

БИОИНФОРМАТИКА

УДК 576.31+519.24
©Коллектив авторов

СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ НЕОДНОРОДНОСТЕЙ БЕЛКОВЫХ И ИНТЕРФАЗНЫХ ХРОМАТИНОВЫХ СЕТЕЙ

Е.И. Гуторов¹, А.Е. Гуторов², Э.М. Коган²

¹ГУ НИИ БМХ РАМН им. Ореховича, Россия, 119121, Москва, Погодинская, 10.
тел : (095)246-7115, факс: (095)245-08-57, эл.почта: fox@ibmh.msk.su
²ГНИВЦ ГТК РФ

Область межклеточных контактов, организация ядерного хроматина, кристы митохондрий, мембраны эндоплазматического ретикулума - все эти структуры на электронных микрофотографиях представляются в виде сложного сетчатого рисунка - текстуры. В работе проводится анализ структур указанного рода с целью выявления общих статистических закономерностей их строения. На основе большого количества экспериментальных данных проанализированы формы гистограмм распределения длин сегментов в белковых и интерфазных хроматиновых сетях различного происхождения: в ядрах и клетках эукариот, в бактерии *E. coli* и вирусах.

В приближении дислокационной модели сетей показано, что экспериментально наблюдаемые гистограммы можно описать следующей формулой:

$$F(x) = (5^3/4!) x^4 e^{-x/l},$$

где $x = l/\langle l \rangle$, l - длина сегмента в сети, а $\langle l \rangle$ - её среднее значение.

В отличие от гауссовского распределения, экспериментальные гистограммы и функция $F(x)$ характеризуются ярко выраженной асимметрией. Распределение $F(x)$ имеет универсальный характер, не зависящий от биохимической природы анализируемых сетей. Универсальность аналитического представления гистограмм распределения длин сегментов свидетельствует о том, что наблюдаемая асимметрия обусловлена фундаментальными свойствами организации монотонных упорядоченных биологических структур.

Ключевые слова: белковая сеть, сеть интерфазного хроматина, сегмент. сеть, распределение длин, асимметрия, дислокационная сеть.

ВВЕДЕНИЕ Наблюдая различные биологические объекты, исследователи достаточно часто сталкиваются с текстурами - сложноупорядоченными рисунками, состоящими из множества мелких повторяющихся элементов. В качестве очевидных примеров можно привести микрорельеф кожи, сосудистую сетку глазного дна и др. Каждую из такого рода текстур можно рассматривать как сеть, в которой узлы соединены с ближайшими соседями сегментами.

Сокращения : БС - белковая сеть, МДМ - морфоденситометрия, СИХ - сеть интерфазного хроматина, ЯМ - ядерный матрикс.

Текстуроподобные структуры наблюдаются и на микрофотографиях, сделанных методом электронной микроскопии. Так, например, выглядит ядерный матрикс (ЯМ), который представляет собой скелетную основу ядра клетки и служит для крепления суперскрученной ДНК.

Биохимически ЯМ состоит из кислых негистоновых белков (90-98%) с небольшой добавкой полисахаридов, липидов и нуклеиновых кислот. ЯМ выполняет функцию обеспечения постоянства пространственной структуры и формы ядра клетки, он также связан с процессами репликации, транскрипции, участвует в процессах транспорта мРНК и во взаимодействии всех содержащихся в ядре компонентов [1].

Белковые сети (БС) также наблюдаются в местах плотных межклеточных контактов [3-7], обуславливающих барьерные функции ткани. Кроме того, БС представлены в той или иной мере в самой цитоплазме клетки и её органеллах [2].

Структурная упорядоченность образований по типу ЯМ и связанных с ним сетей интерфазного хроматина - СИХ и БС играет важную роль в обеспечении нормального функционирования клетки. Пространственному строению ЯМ, БС и СИХ посвящено множество публикаций, как в мировой, так и в отечественной литературе, например, [1-18], однако статистический анализ их неоднородностей не проводился.

В данной работе, на основе многочисленных собственных и литературных данных, полученных методами телевизионной компьютерной морфоденситометрии (МДМ) [8] и электронной микроскопии, определяются статистические характеристики неоднородностей строения ЯМ, БС и СИХ в ядрах и клетках эукариот, в бактерии *E.coli* и вирусах.

МАТЕРИАЛЫ И МЕТОДЫ АНАЛИЗА ИЗОБРАЖЕНИЙ СЕТЕЙ.

Ядерный матрикс гепатоцитов.

Структуры ядерного матрикса получены с применением экстракции изолированных ядер клеток печени крыс неионным детергентом - тритоном X-100 в комбинации с обработкой нуклеазами и экстракцией 2М NaCl.

Затем производился срез ядра и изготавливались платиново-углеродные реплики с последующим просвечиванием на электронном микроскопе с увеличением 15-20 тысяч раз. Исследовались как здоровые, так и опухолевые клетки. Последние отличались от здоровых тем, что имели большую плотность ЯМ [9].

Все микрофотографии с ЯМ были получены в Институте биологии развития АН СССР в лаборатории академика И.Б. Збарского.

Цитоскелет

Для анализа клеточного белкового скелета использовалась микрофотография, полученная с помощью метода электронной микроскопии без заливки среза смолой [10, стр. 312, рис. 11.3 Б]. Увеличение составило 24 тыс. раз.

