты
AKer	Кератоэпителин	Л	П	Роговица

Продолжение на следующей странице

Белковый компонент фибрилл	Предшественник	Системный и/или локализованный	Приобретенный (П) или наследственный (Н)	Место локализации и/или происхождение
ALac	Лактоферрин	Л	П	Роговица
AOAPP	Одонтогенный амебобласт-ассоциированный белок	Л	П	Одонтогенная раковая опухоль
ASem1	Семеногелин 1	Л	П	Семенные сосуды
AEnf	Энфурвитид	Л	П	Ятрогенного происхождения
ACatK****	Катепсин К	Л	П	Онкогенного происхождения

Сейчас в особенности широко изучается вопрос амилоидозов в свете таких заболеваний, как болезни Альцгеймера, Паркинсона, Хантингтона, амиотрофический латеральный склероз, разные деменции (Ke *et al.*, 2020). Более того, помимо центральной нервной системы, найденные амилоидные агрегаты описаны и для заболеваний других систем органов. Среди них наиболее известными являются: диабеты второго и первого типов, апополипротеин А-ассоциированные амилоидозы I, II, IV; амилоидозы легких и тяжелых цепей иммуноглобулина (Westermarck *et al.*, 2018).

### 1.3.2. Функциональные амилоиды

На сегодняшний день функциональные амилоиды наиболее изучены у прокариот и низших эукариот, однако известны и некоторые функциональные амилоиды высших эукариот, в том числе и млекопитающих. Например, из белков млекопитающих выделяют многие пептидные гормоны, аккумулирующиеся в секреторных гранулах (Maji *et al.*, 2009b); (Табл.2).

На данный момент показано, что функциональные амилоиды прокариот являются структурными компонентами многих внеклеточных структур таких, как биопленки, обеспечивающие защиту, адгезию клеток и даже цитотоксичность в случае патогенных видов. Кюрлины (CsgA-CsgF (Chapman *et al.*, 2002)) *E. coli* и чаплины (ChpA-H (Claessen *et al.*, 2003, Yang *et al.*, 2017)) *S. coelicolor* - первые и наиболее известные функциональные амилоиды бактерий.

Образование внеклеточных фимбрий, как оказалось, является достаточно консервативным процессом. Комплексы белков кюрлинов обнаруживаются как минимум в четырех различных филах бактерий (Dueholm *et al.*, 2012). Это позволяет задуматься об эволюционной консервативности и, соответственно, функциональной значимости таких агрегатов. Основным компонентом амилоидных фибрилл фимбрий является белок CsgA. Хотя и в определенных условиях данный белок способен к автономной фибриллизации, в клетках бактерий этот процесс опосредован еще одним заякоренным на внешней мембране белком комплекса - CsgB. Помимо перечисленных участников, система образования фимбрий включает в себя еще 4 белка, особенно важным из которых является белок порового комплекса внешней мембраны CsgG. Вместе с другими белками порового комплекса (CsgF, CsgE) данный белок обеспечивает экспорт уже белков CsgA-B на поверхность клетки, где CsgB в комплексе с CsgA активирует фибриллизацию с образованием видимых в микроскоп внеклеточных структур. Оставшиеся белки, кодируемые данными оперонами, отвечают за: поддержание CsgA и CsgB в мономерной форме (CsgC, CsgE) (Biesecker *et al.*, 2018) и регуляцию транскрипции (CsgD). Сегодня данная бактериальная система используется для проверки амилоидогенных свойств многих других белков и является одним из методов их исследования, получившего название C-DAG (Curli-dependent amyloid generator)

(Sivanathan, Hochschild, 2013). Группа белков *S. coelicolor* - чаплины - участвует в создании гидрофобной мембраны, позволяющей образовываться воздушным гифам и участвующая в расселении спор бактерий. Список на сегодняшний день известных функциональных амилоидов прокариот приведен в таблице 2.

Также первыми описанными функциональными амилоидами эукариот являются белки гетерокариотической несовместимости грибов вида *Podospora anserina* het-s и het-S. Встреча гифов, несущих разные аллели het-s локуса (het-s и het-S), приводит к началу клеточной смерти. Однако клеточная смерть возможна только в случае, если het-s имеет амилоидную конформацию - HET-s. Кроме этого, показаны прионные свойства HET-s, так как он приводит к агрегации нормальной формы белка het-s партнера при слиянии гифов. Клеточная смерть происходит ввиду взаимодействия прионов HET-s и HET-S через структуру  $\beta$ -соленоида и нарушению конформации HeLo домена HET-S белка. HeLo домен приобретает форму спирали и выставление его трансмембранного домена приводит к появлению отверстий в мембране, поэтому способствует образованию трансмембранной спирали и олигомеризации белка. Прионные свойства HET-s и HET-S наблюдаются и у гомологов данных белков. И, кроме этого, HET-s и HET-S *P. anserina* сохраняют прионные свойства в клетках других видов, в том числе и в дрожжах (Otzen, Riek, 2019).

Таким образом, сегодня широко обсуждается вопрос о функциональности не только амилоидов, но и прионов. Однако доказана она в немногих случаях, что делает это направление актуальным для дальнейшего изучения.

**Таблица 2** — Список известных функциональных амилоидов;\* — не показано образование амилоидных структур *in vivo*, \*\* — функциональность связана с прионной формой (Otzen, Riek, 2019)

Виды	Белок	Функция	Источник
<b>Прокариоты, бактерии</b>			
<i>Escherihia coli, Salmonella spp.</i>	Курлины (CsgA, CsgB, CsgF), тафи	Входят в состав биопленок; взаимодействуют с экзополисахаридами, способствуя клеточной адгезии и инвазии в клетки хозяина	Chapman <i>et al.</i> , 2002, Jain, Chapman, 2019
<i>Bacillus subtilis</i>	белок TasA	Входит в состав биопленок, участвуя в адгезии и инвазии в хозяйские клетки, повышает адаптивные способности клеток к изменению внешней среды	Romero <i>et al.</i> , 2010, Romero <i>et al.</i> , 2011, Chai <i>et al.</i> , 2013
<i>Micobacterium tuberculosis</i>	Mtr семейство белков	Пилли-формирующие белки, участвующие в образовании биопленок, клеточной коммуникации; связаны с развитием многочисленных инфекций человека за счет своей адгезионной способности к внеклеточному матриксу клеток человека (иммунным клеткам, эндотелио- и эпителиоцитам и др.)	Alteri <i>et al.</i> , 2007, Pillay, 2015
			Продолжение на следующей странице

Виды	Белок	Функция	Источник
<i>Streptomyces coelicolor</i>	Чаплины и родлины (ChpA-H, ChpD-H, Rdl A и Rdl B)	Входят в состав амфипатической оболочки спор, снижая поверхностное натяжение на границе раздела фаз воздух-жидкость, участвуют в образовании, распространении спор и адгезии к клеткам	Claessen <i>et al.</i> , 2003, Yang <i>et al.</i> , 2017
<i>Klebsiella pneumonia</i>	Микроцин E492 (Mcc)	Антимикробная активность при переходе к олигомерной и/или мономерной структурам в щелочной среде с низкой концентрацией солей, стабилизация биопленок в состоянии амилоидных фибрилл при pH < 8.5 и mM(NaCl)	Destoumieux-Garzón <i>et al.</i> , 2003, Bieler <i>et al.</i> , 2005, Shahnawaz, Soto, 2012
<i>Staphylococcus aureus</i>	Модулин (phenol-soluble modulins $\alpha 3$ , PSM $\alpha 3$ )	Участвует в образовании биопленок, устойчивых к диспергированию сурфактантом и механическим воздействиям	Schwartz <i>et al.</i> , 2012, Schwartz, Boles, 2013, Zheng <i>et al.</i> , 2018

#### Одноклеточные эукариоты, протисты

<i>Plasmodium falciparum</i>	белок MSP2*	Обеспечивает инвазию мерозоитов в эритроциты	Adda <i>et al.</i> , 2009
------------------------------	-------------	--	---------------------------

#### Низшие эукариоты, грибы

Продолжение на следующей странице

Виды	Белок	Функция	Источник
<i>Neurospora crassa</i> , <i>Penicillium spp.</i> и др. филаментозные грибы	Гидрофобины (ESA, DewA и др.)	Образуют амфипатическую оболочку вокруг гиф, способствуют спорообразованию, регулируют поверхностного натяжения, клеточной адгезии	Wösten, Vocht, 2000, Macindoe <i>et al.</i> , 2012, Morris <i>et al.</i> , 2013, Lo <i>et al.</i> , 2014
<i>Podospora anserina</i>	НЕТ-S белки	Регулируют образование гетерокариона и запрограммированную клеточную смерть в зависимости от своего состояния (прионной или растворимой формы)	Coustou <i>et al.</i> , 1997, Riek, Saupe, 2016

#### Высшие многоклеточные эукариоты, Насекомые и пауки

<i>Bombux mori</i>	Белки хориона	Образуют защитную оболочку яйца и ооцитов, участвуют в обеспечении газообмена развивающегося эмбриона и способствуют оплодотворению ооцитов	Iconomidou <i>et al.</i> , 2000, Iconomidou <i>et al.</i> , 2001, Iconomidou <i>et al.</i> , 2006
<i>Drosophila melanogaster</i>	Orb2 (является членом семейства СРЕВ белков)	Участвует в переходе информации из кратковременной в долговременную память в нейронах (т.е. необходим для консолидации памяти)	Oroz <i>et al.</i> , 2020

#### Высшие многоклеточные эукариоты, Рыбы и амфибии

<i>Austrofundulus limnaeus</i>	Белки хориона	Способствуют выживанию эмбрионов в периоды засухи	Podrabsky <i>et al.</i> , 2001
<i>Pachimedusa dactylophora</i>	Анионный дермасептин (aDsp PD3-7)	Образует токсичные агрегаты пептида и участвует во врожденном иммунитете*	Göbller-Schöfberger <i>et al.</i> , 2012

Продолжение на следующей странице

Виды	Белок	Функция	Источник
<b>Высшие многоклеточные эукариоты, Млекопитающие</b>			
<i>Homo sapiens, Mus musculus, Bos taurus, other mammals</i>	С-терминальный домен М $\alpha$ белка Pmel17	В амилоидной форме является субстратом для отложения пигмента в меланосомах во время их созревания и снижает цитотоксическое влияние предшественников меланина (индол-5,6-хинон и его производных)	Fowler <i>et al.</i> , 2006, Watt <i>et al.</i> , 2013, Bissig <i>et al.</i> , 2016
<i>H. sapiens</i>	Белки RIP1, RIP3	Ингибируют факторы апоптоза и каспазы, направляя клетки на гибель путем некроптоза	Li <i>et al.</i> , 2012
<i>M. musculus</i>	Белки группы CRES	Входят в состав оболочки эпидидимиса и участвуют в созревании и защите сперматозоидов	Whelly <i>et al.</i> , 2012, Do <i>et al.</i> , 2019
<i>Rattus norvegicus</i>	Белок FXR1	Препятствует деградации молекул мРНК под воздействием РНКазы А в нейронах	Sopova <i>et al.</i> , 2019

## 1.4. Белки нуклеопорины

### 1.4.1. История открытия и описания ядерно-порового комплекса

Впервые ядерный поровый комплекс (NPC, nuclear pore complex) был описан в 1949 году в исследованиях ядерной мембраны ооцитов амфибий (Callan *et al.*, 1949). Данные исследования были осуществлены с помощью методов электронной микроскопии, разрешающая способность которой не могла дать подробного описания структуры порового аппарата ядра. Сам термин «поровый комплекс» уже использовался некоторыми исследователями, например, М. Ватсоном в своих работах (Watson, 1959), однако на тот момент не был признан официальным. Позднее, в 1967 году, была описана первая ортогональная структура ядерного порового комплекса. Согласно данной модели NPCs имеют сквозную цилиндрическую форму с восьмилучевой симметрией вдоль оси, перпендикулярной мембране. NPCs располагаются в особых участках видоизмененной ядерной мембраны, характеризующиеся множественными слияниями внутренней и внешней ядерных мембран. Данное слияние образует специальную «пору» шириной около 80 нм, внутренняя часть которой включает один комплекс белков NPC (Watson, 1959). Позднее была описана более детальная структура ядерного порового комплекса (Szymborska *et al.*, 2013, Löscherberger *et al.*, 2014).

### 1.4.2. Структурные особенности ядерной поры и её функции

Наиболее важные внутриклеточные процессы в эукариотических клетках напрямую зависят от постоянной передачи множества молекул посредством ядерно-цитоплазматического транспорта. Главным посредником данного процесса являются NPC, которые представляют собой огромный макромолекулярный комплекс. Основными их структурными компонентами являются белки ядерной поры - нуклеопорины.

Ядерный поровый комплекс имеет молекулярную массу примерно в 60 МДа в клетках дрожжей и 120 МДа в клетках человека и состоит из 8-64 копий около 30 разных нуклеопоринов, которые обладают высокой консервативностью в царстве эукариот (Kim *et al.*, 2018). В составе ядерного порового комплекса выделяют следующие принципиальные элементы (Рис.2):

- комплекс нуклеопоринов внешних колец, который включает ядерное и цитоплазматическое кольца поры;
- комплекс внутреннего кольца, включающий внутреннее кольцо поры: нуклеопорины скаффолда (Nup53 (hsNup35), Nic96 (hsNup93), Nup170/157 (hsNup155),

Nup192 (hsNup205), Nup145N (hsNup98) и Nup188 (hsNup188)) и нуклеопорины CNT-комплекса (Nup57 (hsNup54), Nup49 (hsNup58) и Nsp1 (Nup62));

- комплекс нуклеопоринов периферических элементов, состоящий из ядерной корзинки и цитоплазматических филаментов.

Первые три элемента образуют так называемый симметричный кор (скаффолд), для которого помимо восьмилучевой симметрии показано и наличие двулучевой симметрии через ядерную мембрану (Рис.2). А ассоциированные с ядерным и цитоплазматическим кольцами ядерная корзинка и цитоплазматические филаменты, соответственно, образуют асимметричный периферический компонент. Таким образом, ядерный поровый комплекс принято считать псевдосимметричным (Kim *et al.*, 2018).

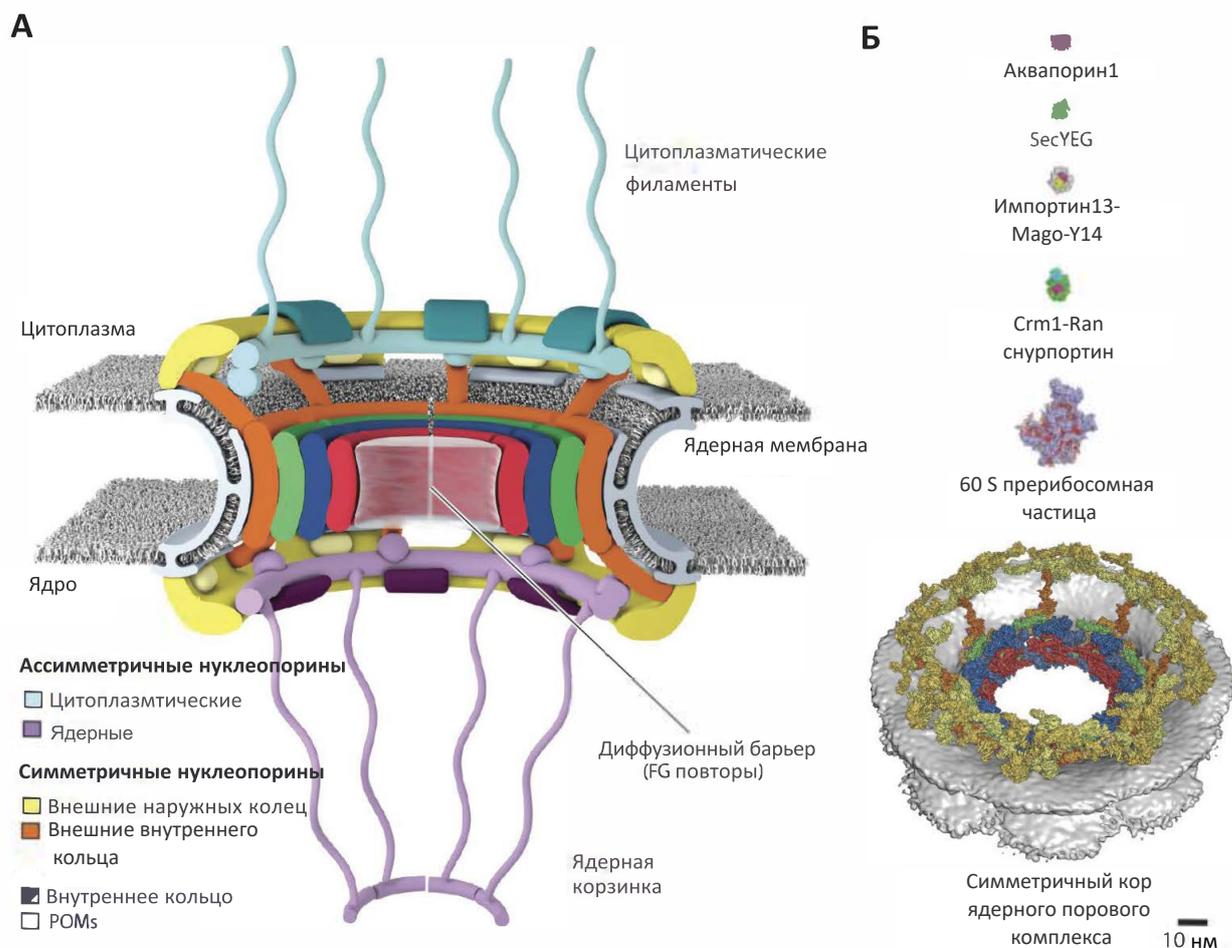
Хотя почти все нуклеопорины входят в состав только конкретного модуля, некоторые, например, NUP98 и NUP62, обладают множественной локализацией среди модулей либо ввиду того, что являются структурными единицами множественных модулей, либо за счет того, что связывают эти множественные модули (Kim *et al.*, 2018).

### **Комплекс внешних нуклеопоринов (CNC/Y-complex)**

CNC (Coat Nucleoporin Complex) или комплекс внешних нуклеопоринов объединяет в себе все внешние кольца порового комплекса, таким образом, являясь местом прикрепления других нуклеопоринов. Впервые структура Y-комплекса была получена с помощью электронной микроскопии изолированного CNC комплекса дрожжей (Siniosoglou *et al.*, 2000, Lutzmann *et al.*, 2002), которая напоминала форму латинской буквы 'Y' и оказалась консервативной среди других организмов, в том числе и человека (Vui *et al.*, 2013). Для обозначения данного комплекса приняты не только приведенные классические названия, но также и основанные на главном структурном нуклеопорине комплекса в зависимости от организма. Так, для дрожжей его также принято называть комплексом Nup84, а для организмов из царства животных - комплекс NUP107-NUP160 (Stuwe *et al.*, 2015a).

Так как структура CNC является наиболее описанной, то и входящие в него нуклеопорины известны для разных видов. И примечательно, что структурный состав Y-комплекса может отличаться не только среди царств, но и среди таксономических групп более низкого уровня, в том числе и среди разных видов (Cronshaw *et al.*, 2002, Fontoura *et al.*, 1999, Belgareh *et al.*, 2001, Franz *et al.*, 2007, Rasala *et al.*, 2006, Loïodice *et al.*, 2004). То есть существуют как и консервативные белки, входящие в состав CNC у разных организмов, так и специфические для определенной группы. Например, NUP43 является строго специфичным компонентом для животных (Neumann *et al.*, 2010).

В Y-комплексе выделяют несколько компонентов. И особо интересным является преобладание  $\beta$ -структур в составе входящих в него белков. Это является любопытным, так как для амилоидных и амилоидоподобных белков основным структурным компонентом является кросс- $\beta$ -структура, в том числе и необычные  $\beta$ -соленоиды. Хотя для нуклеопоринов Y-комплекса и не показана именно такая структурная особенность, но наличие практически



**Рисунок 2 — Структура ядерного порового комплекса.** На схеме представлена полученная с помощью метода криоэлектронной микроскопии молекулярная структура (А) ядерного порового комплекса с указанием элементарного и модульного строения и (Б) центрального кора человеческого ядерного порового комплекса в сравнении с размерами других транспортеров (человеческий Аквaporин-1 и бактериальный SecYEG) и комплексов транспортируемых макромолекул: с рецепторами ядерного транспорта (комплексы Импортин13-Mago-Y14 и Crml-Ran-снурпортин *Drosophila melanogaster*), не связанных с факторами ядерного транспорта макромолекул (прерибосомная 60S частица *Saccharomyces cerevisiae*). Рисунок взят из статьи Д. Лина 2019 г. (Lin, Hoelz, 2019).

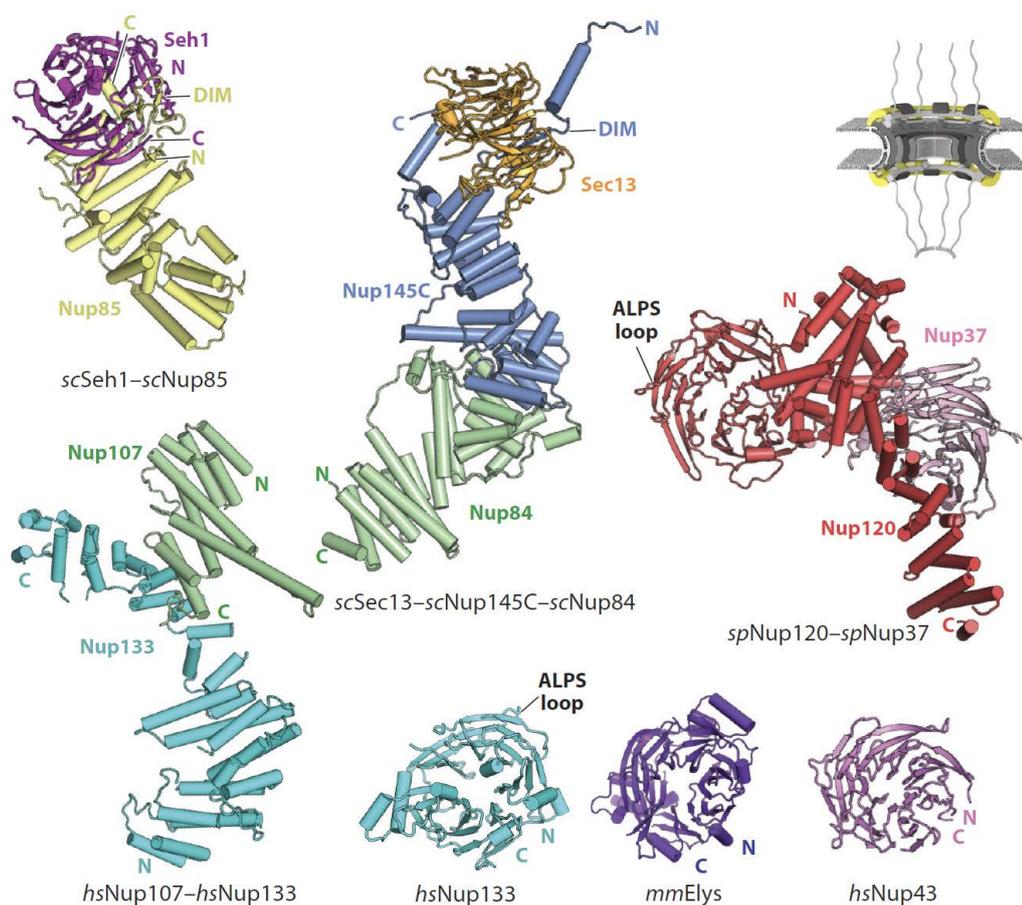
во всех таких  $\beta$ -структур, как  $\beta$ -пропеллеры (например, гетеротример Sec13-Nup145C-Nup84 (hsSec13-hsNup96-hsNup107)), позволяет предполагать возможность этих нуклеопоринов образовывать амилоидоподобные структуры как нативно, так и в искусственных условиях (Brohawn, Schwartz, 2009, Nagy *et al.*, 2009).

