

АГЛЯДЫ

УДК 612.605;616.612.0;61.575

Л. П. ТИТОВ

МЕДИЦИНСКАЯ ГЕНОМИКА: ОРГАНИЗАЦИЯ ГЕНОМА, РЕГУЛЯЦИЯ ЭКСПРЕССИИ ГЕНОВ, ГЕНЕТИЧЕСКАЯ ВАРИАБЕЛЬНОСТЬ

*Республиканский научно-практический центр эпидемиологии и микробиологии, Минск, Беларусь,
e-mail: leonidtitov@tut.by*

Медицинская геномика – одно из наиболее бурно развивающихся направлений биологической науки. Бум в этой области медицины, наблюдаемый последние 25 лет, связан с началом международного проекта «Геном человека». С расшифровкой нуклеотидной последовательности генома, накоплением знаний о генетической вариативности внутри вида, отдельных этносов (выявлением мутаций, генетических перестроек) связывают возможности более глубокого понимания его биологического значения в предрасположенности к заболеваниям, резистентности к факторам внешней среды, включая биологических агентов. На фоне ускоренного развития структурной геномики отмечается существенное отставание в развитии функциональной геномики, представлений о механизмах работы всего генетического аппарата (хроматина, структурных генов и регуляторных элементов) как единого целого в норме и при патологии. Иммуномика – новое направление, изучающее клетки и гены иммунной системы, их роль в биологии человека. Новые технологии – полногеномное секвенирование, микроэрей технологии и методы биоинформатики находят все более широкое применение в практической медицине, что в конечном итоге приближает переход от избирательного генетического тестирования лиц с наследственными моногенными дефектами и заболеваниями к более широкому использованию и доступности таких методов для каждого человека, что создаст научно обоснованную базу для персонализированной профилактики и терапии распространенных моногенных и мультифакториальных заболеваний.

Ключевые слова: медицинская геномика, иммуномика, экспрессия, регуляция, транскрипция, микро-РНК, секвенирование.

L. P. TITOV

MEDICAL GENOMICS: HUMAN GENOME ORGANIZATION, GENE EXPRESSION REGULATION AND GENETIC VARIABILITY

Republican Scientific-Practical Center of Epidemiology and Microbiology, Minsk, Belarus, e-mail: leonidtitov@tut. by

Medical genomics is one of the most rapidly developing areas of biological science. The boom in this area of medicine is observed in the last 25 years and is associated with the beginning of the international project «Human Genome». With the decoding of the nucleotide sequence of the genome, knowledge about genetic variability within the species, specific ethnic groups (detection of mutations, genetic mutations) linked the possibility of a deeper understanding of its biological significance in susceptibility to disease, resistance to environmental factors, including biological agents. Against the backdrop of accelerated development of structural genomics there is a significant delay in the development of functional genomics, ideas and mechanisms of the entire genetic apparatus (chromatin, structural genes and regulatory elements) as a whole in health and disease. Immunomics is a new direction to study the cells and genes of the immune system and their role in human biology. New technologies are the whole genome sequencing, the microarray technology and bioinformatics methods are increasingly used in medical practice, which ultimately brings the transition from selective genetic testing of persons with hereditary monogenic defects and diseases to a wider use and the availability of these methods for each person, which will create a scientifically based framework for personalized prevention and treatment of common monogenic and multifactorial diseases.

Keywords: medical genomics, immunomics, expression, regulation, transcription, miRNA, sequencing.

Введение. Современный человек сформировался в результате многовековой эволюции, а его геном включает не только «работающие» гены, определяющие архитектуру молекул, клеток, органов, их биологические функции, связь с внутренней и внешней средой, но и огромную инфор-

мационную базу эволюционных событий жизни индивидуума, этнических групп и популяций людей, населяющих планету. Геномы людей в популяции варьируются от 0,01 до 0,1 %, но именно эта относительно небольшая часть определяет уникальные характеристики личности индивидуума, формы и фенотип заболеваний, ответные реакции на инфекцию, иммунизацию, лекарственную терапию. Подавляющее большинство генетических вариантов (аллелей) генома селективно нейтральны, указывают не столько на их патогенетическую роль и ассоциацию с заболеваниями, сколько на природную сбалансированность/несбалансированность фиксированных мутаций (полиморфизмов) с противоположными функциями [1–4]. Вместе с тем изменения во многих генах и их регуляторных участках могут не только способствовать возникновению заболеваний, но и защищать от них.