Вирусы

Для статистических вычислений использовались микрофотографии белковых сетей, приведенные в работе [10] на рисунках 2.1; 2.2, 2.4. Увеличение структур составляло от 105 до 234 тысяч раз.

Плотные контакты

Статистические данные по плотным контактам были взяты из работы авторов [5]. Они получены на основе исследования однородно-связанной БС плотного контакта всасывающих клеток эпителия тонкого кишечника головастика лягушки, свободно связанного пояса замыкания (плотный контакт) толстой кишки головастика, густо связанной гибкой сети плотного контакта, обнаруженной между 2-мя клетками мерцательного эпителия пищевода личинки, сложной БС плотного контакта, сформированного между бокаловидной и жгутиковой клеткой личиночного пищевода головастика, а также, равномерно связанной БС плотного контакта толстой кишки послеметаморфозного головастика из работы [6].

Микрофотографии перечисленных БС были получены методами электронной микроскопии с увеличением 50-60 тысяч раз. Были использованы также микрофотографии БС в области плотных контактов из [7], посвященной исследованию вариаций в плотных и щелевых контактах тканей млекопитающих.

Сети интерфазного хроматина в ядрах лимфоцитов

Структуры СИХ выявлялись в интерфазных ядрах лимфоцитов на фельгеновских препаратах клеточных мазков при компьютерной обработке телевизионного изображения ядра клетки на системах анализа изображений IBAS II фирмы "Opton" (ФРГ) и "Диаморф" (Россия).

В процессе работы сначала измерялись оптические показатели светлых (активных) и темных (пассивных) участков интерфазного хроматина. Затем по денсикограмме строилась методом сегментации приблизительная, но устойчивая бинарная модель текстуры. Далее, комбинацией методов математической морфологии с использованием алгоритма "Горомар" строилась геометрически точная и адекватная модель текстуры или граф СИХ [8]. Увеличение достигало ~10 тыс. раз.

СИХ в ядрах гепатоцитов

Для статистических вычислений использовалась микрофотография петель хроматина ядра гепатоцита крысы, полученная с помощью электронной микроскопии с увеличением 20 тыс. раз.

Применялась методика отпечатка клеток печени на подложке. Затем производился лизис бидистиллированной водой и частичное подсушивание препарата. После обработки формальдегид/уриациетатом, производилось напыление окиси вольфрама и фольга поступала на просвечивание [11].

Микрофотография генома E.coli

Для статистической обработки использовалась микрофотография генома из [12] на рис. 28.5, стр. 348. Нуклеонды из клеток бактерии *E.coli* выделялись непосредственно в виде очень быстро седиментирующего комплекса, который на 80% своей массы состоял из ДНК, организованной в виде сети нитевидных петель.

Все микрофотографии, за исключением плотных контактов, полученные методом электронной микроскопии, подвергались дополнительной телевизионной компьютерной обработке методом МДМ с целью получения визуально четкого однородного по толщине линий графа сетей и выявления их текстуры, структурных и топологических особенностей (рис.1). Затем с помощью алгоритма "Grankor" проводился количественный анализ изображения, при этом длина сегмента считалась эквивалентной числу пикселей, образующих сегмент. Более подробно метод МДМ описан в [8]. Для контроля, параллельно с машинным методом, проводились измерения длин сегментов вручную, с помощью линейки. Расхождение, как правило, было статистически незначимым.

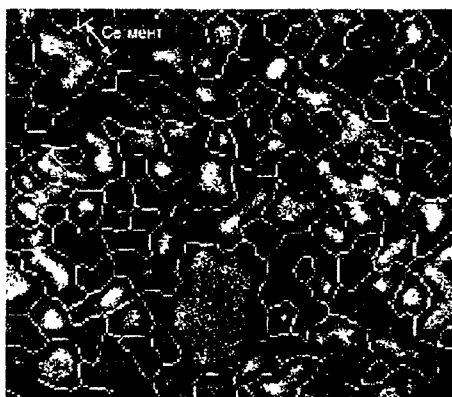


Рисунок 1

Микрофотография ядерного матрикса с наложенной на нее текстурой (белые линии), полученной методом МДМ.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ. Среди множества характеристик сетей, таких как длины сегментов и величина углов между ними в узлах, площадей ячеек сетей, расстояний между центрами ближайших соседних ячеек, плотностью разрывов сплошности сетей и т.д., для статистической обработки были выбраны длины сегментов, как важнейшие, формообразующие элементы сети. Учитывалась также их энергетическая значимость: если сегменты рассматривать как гибкие натянутые струны, то в приближении линейного натяжения, распределение длин сегментов, будучи полученным, представляло бы собой распределение от узла к узлу собственной упругой энергии сегментов для той или иной БС или СИХ.

Исследовалась статистика распределения длин l сегментов с учетом их направления в сетях между узлами-точками пересечения трех и более сегментов. Если сегмент не образовывал узел, его длина не измерялась. Для каждого отдельно взятого ядра, клетки или генома, измеренные в пикселах значения длин l приводились (нормировались) к безразмерным длинам x по формуле:

$$x = l / \langle l \rangle, \quad (1)$$

где $\langle l \rangle$ - средняя длина сегмента в данной выборке.