Данный комплекс состоит из нескольких белков, которые формируют отдельные пространственные мультимерные структуры (Stuwe *et al.*, 2015a):

- «рука» Y-формы, образованная дрожжевыми белками Nup120 и Nup85 (человеческие гомологи hsNup160, hsNup75);
- длинная нижняя «ножка», состоящая из нуклеопоринов Nup145C, Nup84, Nup133 (hsNUP96, hsNup107, hsNup133, соответственно);

- связывающие нуклеопорины Sec13, Seh1;
- мембран-ассоциированные нуклеопорины Nup120 и Nup133 (hsNup160, hsNup133)

Необходимо также заметить, что в данном модуле CNC дрожжей белок Nup145 представлен только С-концевым доменом, тогда как его N-концевой домен участвует в образовании множественных соединений между разными модулями. В том числе и взаимодействует со своим паралогом другого модуля Nup100. Однако эта особенность - функциональная дифференциация комплекса путем дубликаций - характерна только для дрожжей, в то время как у других организмов, даже у других видов грибов, отличающиеся от функций ортологов Nup145 структурные свойства ядерно-цитоплазматического комплекса обеспечиваются иными не паралогичными белками (Lin, Hoelz, 2019).



**Рисунок 3 — Структура нуклеопоринов CNC модуля (Y-комплекса).** На данном рисунке представлены субкомплексы Y-модуля с основными нуклеопориными, входящими в их состав. Локализация данного модуля внутри ядерной поры показана сверху справа на схематическом изображении комплекса ядерной поры. ALSP - амфипатический сенсор липидной упаковки, CNC - комплекс внешних нуклеопоринов, DIM - домен инвазионного мотива, NPC - ядерный поровый комплекс, hs - *Homo sapiens*, sc - *Saccharomyces cerevisiae*, mm - *Mus musculus* (Lin, Hoelz, 2019).

Помимо данных нуклеопоринов для человека показано наличие еще трех нуклеопоринов: hsNup37, hsNup43, ELYS - однако они не встречаются в NPC грибов или же отличаются трехмерной структурой (Рис.3). Эта архитектурная особенность

объясняется отличным от царства грибов митотическим процессом. У организмов царства животных митоз происходит по типу открытого митоза, тогда как для грибов показан процесс закрытого или полуоткрытого митоза. Эта особенность может объяснять отличия белка млекопитающих ELYS от такового некоторых видов грибов, а именно *Schizosaccharomyces pombe* и *Chaetomium thermophilum*. Данный белок прикрепляет CNC к хроматину для инициации постмитотической сборки NPC. И, вероятно, данный процесс является необязательным для видов грибов с закрытым и полуоткрытым митозом (Franz *et al.*, 2007, Rasala *et al.*, 2006).

### Комплекс внутреннего кольца

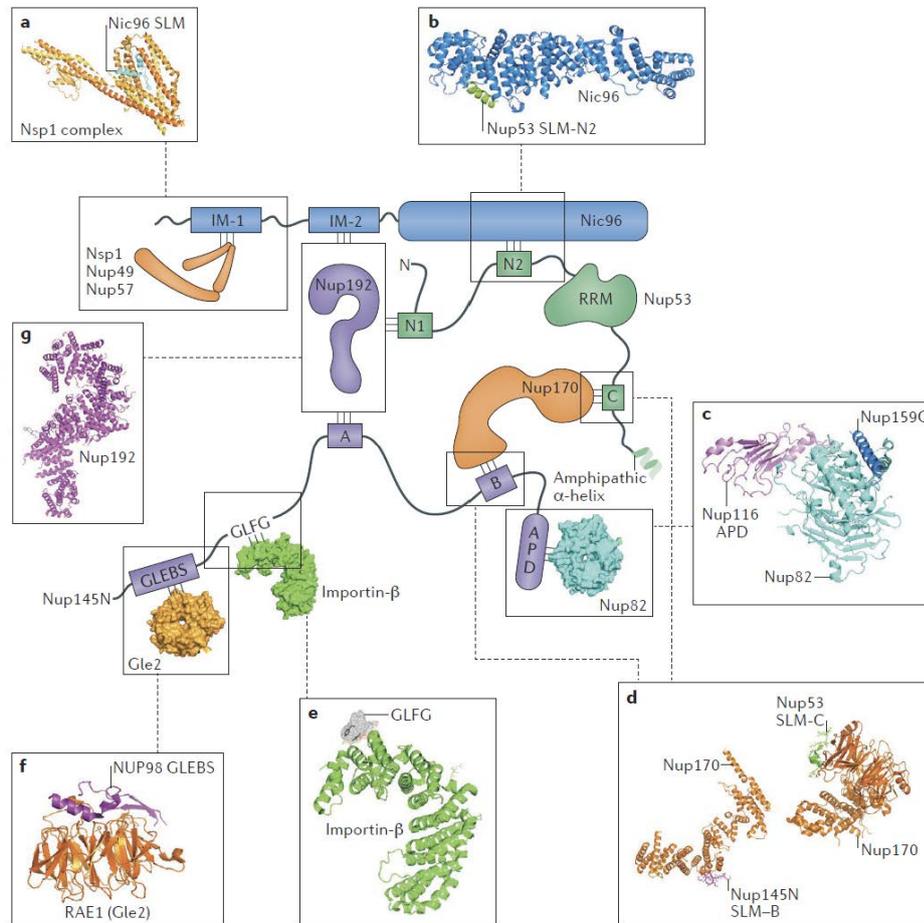
Структура модуля внутреннего кольца является более вариабельной среди разных видов (Lin, Hoelz, 2019). Она впервые была получена для белков *Chaetomium thermophilum* (Lin *et al.*, 2016).

Однако данная архитектура не оказалась универсальной. Полученная после структура модуля внутреннего кольца для *S. cerevisiae* позволила установить необычные особенности. Так, для *S. cerevisiae* показано присутствие видоспецифичных генных дупликаций, кодирующих паралоги с частично перекрывающимися функциями, которые не обнаруживаются у других модельных организмов. Nup157 и Nup170 являются паралогами друг друга и ортологами человеческого нуклеопорина hsNup155, Nup53 и Nup59 - hsNup53, а Nup100, Nup116 и Nup145N - hsNup98. Это, в свою очередь, приводит к значительному усложнению структуры комплекса внутреннего кольца (inner ring complex, IRC) и сложности поиска функциональных гомологов среди разных видов (Stuwe *et al.*, 2015a).

В отличие от Y-комплекса, в модуле внутреннего кольца нуклеопорины обладают множественными сайтами взаимодействия друг с другом, таким образом образуя сеть взаимодействий (Stuwe *et al.*, 2015b, Lin *et al.*, 2016) (Рис.4):

1. с нуклеопорином внешнего комплекса CNC Nup145C;
2. с нуклеопоринами цитоплазматических филаментов Nup82, Gle2.

Причем внутри сети взаимодействий Nup145N образование димеров является конкурирующим и, например, Nup188 конкурирует с Nup192 и Nup170 за сайт связывания с Nup145N (Lin *et al.*, 2016, Amlacher *et al.*, 2011). Однако данное множественное взаимодействие показано лишь для некоторых видов грибов, тогда как ортологи Nup145N млекопитающих, насекомых и даже его гомологи грибов других классов не обладают такой сложной сетью взаимодействий. И, таким образом, другие имеют более упрощенную структуру IRC.



**Рисунок 4 — Структурные взаимодействия нуклеопоринов модуля внутреннего кольца и роль SLM в их стабилизации.** На рисунке схематически показано взаимодействие нуклеопоринов внутри модуля внутреннего кольца. Нуклеопорины скаффолда Nup192, Nup170 и Nsp1-комплекса (CNT) являются основными структурными белками модуля внутреннего кольца ядерной поры. Целостность IRC обеспечивают другие нуклеопорины данного модуля: Nic96, Nup145N и Nup53. Помимо взаимодействий между нуклеопоринами на схеме представлены сайты контакта с белками ядерно-цитоплазматического транспорта: GLEBS (GLE2-binding sequence), GLFG - и сайт узнавания РНК RRM (RNA-recognition motif). Структуры отдельных компонентов (a-f) из статьи Д. Лина 2019 г. (Lin, Hoelz, 2019).

### 1.4.3. Транспорт через ядерную пору

Ядерно-цитоплазматический транспорт способен переносить разные макромолекулы: от малых молекул за счет диффузии до импорта белков, экспорта РНК и рибонуклеопротеинов за счет механизма активного транспорта. Основной функцией комплекса пор ядерной мембраны, как уже было упомянуто выше, является обеспечение ядерно-цитоплазматического транспорта. В отличие от других мембранных транспортеров, ядерные поровые комплексы способны осуществлять транспорт макромолекул, даже имеющих сложную трехмерную структуру. Это позволяет макромолекулам немедленно вовлекаться в разные функциональные процессы внутри ядра и цитоплазмы.

Процесс транспорта макромолекул обеспечивается каналом диффузии NPC. В его формировании участвует особая подгруппа белков ядерной поры - FG-обогащенные нуклеопорины. FG-обогащенные домены белков являются неупорядоченными и ввиду своего аминокислотного состава обладают сродством к NTR (nuclear transport receptors, рецепторам ядерного транспорта) (Bayliss *et al.*, 2000) и NTF2 (nuclear transport factor 2, фактор ядерного транспорта 2) (Milles *et al.*, 2015), позволяя им связываться с каналом диффузии NPC и таким образом осуществлять перенос ассоциированных с NTR макромолекул (Milles *et al.*, 2015).

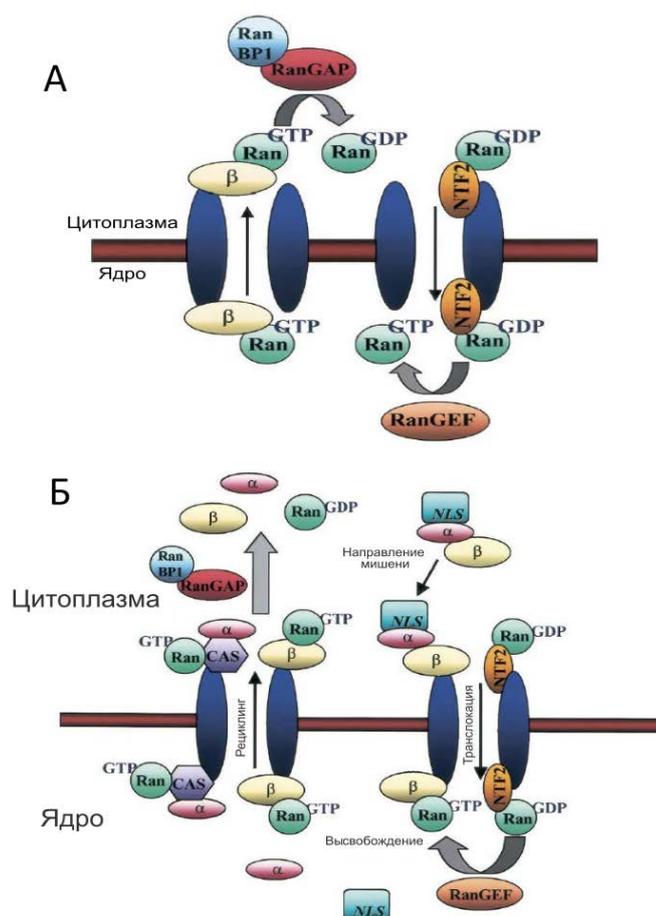
Ядерно-цитоплазматический транспорт осуществляется с помощью нескольких компонентов (Beck, Hurt, 2017):

- рецепторов ядерного транспорта (NTR), связывающихся со специальными сигнальными аминокислотными последовательностями транспортируемых макромолекул, NLS (nuclear localization signal, сигнал ядерной локализации) и NES (nuclear export signal, сигнал ядерного экспорта) в случае белков - белки экспортины и импортины семейства кариоферринов;
- малой ГТФазы Ran;
- факторов рециклинга - второй фактор ядерного транспорта (NTF2) и экспортин CAS.

### **Транспорт молекул за счет рецепторов ядерного транспорта**

NTR импортин- $\alpha$  - адапторный белок. Он связывается с NLS транспортируемого белка NLS-связывающим карманом, который обогащен так называемыми армадилло повторами. После фактор транспорта импортин- $\beta$  связывается с образовавшимся комплексом и ускоряет его перенос через NPC. Причем наличие адаптора не является обязательным и многие рецепторы ядерного транспорта способны напрямую связываться с NLSs таких макромолекул, как коровых гистонов, рибосомных белков или некоторых РНК-связывающих белков. Процесс ядерного транспорта является энергетически затратным и потому зависит от градиента RanGDP. Образующийся комплекс кариоферринов и белков или РНК, прежде чем катализировать транспорт макромолекул через центральный канал, подходит к ядерной мембране и связывается с RanGDP. Это позволяет ему «заякориться» в NPC. С ядерной стороны специальный фактор RanGEF способствует обмену GDP на GTP и образование RanGTP приводит к высвобождению связанного белка в нуклеоплазму. При этом высвобождающийся адаптор импортин- $\alpha$  возвращается в цитоплазму за счет фактора рециклинга CAS, а образующийся димер RanGTP-импортин- $\beta$  сохраняется для последующего экспорта в цитоплазму, где уже диссоциирует за счет гидролиза RanGTP с помощью специального компонента системы поддержания градиента RanGDP RanGAP (Quimby, Corbett, 2001) (Рис.5).

Экспорт макромолекул, несущих NES сигнальные последовательности осуществляется также за счет RanGTP зависимого механизма, только факторами связывания с сигнальными последовательностями выступают белки экспортины, и высвобождение происходит за



**Рисунок 5 — Общий механизм ядерно-цитоплазматического транспорта.** На схеме представлен механизм ядерного транспорта с показанным градиентом RanGTP: с помощью NTR импортина- $\beta$  и фактором рециклинга NTF2 (А) и пути импорта с помощью импортина- $\alpha$  с участием фактора рециклинга CAS (Б). Изображение взято из статьи Б. Квимби (Quimby, Corbett, 2001).

счет гидролиза RanGTP, сопровождающегося полным распадом образующегося комплекса RanGTP-NTR-груз (Matsuura, 2016).

Хотя большинство белков действительно транспортируется за счет связывания с их NLS и NES сигнальными последовательностями соответствующих рецепторов, некоторые макромолекулы переносятся за счет особенностей своих структур (например, обогащенные гидрофобными аминокислотами белки (Christie *et al.*, 2016, Frey *et al.*, 2018) или через белки-адапторы (например, описанный выше механизм транспорта с помощью импортина- $\alpha$ ) (Christie *et al.*, 2016, Matsuura, 2016).

#### 1.4.4. Барьерная функция ядерно-порового комплекса и значение FG-повторов

Помимо участия в ядерно-цитоплазматическом транспорте, NPC также является барьером для крупных белков с молекулярной массой выше 30 кДа. Все молекулы с

меньшей молекулярной массой транспортируются за счет диффузии. Однако крупные белки неспособны независимо проходить через NPC ввиду энергетических ограничений. Показано, что ядерно-цитоплазматический транспорт через диффузионный барьер не зависит от количества FG-повторов в нуклеопоринах канала, а скорость транспорта макромолекул определяется именно степенью их гидрофобности (Lin, Hoelz, 2019).

Домены FG-повторов представляют собой нативно неструктурированные участки (intrinsically disordered regions, IDR). Барьер проницаемости образует собой гидрогель, состоящий из доменов с FG-повторами (Frey, Görlich, 2009, Patel *et al.*, 2007). Данная модель была предложена Д. Горлихом и К. Риббек в 2001 году и была названа «модель селективной фазы» (Ribbeck, Görlich, 2002, Frey, Görlich, 2007, Ribbeck, Görlich, 2001). В качестве «селективной фазы» выступает FG-гидрогель, образованный доменами FG-повторов за счет взаимодействий между гидрофобными кластерами. Селективность обуславливается тем, что размер каждой ячейки внутри геля устанавливает предел размера транспортируемого комплекса 30 кДа. А транспорт макромолекул, превышающих данный размер, осуществляется с помощью NTRs. Данное свойство было описано в модели Д. Горлиха и К. Риббек, и согласно их модели эта способность опосредована индукцией собственного прохождения комплексов за счет более высокого сродства NTRs к гидрофобным FG-повторам, чем интактных макромолекул. Последнее способствует «расширению» ячеек за счет диссоциации образованных FG-повторами сшивок. При этом освобожденное для комплексов NTRs с макромолекулами пространство способно повторно замыкаться после прохождения комплекса, что препятствует нежелательному транспорту инертных макромолекул (Lin, Hoelz, 2019).

Предложенная модель селективной фазы была впоследствии доказана в экспериментах с FG/FxFG доменами Nsp1 и также с GLFG доменами Nup49, Nup57 (Frey *et al.*, 2006, Frey, Görlich, 2007, Frey, Görlich, 2009). Свойства проницаемости восстановленных комплексов были аналогичны проницаемости, показанной для самих комплексов ядерных пор. При этом было продемонстрировано, что именно фенилаланины играют основную роль в формировании FG-гидрогеля и также в связывании с NTRs (Frey *et al.*, 2006).

Механизм импортин-опосредованного проникновения заключается в том, что импортин способны либо усиливать диссоциацию межповторных контактов после предварительного их ослабления, либо напрямую вызывать диссоциацию межповторных контактов (Kustanovich, Rabin, 2004). Результатом обоих механизмов является снижение необходимой энергии активации для последующего разрушения других межповторных контактов.

## 1.4.5. Структурные особенности белка Nsp1 и Nup100

### Функции доменов нуклеопорина Nsp1

Особый интерес среди белков FG-гидрогеля вызывают нуклеопорины Nsp1 и Nup100, так как ранее для них были показаны амилоидные свойства (Ader *et al.*, 2010, Labokha *et al.*, 2013). Дрожжевой нуклеопорин Nsp1 является основным структурным компонентом комплекса гетеротримера нуклеопоринов (CNT) или канала диффузии, также названного по превалирующему белку Nsp1-комплексом. Данный белок согласно разным источникам представлен в виде 32 копий внутри комплекса ядерной поры (Cronshaw *et al.*, 2002, Terry, Wentе, 2009). Аналогично другим нуклеопоринам в составе данного гетеротримера - Nup49 и Nup57 - Nsp1 обладает следующими двумя основными характерными доменами (Devos *et al.*, 2006):

- N-терминальным доменом FG-повторов, разделенного NQ-обогащенными последовательностями;
- C-терминальной суперспиралью.

Каждый из этих модулей выполняет определенные функции как в структуре внутреннего канала, так и в осуществлении транспорта сквозь ядерную пору.

В случае Nsp1 IDR составляют участок с 1 по 600 аминокислоту (Devos *et al.*, 2006). Данные участки обогащены гидрофобными аминокислотами фенилаланином и глицином, которые играют важную роль в образовании селективного барьера и являются сайтами взаимодействия с рецепторами ядерного транспорта кариоферинами (NTRs). Однако хоть для участков FG-повторов и показана гомологичность аминокислотного состава среди разных нуклеопоринов, существуют различия в аминокислотных последовательностях повторов (Denning, Rexach, 2007). Так, в случае Nsp1 данные домены образованы FG/FxFG-повторами (Hurt, 1988, Rout, Wentе, 1994, Buss *et al.*, 1994), в то время как для Nup49 и Nup57 данные домены образованы GLFG-повторами (Frey, Görlich, 2009).

Важно отметить, что была показана зависимость прочности и селективности центрального канала от окружающих повторы NQ-обогащенных спейсерных участков. Образуемая ими особая структура обеспечивает прочность и непроницаемость гидрогеля для так называемых инертных макромолекул, транспорт которых не связан с кариоферинами (Ader *et al.*, 2010). В случае Nsp1 гидрофильные NQ-обогащенные участки регулируют размер центральной части FG-гидрогеля, увеличивая его площадь внутри NPC (Alber *et al.*, 2007, Реуго *et al.*, 2015).

C-терминальный домен Nsp1 является стабилизатором образуемого барьера диффузии и препятствует укладке NQ-обогащенных участков в амилоидоподобные  $\beta$ -структуры (Ader *et al.*, 2010). C-терминальный участок Nsp1 состоит из четырех субдоменов: CCR1 с 630 по 665, CCR2 с 680 по 730, CCR3 с 740 по 785 и CCR4 с 790 по 823 аминокислоты. Каждый из этих доменов образует отличительную вторичную структуру -  $\alpha$ -суперспираль. И, например,

домен суперспирали CCR2 способен взаимодействовать с разными Nsp1-содержащими комплексами, то есть с FG-гидрогелями (Bailer *et al.*, 2001). Однако более подробный механизм их конкурентного связывания с CCR2 остается до конца не изученным.

### **Роль Nsp1 в распределении ядерно-поровых комплексов при делении**

Кроме этого, помимо роли в структуре NPC и ядерно-цитоплазматическом транспорте, для дрожжей *S. cerevisiae* показано, что Nsp1 отвечает за передачу и наследование комплекса ядерной поры при делении (Colombi *et al.*, 2013). Во время деления Nsp1 связывается с эндоплазматическим ретикуломом, ассоциированным с образующейся дочерней клеткой дрожжей, и за счет образования нуклеоподий (вытянутых выростов ядерной мембраны и нуклеоплазмы) обеспечивает передачу NPCs в дочерние клетки. Данный механизм зависит также от актинового цитоскелета и участия дрожжевого миозина Myo2, и в случае делеции как Nsp1, так и Myo2 передача ядра в дочерние клетки не происходит. Предполагается, что Nsp1 способствует преодолению диффузионного барьера (Shcheprova *et al.*, 2008), образующегося на месте перемычки между материнской и дочерней клетками дрожжей.

### **Функции доменов нуклеопорина Nup100**

Нуклеопорин дрожжей Nup100 является, как и Nsp1, неструктурированным FG-обогащенным белком. Он входит в состав модуля внутреннего кольца ядерного порового комплекса, т.е. IRC, и располагается с цитоплазматической стороны. Важно отметить, что у дрожжей рода *Saccharomyces* есть три нуклеопорина, гены которых являются паралогами: Nup100, Nup116, Nup145N. И в ходе эволюции эти белки приобрели разное местоположение внутри комплекса ядерной поры и разные функции. Nup100, Nup116 преимущественно располагаются с цитоплазматической стороны, а Nup145N - со стороны кариоплазмы. Для данных паралогов показаны отличия в доменном составе. В отличие от Nup145N, Nup116 и Nup100 не образуются в ходе протеолитического расщепления предшественника. Nup145N является продуктом протеолиза полноразмерного Nup145. Удивительный факт, что среди этих белков только Nup116 содержит сайт связывания с фактором ядерного транспорта Gle2 - GLEBS (Lin, Hoelz, 2019).

Среди FG-обогащенных белков для многих, в том числе и Nsp1, показано отсутствие упорядоченных структур и их FG-обогащенные участки принято называть неупорядоченными (IDR). В то время как для Nup100, Nup116 и Nup145N в некоторых исследованиях была показана глобулярная структура (Yamada *et al.*, 2010).

N-терминальный домен Nup100 (1-610 аминокислота) обогащен GLFG-повторами (один из типов FG-повторов). Также этот участок характеризуется низким содержанием заряженных аминокислотных остатков (Yamada *et al.*, 2010). Еще одним отличием GLFG-обогащенного домена Nup100 от FG/FxFG-обогащенного участка Nsp1 является то, что ввиду своей высокой гидрофобности он имеет сродство к большему количеству

кариоферинов (Allen *et al.*, 2001, Allen *et al.*, 2002, Huang *et al.*, 2002). Располагаясь с цитоплазматической стороны внутреннего кольца ядерной поры, Nup100 играет важную роль в осуществлении импорта макромолекул. В ряде исследований показано, что GLFG-обогащенный участок Nup100 участвует в кариоферин-опосредованном транспорте макромолекул и разных типов РНК. Кроме этого, было показано, что Nup100 также важен для экспорта молекул именно мРНК, связываясь со специальным рецептором дрожжей Mex67 (Strawn *et al.*, 2001).

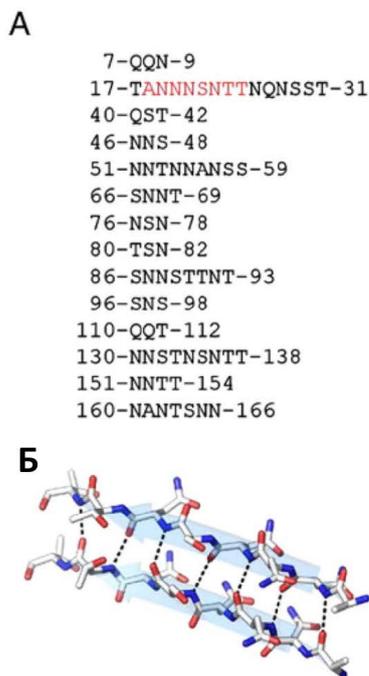
Следующий за ним участок с 611 по 800 аминокислоту имеет заряженный состав аминокислотных остатков и считается неупорядоченным. Этот домен разделяет GLFG-обогащенный участок и NPC-заякоривающий домен. И, таким образом, обеспечивает правильное расположение доменов центрального канала, образованных N-терминальными доменами Nup100, внутри комплекса ядерной поры, диаметр канала и процесс ядерно-цитоплазматического транспорта (Yamada *et al.*, 2010, Peyro *et al.*, 2015). Примечательно, что неструктурированные положительно заряженные участки Nup100, аналогично его GLFG-обогащенному домену, способны взаимодействовать с некоторыми рецепторами ядерного транспорта. Эти взаимодействия не только обеспечивают ядерно-цитоплазматический транспорт, но и способствуют изменению диаметра центрального канала диффузии (Yamada *et al.*, 2010).