Большинство заболеваний человека являются полигенными и, соответственно, установление ведущей роли определенного гена представляется затруднительным, а с учетом их связи с триггерными факторами внешней среды такие болезни рассматриваются как мультифакториальные. В то же время регуляторная функция некоторых генов по сравнению с другими может иметь большее биологическое значение. Например, гены ключевых рецепторов мембран клеток, ряда гормонов и витаминов, белков клеточного цикла могут быть причастны к возникновению многих нарушений и заболеваний. Так, ген рецептора витамина Д регулирует транскрипцию 900 генов, а ген рецептора эстрогена – 6000 генов, или 26 % генома человека. Для многих генов иммунной системы и их продуктов (цитокинов, рецепторов) характерна плеiotропность, или множественность действий, которые часто носят разнонаправленный, компенсаторный и/или антагонистический характер [5–7].

Медицинская геномика – новое научное направление, сформировавшееся на стыке медицинской генетики, молекулярной биологии и биоинформатики и базирующееся на технологиях секвенирования и картирования отдельных генов, полных геномов, оценки экспрессии и регуляции активности генов, создании международных баз данных последовательностей биомолекул (ДНК, РНК, белков), компьютерных программ анализа, использовании результатов в медицине с целью выявления значимых генетических вариаций, определения рисков развития заболеваний, создания генно-адресованных лекарств, прогноза течения и исхода заболеваний. Рост интереса к медицинской геномике наблюдается с 1990-х годов (начало проекта «Геном человека»). В настоящее время это направление входит в стадию стабильного развития, характеризуется поступательным ростом числа публикаций, увеличением научных групп и проектов, а секвенирование генома пациентов и микроэрепей технологии становятся доступными широкому кругу исследователей [7–9], что синергично фундаментальным преобразованиям медико-биологической науки. Специалисты основных областей медицины начинают подходить к исследованию и решению этих актуальных проблем с установления генов, ответственных за их возникновение, исследования генетических вариаций, возможных механизмов нарушения экспрессии и разработки на этой основе способов коррекции. Гены, которые несут определенный тип полиморфизма, тесно связанный с тем или иным заболеванием, определяют как гены предрасположенности [6, 7].

Иммунология и научные исследования в этой области в последнее время также претерпевают значительные концептуальные и методологические преобразования. Однако до настоящего времени, несмотря на быстрый прогресс в этой области, механизмы и функции иммунной системы, определяющие стабильность состояния здоровья на протяжении жизни, равно как и ответную реакцию организма на воздействие неблагоприятных факторов среды и развитие на этой основе иммунопатологии, раскрыты недостаточно [8–10]. Перспективы развития медицинской геномики основываются на постулате, что расшифровка нуклеотидной последовательности ДНК генома человека, анализ и интерпретация результатов дадут ответ на многие вопросы фундаментального характера и позволят решить актуальные проблемы практической медицины, включая генную терапию [11].

Геном человека. В последние десятилетия все больше внимания уделяется изучению структурной организации, молекулярно-генетических механизмов и закономерностей функционирования генома в норме и при патологии [12]. Проект «Геном человека», стартовавший 1 октября 1990 г. и завершившийся в 2004 г., явился важнейшим достижением мировой медицины обще-

биологического значения. Он открыл перспективы использования геномных данных для анализа структуры и функции отдельных генов, их комплексов (сетей) и их белковых продуктов, реализующих многообразные иммунобиологические функции и процессы организма в норме и при патологии [13]. С точки зрения информатики геном человека является огромным текстом ДНК, ошибки в котором могут быть причиной возникновения более или менее серьезной патологии. Полное секвенирование генома (ПСГ) человека является важнейшим методологическим подходом медицинской геномики. Информация о полной последовательности генома человека имеется в свободном доступе (<http://genome.ucsc.edu>; <http://www.ensembl.org>; <http://www.ncbi.nlm.nih.gov/genome/quide/human>). Для более глубокого понимания роли генетической информации в обеспечении здоровья, предрасположенности к заболеваниям необходимо как минимум знать/определить последовательность нуклеотидов текста генома (совокупности генов) индивидуума, а также типы и частоту генетической вариабельности в популяции человека в целом и у лиц с определенной патологией. Целью проекта, стартовавшего в 2008 г., было секвенирование геномов 1000 человек, принадлежащих к разным этническим группам, и создание каталога генетических вариаций [14]. При наследственных заболеваниях вариабельность генома обусловлена дефектом отдельного гена, а при мультифакториальных скорее затрагивает не структуру гена, а регуляторные механизмы экспрессии.