Найденные значения x со своими частотами, группировались в соответствии с рекомендациями в [19] с интервалом 0,2 для построения гистограмм для всех видов БС и СИХ, приводимых в таблице 1. Затем первичные гистограммы группировались, в свою очередь, для построения суммарных гистограмм определенного вида БС или СИХ, например, "ЯМ гепатоцитов" или "СИХ лимфоцитов". И, наконец, суммарные гистограммы объединялись для построения интегральных гистограмм по всем белковым или хроматиновым сетям. Общая гистограмма объединяла все белковые и хроматиновые сети. Всего было измерено 13125 длин сегментов.

Анализируя численные значения приведенных длин сегментов x , которые сгруппированы в таблице 1, а также форму построенных на их основе первичных, суммарных и интегральных гистограмм, а также общей гистограммы - последние два вида гистограмм построены на рисунках 2а, 3а, 4а, можно заметить следующие характерные особенности их строения. Все гистограммы имеют только один максимум, приходящийся на наиболее вероятный интервал значений длин сегментов, от 0,8 до 1,0. Число коротких сегментов резко нарастает до максимума, равного 0,8, а распределение более длинных сегментов наоборот, плавно спадает.

Таблица 1. Частоты приведенных длин сегментов x исследовавшихся белковых и интерфазных хроматиновых сетей.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
интервал x	БС							СИХ							БС+СИХ	
	гепато- циты, ЯМ	цитос- келет	миту- сы	плотн. конт- ракты М	интер- ракты М	доли α - частоты $\Sigma n/N$	кумуля- тивные $\Sigma n/N$	лимфо- циты	гепато- циты	тканев- Е-сели	интерфаз- ные частоты $\Sigma n/N$	доли α - частоты $\Sigma n/N$	кумуля- тивные $\Sigma n/N$	общие частоты $\Sigma n/N$	доли α - частоты $\Sigma n/N$	кумуля- тивные $\Sigma n/N$
0,0-0,19	7	0	8	13	28	0,0026	0,26	2	1	4	7	0,00307	0,307	35	0,00266	0,266
0,2-0,39	124	7	120	77	328	0,0302	3,28	45	29	58	132	0,05789	6,096	460	0,03505	3,77
0,4-0,59	399	50	581	203	1235	0,1139	14,67	161	94	54	309	0,13547	19,64	1544	0,11764	15,54
0,6-0,79	608	122	952	313	1995	0,1840	31,07	140	132	77	349	0,153	34,94	2344	0,17859	33,39
0,8-0,99	935	229	954	446	2564	0,2364	56,71	245	120	78	443	0,19421	54,364	3007	0,22911	56,30
1,0-1,19	727	152	767	239	1885	0,1738	74,09	209	53	98	360	0,15782	70,146	2245	0,17105	73,40
1,2-1,39	425	101	446	234	1206	0,1112	85,21	152	85	52	289	0,1267	82,816	1495	0,11391	84,80
1,4-1,59	235	59	270	113	677	0,0624	91,45	96	41	41	178	0,07804	90,62	855	0,06514	91,31
1,6-1,79	132	25	188	40	385	0,0355	95,0	47	26	17	90	0,03946	94,56	475	0,0362	94,93
1,8-1,99	59	25	110	58	252	0,0232	97,32	21	27	6	54	0,02367	96,23	306	0,02331	97,26
2,0-2,19	27	14	56	23	120	0,0111	98,43	11	14	13	37	0,01622	98,55	157	0,01196	98,46
2,2-2,39	22	8	46	33	99	0,0091	99,34	1	11	2	14	0,00614	99,17	113	0,00861	99,32
2,4-2,59	6	1	13	10	30	0,0028	99,62	2	5	2	9	0,00395	99,56	49	0,00297	99,62
2,6-2,79	8	4	4	7	22	0,0020	99,82	1	2	4	7	0,00307	99,87	29	0,00221	99,84
2,8-2,99	3	3	3	6	15	0,0014	99,96	0	0	1	1	0,00044	99,913	16	0,00122	99,96
3,0-3,19	1	0	1	1	3	0,0003	99,99	0	0	2	2	0,000877	100	5	0,00038	99,96
СУММА	3718	800	4508	1818	10844	1,00	100	1133	639	509	2281	1,00	100	13125	1,00	100

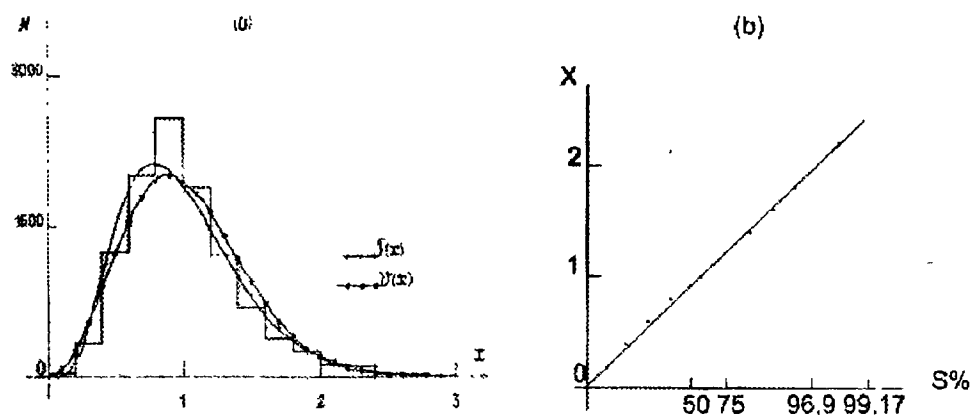


Рисунок 2

Интегральная гистограмма распределения приведенных длин сегментов белковых сетей и её аппроксимация функциями $V(x)$ и $F(x)$ (а) и вероятностной прямой $F(x)$ (б).