Участвуя в повторном экспорте зрелых молекул тРНК, Nup100 играет важную роль в регуляции времени жизни клеток дрожжей. За счет обеспечения экспорта зрелых тРНК Nup100 препятствует накоплению в ядре транскрипционного фактора Gcn4, активирующего гены, ответственные за регуляцию репликативного старения клеток дрожжей (Lord *et al.*, 2017). Однако происходит ли контроль экспорта тРНК напрямую с помощью Nup100 или опосредованно, остается загадкой, и требуются дальнейшие исследования молекулярного механизма экспорта тРНК.

#### 1.4.6. Амилоидные свойства Nsp1

Применение твердотельного ЯМР позволило показать, что внутри FG-гирогеля Nsp1 формируются внутримолекулярные  $\beta$ -листы за счет взаимодействия гидрофильных спейсерных участков внутри доменов FG-повторов, так называемых NQTS-обогащенных последовательностей. Примечательно, что образующаяся структура  $\beta$ -листов NQST-обогащенных участков сходна со структурной единицей амилоидных фибрилл (Chiti, Dobson, 2006) (Рис.6). Кроме этого, для NQ-обогащенных FG-гидрогелей белков Nup100 и Nup116 было показано окрашивание Тиофлавином Т, что так же подтверждает их амилоидную природу (Schmidt, Görlich, 2015). Также стоит отметить, что данные участки способны взаимодействовать с NQ-обогащенными последовательностями амилоидного белка Sup35 *in vitro*, что позволяет предположить амилоидоподобную структуру NQ-обогащенных

участков внутри доменов FG-повторов (Ader *et al.*, 2010). Но до сих пор проверка амилоидных свойств не была осуществлена *in vivo*.



**Рисунок 6** — Структурные модели  $\beta$ -листов NQST-обогащенного участка Nsp1 нуклеопорина, созданные с помощью метода HADDOCK (Frey *et al.*, 2006). На данном рисунке представлены (А) найденные в доменах FG-повторов нуклеопорина Nsp1 последовательности; (Б) полученная модель  $\beta$ -листа. Пунктирными линиями обозначены водородные связи внутри  $\beta$ -листов NQST-обогащенных спейсерных участков. Как показали исследования, в транспортные свойства FG-гидрогеля большой вклад вносят полярные взаимодействия серина (Ser) с аспарагином (Asp). Изображение было взято из работы Д. Горлиха и Х. Адера от 2010 года (Ader *et al.*, 2010).

### **Спорный вопрос о необходимости NQ-обогащенных доменов в образовании FG-гидрогеля белка Nsp1 и его ортолога Nup98**

Участие NQ-обогащенных участков в образовании FG-гидрогеля дрожжевого белка Nsp1 является объектом дискуссии по поводу функциональности амилоидоподобных структур внутри ядерного порового комплекса. С одной стороны, это показано для Nsp1 *S.cerevisiae* (Ader *et al.*, 2010), но эти результаты не подтверждаются на модели *Xenopus*'а (Labokha *et al.*, 2013). В исследовании Х. Адера с соавторами было показано, что внутри FG-гидрогеля, образованного Nsp1, наблюдаются характерные для амилоидов взаимодействия: гидрофобные между фенилаланином и метильными остатками боковых цепей и внутримолекулярное образование  $\beta$ -листов между аспарагин-обогащенными участками. Внутримолекулярные  $\beta$ -листы аспарагин-обогащенных спейсеров являются кинетически наиболее стабильными структурами FG-гидрогеля, а свойства когезии гидрогеля очень сходны с таковыми для аспарагин-глутамин обогащенных доменов дрожжевого приона Sup35p. Однако в работе А. Лабохи с соавторами амилоидоподобные свойства ортолога Nsp1 у *Xenopus* - Nup98 - не были подтверждены. Более того, роль

FG-обогащенных участков Nup98 в барьере обеспечивалась особым типом модификации - о-гликозилированием. Причем даже содержащие спейсеры FG-обогащенные участки Nup98 спокойно образовывали гидрогели и в их отсутствии. И такая модификация вовсе препятствовала образованию амилоидоподобных  $\beta$ -структур белка, согласно данным NMR спектроскопии. Данные результаты были продемонстрированы как для Nup98 *Xenopus*, так и для Nup98 других млекопитающих (Labokha *et al.*, 2013).

Такие противоречия позволяют задуматься об общих паттернах образования амилоидоподобных структур и консервативности функциональности данных структур в свете ядерно-цитоплазматического транспорта. С другой стороны, может быть логическое обоснование таких отличий с точки зрения различий в типе митоза, так как для многоклеточных животных показана дезорганизация ядерной мембраны и NPCs во время митоза. Это связано с тем, что они обладают открытым типом митоза, в ходе которого разрушение NPCs на субкомплексы сопровождается высвобождением FG-доменов. И в случае наличия в них амилоидоподобных NQ-обогащенных участков их высвобождение и распространение было бы токсичным и приводило бы к их агрегации. В то время как для грибов показаны либо закрытый (в частности, для дрожжей), либо полуоткрытый тип, для которых не возникает проблема дезорганизации NPCs во время деления. А предотвращение агрегации NQ-обогащенных участков осуществляется за счет стабилизации C-концевой последовательностью, обогащенной пролинами и заряженными аминокислотными остатками. Такого рода последовательности известны как «разрушители амилоидов», так как для них показана способность препятствовать агрегации в фибриллы за счет своей конформации и биохимических свойств (Monsellier, Chiti, 2007, Paz, Serrano, 2004).

Для более точного доказательства данного предположения требуется провести больше структурных исследований взаимодействий внутри FG-гидрогелей как *in vitro*, так и *in vivo*, что представляется сложной задачей ввиду биохимических и биофизических свойств гидрогелей. Данные открытия, с одной стороны, противоречат NQ-амилоидоподобной теории FG-гидрогелей и ставят под вопрос функциональную роль амилоидоподобных структур в ядерно-цитоплазматическом транспорте.

#### 1.4.7. Амилоидные и прионоподобные свойства Nup100

Если для млекопитающих прионы больше известны как патологические агенты, то для дрожжей прионное состояние белков чаще всего не приводит к гибели клеток. Например, в случае дрожжей прионная форма белка Sup35 [*PSI<sup>+</sup>*] агрегирует и не может нормально взаимодействовать с Sup45 и рибосомами, что и позволяет считываться стоп-кодонам (Wickner, 1994).

FG-повторы Nup100 разделены особыми аспарагин-глутамин обогащенными последовательностями, NQ-обогащенными спейсерами. В работе Р. Халфмана с соавторами об исследовании амилоидных и прионных свойств FG-обогащенных нуклеопоринов было показано, что в образовании амилоидо- и прионоподобных структур нуклеопорина Nup100 важную роль играют не изолированные домены FG-повторов, а гидрофильные NQ-обогащенные спейсерные участки. В работе было показано, что NQ-обогащенный участок дрожжевого нуклеопорина Nup100<sup>201-400</sup> при сверхэкспрессии приводил к образованию агрегатов (Halfmann *et al.*, 2012). Было также продемонстрировано, что образование агрегатов CFP-Nup100<sup>201-400</sup> в отсутствие Hsp104 или уже существующего приона [*RNQ<sup>+</sup>*] не происходило. Последнее свидетельствует о том, что как и в случае многих других прион-образующих белков дрожжей, поддержание агрегатов Nup100 зависит от наличия фрагментирующего фактора Hsp104 и уже существующего прионного белка [*RNQ<sup>+</sup>*] (Halfmann *et al.*, 2012). Более того, было также показано, что сверхэкспрессия NQ-обогащенного участка Nup100 приводила к агрегации других FG- и NQ-обогащенных нуклеопоринов *S. cerevisiae* - Nup116, Nup49 и полноразмерного Nup100. В то время как дислокализация и агрегация не обогащенного FG-повторами и NQ-спейсерами нуклеопорина Nup2 при сверхэкспрессии Nup100<sup>201-400</sup> не наблюдались (Halfmann *et al.*, 2012).

В этой же работе было показано, что NQ-обогащенный участок Nup100 - наиболее близкий к дрожжевым прионам Nup100<sup>300-400</sup> и расширенный вариант Nup100<sup>1-595</sup> - обладают способностью связываться с амилоид-специфичным красителем Тиофлавином Т (Halfmann *et al.*, 2012) и приводят к ускорению образования агрегатов при добавлении уже существующих фибрилл. Амилоидоподобные и прионоподобные свойства были показаны как *in vitro*, так и *in vivo*. В частности, в клетках была показана агрегация Nup100<sup>300-400</sup> (Nup100f) без сверхпродукции (Halfmann *et al.*, 2012). И также была показана устойчивость выделенных из клеток агрегатов Nup100f-Sup35C к обработке додецил-сульфатом натрия (SDS), что является одним из способов проверки агрегатов на амилоидные свойства. В довершение ко всему стоит отметить, что авторам удалось показать агрегацию эндогенного Nup100 внутри клеток при наличии сверхэкспрессированных молекул Nup100, SDS-устойчивость образующихся агрегатов и даже проанализировать менделеевское наследование приона Nup100 (Halfmann *et al.*, 2012). Все вышеперечисленное многократно подтверждает амилоидо- и прионоподобную природу Nup100.

NQ-обогащенные участки FG-повторов были показаны не только для Nsp1 и Nup100, но и для других нуклеопоринов: Nup116, Nup49 и Nup57 (Michelitsch, Weissman, 2000, Frey, Görlich, 2009, Patel *et al.*, 2007). Более того, было также показано, что некоторые другие NQ-обогащенные нуклеопорины обладают амилоидными свойствами *in vivo* (Alberti *et al.*, 2009, Halfmann *et al.*, 2012).

Таким образом, ввиду наличия у некоторых участников NPC амилоидных свойств и в целом структурного консерватизма нуклеопоринов значимым является проверка амилоидных свойств среди разных нуклеопоринов.

## 1.5. Биоинформатические методы анализа амилоидогенных свойств

### 1.5.1. Общие принципы

Исследование физических и биохимических свойств амилоидогенных белков легло в основу алгоритмов предсказания способности белков формировать амилоиды. Сейчас алгоритмы учитывают несколько различных составляющих потенциально-амилоидогенных участков (Ahmed *et al.*, 2015):

- общий аминокислотный состав;
- чередование полярных и неполярных аминокислот  $\beta$ -листов;
- боковые взаимодействия аминокислот внутри  $\beta$ -листов;
- укладка боковых цепей аминокислот между  $\beta$ -листами.

Изначально большая часть алгоритмов принимала, что достаточными для образования амилоидогенных структур являлись участки из, как правило, 6 аминокислот, состав которых соответствовал выше перечисленным требованиям. Существующие тогда алгоритмы делились по методам на две группы: первые определяли склонность белков к агрегации, основываясь на физико-химических свойствах аминокислот; вторые же проводили поиск пептидов, склонных к агрегации (Caffisch, 2006). К последним, например, относятся такие ресурсы как PASTA, AGGRESKAN, Zyggregator и Pafig. Минусом ранних программ было то, что, во-первых, они учитывали только аминокислотный состав, не беря во внимание другие биофизические характеристики; во-вторых, не позволяли предсказывать трехмерные структуры потенциально-амилоидогенных белков; в-третьих, учитывали только способность белков образовывать соответствующие структуры в pH-нейтральных средах (Ahmed *et al.*, 2015).

Впоследствии были разработаны программы, которые основаны на частотах мутаций, ассоциированных с образованием амилоидов, другие оценивают склонность к образованию вторичных структур (например, так называемых  $\beta$ -арок - характерных для амилоидов вторичных структур, представленных последовательно  $\beta$ -тяжем, «участком поворота» и  $\beta$ -тяжем (Ahmed *et al.*, 2015)).

Качество предсказания программ также оценивается по некоторым параметрам. На данный момент, чаще всего результаты методов оценивают по следующим параметрам: чувствительность (SE), специфичность (SP), точность (ACC) и коэффициент Мэтью (MCC, Matthew coefficient) - и в зависимости от способов оценки разные алгоритмы отличаются качеством параметров (Табл.3). Чувствительность алгоритмов оценивалась как отношение положительно определенных последовательностей к предположительно позитивным предсказаниям. Специфичность же отображает долю ложно положительных результатов предсказаний для набора отрицательных контролей, то есть неамилоидных последовательностей. ACC характеристика показывает пропорцию верных предсказаний во

всем наборе данных, а MCC параметр обобщает качество алгоритма в целом (Niu *et al.*, 2018).

**Таблица 3** — Сравнение качеств предсказания существующих алгоритмов классификации на амилоидные и не амилоидные белки (Niu *et al.*, 2018)

Алгоритм	ACC	MCC	SE	SP	Ссылки
iAMY-SCM	82,70	0,570	0,606	0,922	Charoenkwan <i>et al.</i> , 2021
RFAmy	89,19	0,739	0,781	0,927	Niu <i>et al.</i> , 2018
ArchCandy	99,88	0,893	-	-	Ahmed <i>et al.</i> , 2015
Waltz	89,19	0,739	0,781	0,927	Maurer-Stroh <i>et al.</i> , 2010
FoldAmyloid	89,19	-	0,750	0,750	Garbuzynskiy <i>et al.</i> , 2009
Pafig	81,00	0,630	0,820	0,800	Tian <i>et al.</i> , 2009
AmylPred	79,00	0,58	0,740	0,840	Frousios <i>et al.</i> , 2009

## 1.5.2. Краткое описание современных методов биоинформатического предсказания потенциально амилоидогенных белков

### AmylPred

Данный алгоритм сочетает в себе пять разных и независимо опубликованных методов предсказания амилоидогенных участков (Frousios *et al.*, 2009). Их общей чертой является учет информации именно о первичной структуре белка и каждый сочетает в себе следующие принципы:

1. Оценка плотности упаковки (Galzitskaya *et al.*, 2006b, Galzitskaya *et al.*, 2006a);
2. Поиск участков, способных формировать различные вторичные структуры;
3. Поиск идентифицированных ранее амилоидогенных пептидов (Paz, Serrano, 2004);
4. Предсказание амилоидогенных пептидов с помощью алгоритма TANGO (Fernandez-Escamilla *et al.*, 2004);
5. Анализ стабильности амилоидных фибрилл из найденных пептидов (Zhang *et al.*, 2007)

Данный обобщенный алгоритм показал наилучшие результаты предсказания относительно его компонентов по отдельности.

### Pafig

Алгоритм Pafig (**P**rediction of **a**myloid **f**ibril-forming **s**egments) использует в своей основе машинное обучение - метод опорных векторов (SVM, support vector machine) (Tian *et al.*, 2009). Данный алгоритм учитывает физико-химические свойства аминокислот и анализирует короткие гексапептиды.

## FoldAmyloid

Метод FoldAmyloid предсказывает амилоидогенные участки, совмещая два других метода: предсказания формирования водородных связей между различными аминокислотами (боковыми цепями и полипептидной цепью) и анализа объемной плотности взаимодействия остатков, то есть количество взаимодействий на аминокислоту (Garbuzynskiy *et al.*, 2009).

Обе характеристики использовались для построения профиля связей анализируемой белковой последовательности. Участок считался амилоидогенным, если полученный профиль был выше нижней границы и размер участка был равен или больше используемого размера окна, а ожидаемые количества контактов и водородных связей также располагались выше границы. Пептид далее принимался амилоидогенным, если содержал хотя бы один амилоидогенный участок.

## Waltz

Waltz использует в анализе принцип построение позиционной матрицы, которая строится для каждой аминокислотной позиции последовательности пептидов, на которые был разбит белок (Maurer-Stroh *et al.*, 2010). Метод оценивает «амилоидогенность» гексапептидов, так как считается, что данный размер пептида является минимальным для начала агрегации белка при inserции такого гексапептида в последовательность белка.

Позиционная матрица (Position-Specific Scoring Matrix, PSSM) строится на множественном выравнивании последовательностей белков и включает в себя значения долей разных типов аминокислот по каждой позиции выравнивания. Получаемая функция включает в себя выбранные физические свойства, состоящие из суммы произведений частоты встречаемости аминокислот исследуемого набора последовательностей с нормализованным значением свойства соответствующей аминокислоты для каждого положения. Выбираются свойства, обладающие достаточно высокими нормализованными значениями. Позиционная матрица далее используется для ранжирования гексапептидов. Такая функция учитывает структурные характеристики за счет своего компонента - позиционно-зависимой псевдоэнергетической матрицы из структурного моделирования с использованием известных структур амилоидного «скелета».

В алгоритме Waltz PSSM была получена на исходном наборе амилоид-образующих (положительных) и не образующих амилоиды (негативных) гексапептидов из базы данных AmylHex, состоящая из наиболее отличающихся друг от друга гексапептидов. При этом в анализ были взяты только те белки и пептиды, для которых были показаны фибриллярные агрегаты с характерными для амилоидов свойствами.

Согласно полученным данным аминокислоты с низкой гидрофобностью, такие как глутамин и аспарагин, являются наиболее зависимыми от положения. Это позволило объяснить, почему предыдущие алгоритмы, основанные на только физико-химических

свойства и гидрофобности аминокислотной последовательности, с трудом определяли глутамин- и аспарагин-обогащенные амилоиды, характерные для дрожжей. Аналогичная зависимость амилоидогенности от положения была показана и для ароматических аминокислот.

Результаты такого алгоритма также оценивались параметрами чувствительности и специфичности (Табл.3). И было показано, что алгоритм Waltz лучше отличает амилоиды от аморфных агрегирующих последовательностей, чем TANGO (Otoo *et al.*, 2008). Но проигрывает в дифференциации агрегирующих последовательностей от растворимых мотивов.

## ArchCandy

В отличие от других методов, данный алгоритм оценивает не физико-химические свойства аминокислотной последовательности в целом и/или пептидов, а проводит поиск определенного структурного мотива -  $\beta$ -арки, представленной последовательностью из  $\beta$ -тяжа, «участка поворота» и снова  $\beta$ -тяжа (Ahmed *et al.*, 2015). Общие закономерности формирования такой структуры были получены путем совмещения данных о  $\beta$ -арках известных кристаллических структур и характерных параллельных  $\beta$ -структур в составе  $\beta$ -соленоидов.

С использованием программ моделирования GROMACS, Rosetta, FoldX и Modeller удалось идентифицировать дестабилизирующие аминокислоты: заряженные, полярные, пролин.

Каждая последовательность оценивалась по способности формировать  $\beta$ -арки. Для этого была выведена функция, отражающая необходимые для искомой укладки взаимодействия. И данный параметр был нормализован и лежит в диапазоне от 0 (точно не образующая  $\beta$ -арки последовательность) до 1 (наиболее вероятный кандидат). Расчет этого параметра основан на:

1. стерической совместимости внутренних аминокислот  $\beta$ -арки;
2. взаимодействия заряженных остатков, акцент делался на возможных образованиях солевых мостиков (ионных взаимодействий) боковых цепей и/или отсутствии компенсации боковых контактов внутри  $\beta$ -арки;
3. электростатическом напряжении;
4. укладке боковых цепей и образующихся водородных связей внутри  $\beta$ -арок;
5. положении остатков пролина;
6. отношении площади контакта с растворителем и общего объема  $\beta$ -арки;
7. длине  $\beta$ -арок.

Для полученного алгоритма ArchCandy после проведения многократных тестовых анализов было выведено пороговое значение функции, равное 0,578 с погрешностью в 0,019. Данный порог позволяет отделять амилоидогенные последовательности от

неамилоидогенных. И последовательности, формирующие хотя бы одну  $\beta$ -арку, принимаются как потенциально амилоидогенные.

При этом, стоит отметить, что среди наиболее часто использующихся программ (PASTA, TANGO, Waltz, FoldAmyloid и Aggrescan) ArchCandy обладал самой низкой долей ошибочных предсказаний - от 0% до 1,4% - на разных по количеству последовательностей наборов (Roche *et al.*, 2017). Именно ввиду высокой точности результатов предсказания ArchCandy данный алгоритм был выбран для представленной работы.

Однако авторы признают, что данный алгоритм имеет и некоторые ограничения, среди которых лимитирующим является способность программы предсказывать только наиболее распространенные амилоидные структуры - параллельные  $\beta$ -арки. Другие же анти-параллельные  $\beta$ -арки,  $\beta$ -соленоиды и аморфные агрегаты не могут быть проанализированы данным алгоритмом.

## **RFAmy**

Алгоритм предсказания RFAmy основан на анализе физико-химических свойств и аминокислотного состава белковых последовательностей и является методом машинного обучения (Niu *et al.*, 2018). Таким образом, является более пластичным методом в том смысле, что модифицируется автоматически в направлении улучшения предсказания в зависимости от направляемого ему набора данных.

Данный алгоритм обладает несколькими преимуществами относительно предыдущих. Во-первых, он является «обучаемым». Во-вторых, показывает лучшие результаты предсказания, чем другие существующие алгоритмы (Табл.3).

## **iAMY-SCM**

Алгоритм iAMY-SCM является еще одним методом машинного обучения, в основе которого лежат функции взвешенных сумм и статистика с оптимизационным подходом (Charoenkwan *et al.*, 2021). Алгоритм машинного обучения представлен в этом методе SCM-подходом (scoring card method). Функция взвешенной суммы используется для оценки склонности дипептидов образовывать амилоидоподобные структуры.

Дипептидный профиль оценивался с помощью метода DPC (dipeptide protein composition), описывающего нормализованный состав дипептидов белка. В данном методе последовательность белка представлена вектором 400 дипептидов анализируемой последовательности. Склонность последовательности белка образовывать амилоидные структуры оценивалась по физико-химическим свойствам входящих в белок аминокислот и дипептидов с помощью генетического алгоритма. Оценки склонности свойств аминокислот обуславливать формирование амилоидных структур усреднялись относительно каждых 40 аминокислот белка (размера «скользящего окна») и если нормализованное значение

свойств аминокислоты было выше 0,5, то такие свойства принимались значимыми в случае образования амилоидных структур.

По словам авторов данный алгоритм является первым наиболее интерпретируемым для большинства биологов. Точность как при перекрестной проверке, так и при тестировании независимого набора последовательностей была сопоставима с таковой RFAmy (0,895 и 0,827, соответственно) (Табл.3). И также данный алгоритм позволяет оценивать биофизические и биохимические свойства амилоидных белков из оценки предрасположенности аминокислот и дипептидов к образованию характерных для амилоидов структур. Помимо этого, авторы утверждают, что разработанный метод является первым методом машинного обучения, который использует подход на основе SCM для прогнозирования и анализа амилоидных белков.

## 2. Материалы и методы

### 2.1. Материалы

#### 2.1.1. Поиск ортологов нуклеопоринов дрожжей *S.cerevisiae*

Последовательности ортологов были взяты из базы данных EggNOG (Huerta-Cepas *et al.*, 2018). Причина выбора именно этой базы данных ортологов среди спектра других заключается в том, что она является наиболее полной и проверенной. Под полнотой подразумевается, что в базе данных представлены группы ортологов с наибольшим количеством следующих параметров: последовательностей белков, таксономических групп организмов, доменов жизни и разных уровней аннотаций (для иллюстрации в табл.4 представлены сравнения разных баз данных ортологов по количеству таксономических групп и организмов). В выбранной базе данных ортологов проводился поиск ортологов для нуклеопоринов Nsp1, Nup1, Nup2, Nup42, Nup49, Nup50, Nup54, Nup57, Nup58, Nup60, Nup100, Nup153 и Nup159. Для работы было принято решение включить в анализ набор ортологов среди таксономической группы Opisthokonta. Данные по группе Opisthokonta для белков были загружены с базы данных EggNOG (<http://eggnog5.embl.de/#/app/home>) 31.07.2020 в формате .fasta.