В общем виде геном человека представлен сложным комплексом генов ядра клетки и относительно небольшого, но важного набора генов митохондрий. С одной стороны, геном является консервативным и высокостабильным, с другой – весьма изменчивым. Гены ядра клеток сосредоточены в 24 хромосомах (22 аутосомах и X- и Y-хромосомах) (см. таблицу) и содержат от 23 до 25 000 генов, функционирующих в соответствии с центральной догмой молекулярной биологии, т. е. генетическая информация ДНК транскрибируется в мРНК, а мРНК транслируется в полипептиды. Доля таких генов суммарно составляет около 2 % ДНК генома. В геноме имеется 3000–6000 генов, кодирующих молекулы РНК, которые не подчиняются данной концепции, т. е. ДНК транскрибируется в РНК, но РНК не транслируется в полипептиды белков. Общая протяженность генома почти наполовину состоит из однокопийной, или уникальной, ДНК, нуклеотидная последовательность которой представлена однократно. Остальная часть генома характеризуется многократно повторяющимися элементами ДНК, которые поддерживают структурную (хромосомную) организацию генома и в значительной мере определяют генетическую вариабельность популяции. Таким образом, 98 % ДНК генома человека не кодирует белки, а их функциональное значение недостаточно изучено. Приведенная в таблице информация отражает, насколько огромной и сложной системой в количественном, структурном и функциональном плане является геном человека [13–15].

С целью углубления знаний и представлений о механизмах реализации генетической информации, наследовании признаков в норме и при заболеваниях необходимо исследовать, оценивать и интерпретировать основные

Характеристика генома человека
(хромосомы, гены, микро-РНК)
(www.ensembl.org. *Homo sapiens* – whole genome)

Хромосомы (ядро клетки)	Объем информации, Mb	К-во генов	Объем информации генов, Mb	К-во генов <i>miRNA</i>
1	247,25	2153	8,7	68
2	242,95	1313	5,4	60
3	199,50	1105	5,5	57
4	191,27	786	4,1	42
5	180,86	894	4,9	46
6	170,90	1109	6,5	36
7	158,82	1008	6,3	43
8	146,27	743	5,1	38
9	140,27	904	6,4	40
10	135,37	819	6,1	35
11	134,45	1368	10,2	37
12	132,35	1069	8,1	43
13	114,14	356	3,1	23
14	106,37	662	6,2	62
15	100,34	634	6,3	21
16	88,83	902	10,2	20
17	78,77	1217	15,5	40
18	76,12	289	6,0	10
19	63,81	1427	22,4	82
20	62,44	603	9,7	28
21	46,94	283	6,0	10
22	49,69	508	10,2	18
X	145,91	874	5,6	97
Y	57,77	80	1,4	3
Митохондрии (гены, цитоплазма)	0,016	37	–	–

структурно-функциональные элементы генома: а) структуру и функции белок-кодирующих генов; б) строение и функции генов, которые кодируют молекулы РНК, обеспечивающих экспрессию генов и влияющих на формирование предболезни или болезни; в) строение и функции регулирующих транскрипцию и осуществляющих эпигенетический контроль экспрессии генов элементов; г) строение и функцию консервативных элементов, не относящихся к приведенным выше категориям. Ранее предполагалось, что белок-кодирующие гены – это преимущественно репозиторий наследственных признаков. Мутации в белок-кодирующей области гена приводят к продукции атипичных, часто неполноценных белков и к их наследованию, что определяет предрасположенность к патологии. В эпигенетической регуляции генов задействованы не изменения в структуре ДНК (мутации и др.), а процессы химической модификации ДНК (метилование), гистонов и негистонных белков (ацетилирование, фосфорилирование, синтез убиквити-на) [12, 16, 17]. В геноме человека помимо генов предрасположенности и резистентности к заболеваниям выделяют также гены «детоксикации», ассоциированные с реакциями организма на факторы окружающей среды, и «триггерные» гены, обуславливающие инициацию в организме патологических изменений в ответ на воздействие факторов внешней среды.