Практически вся область значений "х" ограничена интервалом от 0 до 4

Визуальный анализ электронно-микроскопических микрофотографий всех исследованных БС и СИХ и изучение распечаток соответствующих графов после компьютерной обработки изображения показывает, что их общим элементом является текстурная и топологическая неоднородность. В то же время распределение "l", как это следует из рассмотрения табличных данных, имеет универсальный характер, который не зависит от биохимической природы объекта исследования. Далее отметим, что толщина платиново-углеродных или из окиси вольфрама реплик, просвечиваемых в методе ПЭМ, была несоизмеримо мала по сравнению с наблюдавшимися длинами сегментов ячеек, так что измеряемые сегменты лежали практически в плоскости сетей и их видимые длины являлись их истинными длинами, а не проекциями длин. То же самое относится и к сканограммам ядер клеток, полученных методом МДМ при исследовании СИХ лимфоцитов и гепатоцитов.

Все семь суммарных гистограмм (колонки 2-5 и 9-11 в табл. 1), интегральные гистограммы (колонки 6 и 12) и общая гистограмма (колонка 15) обладают общим свойством - они асимметричны

В настоящей работе была предпринята попытка найти универсальную функцию распределения длин сегментов. Для этой цели имело смысл пойти двумя путями. С одной стороны, один из канонических подходов базируется исключительно на статистических характеристиках выборки. С другой стороны, отвлекаясь от статистических оценок, можно было бы предложить биофизическую модель и попытаться описать поведение наблюдаемой системы с позиций этой модели. Далее приводятся результаты, показывающие несостоятельность попытки вывести функцию распределения каноническим методом расчета моментов, а затем предлагается квантово-статистическая и дислокационная модели наблюдаемых сетей. Последняя, с точки зрения авторов, более адекватна, как для формального описания, так и для общего понимания законов организации природных биоструктур.

Для аппроксимации полученных гистограмм аналитическими функциями были выбраны кривые Пирсона. Их статистики определялись методом вычисления моментов. Значения статистик приводятся в таблице 2. Как видно из таблицы 2, все искомые кривые выравнивающих частот относятся к типу VI. Выведенные аналитические формы имеют следующий общий вид:

$$f(x) = 10^{b_0} (x-l)^{b_1} \cdot x^{-b_2} \quad (2)$$

Показатели степеней b_0 , b_1 , b_2 , переменной l и точки максимумов X_0 форм (2) даны в таблице 2

Для СИХ в случае кривой типа III ($\alpha > 4$) имеем.

$$f_{\text{СИХ-III}}(x) = 2154(1+x/0,981)^{5,558} \cdot e^{-3,665x} \quad (3)$$

НЕОДНОРОДНОСТИ БЕЛКОВЫХ И ХРОМАТИНОВЫХ СЕТЕЙ

Таблица 2. Значения статистик распределений кривых Пирсона и их коэффициентов.

Статистики и коэффициенты	Сети		
	БС	СИХ	БС+СИХ
Дисперсия, σ^2	4,376	5,115	4,504
Основное отклонение, σ	2,092	2,262	2,122
Поправка Шеппарда на группирование в разряды, Δ %	0,191	0,177	0,198
Косость, α	0,995	0,781	0,951
Крутость, t	1,630	0,936	1,492
Критерий Пирсона, χ^2	3,195	12,554	3,057
Тип кривой Пирсона	VI	VI	VI
b_0	57,481	636,983	51,56
b_1	509,861	45620,406	972,720
b_2	566,481	45954,692	1031,811
l	0,341	0,212	0,316
Положение максимума X_0	3,413	29,253	5,518

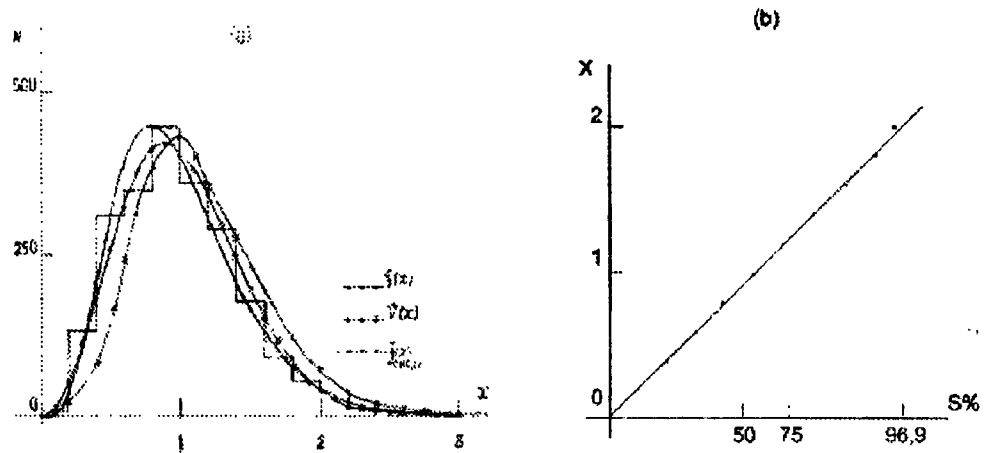


Рисунок 3

Интегральная гистограмма распределения приведенных длин сегментов сетей интерфазного хроматина и её аппроксимация функциями $f(x)$; $V(x)$ и $F(x)$ (а) и вероятностной прямой $F(x)$ (б).