Таблица 4 — Базы данных ортологов

База данных	Текущая версия	Таксономические группы	Количество организмов	Ссылки на статьи	Ссылка на интернет-ресурс
EggNOG	5.0	379 таксономических групп	4445 бактерий, 168 архей, 477 эукариот, 2502 вирусов	Huerta-Cepas <i>et al.</i> , 2018	<a href="http://eggnog5.embl.de/#/app/home">http://eggnog5.embl.de/#/app/home</a>
OrthoDB	v10	624 таксономических групп	5609 бактерий, 404 архей, 1271 эукариот, 6488 вирусов	Kriventseva <i>et al.</i> , 2018	<a href="https://www.orthodb.org/">https://www.orthodb.org/</a>
OMA	2.4.1		1671 бактерия, 141 архея, 327 эукариот, 0 вирусов	Altenhoff <i>et al.</i> , 2017	<a href="https://omabrowser.org/oma/home/">https://omabrowser.org/oma/home/</a>
KEGG-OC			4301 бактерий, 253 археи, 394 эукариот, 0 вирусов	Nakaya <i>et al.</i> , 2012	<a href="http://www.genome.jp/tools/oc/">http://www.genome.jp/tools/oc/</a>

## 2.1.2. Таксономическое положение видов по базе данных NCBI

В качестве таксономической базы данных было принято решение использовать NCBI Taxonomy (<https://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/>), а названия видов были извлечены с помощью инструмента Taxonomy Name/ID Status Report (<https://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/>) 31.07.2020. Из базы данных EggNOG мы получали последовательности с ID, состоящие из идентификатора вида и идентификатора последовательности. Далее видовые названия из таксономической базы данных NCBI были совмещены с ID последовательностей путем обработки данных с помощью языка программирования R (см. раздел «Обработка данных в R»).

## 2.2. Построение выравниваний для предварительной проверки последовательностей для всех ортологов

В качестве программы для построения выравниваний была выбрана UGENE (Okonechnikov *et al.*, 2012) по следующим причинам: 1) более удобный интерфейс и способ визуализации; 2) открытый код и доступ к расширенным настройкам; 3) возможность импорта и экспорта файлов разного формата. Так как алгоритм MUSCLE (Edgar, 2004) позволяет проводить быстрые и точные выравнивания. Также он является наиболее используемым на сегодняшний день, согласно данными о количестве цитирований оригинальной статьи в базе данных Scopus на осень 2020. Потому именно его решили использовать для построения выравнивания набора ортологов. Полученные выравнивания далее были проанализированы вручную на поиск сильно выделяющихся и выпадающих из выравнивания последовательностей. К таким последовательностям было принято относить следующие варианты:

- Сильно отличающиеся от консенсусной последовательности выравнивания или от общего внутригруппового выравнивания;
- Имеющие протяженные (>20 ак) инделы;
- Неточно аннотированные последовательности, согласно базам данных;
- Последовательности, соответствующие фрагментам полноразмерных белков (короткие последовательности и С-концы);
- Повторяющиеся последовательности;
- Последовательности, информация о которых отсутствует в базах данных.

## 2.3. Анализ амилоидогенности с помощью программы ArchCandy

Для анализа амилоидогенности найденных ортологов полученные наборы последовательностей были проанализированы с помощью программ ArchCandy (<https://bioinfo.crbm.cnrs.fr/index.php?route=tools&tool=7>, Ahmed *et al.*, 2015) и IUPRed (<https://iupred2a.elte.hu/>, Dosztányi *et al.*, 2005). Последнее было необходимо для получения информации о положении неструктурированных участков внутри каждой последовательности. Анализ амилоидогенности заключался в оценке биофизических, биохимических, кинетических (то есть характеристик энергетического состояния), структурных параметров каждой аминокислоты, которые бы могли позволить образовать соответствующие структуры, называемые  $\beta$ -арками. На основе всех параметров подсчитывался общий потенциал (кумулятивный скор) возможности образования амилоидогенных структур, величина которого и говорила о том, может ли данная аминокислота принимать участие в формировании амилоидоподобных структур -  $\beta$ -арок. Полученные данные ArchCandy визуализирует в виде гистограмм, в которых значимые столбцы собраны из пиков значений скоров входящих в них аминокислот. Эта гистограмма, в свою очередь, позволяет выделить амилоидогенные участки внутри последовательности белка. Для большей достоверности полученного результата требовалось также профильтровать найденные участки на предмет вхождения в структурные участки белка. Как было сказано выше, данные о неструктурированных участках были получены нами из программы IUPRed. Если на гистограмме пики соответствовали неструктурированным участкам внутри последовательности белка, то данные последовательности было принято считать потенциально амилоидогенными. В противном случае (то есть при вхождении пиков в структурированные участки) данные последовательности не принимались как амилоидогенные. Пороговым значением для ArchCandy было принято 0,578, этот параметр был рекомендован в оригинальной статье (Ahmed *et al.*, 2015). А для программы IUPRed порогом было принято значение, равное 0,3. Данный этап был осуществлен благодаря работе коллег Лаврентия Глебовича Данилова и Станислава Александровича Бондарева.

## 2.4. Список использованных пакетов R

Для анализа полученных данных мы использовали среду R (версии 4.0.5 (2021-03-31)) и графическую оболочку RStudio (версии 1.4.1106), а также набор дополнительных пакетов:

1. пакеты группы `dplyr`, `plyr`, `tidyr` были использованы для оперирования таблицами и использования регулярных выражений (Wickham *et al.*, 2021, Wickham, 2011, Wickham, 2020);
2. пакет `Biostrings` из набора пакетов `BiocManager` — для работы с `.fasta` файлами (Pages *et al.*, 2020);
3. пакет `seqinr` — для экспорта полученных данных в формате `.fasta` (Charif, Lobry, 2007);
4. `ggplot2`, `scales` — построения графиков (Wickham, 2016, Wickham, Seidel, 2020);
5. `muscle` был необходим для выравнивание белковых последовательностей с помощью одноименного алгоритма;
6. `stringr` был использован для работы с любыми текстовыми строками (Wickham, 2019);

Созданный набор функций и тестовые файлы представлены на репозитории GitHub: [Evolutionary\\_conservatism\\_of\\_amyloidogenic\\_properties\\_of\\_nucleoporins\\_with\\_FG\\_repeats](#).

## 2.5. Оценка долей потенциальных амилоидов среди наборов ортологов

Для оценки амилоидогенности последовательностей внутри выбранных таксономических групп (царств, типов и классов) требовалось посчитать следующие параметры:

- Общее количество последовательностей, принадлежащих данной группе;
- Общее количество потенциально амилоидогенных последовательностей;
- Долю потенциально амилоидогенных белков, являющуюся частным от деления количества амилоидогенных белков на общее количество последовательностей.

На их основе далее планировалось проводить анализ данных и соответствующую группировку по нужной переменной. С помощью нового набора данных кумулятивных значений проводился сбор сводных данных по крупным таксономическим группам: классам, филумам, порядкам и царствам.

Используемая функция описана в скрипте [tree\\_creation.R](#).

## 2.6. Поиск потенциальных консервативных амилоидогенных участков, не входящих в структурные домены

### 2.6.1. Получение подвыборок набора ортологов и их последующее выравнивание

В ходе работы было принято решение, что последующие анализы будут включать таксономические группы, для которых известно хотя бы несколько последовательностей ортологов. Отбор осуществлялся по параметру количества последовательностей на класс. Нами отбирались только те классы, для которых выполнялись условия:

- наличие от 3 до 10 (включительно) последовательностей;
- если класс представлен больше, чем 10 последовательностями, то случайным образом отбирались 10 из них.

Далее полученный набор последовательностей выравнивался с использованием пакета *muscle* (Edgar, 2004) в два этапа:

1. Первое выравнивание с использованием настроек по умолчанию;
2. Второе выравнивание с опцией «refine».

Для анализа большого количества случайных выборок мы использовали R, что позволило автоматизировать процесс и выполнять автоматические один и тот же анализ 10 раз для каждого белка.

### 2.6.2. Анализ выравнивания белковых последовательностей

Необходимо было сопоставить полученные с помощью программы ArchCandy значения кумулятивных скоров с новыми выравниваниями. Значения кумулятивных скоров — мера способности аминокислоты в конкретном положении в белке входить в состав  $\beta$ -арок. Мы соотносили значения кумулятивного скоры с позициями соответствующих аминокислот в выравнивании. В случае, если в строке последовательности из выравнивания обнаруживалась какая-либо из аминокислот, то по данной позиции ставился соответствующий кумулятивный скор. Если же по порядковым позициям в строке выравнивания встречалось значение гэпа («-»), то в таблице кумулятивных скоров по данной позиции присваивалось отсутствующее значение (NA).

Помимо сбора суммарных значений по таксонам, важным этапом этого анализа было получение и визуализация в каждой позиции выравнивания следующих параметров:

- доля гэпов;
- доля амилоидогенности, то есть доля последовательностей со значением метрики ArchCandy выше 0,578;

- доля консервативности, то есть доля наиболее часто встречаемой аминокислоты.

На данном этапе важно было сделать акцент на определенные аминокислоты. Так, нами анализировались доли следующих групп аминокислот:

- полярных аминокислот: аспарагина и глутамина;
- ароматических аминокислот: фенилаланина, триптофана, тирозина;
- гидрофобных аминокислот: аланина, лейцина, метионина, изолейцина, валина.

Причина выбора из полярных аминокислот только аспарагина и глутамина - многочисленные исследования и данные по поводу участия именно этих аминокислот в образовании амилоидных структур некоторых белков (Ahmed, KaJava, 2013).

Подсчет необходимых значений был осуществлен с использованием функции `consensusMatrix` из пакета `Biostrings` (Pages *et al.*, 2020). Данная функция позволяла подсчитывать частоту встречаемости каждой аминокислоты в выравнивании. Доля амилоидогенности подсчитывалась следующим способом. Под параметром амилоидогенности принимались значения кумулятивного сора `ArchCandy` для каждой позиции. Аминокислота в белке считалась потенциально амилоидогенной, если полученное значение было больше 0,578. Затем для каждой позиции выравнивания проводился подсчет количества строк, в которых она приходится на амилоидогенный участок. Это число нормировалось на общее число последовательностей. Данный анализ осуществляется с использованием общей функции (скрипт `MAIN_function.R`, включающей код трех отдельно доступных функций: `filter_plus_alignment.R`, `ali_aa_statistics_plot_R.R`, `CS_subset_final.R`).

### 2.6.3. Картирование участков с известной структурой в выравнивании

Информация о структурных доменах анализируемых белков была взята из баз данных UniProt (Consortium, 2020) и PDB (Velankar *et al.*, 2010). Впоследствии полученные последовательности структурных участков были пересчитаны с помощью R в соответствии с выравниванием. Функция и её описание представлены в скрипте `str_function.R`, доступного на репозитории.

Далее полученные координаты структурных доменов наносились на графики распределения параметров амилоидогенности, аминокислотного состава и консервативности по позициям выравнивания. Так как ранее была замечена проблема смещения результатов в отсутствии «гэпов», то функции сообщалась таблица параметров по позициям выравниваний с учетом «гэпов». Помимо этого функция на вход принимала вектора начальных и конечных координат, а также название создаваемого графика.

### 3. Результаты и обсуждение

#### 3.1. Проверка набора ортологов нуклеопоринов

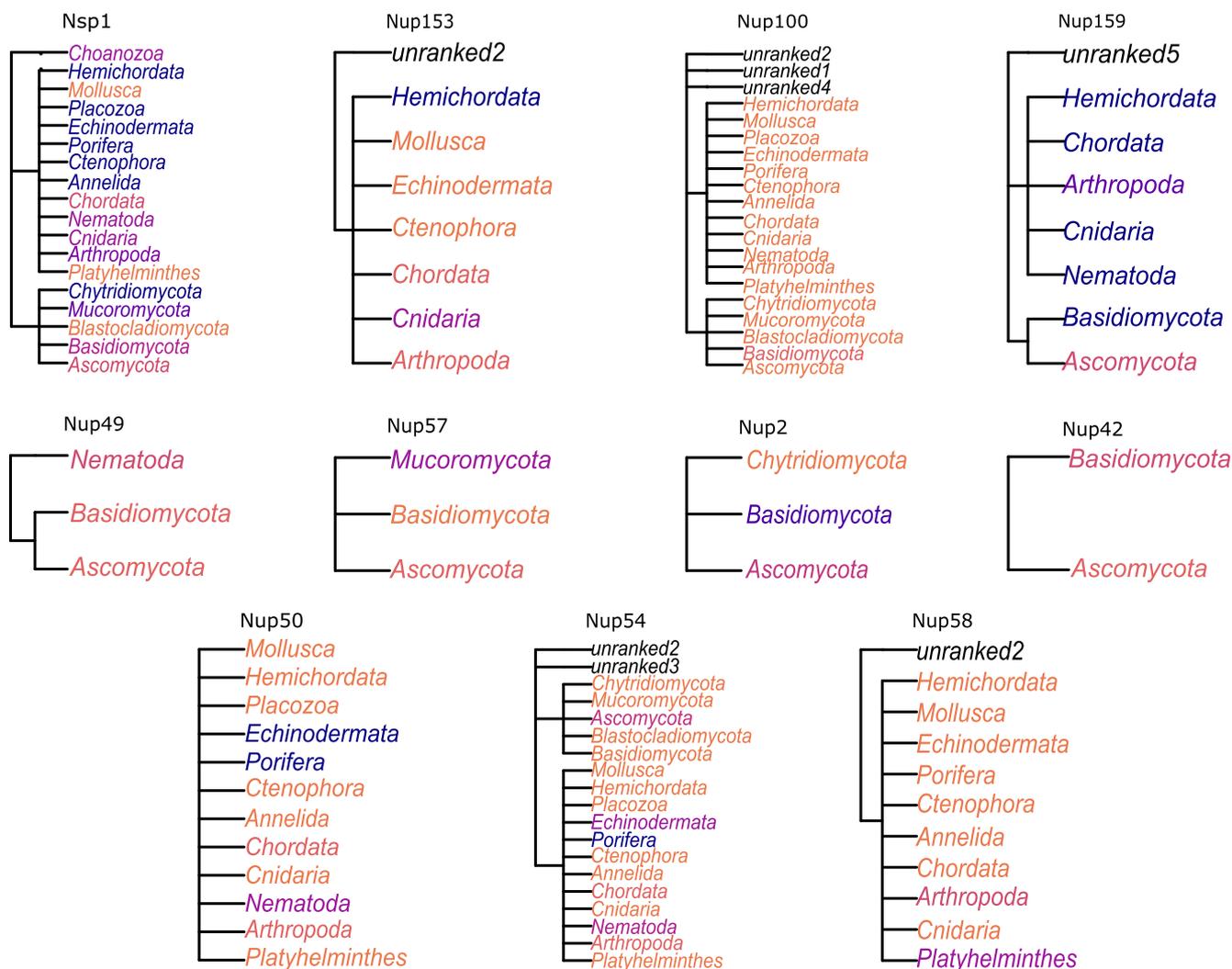
Наборы ортологов для нуклеопоринов Nsp1, Nup1, Nup2, Nup42, Nup49, Nup50, Nup54, Nup57, Nup58, Nup60, Nup100, Nup153 и Nup159, экспортированных из базы данных EggNOG, далее подвергали предобработке для более удобной интерпретации данных. Для последующего анализа требовалось провести замену ID видов последовательностей на их полные латинские названия с сохранением ID последовательности, что было проведено нами с помощью языка программирования R. После этого были построены выравнивания наборов ортологов с использованием программы UGENE. Во время данного анализа перед нами стоял вопрос выбора наиболее достоверного, быстрого и точного алгоритма выравнивания. На сегодняшний день наиболее часто применяемыми алгоритмами для построения выравниваний являются MUSCLE (Edgar, 2004), MAFFT (Katoh *et al.*, 2002) и ClustalW (Thompson *et al.*, 1994). Каждый из них отличается соотношением таких параметров как скорости процесса и качества результата и самим алгоритмом построения выравниваний. Так как в ходе первичного анализа требовалось провести быстрое выравнивание с наиболее низкой потерей качества, то после сравнения тестовых выравниваний всех алгоритмов было принято решение остановиться на MUSCLE.

Полученные выравнивания мы отфильтровали вручную по указанным в разделе «Материалы и методы» критериям. В результате из последовательностей ортологов белков Nsp1, Nup49, Nup54, Nup98, Nup100 и Nup159 мы исключили от двух до 37 последовательностей (см. «Приложения», Табл.I).

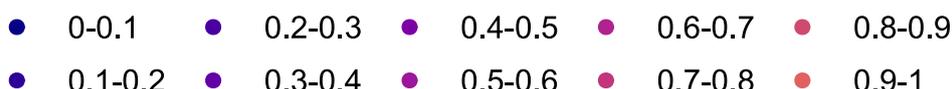
#### 3.2. Анализ амилоидогенности последовательностей набора ортологов нуклеопоринов

Амилоидогенность оценивалась с использованием двух биоинформатических программ: ArchCandy (Ahmed *et al.*, 2015) и IUPRed (Dosztányi *et al.*, 2005). Затем была оценена доля потенциальных амилоидов внутри следующих таксономических групп: царств, типов и классов. Результаты представлены в таблицах (см. «Приложения» Табл.II-XIV).

По результатам анализов доля потенциальных амилоидов (ПА) для ортологов дрожжевых нуклеопоринов сильно варьирует среди царств и типов (Рис.7). Чаще всего вклад в высокую долю ПА вносят конкретные классы, для которых данный показатель является очень высоким. Для ортологов белка Nsp1 в случае типа Ascomycota такими

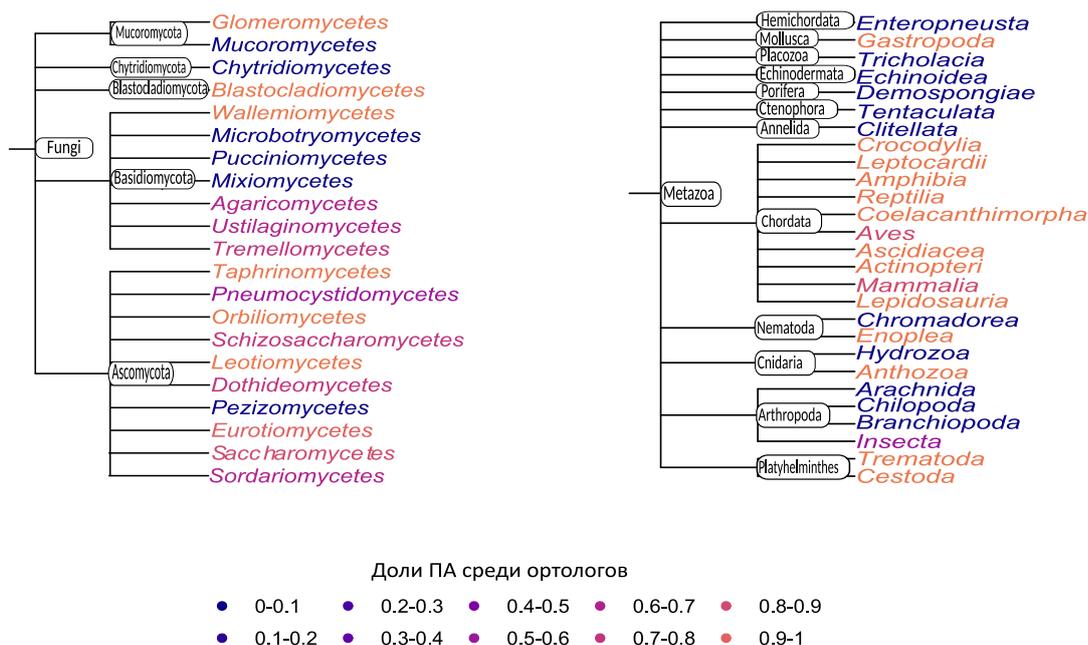


Доли ПА среди ортологов



**Рисунок 7 — Встречаемость ПА среди анализируемых нуклеопоринов из разных таксономических групп.** Дендрограммы показывают филогенетическое родство различных типов. Цвета отражают доли ПА среди последовательностей, относящихся к разным таксономическим группам. unranked - группы с неопределенным таксономическим положением согласно базе данны NCBI Taxonomy.

классами являются: Dothidiomycetes, Eurotiomycetes, Saccharomycetes и Sordariomycetes (Рис.8). Причем любопытно, что для класса Saccharomycetes, включающего вид анализируемых белков, доля ПА является не 100%. Последнее особенно любопытно в случае Nsp1 *S. cerevisiae*, так как ранее А. Лабохой и соавторами была показана его способность формировать амилоиды. Разброс показателей среди классов, в свою очередь, позволяет предположить, что какой-либо эволюционный тренд на приобретение или утрату амилоидных свойств отсутствует. Возможно, такая вариация обуславливается разными условиями среды, образом жизни или особенностями жизненного цикла и формой организма.

Доли ПА среди ортологов Nsp1 среди *Fungi*Доли ПА среди ортологов Nsp1 среди *Metazoa*

**Рисунок 8** — Сравнение встречаемости ПА среди наборов ортологов дрожжевого белка Nsp1 для групп *Fungi* и *Metazoa*. На графике представлены деревья царств с входящими в них классами. Доли ПА выделены цветом, а цветовые обозначения долей приведены внизу рисунка.

Среди других типов организмов количество классов, влияющих на долю потенциальных амилоидов типов, значительно меньше. И в целом значения долей ПА ниже. Хотя некоторые данные являются интересными: так, например, примечателен большой и единственный вклад класса *Insecta* в значение доли ПА нуклеопоринов, относящихся к типу *Arthropoda* (Рис.8). Ранее А. Лабохой и соавторами было показано, что ортолог *X. laevis* дрожжевого Nsp1, Nsp62, не обладает амилоидными свойствами (Labokha *et al.*, 2013). Однако, согласно нашему анализу, доля ПА класса *Amphibia*, куда таксономически входит *X. laevis*, является 100%.

В то же время следует отметить, что доля ПА *Mammalia* и *Saccharomycetes* является сходной (Рис.8). Помимо того, что млекопитающие вносят основной вклад в значение анализируемого показателя для типа *Chordata*, такое совпадение между достаточно далекими филогенетически группами позволяет задуматься о важности приобретения амилоидогенности нуклеопоринов в ходе эволюции. Однако причины и факторы, которые могли повлиять на такое развитие событий, невозможно выявить на данном этапе анализа и в целом высказанное предположение требует тщательной проверки. Возможно, наличие потенциально амилоидогенных белков в составе NPC связано с типом митоза у разных организмов более мелких таксономических групп. Последнее следует из того, что такой признак как тип митоза сильно влияет на организацию ядерного порового комплекса, является сильно варибельным признаком и может значительно отличаться уже на уровне видов. Потому возможная зависимость будет прослежена среди мелких таксономических групп и

потребуется глубокого анализа количества ПА и типа митоза на уровне видов (Sazer *et al.*, 2014).

Кроме этого, таблицы говорят о том, что ПА среди ортологов Nsp1, Nup1, Nup42, Nup50, Nup54, Nup58, Nup153 и Nup159 довольно много, во всяком случае у некоторых групп (см. «Приложения» Табл.II-XIV, рис.7).

Ранее предполагалось, что способность к образованию амилоидоподобных структур тесно связана с типом митоза (Labokha *et al.*, 2013). Но согласно проведенному анализу такая тенденция наблюдается лишь для Nup159 (Рис.7). А нуклеопорины, для которых изначально была предположена такая гипотеза (Nsp1, Nup100), и других из анализируемого набора - нет. Потому нельзя точно говорить о связи потери и сохранения амилоидных свойств нуклеопоринов с типом митоза.

Также можно заметить, что есть характерные только для многоклеточных эукариот белки, а именно Nup50, Nup54 и Nup58. Например, доли ПА белка Nup58 минимальны в случае наиболее просто устроенных высших эукариот - типа Platyhelminthes - и максимальны в случае типа хордовых (Рис.7). И, возможно, доли ПА увеличиваются в зависимости от усложнения организации организмов.

Так, можно отметить, что в нашем анализе наборы ортологов белков Nup50, Nup54, Nup58, Nup100, Nup153 и Nup159 включают большее количество типов организмов, чем наборы ортологов других анализируемых нуклеопоринов (Табл.VIII, X, XII, XIII, XIV). Набор Nup54 содержит 18 типов, Nup58 - 10, Nup100 - 19, Nup153 и Nup159 - 8. И в целом, для данных белков показаны высокие доли ПА среди разных типов, начиная низшими эукариотами - Ascomycota, Basidiomycota, Mucoromycota, Chytridiomycota - и заканчивая высшими эукариотами - Chordata.