Хромосомные, или ядерные, гены. Геном человека неравномерно представлен в хромосомах (см. таблицу). Хромосомы 1, 2, 3, 6, 7, 11, 12, 17 и 19 необычайно богаты генами, тогда как хромосомы 13, 18 и 21 относительно бедны ими. Размер генов и число экзонов в них существенно варьируются. Размер гена в среднем составляет 27 kb, но имеются малые (<1 kb) и очень крупные (2400 kb и более) гены. Все клетки организма содержат гены «домашнего хозяйства», детерминирующие совокупность белков, выполняющих общие для всех типов клеток функции (клеточного цикла, метаболизма, цитоскелета, эндоплазматического ретикулума, рецепторов и др.). В механизмах регуляции экспрессии генов участвуют цис-действующие последовательности ДНК и транс-действующие белки. В структуре гена различают кодирующие белок и регуляторные последовательности. Кодирующий участок гена состоит из экзонов (кодируют мРНК) и интронов (не кодируют мРНК). Экзоны содержат точную инструкцию, какой по структуре должна быть мРНК и, в итоге, тот или иной белок. Регуляторные последовательности (промоторы, энхансеры и др.) определяют, какие регуляторные белки, например транскрипционные факторы, способны взаимодействовать с ними, изменять активность гена и скорость биосинтеза полипептидов. Суммарно в геноме выявляется более 300 000 экзонов. Среднее число экзонов, приходящихся на 1 ген, равно 9 (7,7–10,9), а среднее число нуклеотидов в экзоне – 112. Общий объем информации всех экзонов составляет более 78 Mb. Некоторые гены не содержат интронов (гены гистонов), а другие могут содержать их в большом количестве (более 300). Вследствие механизмов альтернативного сплайсинга один и тот же ген может иметь различные изоформы и, соответственно, разные по структуре белковые продукты [12, 18, 19]. Для определенных типов клеток характерен специфичный паттерн экспрессии генов, который рассматривается в контексте специализированной регуляторной сети, а не как экспрессия отдельных генов.

Гены митохондрий. Митохондрии являются важнейшими органеллами клеток человека. В каждой клетке их может содержаться до 1000. Митохондрии, как и бактерии, делятся простым делением. Помимо молекул ДНК в них содержатся и рибосомы. В каждой митохондрии имеется от 2 до 10 молекул ДНК (мтДНК). В яйцеклетках количество молекул ДНК на несколько порядков больше, чем в сперматозоидах. После оплодотворения яйцеклетки мтДНК сперматозоидов разрушается, так как помечена убиквитином. Несколько иначе в сравнении с ядерной ДНК организована мтДНК, которая представляет собой фрагменты двуцепочечной ДНК (тяжелой и легкой цепей – H и L соответственно) циркулярной формы длиной 16 569 пар оснований (п. о.), ассоциирована с митохондриальным матриксом, не защищена гистонами и характеризуется высокой скоростью мутаций. Цепи молекулы мтДНК неравноценны по составу: H-цепь содержит больше пурина, а L- цепь – больше пиримидина. В мтДНК локализуется 37 генов (13 – кодируют белки, 24 – РНК), на долю которых приходится 93 % ДНК. Отличительной особенностью генов митохондрий является отсутствие в их структуре интронов и наследование генетической информации исключительно по материнской линии (оба пола наследуют ДНК матери). Субъединицы белков мультиферментных комплексов транспорта электронов и синтеза АТФ кодируют 13 ге-

нов, которые синтезируются рибосомами митохондрий, а 24 РНК-кодирующих гена определяют структуру и функцию двух типов рРНК (16S и 23S) и 22 типов тРНК (соответственно для каждой из аминокислот) [18, 19].