График зависимости 3, нормированный по площади гистограммы, приводится на рисунке 3а. Как видно из таблицы 2 и рисунков, зависимости (2) и (3) не описывают экспериментальные гистограммы как по положению максимумов, так и по протабулированным значениям этих функций.

Поскольку исследуемые сети относятся к нанобиологическим объектам, целесообразно рассмотреть решение задачи аппроксимации методом квантовой статистики.

Сопоставим каждому сегменту, участвующему в образовании БС или СИХ, волновую функцию $\Psi(t_i)$, где t_i - временная координата, $i = \{1, 2, 3, \dots, n\}$. Поскольку онтогенез биологических объектов принципиально необратим во времени, т.е. их энтропия постоянно возрастает, то для описания взаимодействия сегментов в процессе вязания сетей можно использовать только Гамильтониан $H(t)$, не инвариантный относительно операции обращения времени, т.е.:

$$H(t) \neq H(-t) \quad (4)$$

В работе [20] показано, что уровни энергии E_j такого Гамильтониана распределены по следующему закону:

$$V(x) = (\sqrt{\pi} \cdot x)^2 \cdot \exp(\pi x / 2 \sqrt{2})^2 \quad (5)$$

где $x_j = E_j / \langle E \rangle$; $x = \{x_j\}$; $\langle E \rangle$ - среднее значение энергии.

Учитывая принятое выше приближение струнной модели сегментов БС и

СИХ, при котором упругая энергия сегментов $E_s \approx l_s$, распределение (5) можно интерпретировать как распределение длин сегментов от узла к узлу.

Функция $V(x)$, нормированная по площади гистограмм, построена на рисунках 2а, 3а, 4а. При сравнении гистограмм и функции $V(x)$ замечаем их хорошее соответствие между собой по форме, моде и размаху. Однако их максимумы не совпадают. у $V(x)$ они несколько ниже. Поэтому функция $V(x)$, по видимому, также не подходит для описания гистограмм.

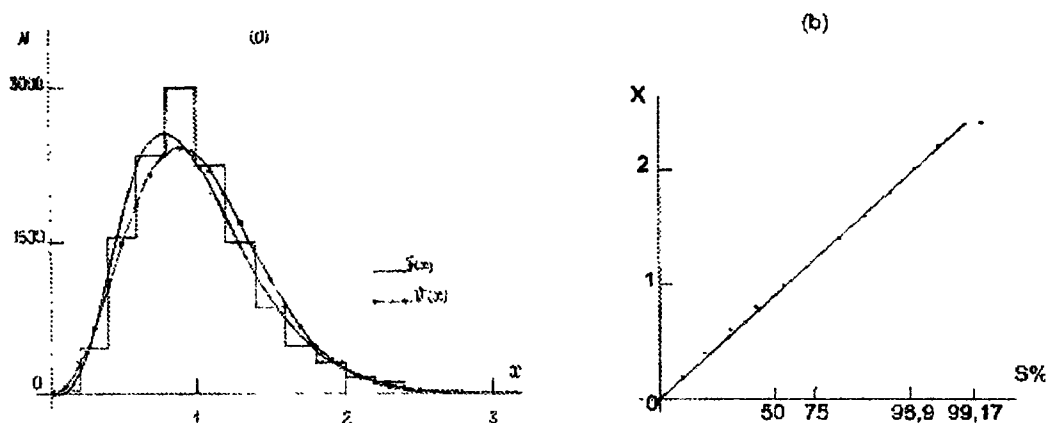


Рисунок 4

Общая гистограмма распределения приведенных длин сегментов белковых и интерфазных хроматиновых сетей и её аппроксимация функциями $V(x)$ и $F(x)$ (а) и вероятностной прямой $F(x)$ (б).

Теперь перейдем к дислокационной модели описания текстур. Модель основывается на проведении аналогии между пространственной организацией фиксирующего аппарата клетки, представленного в виде мембран, включающих в себя белковые и хроматиновые сети, и пространственной организацией жидкого кристалла. В данном контексте узлам сетей могут быть сопоставлены узлы дислокационной сетки в жидком кристалле, а длины сегментов - сегментам дислокаций в жидком кристалле. Указанная аналогия позволяет естественным образом объяснить и кажущуюся неупорядоченность сетчатых структур ЯМ, БС и СИХ.

Как следует из результатов исследования, все рассматриваемые в таблице 1, сети связаны с мембранами. Белковые сети локализованы непосредственно в мембранах, а прицентромержные гетерохроматиновые ядерные сети простираются, в основном, в примембранном слое ядра в виде узкого периферического ободка [8, 21]. В то же время хорошо известно, что строение мембраны аналогично жидкокристаллической смектической фазе типа "А" [22]. Тогда, учитывая принцип Робертсона об универсальности строения мембран, для анализа топологии БС и СИХ применим континуальную теорию дислокаций, разработанную в [23] для смектиков типа "А".