Белки Nup50, Nup153 и Nup100 характеризуются высокими долями ПА в разных типах организмов. Причем для белка Nup100 показаны самые высокие значения долей практически во всех типах - от низших до высших эукариот - присутствующих в наборе ортологов этого белка (Рис.7 и Табл.XII). Однако такие высокие доли могут объясняться вкладом последовательностей паралога данного белка - Nup116 и Nup145 - которые заменяют функционально и структурно Nup100 в ядерно-поровом комплексе в эволюции организмов. Но интересны и полученные результаты для белка Nup159. Так, мы видим, что данный белок обладает низкими долями ПА практически во всех типах организмов, у которых наблюдается наличие ортологов данного белка. И единственным типом, для которого показаны высокие доли ПА белка Nup159, является Ascomycota. Последнее может говорить о том, что амилоидные свойства Nup159 являются специфичными для типа Ascomycota.

Важно отметить, что в случае белков Nup54, Nup58 мы не наблюдаем наличия ортологов для типов грибов, однако видим высокие доли ПА среди типа хордовых (Рис.7). При этом Nup54 представлен как среди типов царства грибов (Ascomycota, Chytridiomycota, Basidiomycota и Mucoromycota), так и среди множества типов царства животных. И доли ПА для животных, а именно типа Chordata, выше таковых для грибов, конкретно типа

Ascomycota. Как уже было отмечено выше, наборы ортологов белков Nup2, Nup42, Nup49 и Nup57 найдены только среди некоторых типов организмов: Ascomycota, Basidiomycota, Chytridiomycota, Mucoromycota и Nematoda (Рис.7), — согласно базе данных EggNOG. На основе этого можно предположить особую роль этих белков в ядерно-цитоплазматическом транспорте низших эукариот, однако такие результаты могут быть обоснованы неполнотой используемой базы данных EggNOG. При этом если Nup2 не обладает высокой долей ПА в случае типа Basidiomycota, то другие белки являются высоко амилоидогенными у данного типа. И также является интересным, что среди этих белков только Nup49 обнаруживается в типе Nematoda. Возможно, с эволюцией организмов и изменениями структуры ядерно-порового комплекса данные белки перестают быть функционально необходимыми и большая часть выполняемых ими функций в структуре NPC и процессе транспорта распределяется между другими участниками порового комплекса и/или между новыми белками, возникшим под действием факторов микроэволюции.

Но также возможно, что такие данные могут быть связаны и с корреляцией доли ПА с количеством взятых в анализ последовательностей на таксон, так как в целом более высокие значения долей ПА были получены для таксонов с большим количеством проанализированных последовательностей.

Тем не менее, в целом, если просуммировать доли ПА всех нуклеопоринов, кроме Nup60 и Nup159, то у нас получается разброс от 0,717 до 0,999. А для Nup60 и Nup159 общая доля ПА всех наборов - 0,353 и 0,345, соответственно. И можно сделать вывод, что среди нуклеопоринов с FG повторами высока доля потенциально амилоидогенных белков.

### **3.3. Поиск участков с консервативными амилоидными свойствами**

#### **3.3.1. Подсчет параметров амилоидогенности, консервативности и аминокислотного состава**

Количество последовательностей ортологов исследуемых дрожжевых нуклеопоринов для разных таксономических групп сильно отличалось. И большое количество близкородственных последовательностей могло сильно исказить результаты. Поэтому требовалось сделать отбор случайных подвыборок, сохраняя сопоставимость по количеству последовательностей внутри каждой подвыборки (см. «Материалы и методы»). И всего нами было проанализировано 10 случайных подвыборок из всего набора ортологов.

Для решения поставленной задачи был создан специальный набор функций в R, который позволял автоматизировать необходимый алгоритм. И созданный набор позволил провести поиск фрагментов нуклеопоринов с консервативными амилоидными свойствами (т.е. участков белков, обладающих склонностью к агрегации у большинства

проанализированных белков) для белков: Nsp1, Nup1, Nup2, Nup42, Nup49, Nup50, Nup54, Nup57, Nup58, Nup60, Nup153 и Nup159.

Анализировались именно эти белки, так как по предварительным данным ArchCandy они являлись потенциальными амилоидами. Однако анализ Nup100 невозможно было провести, так как этот белок входит в группу паралога с Nup116 и Nup145 и ввиду этого в анализируемом нами наборе присутствовали последовательности Nup116 и Nup145. Поэтому для избежания смещения наших оценок амилоидогенности было принято исключить Nup100 из анализа.

Во время предыдущего анализа мы лишь анализировали наличие или отсутствие предсказанных потенциально амилоидогенных участков в белке. Однако важно было учесть и конкретное положение таких фрагментов. Для этого был создан набор функций, который позволял проанализировать положения амилоидогенных регионов внутри множественного выравнивания последовательностей ортологов. Это дало возможность проследить консервативность амилоидных свойств и аминокислотный состав выравненных фрагментов белков. Консервативность амилоидных свойств по позиции выравнивания мы рассчитывали как долю последовательностей, в которых соответствующая аминокислота включена в амилоидогенный участок, от всех последовательностей. Чтобы избежать смещения расчетов в сторону таксономических групп с большим набором последовательностей, мы проводили случайный отбор только части последовательностей из групп (см. «Материалы и методы»). Кроме этого, нами также была собрана информация о наличии третичных структур в нуклеопоринах из базы данных PDB (<https://www.ebi.ac.uk/pdbe/>), данные о которой были наложены на построенные выравнивания. Это было необходимо, так как в случае образования амилоидных структур, склонных к агрегации, важным является отсутствие упорядоченности на этом же участке.

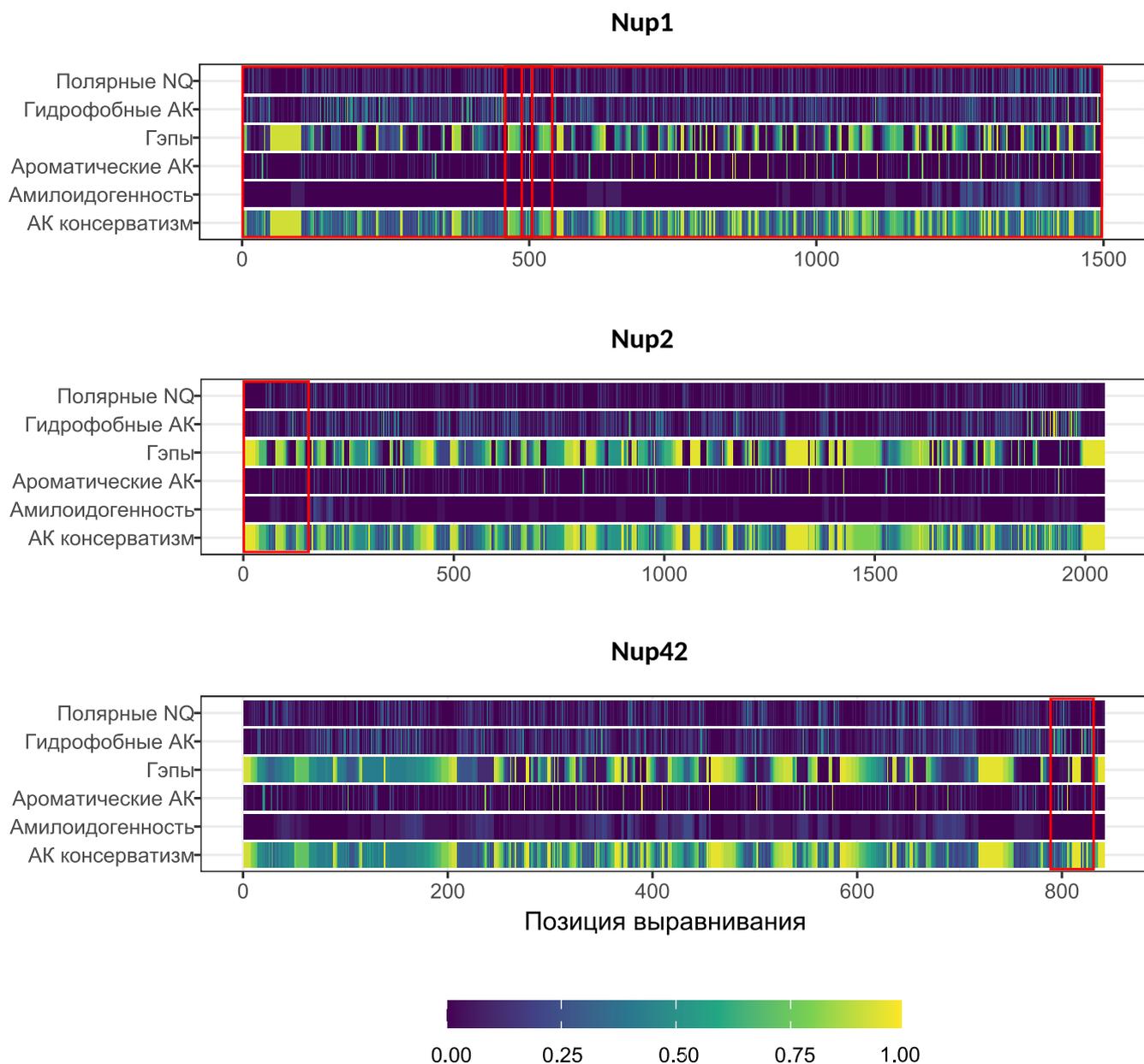
### **3.3.2. Поиск участков с консервативными амилоидными свойствами**

В ходе поиска участков с консервативными амилоидными свойствами у белков: Nsp1, Nup1, Nup2, Nup42, Nup49, Nup50, Nup54, Nup57, Nup58, Nup60, Nup153 и Nup159 - удалось обнаружить, что исследованный набор белков по анализируемым параметрам можно разделить на две группы:

- белки, не обладающие консервативными амилоидогенными участками;
- белки с консервативными амилоидогенными участками.

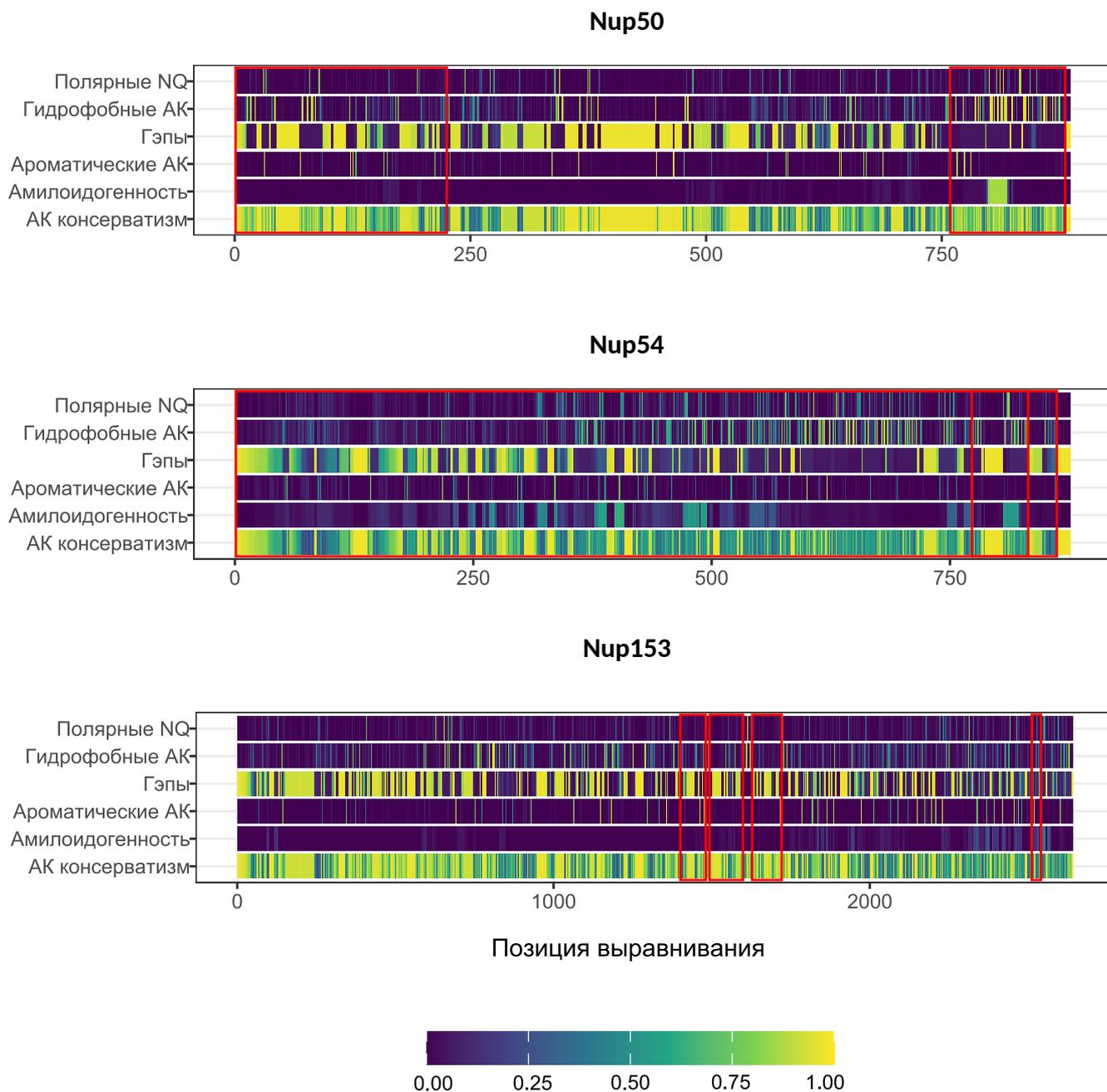
Значимыми принимались только участки, которые более чем у половины последовательностей были классифицированы как ПА. К сожалению, в анализ расположения структурных участков относительно потенциально амилоидогенных не удалось включить белки Nup60 и Nup159, так как для них отсутствовала информация о структуре доменов в базе данных UniProt.

Так, судя по полученным результатам, белки Nup1, Nup2, Nup42, Nup50, Nup54 и Nup153 не содержат подобных участков. В случае Nup50 и Nup54 фрагменты, которые предсказаны как амилоидогенные для большинства последовательностей, локализованы в участке с известной трехмерной структурой. Поэтому маловероятно, что они в действительности могут агрегировать. Из 10 повторностей каждого белка представлен наиболее типичный результат (Рис.9, 10). И согласно имеющимся литературным данным данные белки не имеют общего расположения внутри NPC, но все участвуют в обеспечении импортин-опосредованного транспорта.



**Рисунок 9** — В белках Nup1, Nup2 и Nup42 отсутствуют участки с консервативными амилоидогенными свойствами. Доли параметров обозначены цветом, где максимуму соответствует желтый цвет, а минимуму — темно-фиолетовый. Красными рамками выделены участки, для которые есть данные о третичной структуре, информация о которых была получена из базы данных UniProt (<https://www.uniprot.org/>). Для Nup1 есть данные о третичной структуре для трех участков, которые перекрываются: с 1 по 1076, с 316 по 340 и с 322 по 355 ак.

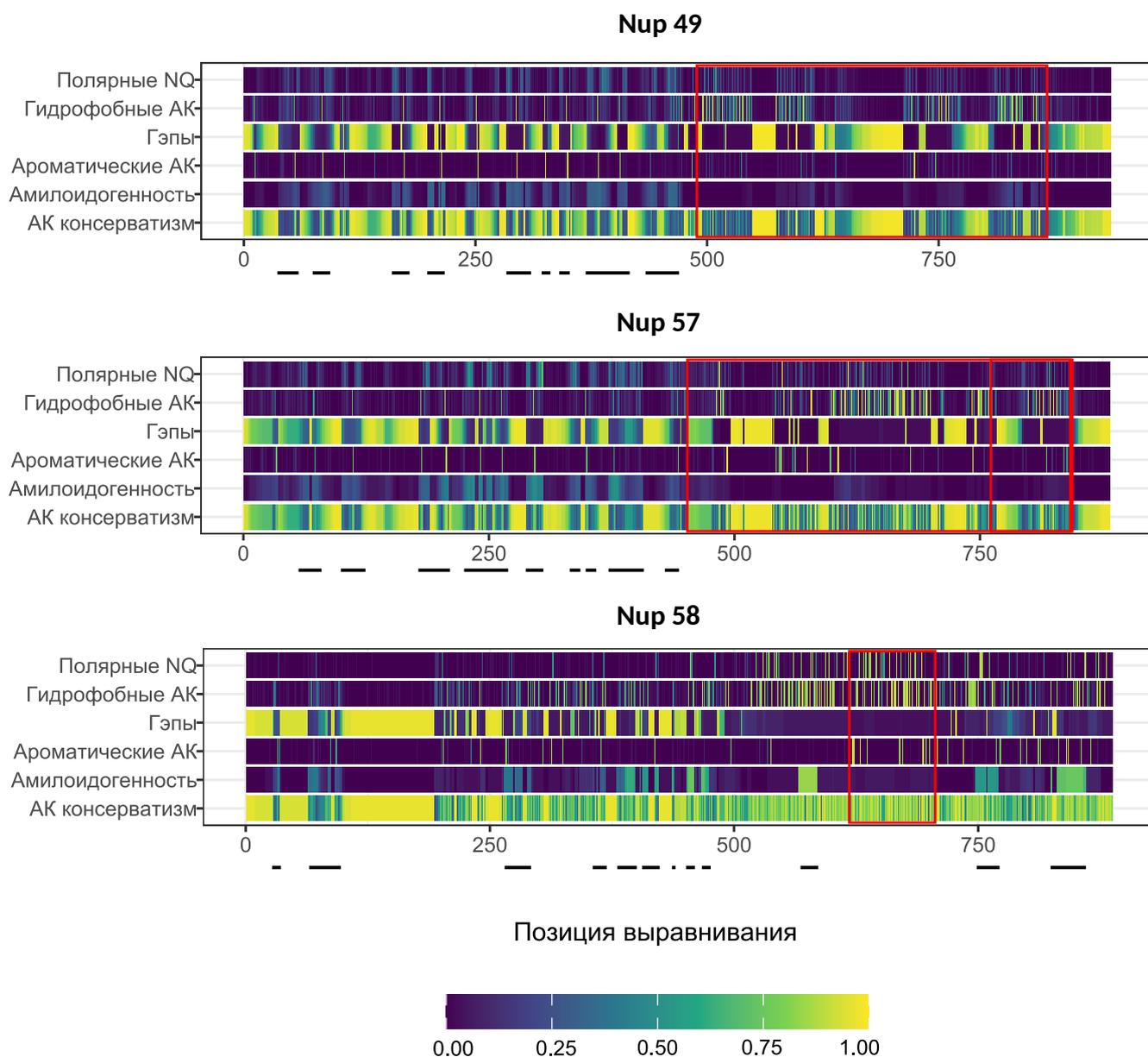
Нельзя исключить, что при анализе консервативности амилоидных свойств последовательностей этих белков, относящихся к более мелким таксономическим группам, могут быть обнаружены консервативные участки с амилоидными свойствами. И в дальнейшем этот анализ будет проведен.



**Рисунок 10** — В белках Nup50, Nup54 и Nup153 отсутствуют участки с консервативными амилоидогенными свойствами. Доли параметров обозначены цветом, где максимуму соответствует желтый цвет, а минимуму — темно-фиолетовый. Красными рамками выделены структурные участки, информация о которых была получена из базы данных UniProt (<https://www.uniprot.org/>). Для Nup54 есть данные о третичной структуре для двух участков, которые перекрываются: с 1 по 507 и с 453 по 491 ак.

В то же время для белков Nup49, Nup57 и Nup58 были обнаружены консервативные амилоидогенные участки (Рис.11). Для Nup49 всего было обнаружено 9 таких участков (20-35 ак, 50-75 ак, 180-220 ак, 270-310 ак, 375-420 ак и 470-490 ак). Для Nup57 наблюдается

так же 9 аналогичных участков. А для Nup58 было найдено 12 участков по выравниванию (2 участка в районе 25-70 ак, участки в интервале 260-600 ак и 750-850 ак). Таким образом, данные белки являются ценными для дальнейшего анализа консервативности и функциональности амилоидогенных структур.



**Рисунок 11 — Участки с консервативными амилоидогенными свойствами, найденные в белках Nup49, Nup57 и Nup58.** Доли параметров обозначены цветом, где максимуму соответствует желтый цвет, а минимуму — темно-фиолетовый. Красными рамками выделены структурные участки, информация о которых была получена из базы данных UniProt (<https://www.uniprot.org/>), а потенциальные консервативные амилоидогенные участки обозначены черными линиями вдоль оси ОХ под каждым графиком. Для Nup57 есть данные о третичной структуре для двух участков, которые перекрываются: с 74 по 319 и с 265 по 317 ак.

Ранее в работе Лаврентия Глебовича Данилова и коллег были показаны *in vitro* и *in vivo* амилоидные свойства белка Nup58. И наш проведенный биоинформатический анализ также подтверждает потенциальную амилоидогенность данного белка (Рис.11). Кроме этого, высокие доли ПА характерны для разных типов организмов в случае обоих белков (Рис.7).

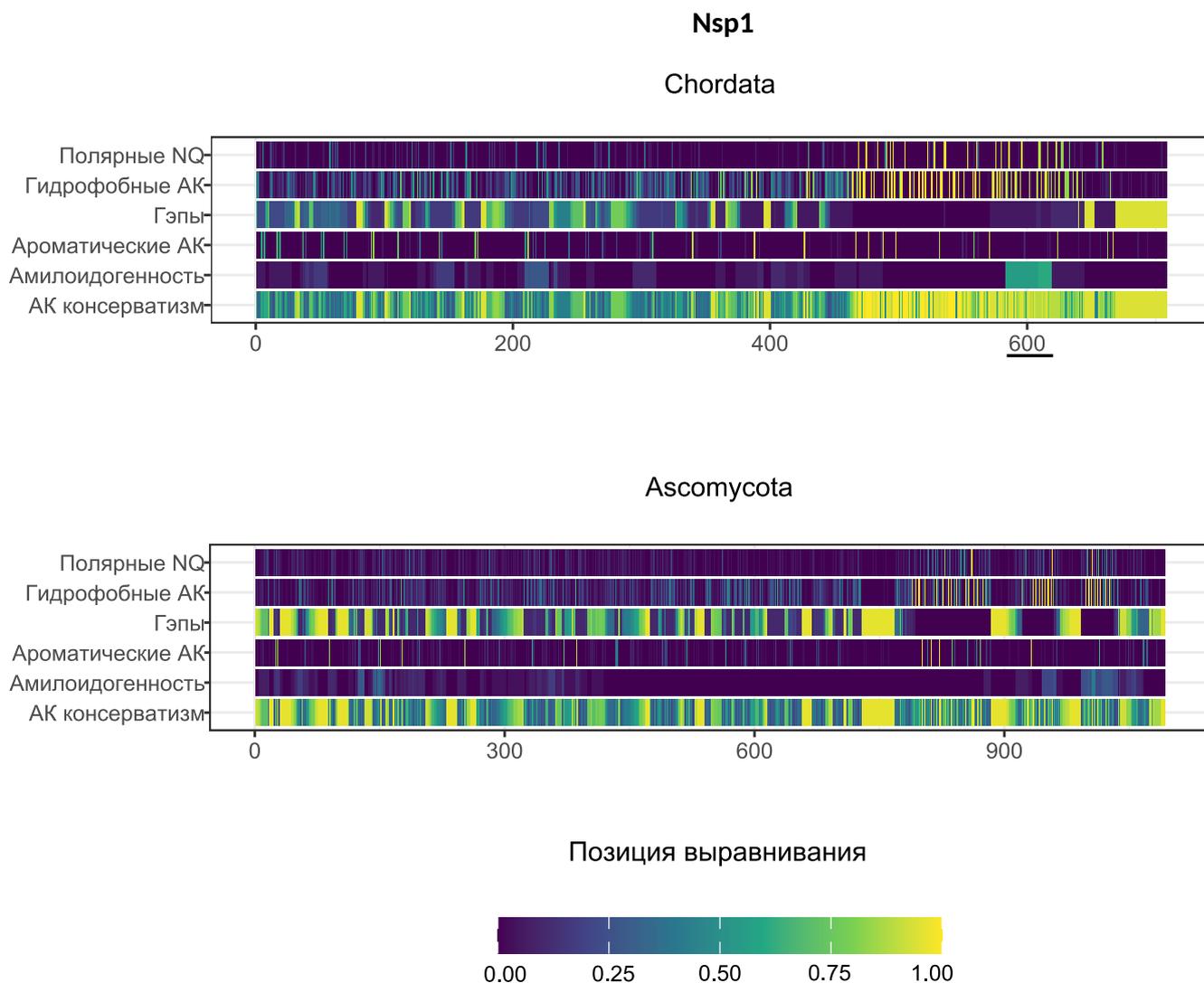
Nup58 вместе с Nup57 и Nup49 являются важными участниками CNC-модуля ядерной поры, в который также входит и анализируемый белок дрожжей Nsp1 (hNup62). При этом Nup57 дрожжей функционально соответствует человеческому белку Nup54 в NPC, Nup49 дрожжей — Nup58 человека. Наличие консервативных амилоидных участков у двух белков из далеких друг от друга эволюционно организмов позволяют предполагать функциональность амилоидогенных свойств. Причем эти свойства сохраняются даже после дивергенции последовательностей. К сожалению, точная роль Nup58, Nup57 и Nup49 в данном комплексе пока остается не изученной. Но есть информация, что они необходимы для обеспечения транспорта макромолекул и, возможно, также входят в состав FG-гидрогеля. Обнаруженные потенциально амилоидогенные участки белков Nup58, Nup57 и Nup49 попадают в FG-обогащенные домены, которые лежат с 7 по 579, с 2 по 223 и с 2 по 236 аминокислоты белков, соответственно. И последнее наблюдение позволяет предположить, что амилоидные свойства являются функциональными в свете обеспечения ядерно-цитоплазматического транспорта.