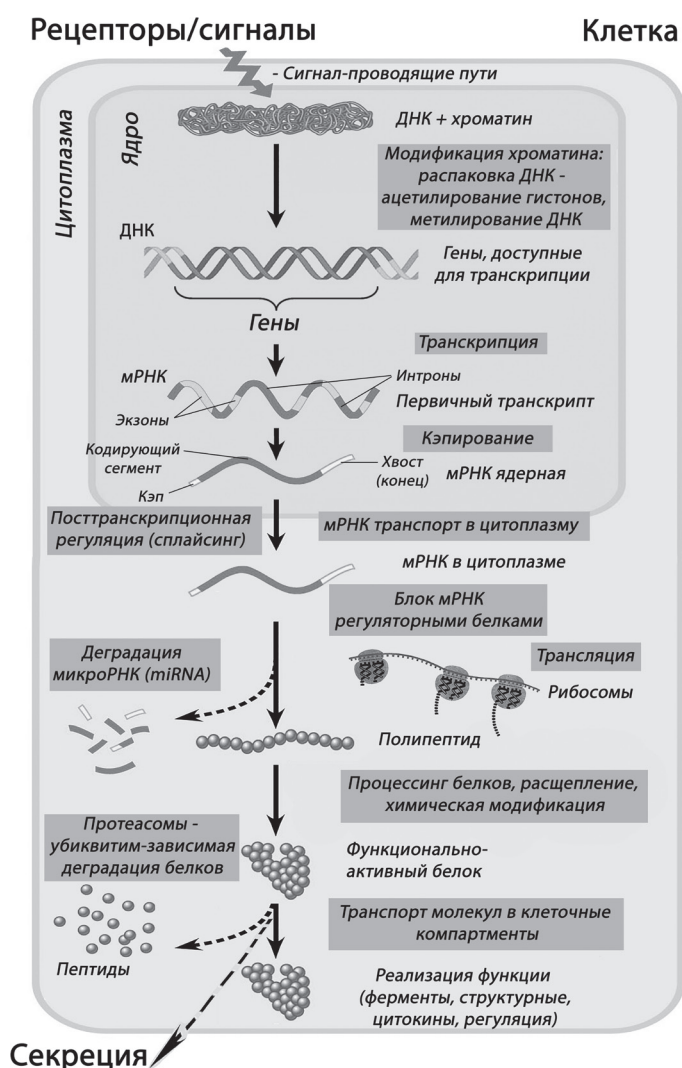
Псевдогены. Псевдогены – инактивированные, «убитые» или нефункционирующие гены. Сиквенсы этих генов могут быть транскрибируемы, но содержат мутации, вследствие которых конечный продукт может утрачивать функциональную активность. Псевдогены могут возникать не только в результате дупликаций гена, когда одна из копий аккумулирует инактивирующие мутации (непроцессированные псевдогены), но и из-за нарушений процессов транскрипции [18, 19].

Гены иммунной системы (ГИС). ГИС характеризуются полиморфизмом, высокой частотой мутаций и соматических гипермутаций (гены иммуноглобулинов), что обеспечивает адаптацию организма к внешней среде и способность развивать высокоспецифичный противомикробный и противоопухолевый иммунный ответ [10]. Уровень и динамика экспрессии ГИС, концентрация их пептидных продуктов в тканях и жидкостях организма варьируются в физиологических пределах, но значительно изменяются при инфекционных заболеваниях и патологических иммуновоспалительных состояниях (при одних усиливаются в 1,5–5,0 раза, при других – в 100 раз и более, а при третьих снижаются) [6, 10, 11, 20]. Весьма заметные изменения в спектре и уровнях экспрессии генов клетками иммунной системы отмечаются в период внутриутробного развития, в детстве и при старении организма [20]. В 2007 г. впервые инициирован проект «Иммунологический геном (ImmGen), объединяющий ученых 15 иммунологических и киберлабораторий. Целью этого проекта являлась разработка «дорожной карты» по изучению экспрессии генов популяциями и субпопуляциями иммунокомпетентных клеток (<http://ImmGen.org/protocols>) [17, 21]. Это послужило основой развития нового направления иммунологии – иммуномики, изучающей регуляцию клеток иммунной системы с использованием подходов структурной и функциональной геномики, анализа и экспрессии генов. Новые методы геномики и биоинформатики дают исследователям все больше возможностей для выявления и визуализации биологических сетей, понимания их функционирования. Термин «**иммунот**» относится к клеткам, генам и белкам, участвующим в иммунном ответе, за исключением экспрессируемых в других типах клеток. Реакции иммунного ответа – «**реакции иммунот**». Совокупность генов «иммунот» объединяет 893 гена, участвующих в генезе 173 заболеваний. Из них 181 ген кодирует белки, интегрированные с мембранами иммунокомпетентных клеток (белки сигнальных путей, ядерных факторов), 91 – внеклеточные белки (иммуноглобулины, компоненты системы комплемента, цитокины, хемокины), а 119 – рецепторы (CD молекулы, Fc-рецепторы, рецепторы цитокинов и др.) [5, 21–23]. В норме около 2/3 всей совокупности генов человека находятся в активном состоянии. В процессе онтогенеза гены иммунной системы, ответственные за дифференцировку клеток, экспрессию поверхностных рецепторов (ВКР, ТКР, цитокинов и др.), включаются в разные периоды внутриутробного развития и созревания в постнатальном периоде [24–26].