В связи с этим будем полагать, что БС и СИХ имеют дислокационную природу и располагаются в виде плоской, обладающей минимальной поверхностной энергией, границы и образуются системами нерасщепленных дислокаций l_s с векторами Бюргерса b_s в виде некоторой кусочно-прямолинейной конфигурации.

Генерирование дислокационных петель и их вязание в сеть может быть обусловлено за счет флуктуаций при гидрофильно-гидрофобных взаимодействиях в процессе развития и жизнедеятельности организма. Тогда, чтобы обеспечить полную компенсацию возникшего сложного напряженного состояния на границе, достаточно, как это следует из теоремы Тейлора, действия пяти систем скольжения дислокаций [24].

Допустим, что все пять систем скольжения статистически независимы и в каждой из них распределение источников генерации дислокационных петель описывается экспонентой:

$$F(l_j) = (1/\langle l \rangle) \cdot e^{-l/\langle l \rangle} \quad (6),$$

где $\langle l \rangle$ - среднее значение расстояния между центрами генерации. Тогда, учитывая все сделанные допущения и, используя метод характеристической функции [19, 25], можно показать, что плотность распределения вероятностей приведенных длин сегментов x в БС и СИХ подчиняется следующему универсальному закону:

$$F_s(x) = (5^3/4!) \cdot x^4 e^{-5x} \quad (7),$$

В случае действия на гетерогранице меньшего числа систем скольжения n , распределение x можно, при необходимости, сравнить с расчетной формулой общего вида:

$$F_n(x) = (n(nx)^{n-1} e^{-nx}) / (n-1)! \quad (8).$$

При $n=1$ мы получаем экспоненциальный рост коротких, стремящихся к нулю, сегментов. В результате, основной массив выборки, равный $N_0(x \rightarrow 0) = N f_{x \rightarrow 0}(1-x)$, сгруппируется вблизи нулевой области "х", быстро спадая на расстоянии от $x=0$ до $x=5$ до 5% мощности выборки, т.е. по экспоненте. С другой стороны, в исследовавшихся сетях наблюдается только одномодальное распределение числа сегментов вблизи $x=0,8$, т.е. БС и СИХ имеют некоторую регулярность или трансляционную симметрию с периодом решетки, равным наиболее вероятной длине сегмента $l_{ns} = 0,8 \langle l \rangle$. Причина и характер этой регулярности представляют несомненный интерес для последующих исследований.

Распределение (7) имеет максимум в точке $x=0,8$, его форма почти идентична формам экспериментальных гистограмм, включая асимметрию и правую скошенность.

Построенные на рисунках 2а, 3а, 4а теоретические кривые $F(x)$, нормированные по площади гистограмм, практически совпадают с экспериментальными линиями гистограмм.

Для доказательства того, что найденные гистограммы распределения длин сегментов в сетях описываются функцией $F(x)$, была проведена графическая оценка этого распределения. Согласно процедуре такой оценки, описанной в [19], была построена вероятностная бумага для распределения (7) и на нее, против упорядоченных от наименьшего до наибольшего значений $x \{x_1, x_2, \dots, x_n\}$, были проставлены суммарные проценты наблюдений (рис. 2б, 3б, 4б). В соответствии с сущностью этого метода оценки, аппроксимирующее распределение считается выбранным правильно, если точки, нанесенные описанным способом, сгруппируются вблизи прямой линии, хотя и будут наблюдаться некоторые отклонения вследствие случайных колебаний в выборках.

Как видно из рисунков 2б, 3б, 4б, все экспериментальные точки, приводимые в колонках 8, 14, 17 таблицы 1, попадают на вероятностные прямые почти во всем интервале суммарного процента наблюдений или кумулятива S - от 0 до 97% - на логарифмической вероятностной бумаге, что, таким образом, и подтверждает справедливость выбранной модели.

Помимо графической оценки полученных распределений длин сегментов сетей с помощью вероятностной бумаги, была произведена численная оценка уровней значимости о принадлежности каждой пары выборок одной генеральной совокупности, а также согласия между экспериментальными гистограммами и аппроксимирующими их функциями $F(x)$. Ко всем 21 парам из 7 выборок, перечисленных в таблице 1, был применен критерий Колмогорова. В приближении применения критерия Колмогорова мы исходили из того, что рассматриваемые объемы выборок БС и СИХ и их суммы достаточно велики, так что они выступают как представительные классы общей совокупности. В этом случае влияние объединения в разряды можно в значительной степени устранить так называемыми поправками Шеппарда. Данная поправка устраняет смещение выборочной дисперсии σ^2 , вычисленной по группированным данным и, как следует из таблицы 2, во всех 3-х случаях поправка менее 0,2%. Очевидно, что

ошибка менее 0,2% при производстве дальнейших вычислений является несущественной. Результаты вычислений приводятся в таблице 3. Если принять 1%-ный уровень значимости, то из таблицы 3 видно, что в 13 случаях наблюдается расхождение от 0,001% до 0,54%, что составляет 61,9% от всего числа пар выборок. Для остальных 38,1% выборок критерий согласия Колмогорова превышает принятый уровень значимости, доходя до 15% для БС. Таким образом, из таблицы 3 следует, что более 38% рассмотренных пар выборок принадлежат одной и той же генеральной совокупности.