Также любопытно, что результаты по наличию консервативных амилоидных доменов противоположены для двух функционально соответствующих белков - дрожжевого нуклеопорина Nup57 и нуклеопорина человека Nup54. По проведенному биоинформатическому анализу Nup57 содержит потенциально консервативные амилоидные домены, а Nup54 — нет. Ранее в работе К. Чорпа было показано, что Nup57 является более консервативным нуклеопорином, чем Nup54. Авторы объясняют снижение консервативности Nup54 действием очищающего отбора на приобретенную им  $\beta$ -область структурного домена (Chorpa *et al.*, 2019). И, действительно, согласно нашему анализу для Nup54 не было найдено консервативных участков и доля ПА снижается в направлении от низших эукариот к высшим (Рис.7, 10, 11). Таким образом, такие результаты для Nup54 могут объясняться гипотезой К. Чорпа о влиянии очищающего отбора на этот белок в ходе эволюции.

Отдельно мы бы хотели рассмотреть белки Nsp1 и Nup159, поскольку для них мы наблюдали наличие консервативных участков с амилоидогенными свойствами только среди отдельных таксономических групп: Chordata и Ascomycota, соответственно. Согласно данным К. Адера нуклеопорин Nsp1 обладает амилоидными свойствами и соответствующие структуры были показаны для FG-обогащенных участков во время образования гидрогелей (Ader *et al.*, 2010). И предполагалось, что проведенный биоинформатический анализ должен показать сходные результаты. Однако, хотя по анализируемой метрике ArchCandy многие ортологи Nsp1 являлись потенциальными амилоидами (Рис.7 и «Приложения», Табл.II), при анализе аминокислотного состава, консервативности и склонности к агрегации участков полученных наборов множественных выравниваний не были найдены участки с консервативными амилоидными свойствами для всего набора ортологов (см. «Приложения» и Рис.9).

Однако в предыдущем анализе было получено, что доли ПА у типа Chordata выше, чем у типа Ascomycota, для набора ортологов Nsp1 и является 100% в случае класса

амфибий (Рис.7, Табл.II). Такие результаты противоречили результатам А. Лабохи с соавторами. И после проведения анализа консервативности склонности к агрегации участков наборов отдельно для типов Ascomycota и Chordata было замечено, что у хордовых имеется консервативный участок с высокой долей амилоидогенности в районе 800-850 ак выравнивания (Рис.12).

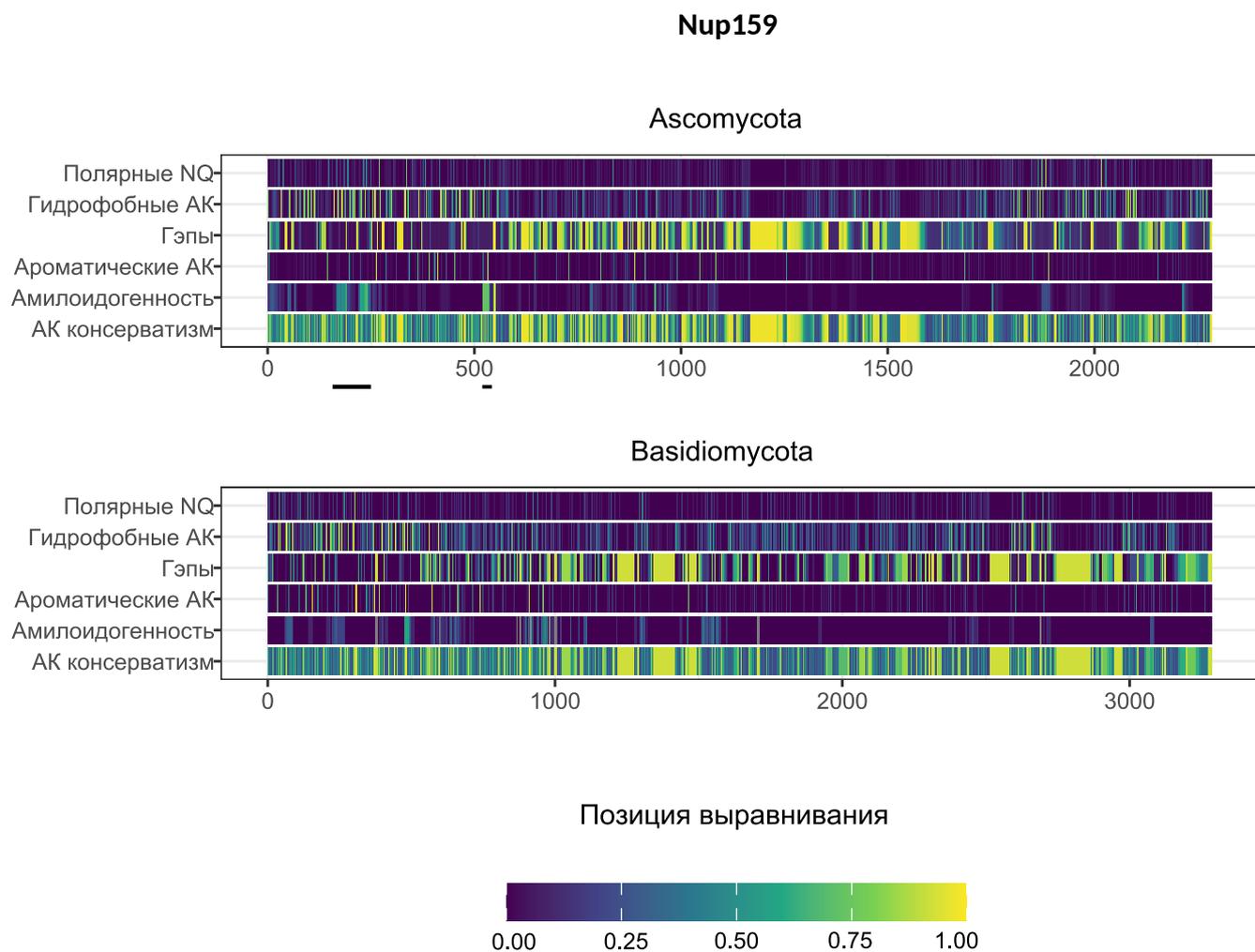


**Рисунок 12 — Ортологи Nsp1 среди типа Ascomycota содержат консервативный участок с высокой долей амилоидогенности.** Для типа Ascomycota не было обнаружено высоко консервативных амилоидогенных участков среди набора ортологов Nsp1, в то время как для типа Chordata найден такой участок в районе 800-850 ак выравнивания. Доли параметров обозначены цветом, где максимуму соответствует желтый цвет, а минимуму — темно-фиолетовый. Консервативные амилоидогенные участки обозначены черными линиями вдоль оси ОХ под каждым графиком.

Таким образом, нельзя говорить о консервативной роли агрегации основного участника FG-гидрогеля, Nsp1, в обеспечении селективности транспорта. И, скорее всего, такая тенденция к амилоидизации находится под сильным давлением эволюции и/или среды.

Также в предыдущем анализе нами было получено, что высокие доли ПА для Nsp159 наблюдаются только в случае типа Ascomycota (Рис.7). Из этого можно

было предположить, что амилоидные свойства Nup159 могут быть специфичными для определенных таксономических групп, а именно для Ascomycota. Для проверки предположения мы провели анализ амилоидогенности и аминокислотного консервативности отдельно для типов Ascomycota и Basidiomycota. И консервативные участки с высокими долями амилоидогенности наблюдались только для типа Ascomycota (Рис.13).



**Рисунок 13** — Амилоидогенные свойства белка Nup159 специфичны для типа *Ascomycota*. Для ортологов Nup159 среди типа *Ascomycota* были обнаружены фрагменты с консервативными амилоидогенными свойствами, в то время как для типа *Basidiomycota* — нет. Доли параметров обозначены цветом, где максимуму соответствует желтый цвет, а минимуму — темно-фиолетовый. Консервативные амилоидогенные участки обозначены черными линиями вдоль оси ОХ под каждым графиком.

В ходе проведенного анализа нам удалось обнаружить участки с консервативными амилоидогенными свойствами среди белков CNC-модуля, которые отвечают за барьерную функцию NPC. Среди них Nup49, Nup57, Nup58, а также Nsp1 (*Chordata*). Как можно было ожидать эти участки перекрываются с неструктурированными регионами FG-повторов в этих белках. В то же время нам не удалось найти аналогичные по своим свойствам фрагменты среди других нуклеопоринов с FG-повторами (Nup1, Nup2, Nup42, Nup50, Nup54 и Nup153), которые не ассоциированы с CNC-модулем. Это позволяет заключить, что амилоидогенные свойства не связаны с самими повторами, а участки с консервативными

амилоидогенными свойствами в Nup49, Nup57, Nup58, а также Nsp1 могут играть важную роль в формировании селективного барьера внутри ядерной поры. Еще одним важным результатом работы стал набор функций, который позволяет проводить автоматизированный анализ для любого другого белка.

#### 4. Выводы

1. Большинство нуклеопоринов с FG-повторами являются потенциальными амилоидами согласно биоинформатическим предсказаниям.
2. Нуклеопорины, образующие FG-гидрогель внутри ядерной поры, обладают участками с консервативными амилоидогенными свойствами.

## Список литературы

1. *Caflisch A.* Computational models for the prediction of polypeptide aggregation propensity // *Current Opinion in Chemical Biology*. — 2006. — т. 10, № 5. — с. 437–444.
2. *Adda C. G. et al.* *Plasmodium falciparum* merozoite surface protein 2 is unstructured and forms amyloid-like fibrils // *Molecular and biochemical parasitology*. — 2009. — Vol. 166, no. 2. — P. 159–171.
3. *Ader C., Frey S., Maas W., Schmidt H. B., Görlich D., Baldus M.* Amyloid-like interactions within nucleoporin FG hydrogels // *PNAS U. S. A.* — 2010. — Vol. 107, no. 14. — P. 6281–6285.
4. *Ahmed A. B., Kajava A. V.* Breaking the amyloidogenicity code: Methods to predict amyloids from amino acid sequence // *FEBS Letters*. — 2013. — Vol. 587, no. 8. — P. 1089–1095.
5. *Ahmed A. B., Znassi N., Château M.-T., Kajava A. V.* A structure-based approach to predict predisposition to amyloidosis // *Alzheimer's & Dementia*. — 2015. — Vol. 11, no. 6. — P. 681–690.
6. *Alber F. et al.* The molecular architecture of the nuclear pore complex // *Nature*. — 2007. — Vol. 450, no. 7170. — P. 695–701.
7. *Alberti S., Halfmann R., King O., Kapila A., Lindquist S.* A systematic survey identifies prions and illuminates sequence features of prionogenic proteins // *Cell*. — 2009. — Vol. 137, no. 1. — P. 146–58.
8. *Alexander J. H., Douglas B. B.* Optical properties of amyloid stained by Congo red: History and mechanisms // *Micron (Oxford, England : 1993)*. — 2008. — Vol. 40. — P. 285–301.
9. *Allen N. P. C., Huang L., Burlingame A., Rexach M.* Proteomic Analysis of Nucleoporin Interacting Proteins // *Journal of Biological Chemistry*. — 2001. — Vol. 276, no. 31. — P. 29268–29274.
10. *Allen N. P. C., Patel S. S., Huang L., Chalkley R. J., Burlingame A., Lutzmann M., Hurt E. C., Rexach M.* Deciphering Networks of Protein Interactions at the Nuclear Pore Complex // *Molecular & Cellular Proteomics*. — 2002. — Vol. 1, no. 12. — P. 930–946.
11. *Altenhoff A. M. et al.* The OMA orthology database in 2018: retrieving evolutionary relationships among all domains of life through richer web and programmatic interfaces // *Nucleic Acids Research*. — 2017. — Vol. 46, no. D1. — P. D477–D485.
12. *Alteri C. J., Xicohténcatl-Cortes J., Hess S., Caballero-Olín G., Girón J. A., Friedman R. L.* Mycobacterium tuberculosis produces pili during human infection // *PNAS U. S. A.* — 2007. — Vol. 104, no. 12. — P. 5145–5150.
13. *Amlacher S., Sarges P., Flemming D., Noort V. van, Kunze R., Devos D. P., Arumugam M., Bork P., Hurt E.* Insight into Structure and Assembly of the Nuclear Pore Complex

- by Utilizing the Genome of a Eukaryotic Thermophile // *Cell*. — 2011. — Vol. 146, no. 2. — P. 277–289.
14. *Astbury W. T., Dickinson S., Bailey K.* The X-ray interpretation of denaturation and the structure of the seed globulins // *Biochemical Journal*. — 1935. — Vol. 29, no. 10. — P. 2351–2360.1.
  15. *Bailer S. M., Balduf C., Hurt E.* The Nsp1p Carboxy-Terminal Domain Is Organized into Functionally Distinct Coiled-Coil Regions Required for Assembly of Nucleoporin Subcomplexes and Nucleocytoplasmic Transport // *Molecular and Cellular Biology*. — 2001. — Vol. 21, no. 23. — P. 7944–7955.
  16. *Baxa U.* Structural Basis of Infectious and Non-Infectious Amyloids // *Current Alzheimer Research*. — 2008. — Vol. 5, no. 3. — P. 308–318.
  17. *Bayliss R., Littlewood T., Stewart M.* Structural Basis for the Interaction between FxFG Nucleoporin Repeats and Importin- $\beta$  in Nuclear Trafficking // *Cell*. — 2000. — Vol. 102, no. 1. — P. 99–108.
  18. *Beck M., Hurt E.* The nuclear pore complex: understanding its function through structural insight // *Nature Reviews Molecular Cell Biology*. — 2017. — Vol. 18, no. 2. — P. 73–89.
  19. *Belgareh N. et al.* An evolutionarily conserved NPC subcomplex, which redistributes in part to kinetochores in mammalian cells // *Journal of Cell Biology*. — 2001. — Vol. 154, no. 6. — P. 1147–1160.
  20. *Bennhold H.* Specific staining of amyloid by Congo red // *Muenchen. Med. Wochenschr.* — 1922. — Vol. 30, no. 1537/1538. — P. 31.
  21. *Benson M. D., Buxbaum J. N., Eisenberg D. S., Merlini G., Saraiva M. J. M., Sekijima Y., Sipe J. D., Westermarck P.* Amyloid nomenclature 2018: recommendations by the International Society of Amyloidosis (ISA) nomenclature committee // *Amyloid*. — 2018. — Vol. 25, no. 4. — P. 215–219.
  22. *Bieler S., Estrada L., Lagos R., Baeza M., Castilla J., Soto C.* Amyloid Formation Modulates the Biological Activity of a Bacterial Protein // *Journal of Biological Chemistry*. — 2005. — Vol. 280, no. 29. — P. 26880–26885.
  23. *Biesecker S. G., Nicastro L. K., Wilson R. P., Tükel Ç.* The Functional Amyloid Curli Protects *Escherichia coli* against Complement-Mediated Bactericidal Activity // *Biomolecules*. — 2018. — Vol. 8, no. 1.
  24. *Bissig C., Rochin L., Niel G. van.* PMEL Amyloid Fibril Formation: The Bright Steps of Pigmentation // *International Journal of Molecular Sciences*. — 2016. — Vol. 17, no. 9. — P. 1438.
  25. *Bonar L., Cohen A. S., Skinner M. M.* Characterization of the Amyloid Fibril as a Cross- $\beta$  Protein // *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)* — 1969. — Vol. 131, no. 4. — P. 1373–1375.

26. *Brohawn S. G., Schwartz T. U.* Molecular architecture of the Nup84–Nup145C–Sec13 edge element in the nuclear pore complex lattice // *Nature Structural & Molecular Biology*. — 2009. — Vol. 16, no. 11. — P. 1173–1177.
27. *Bui K. H. et al.* Integrated Structural Analysis of the Human Nuclear Pore Complex Scaffold // *Cell*. — 2013. — Vol. 155, no. 6. — P. 1233–1243.
28. *Buss F., Kent H., Stewart M., Bailer S., Hanover J.* Role of different domains in the self-association of rat nucleoporin p62 // *Journal of Cell Science*. — 1994. — Vol. 107, no. 2. — P. 631–638.
29. *Callan H. G., Randall J. T., Tomlin S. G.* An Electron Microscope Study of the Nuclear Membrane // *Nature*. — 1949. — Vol. 163, no. 4138. — P. 280–280.
30. *Chai L., Romero D., Kayatekin C., Akabayov B., Vlamakis H., Losick R., Kolter R.* Isolation, characterization, and aggregation of a structured bacterial matrix precursor // *The Journal of Biological Chemistry*. — 2013. — Vol. 288, no. 24. — P. 17559–17568.
31. *Chakravarty A. K., Jarosz D. F.* More than Just a Phase: Prions at the Crossroads of Epigenetic Inheritance and Evolutionary Change // *Journal of Molecular Biology*. — 2018. — Vol. 430, no. 23. — P. 4607–4618.
32. *Chakravarty A. K., Smejkal T., Itakura A. K., Garcia D. M., Jarosz D. F.* A Non-amyloid Prion Particle that Activates a Heritable Gene Expression Program // *Molecular Cell*. — 2020. — Vol. 77, no. 2. — 251–265.e9.
33. *Chapman M. R., Robinson L. S., Pinkner J. S., Roth R., Heuser J., Hammar M., Normark S., Hultgren S. J.* Role of *Escherichia coli* curli operons in directing amyloid fiber formation // *Science (New York, N.Y.)* — 2002. — Vol. 295, no. 5556. — P. 851–855.
34. *Charif D., Lobry J.* SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. // *Structural approaches to sequence evolution: Molecules, networks, populations* / ed. by U. Bastolla, M. Porto, H. Roman, M. Vendruscolo. — New York : Springer Verlag, 2007. — P. 207–232. — (Biological and Medical Physics, Biomedical Engineering).
35. *Charoenkwan P., Kanthawong S., Nantasenamat C., Hasan M. M., Shoombuatong W.* iAMY-SCM: Improved prediction and analysis of amyloid proteins using a scoring card method with propensity scores of dipeptides // *Genomics*. — 2021. — Vol. 113, 1, Part 2. — P. 689–698.
36. *Chiti F., Dobson C. M.* Protein Misfolding, Functional Amyloid, and Human Disease // *Annual Review of Biochemistry*. — 2006. — Vol. 75, no. 1. — P. 333–366.
37. *Chiti F., Dobson C. M.* Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade // *Annual review of biochemistry*. — 2017. — Vol. 86, no. 1. — P. 27–68.
38. *Chopra K., Bawaria S., Chauhan R.* Evolutionary divergence of the nuclear pore complex from fungi to metazoans // *Protein Science*. — 2019. — Vol. 28, no. 3. — P. 571–586.

39. *Christie M. et al.* Structural Biology and Regulation of Protein Import into the Nucleus // Journal of Molecular Biology. — 2016. — Vol. 428, 10, Part A. — P. 2060–2090.
40. *Claessen D., Rink R., Jong W. de, Siebring J., Vreugd P. de, Boersma F. H., Dijkhuizen L., Wósten H. A.* A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in *Streptomyces coelicolor* by forming amyloid-like fibrils // Genes & Development. — 2003. — Vol. 17, no. 14. — P. 1714–1726.
41. *Colombi P., Webster B. M., Fröhlich F., Lusk C. P.* The transmission of nuclear pore complexes to daughter cells requires a cytoplasmic pool of Nsp1 // Journal of Cell Biology. — 2013. — Vol. 203, no. 2. — P. 215–232.
42. *Consortium T. U.* UniProt: the universal protein knowledgebase in 2021 // Nucleic Acids Research. — 2020. — Vol. 49, no. D1. — P. D480–D489.
43. *Coustou V., Deleu C., Saupe S., Begueret J.* The protein product of the het-s heterokaryon incompatibility gene of the fungus *Podospora anserina* behaves as a prion analog // PNAS U. S. A. — 1997. — Vol. 94, no. 18. — P. 9773–9778.
44. *Cronshaw J. M., Krutchinsky A. N., Zhang W., Chait B. T., Matunis M. J.* Proteomic analysis of the mammalian nuclear pore complex // Journal of Cell Biology. — 2002. — Vol. 158, no. 5. — P. 915–927.
45. *Denning D. P., Rexach M. F.* Rapid Evolution Exposes the Boundaries of Domain Structure and Function in Natively Unfolded FG Nucleoporins // Molecular & Cellular Proteomics. — 2007. — Vol. 6, no. 2. — P. 272–282.
46. *Derkatch I. L., Bradley M. E., Hong J. Y., Liebman S. W.* Prions affect the appearance of other prions: the story of [PIN+] // Cell. — 2001. — Vol. 106, no. 2. — P. 171–182.
47. *Destoumieux-Garzón D. et al.* Microcin E492 antibacterial activity: evidence for a TonB-dependent inner membrane permeabilization on *Escherichia coli* // Molecular Microbiology. — 2003. — Vol. 49, no. 4. — P. 1031–1041.
48. *Devos D., Dokudovskaya S., Williams R., Alber F., Eswar N., Chait B. T., Rout M. P., Sali A.* Simple fold composition and modular architecture of the nuclear pore complex // PNAS U. S. A. — 2006. — Vol. 103, no. 7. — P. 2172–2177.
49. *Do H. Q., Hewetson A., Myers C., Khan N. H., Hastert M. C., Harsini F. M., Latham M. P., Wylie B. J., Sutton R. B., Cornwall G. A.* The Functional Mammalian CRES (Cystatin-Related Epididymal Spermatogenic) Amyloid is Antiparallel  $\beta$ -Sheet Rich and Forms a Metastable Oligomer During Assembly // Scientific Reports. — 2019. — Vol. 9, no. 1. — P. 9210–9210.
50. *Dobson C. M., Knowles T. P., Vendruscolo M.* The Amyloid Phenomenon and Its Significance in Biology and Medicine // Cold Spring Harbor Perspectives in Biology. — 2019.
51. *Dosztányi Z., Csizmok V., Tompa P., Simon I.* IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content // Bioinformatics. — 2005. — Vol. 21, no. 16. — P. 3433–3434.