Геномная карта. Геном человека разнообразен структурно и функционально. В нем имеется около 100 000 участков, которые незначительно, но наследственно различаются по структуре, хотя и выполняют одинаковые функции. Их используют в качестве маркеров – меток для картирования генома. Поэтому в геномной карте (представлена на веб-сайте номенклатуры генома – <http://genenames.org>) приводятся положение, размер и названия генов. В настоящее время картировано более 1000 генов человека, связанных с заболеваниями. Предполагается наличие в геноме около 8000 участков, имеющих потенциальные изменения для возникновения патологии. Гены, кодирующие белки, структурированы в семейства на основе их филогенетического подобия. Семейства генов являются результатом длительного процесса эволюции, сформировавшего геномы человека и других видов животных, и характеризуются разным уровнем гомологии. Члены семейства генов, как правило, образуют один или несколько кластеров, но могут быть и рассредоточены по геному. Например, выделяют гены иммуноглобулинов, глобина, цитокинов, молекул главного комплекса гистосовместимости, кластера дифференцировки и др. База данных Pfam (<http://pfam.sanger.ac.uk>, <http://www.uniprot.org>) является коллекцией доменов белков [48]. Идентификация молекулярных доменов современными методами биоинформатики по-

звolyет предсказывать и анализировать структуру и функции разнообразных молекул, а прочтение последовательностей нуклеотидов определенных участков генома – диагностировать многие заболевания. Например, в настоящее время в Японии все новорожденные дети проходят тестирование на наличие 11 наследственных генетических заболеваний, в США – 7, в России – 2 (фенилкетонурия и гипотироз). За исключением однойцевых близнецов, у всех людей в геноме имеются значительные генетические вариации. В качестве стандарта для сравнительных исследований последовательностей полных геномов используют данные референтного генома человека. В перспективе создание «идеального» референтного генома позволит существенно повысить качество интерпретации выполняемых исследований.

Структурная организация генома. Как известно, генетическая информация кодируется и хранится в форме последовательностей азотистых оснований (сиквенсов) молекулы ДНК. Основной объем ее содержится в ядре, т. е. в хромосомах, небольшая часть находится в цитоплазме в виде 2–10 молекул ДНК, связанных с каждой митохондрией. В типичном ядре клеток содержится примерно 2 м ДНК, разделенных между 23 парами хромосом (в среднем по 5–6 см на 1 хромосому). Перед делением клетки ее длина уменьшается посредством суперспирализации и упаковки [12]. Упаковка ДНК в хроматин обеспечивает многократное сокращение линейных размеров молекулы. На рисунке схематично представлены строение генетического аппарата клетки, рецепция мембранных сигналов, передача их в ядро клетки, процессы транскрипции и трансляции генетической информации, посттранскрипционной регуляции и биосинтез белковых продуктов.



Основные этапы реализации генетической программы клетки (E. Barley, K. Fitzpatrick. Campbell Biology. Pearson Education, Inc., 2011)

Нуклеосомы. Нуклеосома – структурный элемент хромосомы, образованный цепью молекулы ДНК в 1,65 оборота вокруг глобулярных агрегатов (ядра) из 8 гистоновых белков (по 2 молекулы гистонов H2a, H2b, H3 и H4). Она формируется 147–200 п. о. ДНК. Бусинкоподобные образования спирально скручиваются в соленоид или спиралевидную катушку, состоящую из 5 или 6 нуклеосом на один оборот и связаны одной молекулой гистона H1. Молекулы гистонов несут положительный заряд и образуют ионные связи с отрицательно заряженными фосфатными группами ДНК. Внутри вида последовательности аминокислот гистонов близки к 100 % гомологии. Такая консервативность гистонов указывает на их чрезвычайную важность как в обеспечении структуры и функции хроматина, так и для функционирования всего генома [6, 7, 14].