Перейдем теперь к оценке критериев согласия Колмогорова между экспериментальными гистограммами и аппроксимацией их функцией $F(x)$ для полученной одной и той же генеральной совокупности выборок, составляющих 38% от всего числа выборок. Вычисленные значения критериев приводятся в таблице 4. Для БС значение критерия Колмогорова выше и ниже 1% по уровню значимости разделились поровну. Для СИХ же согласие, как видно из таблицы 4, между экспериментальными гистограммами и аппроксимирующей их функцией $F(x)$ является полным, что показано ранее на рисунке 3а. При сложении БС и СИХ расхождение наблюдается в 2-х случаях из 6, что составляет 1/3 от всех выборок БС+СИХ. В остальных 4-х случаях (2/3) значение критерия $[1-K(\lambda)]$ равно от 1,12% до 5%. Таким образом, из таблицы 4 следует, что для 50% БС, 100% СИХ и ~67% их суммы БС+СИХ эмпирические гистограммы распределения нормированных длин сегментов описываются теоретической функцией $F(x)$. В качестве иллюстрации на рисунке 5 приводится пример гистограммы выборки 3 из таблицы 4, содержащей 6326 длин сегментов БС и 509 длин сегментов СИХ и аппроксимирующей её функции $F(x)$. Из рисунка видно, что геометрическое соответствие между ними достаточно хорошее при значении критерия согласия Колмогорова $[1-K(\lambda)] > 1,16\%$. Следовательно, для всех приведенных в таблице 4 эмпирических гистограмм распределения 6835 длин сегментов БС, СИХ и

Таблица 3. Значения критерия согласия Колмогорова $[1-K(\lambda)]$ (%) между парами выборок БС и СИХ.

номер выборки	БС и СИХ	цитоскелет	вирусы	плотные контакты	лимфоциты	гепатоциты (СИХ)	геном бак г e-coli
		Критерий согласия Колмогорова. Значения $[1-K(\lambda)]$ (%) для 21-ой пары 7 выборок.					
1	гепатоциты (БС)	15	$2 \cdot 10^{-2}$	$9 \cdot 10^{-3}$	12	10^{-3}	$3,8 \cdot 10^{-2}$
2	цитоскелет	-	10^{-1}	4,7	$5,4 \cdot 10^{-1}$	$5 \cdot 10^{-2}$	$3 \cdot 10^{-3}$
3	вирусы	-	-	3,75	10^{-3}	$4 \cdot 10^{-2}$	1,86
4	плотные контакты	-	-	-	4,0	8,8	$2,5 \cdot 10^{-1}$
5	лимфоциты	-	-	-	-	$2 \cdot 10^{-2}$	4,20
6	гепатоциты (СИХ)	-	-	-	-	-	$2,3 \cdot 10^{-1}$

Таблица 4. Значения критериев согласия Колмогорова $[1-K(\lambda)]$ (%) между гистограммами и аппроксимирующей их функцией $F(x)$ для выборок, принадлежащих одной генеральной совокупности. N - число сегментов

Номер выборки одной генеральной совокупности	БС		СИХ		БС+СИХ	
	$[1-K(\lambda)]\%$	N	$[1-K(\lambda)]\%$	N	$[1-K(\lambda)]\%$	N
1	10^{-3}	4518	9,24	1133	10^{-3}	5671
2	$8 \cdot 10^{-3}$	2618	-	-	$8 \cdot 10^{-3}$	2618
3	1,26	6326	1,45	509	1,16	6835
4	2,2	1818	1,2	1772	1,2	3590
5	-	-	5,0	1642	5,0	1642
6	-	-	1,12	639	1,12	639

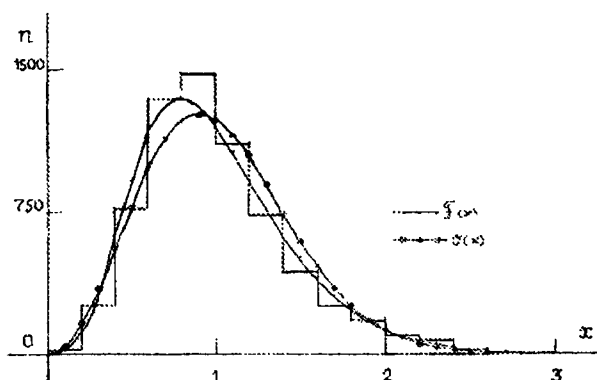


Рисунок 5

Гистограмма генеральной выборки 3 из таблицы 4, содержащей 6326 длин сегментов БС и 509 длин сегментов СИХ и её аппроксимация функциями $F(x)$ и $V(x)$.

БС+СИХ их описание функцией $F(x)$ допустимо с учетом взятого уровня значимости. Отметим также, что, как по положению максимума и области существования, так и по характеру подъема и спада ветвей всех рассмотренных в работе первичных, интегральных и общих гистограмм распределения нормированных длин сегментов ЯМ, БС и СИХ и аппроксимирующей их аналитической функции $F(x)$, наблюдается идентичное соответствие.