52. *Du Z., Park K.-W., Yu H., Fan Q., Li L.* Newly identified prion linked to the chromatin-remodeling factor Swi1 in *Saccharomyces cerevisiae* // *Nature genetics*. — 2008. — Vol. 40, no. 4. — P. 460–5.
53. *Dueholm M. S., Albertsen M., Otzen D., Nielsen P. H.* Curli Functional Amyloid Systems Are Phylogenetically Widespread and Display Large Diversity in Operon and Protein Structure // *PLoS ONE*. — 2012. — Vol. 7, no. 12. — P. 1–10.
54. *Eanes E. D., Glenner G. G.* X-ray diffraction studies on amyloid filaments // *Journal of Histochemistry & Cytochemistry*. — 1968. — Vol. 16, no. 11. — P. 673–677.
55. *Edgar R. C.* MUSCLE: multiple sequence alignment with high accuracy and high throughput // *Nucleic Acids Research*. — 2004. — Vol. 32, no. 5. — P. 1792–1797.
56. *Elghetany M. T., Saleem A.* Methods for Staining Amyloid in Tissues: A Review // *Stain Technology*. — 1988. — Vol. 63, no. 4. — P. 201–212.
57. *Fernandez-Escamilla A.-M., Rousseau F., Schymkowitz J., Serrano L.* Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins // *Nature Biotechnology*. — 2004. — Vol. 22, no. 10. — P. 1302–1306.
58. *Fontoura B. M., Blobel G., Matunis M. J.* A Conserved Biogenesis Pathway for Nucleoporins: Proteolytic Processing of a 186-Kilodalton Precursor Generates Nup98 and the Novel Nucleoporin, Nup96 // *Journal of Cell Biology*. — 1999. — Vol. 144, no. 6. — P. 1097–1112.
59. *Fowler D. M., Koulov A. V., Alory-Jost C., Marks M. S., Balch W. E., Kelly J. W.* Functional amyloid formation within mammalian tissue // *PLoS Biology*. — 2006. — Vol. 4, no. 1. — e6–e6.
60. *Franz C., Walczak R., Yavuz S., Santarella R., Gentzel M., Askjaer P., Galy V., Hetzer M., Mattaj I. W., Antonin W.* MEL-28/ELYS is required for the recruitment of nucleoporins to chromatin and postmitotic nuclear pore complex assembly // *EMBO reports*. — 2007. — Vol. 8, no. 2. — P. 165–172.
61. *Frey S., Görlich D.* A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability Properties of Nuclear Pore Complexes // *Cell*. — 2007. — Vol. 130, no. 3. — P. 512–523.
62. *Frey S., Görlich D.* FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties // *The EMBO Journal*. — 2009. — Vol. 28, no. 17. — P. 2554–2567.
63. *Frey S., Rees R., Schünemann J., Ng S. C., Fünfgeld K., Huyton T., Görlich D.* Surface Properties Determining Passage Rates of Proteins through Nuclear Pores // *Cell*. — 2018. — Vol. 174, no. 1. — 202–217.e9.
64. *Frey S., Richter R. P., Görlich D.* FG-Rich Repeats of Nuclear Pore Proteins Form a Three-Dimensional Meshwork with Hydrogel-Like Properties // *Science*. — 2006. — Vol. 314, no. 5800. — P. 815–817.
65. *Friedreich N., Kekulé A.* Zur Amyloidfrage // *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin*. — 1859. — Vol. 16, no. 1. — P. 50–65.

66. *Frousios K. K., Iconomidou V. A., Karletidi C.-M., Hamodrakas S. J.* Amyloidogenic determinants are usually not buried // *BMC Structural Biology*. — 2009. — Vol. 9, no. 1. — P. 44.
67. *Galzitskaya O. V., Garbuzynskiy S. O., Lobanov M. Y.* A search for amyloidogenic regions in protein chains // *Molecular Biology*. — 2006a. — Vol. 40, no. 5. — P. 821–828.
68. *Galzitskaya O. V., Garbuzynskiy S. O., Lobanov M. Y.* Prediction of Amyloidogenic and Disordered Regions in Protein Chains // *PLoS Computational Biology*. — 2006b. — Vol. 2, no. 12. — P. 1–10.
69. *Garbuzynskiy S. O., Lobanov M. Y., Galzitskaya O. V.* FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence // *Bioinformatics*. — 2009. — Vol. 26, no. 3. — P. 326–332.
70. *Goldsbury C. S., Wirtz S., Müller S. A., Sunderji S., Wicki P., Aebi U., Frey P.* Studies on the in Vitro Assembly of A $\beta$  1-40: Implications for the Search for A $\beta$  Fibril Formation Inhibitors // *Journal of Structural Biology*. — 2000. — Vol. 130, no. 2. — P. 217–231.
71. *Gößler-Schöfberger R., Hesser G., Reif M. M., Friedmann J., Duscher B., Toca-Herrera J. L., Oostenbrink C., Jilek A.* A stereochemical switch in the aDrs model system, a candidate for a functional amyloid // *Archives of Biochemistry and Biophysics*. — 2012. — Vol. 522, no. 2. — P. 100–106.
72. *Greenwald J., Kwiatkowski W., Riek R.* Peptide Amyloids in the Origin of Life // *Journal of Molecular Biology*. — 2018. — Vol. 430, no. 20. — P. 3735–3750.
73. *Greenwald J., Riek R.* On the Possible Amyloid Origin of Protein Folds // *Journal of Molecular Biology*. — 2012. — Vol. 421, no. 4. — P. 417–426.
74. *Halfmann R., Wright J. R., Alberti S., Lindquist S., Rexach M.* Prion formation by a yeast GLFG nucleoporin // *Prion*. — 2012. — Vol. 6, no. 4. — P. 391–399.
75. *Huang L. et al.* The Identification of Protein-Protein Interactions of the Nuclear Pore Complex of *Saccharomyces cerevisiae* Using High Throughput Matrix-assisted Laser Desorption Ionization Time-of-Flight Tandem Mass Spectrometry // *Molecular & Cellular Proteomics*. — 2002. — Vol. 1, no. 6. — P. 434–450.
76. *Huerta-Cepas J. et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses // *Nucleic Acids Research*. — 2018. — Vol. 47, no. D1. — P. D309–D314.
77. *Hurt E. C.* A novel nucleoskeletal-like protein located at the nuclear periphery is required for the life cycle of *Saccharomyces cerevisiae*. // *The EMBO Journal*. — 1988. — Vol. 7, no. 13. — P. 4323–4334.
78. *Iadanza M. G., Jackson M. P., Hewitt E. W., Ranson N. A., E.Radford S.* A new era for understanding amyloid structures and disease // *Nature Reviews. Molecular cell biology*. — 2018. — Vol. 19, no. 12. — P. 755–773.
79. *Iconomidou V. A., Chryssikos G. D., Gionis V., Galanis A. S., Cordopatis P., Hoenger A., Hamodrakas S. J.* Amyloid fibril formation propensity is inherent into the hexapeptide

- tandemly repeating sequence of the central domain of silkworm chorion proteins of the A-family // *Journal of Structural Biology*. — 2006. — Vol. 156, no. 3. — P. 480–488.
80. *Iconomidou V. A., Chryssikos G. D., Gionis V., Vriend G., Hoenger A., Hamodrakas S. J.* Amyloid-like fibrils from an 18-residue peptide analogue of a part of the central domain of the B-family of silkworm chorion proteins // *FEBS Letters*. — 2001. — Vol. 499, no. 3. — P. 268–273.
  81. *Iconomidou V. A., Vriend G., Hamodrakas S. J.* Amyloids protect the silkworm oocyte and embryo // *FEBS Letters*. — 2000. — Vol. 479, no. 3. — P. 141–145.
  82. *Jain N., Chapman M. R.* Bacterial functional amyloids: Order from disorder // *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. — 2019. — Vol. 1867, no. 10. — P. 954–960.
  83. *Kajava A. V., Baxa U., Steven A. C.* Beta arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils // *The FASEB Journal*. — 2010. — Vol. 24, no. 5. — P. 1311–1319.
  84. *Katoh K., Misawa K., Kuma K., Miyata T.* MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform // *Nucleic Acids Research*. — 2002. — Vol. 30, no. 14. — P. 3059–3066.
  85. *Ke P. C. et al.* Half a century of amyloids: past, present and future // *Chemical Society Reviews*. — 2020. — Vol. 49, issue 15. — P. 5473–5509.
  86. *Kim S. J. et al.* Integrative structure and functional anatomy of a nuclear pore complex // *Nature*. — 2018. — Vol. 555, no. 7697. — P. 475–482.
  87. *Kriventseva E. V., Kuznetsov D., Tegenfeldt F., Manni M., Dias R., Simão F. A., Zdobnov E. M.* OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs // *Nucleic Acids Research*. — 2018. — Vol. 47, no. D1. — P. D807–D811.
  88. *Kustanovich T., Rabin Y.* Metastable Network Model of Protein Transport through Nuclear Pores // *Biophysical Journal*. — 2004. — Vol. 86, no. 4. — P. 2008–2016.
  89. *Kyle R. A.* Amyloidosis: a brief history // *Amyloid*. — 2011. — Vol. 18, sup1. — P. 6–7.
  90. *Kyle R. A.* Amyloidosis: a convoluted story // *British Journal of Haematology*. — 2001. — Vol. 114, no. 3. — P. 529–538.
  91. *Labokha A. A., Gradmann S., Frey S., Hülsmann B. B., Urlaub H., Baldus M., Görlich D.* Systematic analysis of barrier-forming FG hydrogels from *Xenopus* nuclear pore complexes // *The EMBO Journal*. — 2013. — Vol. 32, no. 2. — P. 204–218.
  92. *Lendel C., Bjerring M., Dubnovitsky A., Kelly R. T., Filippov A., Antzutkin O. N., Nielsen N. C., Härd T.* A Hexameric Peptide Barrel as Building Block of Amyloid- $\beta$  Protofibrils // *Angewandte Chemie (International ed. in English)*. — 2014. — Vol. 53, no. 47. — P. 12756–12760.
  93. *Li J. et al.* The RIP1/RIP3 necrosome forms a functional amyloid signaling complex required for programmed necrosis // *Cell*. — 2012. — Vol. 150, no. 2. — P. 339–350.

94. *Lin D. H., Hoelz A.* The Structure of the Nuclear Pore Complex (An Update) // Annual Review of Biochemistry. — 2019. — Vol. 88, no. 1. — P. 725–783.
95. *Lin D. H. et al.* Architecture of the symmetric core of the nuclear pore // Science. — 2016. — Vol. 352, no. 6283.
96. *Lo V. C., Ren Q., Pham C. L. L., Morris V. K., Kwan A. H., Sunde M.* Fungal Hydrophobin Proteins Produce Self-Assembling Protein Films with Diverse Structure and Chemical Stability // Nanomaterials (Basel, Switzerland). — 2014. — Vol. 4, no. 3. — P. 827–843.
97. *Loïodice I., Alves A., Rabut G., Overbeek M. van, Ellenberg J., Sibarita J.-B., Doye V.* The Entire Nup107-160 Complex, Including Three New Members, Is Targeted as One Entity to Kinetochores in Mitosis // Molecular Biology of the Cell. — 2004. — Vol. 15, no. 7. — P. 3333–3344.
98. *Lord C. L., Ospovat O., Wentz S. R.* Nup100 regulates *Saccharomyces cerevisiae* replicative life span by mediating the nuclear export of specific tRNAs // RNA. — 2017. — Vol. 23, no. 3. — P. 365–377.
99. *Löschberger A., Franke C., Krohne G., Linde S. van de, Sauer M.* Correlative super-resolution fluorescence and electron microscopy of the nuclear pore complex with molecular resolution // Journal of Cell Science. — 2014. — Vol. 127, no. 20. — P. 4351–4355.
100. *Lutzmann M., Kunze R., Buerer A., Aebi U., Hurt E.* Modular self-assembly of a Y-shaped multiprotein complex from seven nucleoporins // The EMBO Journal. — 2002. — Vol. 21, no. 3. — P. 387–397.
101. *Macindoe I., Kwan A. H., Ren Q., Morris V. K., Yang W., Mackay J. P., Sunde M.* Self-assembly of functional, amphipathic amyloid monolayers by the fungal hydrophobin EAS // PNAS U. S. A. — 2012. — Vol. 109, no. 14. — P. 5152–5153.
102. *Maji S. K., Wang L., Greenwald J., Riek R.* Structure-activity relationship of amyloid fibrils // FEBS Letters. — 2009a. — Vol. 583, no. 16. — P. 2610–2617.
103. *Maji S. K. et al.* Functional amyloids as natural storage of peptide hormones in pituitary secretory granules // Science (New York, N.Y.) — 2009b. — Vol. 325, no. 5938. — P. 328–332.
104. *Matsuura Y.* Mechanistic Insights from Structural Analyses of Ran-GTPase-Driven Nuclear Export of Proteins and RNAs // Journal of Molecular Biology. — 2016. — Vol. 428, 10, Part A. — P. 2025–2039.
105. *Maurer-Stroh S. et al.* Exploring the sequence determinants of amyloid structure using position-specific scoring matrices // Nature Methods. — 2010. — Vol. 7, no. 3. — P. 237–242.
106. *Michelitsch M. D., Weissman J. S.* A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions // PNAS U. S. A. — 2000. — Vol. 97, no. 22. — P. 11910–5.

107. *Milles S. et al.* Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors // *Cell*. — 2015. — Vol. 163, no. 3. — P. 734–745.
108. *Monsellier E., Chiti F.* Prevention of amyloid-like aggregation as a driving force of protein evolution // *EMBO reports*. — 2007. — Vol. 8, no. 8. — P. 737–742.
109. *Morris V. K., Kwan A. H., Sunde M.* Analysis of the Structure and Conformational States of DewA Gives Insight into the Assembly of the Fungal Hydrophobins // *Journal of Molecular Biology*. — 2013. — Vol. 425, no. 2. — P. 244–256.
110. *Nagy V., Hsia K.-C., Debler E. W., Kampmann M., Davenport A. M., Blobel G., Hoelz A.* Structure of a trimeric nucleoporin complex reveals alternate oligomerization states // *PNAS U. S. A.* — 2009. — Vol. 106, no. 42. — P. 17693–17698.
111. *Nakaya A. et al.* KEGG OC: a large-scale automatic construction of taxonomy-based ortholog clusters // *Nucleic Acids Research*. — 2012. — Vol. 41, no. D1. — P. D353–D357.
112. *Neumann N., Lundin D., Poole A. M.* Comparative Genomic Evidence for a Complete Nuclear Pore Complex in the Last Eukaryotic Common Ancestor // *PLoS ONE*. — 2010. — Vol. 5, no. 10. — P. 1–14.
113. *Niu M., Li Y., Wang C., Han K.* RFAmyloid: A Web Server for Predicting Amyloid Proteins // *International Journal of Molecular Sciences*. — 2018. — Vol. 19, no. 7.
114. *Okonechnikov K., Golosova O., Fursov M.* Unipro UGENE: a unified bioinformatics toolkit // *Bioinformatics*. — 2012. — Vol. 28, no. 8. — P. 1166–7.
115. *Oroz J., Félix S. S., Cabrita E. J., Laurents D. V.* Structural transitions in Orb2 prion-like domain relevant for functional aggregation in memory consolidation // *Journal of Biological Chemistry*. — 2020. — Vol. 295, no. 52. — P. 18122–18133.
116. *Otoo H. N., Lee K. G., Qiu W., Lipke P. N.* *Candida albicans* Als Adhesins Have Conserved Amyloid-Forming Sequences // *Eukaryotic Cell*. — 2008. — Vol. 7, no. 5. — P. 776–782.
117. *Otzen D., Riek R.* Functional Amyloids // *Cold Spring Harbor Perspectives in Biology*. — 2019.
118. *Pages H., Aboyoun P., Gentleman R., DebRoy S.* Biostrings: Efficient manipulation of biological strings. — 2020.
119. *Patel S. S., Belmont B. J., Sante J. M., Rexach M. F.* Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex // *Cell*. — 2007. — Vol. 129, no. 1. — P. 83–96.
120. *Paz M. L. de la, Serrano L.* Sequence determinants of amyloid fibril formation // *PNAS U. S. A.* — 2004. — Vol. 101, no. 1. — P. 87–92.
121. *Peyro M., Soheilypour M., Ghavami A., Mofrad M. R. K.* Nucleoporin's Like Charge Regions Are Major Regulators of FG Coverage and Dynamics Inside the Nuclear Pore Complex // *PLoS ONE*. — 2015. — Vol. 10, no. 12. — P. 1–17.

122. *Pillay S. R. M.* Pili of *Mycobacterium tuberculosis*: current knowledge and future prospects // *Archives of Microbiology*. — 2015. — Vol. 197, no. 6. — P. 737–744.
123. *Podrabsky J. E., Carpenter J. F., Hand S. C.* Survival of water stress in annual fish embryos: dehydration avoidance and egg envelope amyloid fibers // *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*. — 2001. — Vol. 280, no. 1. — R123–R131.
124. *Prusiner S. B.* Novel proteinaceous infectious particles cause scrapie // *Science*. — 1982. — Vol. 216, no. 4542. — P. 136–144.
125. *Quimby B. B., Corbett A. H.* Nuclear transport mechanisms // *Cellular and Molecular Life Sciences : CMLS*. — 2001. — Vol. 58, no. 12. — P. 1766–1773.
126. *Rasala B. A., Orjalo A. V., Shen Z., Briggs S., Forbes D. J.* ELYS is a dual nucleoporin/kinetochore protein required for nuclear pore assembly and proper cell division // *Proceedings of the National Academy of Sciences*. — 2006. — Vol. 103, no. 47. — P. 17801–17806.
127. *Ribbeck K., Görlich D.* Kinetic analysis of translocation through nuclear pore complexes // *The EMBO Journal*. — 2001. — Vol. 20, no. 6. — P. 1320–1330.
128. *Ribbeck K., Görlich D.* The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion // *The EMBO Journal*. — 2002. — Vol. 21, no. 11. — P. 2664–2671.
129. *Riek R., Saube S. J.* The HET-S/s Prion Motif in the Control of Programmed Cell Death // *Cold Spring Harbor Perspectives in Biology*. — 2016. — Vol. 8, no. 9. — a023515.
130. *Roberts B., Wickner R. B.* Heritable activity: a prion that propagates by covalent autoactivation // *Genes & Development*. — 2003. — Vol. 17, no. 17. — P. 2083–2087.
131. *Roche D. B., Villain E., Kajava A. V.* Usage of a dataset of NMR resolved protein structures to test aggregation versus solubility prediction algorithms // *Protein Science*. — 2017. — Vol. 26, no. 9. — P. 1864–1869.
132. *Romero D., Aguilar C., Losick R., Kolter R.* Amyloid fibers provide structural integrity to *Bacillus subtilis* biofilms // *PNAS U. S. A.* — 2010. — Vol. 107, no. 5. — P. 2230–2234.
133. *Romero D., Vlamakis H., Losick R., Kolter R.* An accessory protein required for anchoring and assembly of amyloid fibres in *B. subtilis* biofilms // *Molecular microbiology*. — 2011. — Vol. 80, no. 5. — P. 1155–1168.
134. *Rout M. P., Wente S. R.* Pores for thought: nuclear pore complex proteins // *Trends in Cell Biology*. — 1994. — Vol. 4, no. 10. — P. 357–365.
135. *Sawaya M. R. et al.* Atomic structures of amyloid cross- $\beta$  spines reveal varied steric zippers // *Nature*. — 2007. — Vol. 447. — P. 453–457.
136. *Sazer S., Lynch M., Needleman D.* Deciphering the Evolutionary History of Open and Closed Mitosis // *Current Biology*. — 2014. — Vol. 24, no. 22. — R1099–R1103.

137. *Schmidt C.* Ueber das sogenannte "thierische Amyloid" (Substanz der corpuscula amylacea) // *Justus Liebigs Annalen der Chemie.* — 1859. — Vol. 110, no. 2. — P. 250–254.
138. *Schmidt H. B., Görlich D.* Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity // *eLife* / ed. by K. Weis. — 2015. — Vol. 4. — e04251.
139. *Schwartz K., Boles B. R.* Microbial amyloids - functions and interactions within the host // *Current Opinion in Microbiology.* — 2013. — Vol. 16, no. 1. — P. 93–99.
140. *Schwartz K., Syed A. K., Stephenson R. E., Rickard A. H., Boles B. R.* Functional Amyloids Composed of Phenol Soluble Modulins Stabilize *Staphylococcus aureus* Biofilms // *PLoS Pathogens.* — 2012. — Vol. 8, no. 6. — P. 1–11.
141. *Shahnawaz M., Soto C.* Microcin Amyloid Fibrils A Are Reservoir of Toxic Oligomeric Species // *Journal of Biological Chemistry.* — 2012. — Vol. 287, no. 15. — P. 11665–11676.
142. *Shcheprova Z., Baldi S., Frei S. B., Gonnet G., Barral Y.* A mechanism for asymmetric segregation of age during yeast budding // *Nature.* — 2008. — Vol. 454, no. 7205. — P. 728–734.
143. *Si K., Lindquist S., Kandel E. R.* A Neuronal Isoform of the *Aplysia* CPEB Has Prion-Like Properties // *Cell.* — 2003. — Vol. 115, no. 7. — P. 879–891.
144. *Siniosoglou S., Lutzmann M., Santos-Rosa H., Leonard K., Mueller S., Aebi U., Hurt E.* Structure and Assembly of the Nup84p Complex // *Journal of Cell Biology.* — 2000. — Vol. 149, no. 1. — P. 41–54.
145. *Sivanathan V., Hochschild A.* A bacterial export system for generating extracellular amyloid aggregates // *Nature Protocols.* — 2013. — Vol. 8, no. 7. — P. 1381–1390.
146. *Slotta U., Hess S., Spieß K., Stromer T., Serpell L., Scheibel T.* Spider Silk and Amyloid Fibrils: A Structural Comparison // *Macromolecular Bioscience.* — 2007. — Vol. 7, no. 2. — P. 183–188.
147. *Sopova J. et al.* Amyloid conformers of the FXR1 protein prevent mRNA degradation in cortical neurons // *bioRxiv.* — 2019.
148. *Strawn L. A., Shen T., Wentz S. R.* The GLFG Regions of Nup116p and Nup100p Serve as Binding Sites for Both Kap95p and Mex67p at the Nuclear Pore Complex // *Journal of Biological Chemistry.* — 2001. — Vol. 276, no. 9. — P. 6445–6452.
149. *Stuwe T., Correia A. R., Lin D. H., Paduch M., Lu V. T., Kossiakoff A. A., Hoelz A.* Architecture of the nuclear pore complex coat // *Science.* — 2015a. — Vol. 347, no. 6226. — P. 1148–1152.
150. *Stuwe T. et al.* Architecture of the fungal nuclear pore inner ring complex // *Science.* — 2015b. — Vol. 350, no. 6256. — P. 56–64.

151. *Szymborska A., Marco A. de, Daigle N., Cordes V. C., Briggs J. A. G., Ellenberg J.* Nuclear Pore Scaffold Structure Analyzed by Super-Resolution Microscopy and Particle Averaging // *Science*. — 2013. — Vol. 341, no. 6146. — P. 655–658.
152. *Taylor D. L., Allen R. D., Benditt E. P.* Determination of the polarization optical properties of the amyloid-congo red complex by phase modulation microspectrophotometry // *Journal of Histochemistry & Cytochemistry*. — 1974. — Vol. 22, no. 12. — P. 1105–1112.
153. *Terry L. J., Wente S. R.* Flexible Gates: Dynamic Topologies and Functions for FG Nucleoporins in Nucleocytoplasmic Transport // *Eukaryotic Cell*. — 2009. — Vol. 8, no. 12. — P. 1814–1827.
154. *Thompson J. D., Higgins D. G., Gibson T. J.* CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice // *Nucleic Acids Research*. — 1994. — Vol. 22, no. 22. — P. 4673–4680.
155. *Tian J., Wu N., Guo J., Fan Y.* Prediction of amyloid fibril-forming segments based on a support vector machine // *BMC Bioinformatics*. — 2009. — Vol. 10, no. 1. — S45.
156. *Vassar P. S., Culling C. F.* Fluorescent stains, with special reference to amyloid and connective tissues. // *Archives of pathology & laboratory medicine*. — 1959. — Vol. 68. — P. 487.
157. *Velankar S. et al.* PDBe: Protein Data Bank in Europe // *Nucleic Acids Research*. — 2010. — Vol. 39, suppl.1. — P. D402–D410.
158. *Watson M. L.* Further Observations on the Nuclear Envelope of the Animal Cell // *The Journal of Biophysical and Biochemical Cytology*. — 1959. — Vol. 6, no. 2. — P. 147–156.
159. *Watt B., Niel G. van, Raposo G., Marks M. S.* PMEL: a pigment cell-specific model for functional amyloid formation // *Pigment Cell & Melanoma Research*. — 2013. — Vol. 26, no. 3. — P. 300–315.
160. *Westermarck G. T., Fändrich M., Lundmark K., Westermarck P.* Noncerebral Amyloidoses: Aspects on Seeding, Cross-Seeding, and Transmission // *Cold Spring Harbor Perspectives in Medicine*. — 2018. — Vol. 8, no. 1.
161. *Whelley S., Johnson S., Powell J., Borchardt C., Hastert M. C., Cornwall G. A.* Nonpathological extracellular amyloid is present during normal epididymal sperm maturation // *PloS ONE*. — 2012. — Vol. 7, no. 5. — e36394–e36394.
162. *Wickham H.* The Split-Apply-Combine Strategy for Data Analysis // *Journal of Statistical Software*. — 2011. — Vol. 40, no. 1. — P. 1–29.
163. *Wickham H.* ggplot2: Elegant Graphics for Data Analysis. — Springer-Verlag New York, 2016.
164. *Wickham H.* stringr: Simple, Consistent Wrappers for Common String Operations. — 2019.