Упаковка ДНК в хроматин многократно сокращает ее линейные размеры: формирование соленоидов делает ее суммарно короче в 50 раз. Перед фазой митоза хромосомы конденсируются, ДНК компактизуется, сокращаясь в длину в 100 раз [14, 15]. В интерфазе для обеспечения процессов транскрипции и репликации происходит декомпактизация

хромосом. В структуре хроматина различают участки с более конденсированным (гетерохроматин) и деспирализованным (эухроматин) состоянием элементарных дезоксирибонуклеопротеидных нитей [7, 12].

Гетерохроматин сконцентрирован преимущественно вокруг периферии ядра и относительно неактивен в процессах транскрипции. ДНК этих областей сильно метилирована, гистоны деацетилованы, а гены выключены. Таким образом, гетерохроматин выполняет барьерную функцию, препятствует связыванию факторов транскрипции с участками ДНК и, соответственно, транскрипции. Конститутивный гетерохроматин является общим для всех клеток организма, тогда как факультативный варьируется, представляя области генома, дифференцированно экспрессируемые разными типами клеток [6, 7, 14].

Эухроматин, или активный хроматин, – участки хроматина, сохраняющие деспирализованное состояние элементарных дезоксирибонуклеопротеидных нитей в покое ядре, т. е. в интерфазе. ДНК этой области менее метилирована, гистоны более ацетилованы, соответственно, гены более доступны для взаимодействия с регуляторными белками и экспрессии. Эухроматин отличается от гетерохроматина интенсивным синтезом РНК и большим содержанием негистоновых белков. В упакованных хромосомах это проявляется наличием тускло окрашенных R-полос, в которых представлено большинство структурных генов организма. Участки, свободные от нуклеосом (открытый хроматин – open chromatin), – сайты связывания факторов транскрипции (ФТ) и других регуляторных белков. Посадке нуклеосом, как правило, препятствуют связанные с хромосомами белковые факторы, узнающие определенные последовательности ДНК – ФТ, ДНК и РНК-полимеразы. Эти участки, обычно совпадающие с цис-регуляторными последовательностями – промоторами, энхансерами, инсуляторами, сайленсерами и участками начала репликации ДНК, получили название участков, гиперчувствительных к ДНК-азе, или гиперчувствительных сайтов. Их длина составляет примерно 300 п. о. Области открытого хроматина окружены гистонами H3, ацетилованными по 27-му лизину (H3K27Ac), что является своеобразной меткой активных регуляторных областей хроматина. Области открытого хроматина возникают под действием факторов, ответственных за сборку, разборку и перемещение нуклеосом (так называемое ремоделирование хромосом) [25].

В нуклеопротеидных комплексах с молекулой ДНК связано множество негистоновых белков – РНК-полимераз, ДНК-топоизомераз, белков-активаторов и репрессоров транскрипции. Петли хроматина отдельных и разных хромосом располагаются в участках ядра, обогащенных ФТ и РНК-полимеразой. Одним из таких белков является топоизомераза II, необходимая для релаксации суперспирализованной ДНК при репликации или транскрипции. Она связывается с участками, обогащенными АТ мотивами (т. е. содержащими более 65 % оснований А и Т). Это означает, что каждая такая петля может действовать как функционально независимый домен. Молекулы РНК-полимераз катализируют синтез всех типов РНК. Разнообразие клеток, множество генов, сложность клеточных процессов и межклеточного взаимодействия нуждаются и в большом разнообразии механизмов регуляции [15, 16].

Эпигенетическая регуляция экспрессии генов. Функция гистонов при конденсации-деконденсации молекулы ДНК регулируется биохимически. Ацетилирование гистонов осуществляется присоединением ацетильных групп к положительно заряженному лизину, что инициирует транскрипцию. Метилирование оснований ДНК, наоборот, конденсирует хроматин и тормозит экспрессию генов. В ДНК животных метилировано от 2 до 7 % остатков цитозина. Фосфорилирование аминокислот гистонов также приводит к изменению структуры хроматина. Ферменты, модифицирующие хроматин, контролируют экспрессию генов в участках эухроматина [24, 25].