ЗАКЛЮЧЕНИЕ. Проведенный статистический анализ показывает, что белковые и хроматиновые сети описываются сходным законом распределения длин сегментов

В практическом плане, наиболее важным применением выявленной закономерности является проведение сравнительного анализа текстурных сетей в нормальных и патологически измененных тканях. При этом можно предполагать, что будет меняться численное значение характерного параметра распределения $I_{н.к.}$ и, соответственно, можно будет статистически достоверно принимать и отвергать гипотезы о наличии изменений. С другой стороны, проведенная работа открывает широкие возможности для моделирования текстурных сетей. С помощью методологии моделирования можно будет прогнозировать процессы восстановления или разрушения скелетных структур клеток (или тканей) под действием лекарства или патогенного микроорганизма, либо физического процесса, например, ультра-, гиперзвукового и радиоволнового облучения.

В заключение статьи авторы выражают благодарность и признательность И.Б. Збарскому - за обсуждение и поддержку работы, А.В. Жукоцкому - за получение им графов сетей.

ЛИТЕРАТУРА

- 1 Збарский И.Б. (1998) Организация клеточного ядра. Москва, Медицина, 368.
- 2 Bershadsky A.D., Vasiliev J.M. (1989) Cytoskeleton. N.-Y. and London. Plenum Press, 151.
- 3 Снигиревская Е.С., Коммисарчик Я.Ю. (1980) Цитология, XXII, 1011-1036.
- 4 Дашевский В.Г. (1987) Конформационный анализ макромолекул. Москва, Наука, 288.
- 5 Гуторов Е.И., Белоусова С.В., Жукоцкий А.В., Коган Э.М. (1986) Биофизика, XXXI, 538.
- 6 Hull B.E., Staehelin L.A. (1976) J. Cell Biol., 68, 688-704.
- 7 Friend D.S., Giluba X.B. (1972) J. Cell Biol., 53, 758-776.
- 8 Жукоцкий А.В. (1992) Компьютерная телевизионная морфоденситометрия нормальных и патологических структур клеток и тканей. Автореферат на

- соискание уч. ст. д. м. н., РГМУ, Москва, 40.
- 9 Збарский И.Б. (1993) Цитология, XXXV, 69-74.
 - 10 Вирусология (1989) (пер. с англ.). Москва, Мир, 429.
 - 11 Эренпрайса Е.А. (1990) Организация хроматина в ядре интерфазной клетки. Рига, Зинантне.
 - 12 Льюин Б. (1987) Гены (пер. с англ.). Москва, Мир, 615.
 - 13 Шерман С.Л., Андрианов А.М., Ахреш А.А. (1989) Конформационный анализ и установление пространственной структуры белковых молекул. Минск, Наука и техника.
 - 14 Brasck K. (1982) *Exp. Cell Res.*, **140**, 161-171.
 - 15 Capko D., Wan K., Penman S. (1982) *Cell*, **29**, 847-858.
 - 16 Hoxars R. Petry (1993) *Molecular biology of Membranes. Structure and functions*. Plenum Press, N.-Y. and London.
 - 17 Горгидзе Л.А., Воробьева И.А. (2004) Биологические мембраны, **21**, 83-93.
 - 18 Бульдяева Т.В., Кузьмина С.Н., Збарский И.Б. (1978) Докл. АН СССР, **241**, 1461-1464.
 - 19 Прохоров Ю.В. (Гл. ред.) (1999) Вероятность и математическая статистика. Энциклопедия. Москва, Большая Российская энциклопедия, 910.
 - 20 Дайсон Ф. (1963) Статистическая теория энергетических уровней сложных систем. Атомиздат, Москва, 124.
 - 21 Ченцов Ю.С. (2004) Введение в клеточную биологию. ИКЦ Академкнига, Москва.
 - 22 Де Жен П. (1977) Физика жидких кристаллов. Мир, Москва.
 - 23 Kleman М. (1974) *J. de Phys.*, **35**, 595-600.
 - 24 Хирт Дж., Лоте И. (1972) Теория дислокаций. Москва, Атомиздат, 600.
 - 25 Королюк В.С. (1978) Справочник по теории вероятностей математической статистики. Наук. Думка, Киев.

Поступила 18.02.2003

STATISTICAL CHARACTERISTICS OF INHOMOGENEITIES OF PROTEIN AND CHROMATIN NETWORKS

E.J. Gutorov¹, A.E. Gutorov¹, E.M. Kogan²

¹Institute of Biomedical Chemistry, Pogodinskaya, 10, Moscow, 119121 Russia.
tel.: (095)246-33-74, fax: (095)245-08-57, e-mail: fox@ibmh.msk.su

²GNIVC GTK RF

Natural textures (networks) are observed in many cases: the inter-cellular contact sites, endoplasmic reticulum membranes etc.

The vast amount of experimental data was analyzed to produce the distribution histograms for the length of the segments in the protein and chromatin networks of different origin. The networks both from the eukaryotic cells and nuclei, as well as from *E.coli* and viruses are presented.

Statistical analysis demonstrated that all experimentally observed histograms fit to the following formula:

$$F(x) = (5^4/4!) x^4 \exp(-5x)$$

where $x = l/\langle l \rangle$, l - length of the network segment, and $\langle l \rangle$ is the average length of the segment.

In contrast to the Gaussian distribution, the distribution of the segments' lengths is markedly asymmetrical. The shape of the distribution does not depend on the origin of the analyzed network.

Keywords: protein, chromatin, networks, segment, histogram, asymmetry, dislocation net.