165. *Wickham H.* tidy: Tidy Messy Data. — 2020.
166. *Wickham H., Francois R., Henry L., Muller K.* dplyr: A Grammar of Data Manipulation. — 2021.
167. *Wickham H., Seidel D.* scales: Scale Functions for Visualization. — 2020.
168. *Wickner R. B.* [URE3] as an altered URE2 protein: evidence for a prion analog in *Saccharomyces cerevisiae* // Science. — 1994. — Vol. 264, no. 5158. — P. 566–9.
169. *Wösten H. A., Vocht M. L. de.* Hydrophobins, the fungal coat unravelled // Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes. — 2000. — Vol. 1469, no. 2. — P. 79–86.
170. *Wulf M.-A., Senatore A., Aguzzi A.* The biological function of the cellular prion protein: an update // BMC Biology. — 2017. — Vol. 15, no. 1. — P. 34.
171. *Yakupova E. I., Bobyleva L. G., Vikhlyantsev I. M., Bobylev A. G.* Congo Red and amyloids: history and relationship // Bioscience Reports. — 2019. — Vol. 39, no. 1. — BSR20181415.
172. *Yamada J. et al.* A Bimodal Distribution of Two Distinct Categories of Intrinsically Disordered Structures with Separate Functions in FG Nucleoporins // Molecular & Cellular Proteomics. — 2010. — Vol. 9, no. 10. — P. 2205–2224.
173. *Yang W., Willemse J., Sawyer E. B., Lou F., Gong W., Zhang H., Gras S. L., Claessen D., Perrett S.* The propensity of the bacterial rodlin protein RdlB to form amyloid fibrils determines its function in *Streptomyces coelicolor* // Scientific Reports. — 2017. — Vol. 7. — 42867 EP.
174. *Zhang Z., Chen H., Lai L.* Identification of amyloid fibril-forming segments based on structure and residue-based statistical potential // Bioinformatics. — 2007. — Vol. 23, no. 17. — P. 2218–2225.
175. *Zheng Y., Joo H.-S., Nair V., Le K. Y., Otto M.* Do amyloid structures formed by *Staphylococcus aureus* phenol-soluble modulins have a biological function? // International Journal of Medical Microbiology. — 2018. — Vol. 308, no. 6. — P. 675–682.

## А. Приложения

**Таблица I** — Список последовательностей, исключенных из наборов ортологов белков Nsp1, Nup49, Nup54, Nup98, Nup100 и Nup159

Название нуклеопорина	Кол-во последовательностей	ID последовательности	Название вида
Nsp1	2	XP_004861521.1	<i>Heterocephalus glaber</i>
		ELA29607	<i>Colleotrichum gloeosporioides</i>
Nup49	7	XP_004201577.1	<i>Millerozyma farinosa</i>
		XP_004205313.1	<i>Millerozyma farinosa</i>
		GMQ_06001T0	<i>Puccinia graminis</i>
		GMQ_06287T0	<i>Puccinia graminis</i>
		Csp11.Scaffold630.g18990.t2	<i>Caenorhabditis tropicalis</i>
		Csp11.Scaffold630.g18991.t1	<i>Caenorhabditis tropicalis</i>
		Csp11.Scaffold630.g18995.t1	<i>Caenorhabditis tropicalis</i>
Nup54	17	XP_004850144.1	<i>Heterocephalus glaber</i>
		CBN03865	<i>Caenorhabditis brenneri</i>
		CBN016208	<i>Caenorhabditis brenneri</i>
		AMAG_01830T0	<i>Allomyces macrogynus</i>
		AMAG_04333T0	<i>Allomyces macrogynus</i>
		EDO29637	<i>Nematostella vectensis</i>
		XP_008565534.1	<i>Galeopterus variegatus</i>
		ENSCSAVP00000017784	<i>Ciona savignyi</i>
		XP_004208261.1	<i>Hydra vulgaris</i>
		ENSMLUP00000017101	<i>Myotis lucifigis</i>
		ACYP001562-PA	<i>Acyrtosiphon pisum</i>

Продолжение на следующей странице

Название нуклеопорина	Кол-во последовательностей	ID последовательности	Название вида
		SPU_011789-tr	<i>Mycolicibacterium peregrinum</i>
		XP_006049795.1	<i>Bubalus bubalis</i>
		XP_008051116.1	<i>Carlito syrichita</i>
		XP_004316692.1	<i>Tursiops truncatus</i>
		XP_004329063.1	<i>Tursios truncatus</i>
		ENSBTAP00000053680	<i>Bos taurus</i>
Nup98	4	NSGGOP00000028005	Gorilla gorilla gorilla
		ENSMPUP0000008703	Mustela putorius furo
		ENSMPUP0000006735	Mustela putorius furo
		XP_008563391.1	Galeopterus variegatus
Nup100	23	ENSACAP00000013807	<i>Anolis carolinensis</i>
		XP_006785316.1	<i>Neolamprologus bruchardi</i>
		XP_001544558.1	<i>Histoplasma capsulatum</i>
		XP_003037569.1	<i>Schizophyllum commune</i>
		XP_003037570.1	<i>Schizophyllum commune</i>
		FOXG_06462P0	<i>Fusarium oxysporum</i>
		FOXG_07361P0	<i>Fusarium oxysporum</i>
		FOXG_06702P0	<i>Fusarium oxysporum</i>
		FOXG_06802P0	<i>Fusarium oxysporum</i>
		FOXG_07086P0	<i>Fusarium oxysporum</i>
		FOXG_07266P0	<i>Fusarium oxysporum</i>
		FOXG_07361P0	<i>Fusarium oxysporum</i>
		FOXG_16385P0	<i>Fusarium oxysporum</i>
FOXG_17481P0	<i>Fusarium oxysporum</i>		
Продолжение на следующей странице			

Название нуклеопорина	Кол-во последовательностей	ID последовательности	Название вида
		XP_004213012.1	<i>Hydra vulgaris</i>
		EFV616151	<i>Trichinella spiralis</i>
		EFX69233	<i>Daphnia pules</i>
		AAEL0132903-PA	<i>Aedes aegypti</i>
		NV23669-PA	<i>Nasonia vitripennis</i>
		NV24349-PA	<i>Nasonia vitripennis</i>
		ENSGALP00000042128	<i>Gallus gallus</i>
		ENSOANP00000029066	<i>Ornithorhynchus anatinus</i>
		ENSOANP00000030189	<i>Ornithorhynchus anatinus</i>
		XP_004990479.1	<i>Salpingoeca rosetta</i>
		XP_009543518.1	<i>Heterobasidion irregulare</i>
Nup159	37	XP_004850144.1	<i>Rattus norvegicus</i>
		CBN08746	<i>Caenorhabditis brenneri</i>
		CBN16954	<i>Caenorhabditis brenneri</i>
		Csp11.Scaffold629.g14661.t1	<i>Caenorhabditis tropicalis</i>
		Csp11.Scaffold629.g14656.t1	<i>Caenorhabditis tropicalis</i>
		Csp11.Scaffold629.g14658.t1	<i>Caenorhabditis tropicalis</i>
		XP_010750506.1	<i>Larimichthys crocea</i>
		XP_010750522.1	<i>Larimichthys crocea</i>
		XP_007851919.1	<i>Moniliophthora roreri</i>
		XP_007851918.1	<i>Moniliophthora roreri</i>
		XP_007913171.1	<i>Phaeoacremonium minimum</i>
		XP_006754266.1	<i>Myotis davidii</i>
		CJA40505	<i>Streptococcus pneumoniae</i>
Продолжение на следующей странице			

Название нуклеопорина	Кол-во последовательностей	ID последовательности	Название вида
Nup159	37	CJA41762b	<i>Streptococcus pneumoniae</i>
		CJA17179	<i>Streptococcus pneumoniae</i>
		ENSACAP00000020948	<i>Anolis carolinensis</i>
		ENSOGAP00000015017	<i>Otolemur garnettii</i>
		S2JL05	<i>Calyptromyces circinelloides</i>
		S2JLR4	<i>Calyptromyces circinelloides</i>
		S8E9H8	<i>Fomitopsis pinicola</i>
		S8FIP3	<i>Fomitopsis pinicola</i>
		M3IRC0	<i>Candida maltosa</i>
		XP_005975869.1	<i>Pantholops hodgsonii</i>
		XP_005982792.1	<i>Pantholops hodgsonii</i>
		BGIMGA007048-TA	<i>aspergillus parasiticus</i>
		XP_002589550.1	<i>Branchiostoma floridae</i>
		XP_002589549.1	<i>Branchiostoma floridae</i>
		CCF47313	<i>Colleotrichum higginsianum</i>
		CCF45564	<i>Colleotrichum higginsianum</i>
		K1WA41	<i>Trichosporon asahii var. asahii</i>
		ENSOANP00000023273	<i>Ornithorhynchus anatinus</i>
		I1C9B5	<i>Rhizopus delemar</i>
		I1CIE2	<i>Rhizopus delemar</i>
XP_006734685.1	<i>Leptonychotes weddellii</i>		
XP_006734684.1	<i>Leptonychotes weddellii</i>		
XP_006748987.1	<i>Leptonychotes weddellii</i>		

**Таблица II** — Многие ортологи Nsp1 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Fungi	174	0,76	Ascomycota	140	0,8	Dothideomycetes	17	0,795
						Eurotiomycetes	35	0,971
						Leotiomycetes	5	1,0
						Orbiliomycetes	2	1,0
						Peziziomycetes	2	0,0
						Pneumocystidomycetes	2	0,5
						Saccharomycetes	33	0,848
						Schizosaccharomycetes	4	0,75
						Sordariomycetes	39	0,641
						Taphrinomycetes	1	1,0
			Basidiomycota	27	0,63	Agaricomycetes	12	0,667
						Microbotryomycetes	1	0,0
						Mixiomycetes	1	0,0
						Pucciniomycetes	1	0,0
						Tremellomycetes	4	0,75
						Ustilaginomycetes	6	0,667
						Wallemiomycetes	2	1,0
			Blastocladiomycota	2	1,0	Blastocladiomycetes	2	1,0
			Chytridiomycota	2	0,0	Chytridiomycetes	2	0,0
Mucoromycota	3	0,33	Glomeromycetes	1	1,0			
			Mucoromycetes	2	0,0			
Продолжение на следующей странице								

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Metazoa	169	0,76	Annelida	1	0,0	Clitellata	1	0,0
			Arthropoda	35	0,486	Arachnida	1	0,0
						Branchiopoda	1	0,0
						Chilopoda	1	0,0
						Insecta	32	0,531
			Chordata	120	0,883	Actinopteri	19	1,0
						Amphibia	1	1,0
						Ascidiacea	2	1,0
						Aves	7	0,857
						Coelacanthimorpha	1	1,0
						Crocodylia	2	1,0
						Lepidosauria	1	1,0
						Leptocardii	1	1,0
						Mammalia	83	0,843
						Reptilia	3	1,0
			Cnidaria	2	0,5	Anthozoa	1	1,0
						Hydrozoa	1	0,0
			Ctenophora	1	0,0	Tentaculata	1	0,0
			Echinodermata	2	0,0	Echinoidea	2	0,0
			Hemochordata	1	0,0	Enteropneusta	1	0,0
Mollusca	1	1,0	Gastropoda	1	1,0			
Nematoda	2	0,5	Chromadorea	1	0,0			
			Enoplea	1	1,0			

Продолжение на следующей странице

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
			Placozoa	1	0,0	Tricholacia	1	0,0
			Platyhelminthes	2	1,0	Cestoda	1	1,0
						Trematoda	1	1,0
			Porifera	1	0,0	Demospongiae	1	0,0
Protozoa	2	0,5	Choanozoa	2	0,5	Choanoflagellata	1	0,0
						Ichtyosporea	1	1,0
unranked 1	1	1,0	unranked 1	1	1,0	Filastera	1	1,0
unranked 2	1	1,0	unranked 2	1	1,0	Cristidiscoidea	1	1,0

**Таблица III** — Многие ортологи Nup1 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Fungi	34	0,794	Ascomycota	34	0,7941	Pneumocystidomycetes	1	0,0
						Eurotiomycetes	29	0,828
						Schizosaccharomycetes	4	0,75

**Таблица IV** — Немногие ортологи Nup2 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Fungi	46	0,717	Ascomycota	41	0,756	Saccharomycetes	25	0,680
						Eurotiomycetes	9	1,0
						Sordariomycetes	7	0,714
			Basidiomycota	4	0,25	Agaricomycetes	3	0,333
						Wallemiomycetes	1	0,0
						Chytridiomycetes	1	1,0

**Таблица V** — Многие ортологи Nup42 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Fungi	42	0,928	Ascomycota	32	0,968	Saccharomycetes	22	0,954
						Sordariomycetes	5	1,0
						Schizosaccharomycetes	2	1,0
						Dothideomycetes	2	1,0
						Pneumocystidomycetes	1	1,0
			Basidiomycota	10	0,80	Agaricomycetes	4	1,0
						Tremellomycetes	2	0,50
						Ustilaginomycetes	2	0,50
						Mixiomycetes	1	1,0
						Pucciniomycetes	1	1,0

**Таблица VI** — Немногие ортологи Nup49 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Fungi	164	0,951	Ascomycota	136	0,956	Sordariomycetes	39	0,914
						Eurotiomycetes	35	0,886
						Saccharomycetes	30	1,0
						Dothideomycetes	17	0,941
						Leotiomycetes	4	1,0
						Schizosaccharomycetes	4	1,0
						Orbiliomycetes	2	1,0
						Pezizomycetes	2	1,0
						Taphrinomycetes	1	1,0
			Basidiomycota	28	0,928	Agaricomycetes	13	0,923
						Ustilaginomycetes	6	1,0
						Tremellomycetes	4	1,0
						Wallemiomycetes	2	1,0
						Pucciniomycetes	1	1,0
Mixiomycetes	1	1,0						
Microbotryomycetes	1	0,0						
Metazoa	6	0,833	Nematoda	6	0,833	Secernentea	6	0,833

**Таблица VII** — Многие ортологи Nip50 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Metazoa	134	0,947	Chordata	101	0,98	Mammalia	65	0,985
						Actinopteri	19	0,947
						Sarcopterygii	6	1,0
						Aves	5	1,0
						Ascidiacea	2	1,0
						Lepidosauria	2	1,0
						Leptocardii	1	1,0
			Amphibia	1	1,0			
			Arthropoda	21	0,904	Insecta	19	0,895
						Chilopoda	1	1,0
						Branchiopoda	1	1,0
			Nematoda	2	0,5	Enoplea	1	0,0
						Secernentea	1	1,0
			Annelida	1	1,0	Clitellata	1	1,0
			Cnidaria	2	1,0	Hydrozoa	1	1,0
						Anthozoa	1	1,0
			Ctenophora	1	1,0	Tentaculata	1	1,0
Echinodermata	1	0,0	Echinodea	1	0,0			
Hemichordata	1	1,0	Enteropneusta	1	1,0			
Mollusca	1	1,0	Gastropoda	1	1,0			
Porifera	1	0,0	Demospongiae	1	0,0			
Продолжение на следующей странице								

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
			Platyhelminthes	1	1,0	Cestoda	1	1,0
			Placozoa	1	1,0	unranked 1	1	1,0

**Таблица VIII** — Многие ортологи Nup54 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Fungi	37	0,946	Basidiomycota	20	1,0	Agaricomycetes	7	1,0
						Ustilaginomycetes	6	1,0
						Tremellomycetes	3	1,0
						Wallemiomycetes	1	1,0
						Mixiomycetes	1	1,0
						Microbotryomycetes	1	1,0
						Pucciniomycetes	1	1,0
			Blastocladiomycota	2	1,0	Blastocladiomycetes	2	1,0
			Ascomycota	9	1,0	Schizosaccharomycetes	4	0,50
						Pneumocystidomycetes	2	1,0
						Saccharomycetes	1	1,0
						Pezizomycetes	1	1,0
						Taphrinomycetes	1	1,0
			Chytridiomycota	2	1,0	Chytridiomycetes	2	1,0
			Mucoromycota	4	1,0	Mucoromycetes	2	1,0
Glomeromycetes	1	1,0						
Mortierellomycetes	1	1,0						
Metazoa	178	0,938	Chordata	118	0,983	Mammalia	79	0,975
						Actinopteri	20	1,0
						Leptocardii	3	1,0
						Aves	4	1,0

Продолжение на следующей странице

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Metazoa	178	0,938	Chordata	118	0,983	Ascidiacea	3	1,0
						Reptilia	7	1,0
						Sarcopterygii	1	1,0
						Amphibia	1	0,0
			Arthropoda	33	0,940	Insecta	30	0,933
						Brachiopoda	1	1,0
						Chilopoda	1	1,0
						Arachnida	1	1,0
			Annelida	1	1,0	Clitellata	1	1,0
			Nematoda	13	0,615	Chromadorea	12	0,667
						Enoplea	1	0,0
			Platyhelminthes	2	1,0	Cestoda	1	1,0
						Trematoda	1	1,0
			Echinodermata	2	0,5	Echinoidea	2	0,5
			Mollusca	1	1,0	Gastropoda	1	1,0
			Ctenophora	1	1,0	Tentaculata	1	1,0
			Porifera	1	0,0	Denospongiae	1	0,0
Cnidaria	4	1,0	Hydrozoa	2	1,0			
			Anthozoa	2	1,0			
Hemichordata	1	1,0	Enteropneusta	1	1,0			
unranked 1	2	1,0	unranked 1	2	1,0	Choanoflagellata	2	1,0

**Таблица IX** — Многие ортологи Nip57 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Fungi	144	0,903	Ascomycota	133	0,902	Sordariomycetes	40	0,725
						Eurotiomycetes	35	0,971
						Saccharomycetes	32	1,0
						Dothideomycetes	17	0,941
						Leotiomycetes	6	1,0
						Orbiliomycetes	2	1,0
						Pezizomycetes	1	1,0
			Basidiomycota	9	1,0	Agaricomycetes	7	1,0
						Tremellomycetes	1	1,0
						Wallemiomycetes	1	1,0
Mucoromycota	2	1,0	Mucoromycetes	2	0,50			

**Таблица X** — Многие ортологи Nip58 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Metazoa	123	0,983	Chordata	107	1,0	Mammalia	71	1,0
						Actinopteri	19	1,0
						Aves	5	1,0
						Reptilia	5	1,0
						Ascidiacea	2	1,0
						Lepidosauria	2	1,0
						Sarcopterygii	1	1,0
			Amphibia	1	1,0			
			Arthropoda	6	0,833	Insecta	4	0,75
						Chilopoda	1	1,0
						Branchiopoda	1	1,0
			Platyhelminthes	2	0,5	Cestoda	1	1,0
						Trematoda	1	0,0
			Cnidaria	2	1,0	Anthozoa	1	1,0
						Hydrozoa	1	1,0
			Ctenophora	1	1,0	Tentaculata	1	1,0
Echinodermata	1	1,0	Echinoidea	1	1,0			
Hemichordata	1	1,0	Enteropneusta	1	1,0			
Mollusca	1	1,0	Gastropoda	1	1,0			
Porifera	1	1,0	Demospongiae	1	1,0			
unranked 2	1	1,0	unranked 2	1	1,0	Filasterea	1	1,0

**Таблица XI** — Ортологи Nup60 не являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

<b>Царства</b>	<b>Кол-во белков</b>	<b>Доля ПА</b>	<b>Типы</b>	<b>Кол-во белков</b>	<b>Доля ПА</b>	<b>Классы</b>	<b>Кол-во белков</b>	<b>Доля ПА</b>
Fungi	17	0,353	Ascomycota	17	0,353	Saccharomycetes	17	0,353

**Таблица XII** — Все ортологи Nup100 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Fungi	226	0,995	Ascomycota	183	1,0	Saccharomycetes	72	1,0
						Sordariomycetes	42	1,0
						Eurotiomycetes	35	1,0
						Dothideomycetes	18	1,0
						Leotiomycetes	7	1,0
						Schizosaccharomycetes	4	1,0
						Orbiliomycetes	2	1,0
						Pezizomycetes	2	1,0
						Pneumocystidomycetes	1	1,0
			Basidiomycota	37	0,973	Agaricomycetes	20	0,95
						Ustilaginomycetes	5	1,0
						Wallemiomycetes	4	1,0
						Tremellomycetes	4	1,0
						Pucciniomycetes	2	1,0
						Microbotryomycetes	1	1,0
Mixiomycetes	1	1,0						
Chtridiomycota	2	1,0	Chitridiomycetes	2	1,0			
Mucoromycota	3	1,0	Mucoromycetes	2	1,0			
			Mortierellomycetes	1	1,0			
Blastocladiomycota	1	1,0	Blastocladiomycetes	1	1,0			
Metazoa	165	1,0	Chordata	110	1,0	Mammalia	73	1,0

Продолжение на следующей странице

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Metazoa	165	1,0	Chordata	110	1,0	Sarcopterygii	1	1,0
						Actinopteri	19	1,0
						Reptilia	7	1,0
						Aves	6	1,0
						Ascidiacea	2	1,0
						Amphibia	1	1,0
						Leptocardii	1	1,0
			Arthropoda	33	1,0	Insecta	30	1,0
						Branchiopoda	1	1,0
						Chilopoda	1	1,0
						Arachnida	1	1,0
			Cnidaria	2	1,0	Anthozoa	2	1,0
			Nematoda	9	1,0	Rhabditida	8	1,0
						Enoplea	1	1,0
			Platyhelminthes	3	1,0	Trematoda	3	1,0
			Ctenophora	2	1,0	Tentaculata	2	1,0
Echinodermata	1	1,0	Echinoidea	1	1,0			
Hemichordata	1	1,0	Enteropneusta	1	1,0			
Mollusca	1	1,0	Gastropoda	1	1,0			
Placozoa	1	1,0	unranked 1	1	1,0			
Porifera	1	1,0	Demospongiae	1	1,0			
unranked 2	1	1,0	unranked 2	1	1,0	Filasterea	1	1,0
unranked 3	1	1,0	unranked 3	1	1,0	Choanoflagellata	1	1,0

Продолжение на следующей странице

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
unranked 4	2	1,0	unranked 4	2	1,0	Ichthyosporrea	2	1,0

**Таблица XIII** — Многие ортологи Nup153 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Metazoa	141	0,915	Chordata	104	0,933	Mammalia	69	0,98
						Actinopteri	19	0,74
						Reptilia	7	1,0
						Aves	5	1,0
						Sarcopterygii	1	1,0
						Leptocardii	1	1,0
						Ascidiacea	1	1,0
						Amphibia	1	1,0
			Arthropoda	31	0,903	Insecta	29	0,89
						Chilopoda	1	1,0
						Arachnida	1	1,0
			Cnidaria	2	0,5	Hydrozoa	1	1,0
						Anthozoa	1	0,0
			Ctenophora	1	1,0	Tentaculata	1	1,0
Echinodermata	1	1,0	Echinoidea	1	1,0			
Hemichordata	1	0,0	Enteropneusta	1	0,0			
Mollusca	1	1,0	Gastropoda	1	1,0			
unranked 1	1	1,0	unranked 1	1	1,0	Filasterea	1	1,0

**Таблица XIV** — Некоторые ортологи Nup159 являются потенциальными амилоидами; ПА — потенциальный амилоид; unranked 1, unranked 2 — условное обозначение для групп с неопределенным систематическим положением среди царств и типов

Царства	Кол-во белков	Доля ПА	Типы	Кол-во белков	Доля ПА	Классы	Кол-во белков	Доля ПА
Fungi	28	0,785	Ascomycota	25	0,88	Saccharomycetes	21	0,905
						Dothideomycetes	2	1,0
						Sordariomycetes	2	0,5
			Basidiomycota	3	0,0	Pucciniomycetes	3	0,0
Metazoa	24	0,25	Arthropoda	16	0,375	Insecta	16	0,375
			Nematoda	4	0,0	Rhabditida	4	0,0
			Chordata	2	0,0	Mammalia	2	0,0
			Hemichordata	1	0,0	Enteropneusta	1	0,0
			Cnidaria	1	0,0	Anthozoa	1	0,0
unranked 1	1	0,0	unranked 1	1	0,0	Ichthyosporea	1	0,0