В многоклеточном организме регуляция экспрессии генов необходима для своевременной и полной реализации и поддержания на необходимом уровне жизненно важных общих процессов и этапов развития специализированных клеток (от стволовой полипотентной клетки до зрелых форм нейтрофилов, макрофагов, Т- и В-лимфоцитов) и последующей программированной их гибели. Морфологические и функциональные различия между основными типами клеток объясняются дифференцированной экспрессией определенного паттерна генов. При этом нарушение экспрессии генов, отвечающих за определенный тип дифференцированных клеток, может

привести к изменению их функциональной роли, быть причиной возникновения разнообразных патологических состояний (сердечно-сосудистых, нейродегенеративных, злокачественных новообразований и др.) [26–28]. В организме существуют механизмы, поддерживающие стабильную и цикличную репрессию одних генов и дерепрессию других на протяжении всей жизни. В то же время адаптивная регуляция активности генов обеспечивает своевременное реагирование организма в ответ на меняющиеся условия среды [16, 17, 26]. Известно, что ряд дифференцированных клеток продуцируют весьма специализированные белки. Например, гепатоциты продуцируют фактор VIII системы свертывания крови, бета-клетки поджелудочной железы – инсулин, лимфоцитарные плазматические клетки – иммуноглобулины (антитела). Закономерности регуляции экспрессии таких органоспецифичных генов и регуляции синтеза их продуктов изучены недостаточно [29, 30].

Промоторы – фрагменты ДНК, состоящие примерно из 500 п. о. и располагающиеся перед белок-кодирующим сиквенсом (5') или более дистанционно. Они включают множество различных регуляторных элементов (мотивов), которые привлекают и связывают молекулы транскрипционных факторов, обеспечивают сборку РНК полимеразного комплекса. Имеются различия в количестве, ориентации и расстоянии между промоторами разных генов. Описано несколько типов цис-элементов, например ТАТА бокс (25–30 п. о., расположен вблизи участка инициации транскрипции), СААТбокс (70–80 п. о.), GC box (110 п. о., локализуется проксимально) и октамер (120–130 п. о.). С ТАТА боксом взаимодействует специальный ТАТА-связывающий белок (ТВР), вокруг которого собираются компоненты преинициаторного комплекса РНК – Pol-II. Плавление ДНК и формирование открытого комплекса РНК – Pol-II происходит на некотором расстоянии от промотора. Для развития патологических состояний наличие мутаций и полиморфизм в промоторной области генов имеют большее значение, чем в белок-кодирующей [16, 17, 30–34].

РНК-полимеразы. Процесс транскрипции осуществляется при участии трех различных типов РНК полимераз (RNA Pol), каждая из которых ответственна за разные классы транскрипции. Pol-I транскрибирует гены рибосомальной РНК (*rRNA*), Pol-II – гены матричной РНК (*mRNA*), т. е. фактически гены мРНК всех белок-кодирующих генов, включая гены микро-РНК, а Pol-III – гены транспортной РНК (*tRNA*) и другие малые РНК. Транскрипция представляется наиболее важным этапом жизненного цикла мРНК и ответственна не только за синтез транскриптов, но и за 5' кэпирование, сплайсинг, формирование 3' поли-А хвоста, конвертацию пре-мРНК, трансляцию и деградацию, а кроме того, она регулирует длину 5' и 3' нетранслируемых областей. Различные гены в разное время суток экспрессируются с неодинаковой интенсивностью. Наивные В- и Т-лимфоциты, моноциты, макрофаги и дендритные клетки активируются сигналами, вызывающими пролиферацию и дифференциацию субпопуляций клеток, в основе которых лежит усиление или угнетение экспрессии эффекторных и регуляторных генов [24, 25]. Для начала транскрипции белок-кодирующих генов РНКII-полимераза нуждается в участии ФТ. Последние осуществляют свои функции самостоятельно или в комплексе с другими белками, усиливающими (активаторы) или блокирующими (репрессоры) привлечение молекул РНК-полимеразы [25, 30].

Факторы транскрипции. ФТ – совокупность белков или транс-действующих элементов, обеспечивающих прочтение и интерпретацию генетической информации, содержащейся в ДНК генов [27, 28]. Они связываются не со строго специфичной первичной структурой ДНК, а с группой гомологичных структур разной степени родства. Взаимодействие осуществляется за счет электростатических сил, водородных связей и сил Ван-дер-Ваальса. Участки ДНК, взаимодействующие с ФТ, называются сайтами связывания. Они имеются во всех живых организмах. Крупные геномы содержат большее их количество. Примерно 10 % генов человека кодирует регуляторные белки (около 2600), большинство из которых (примерно 2000) выполняют функцию ФТ. Эти белки кодируются генами хромосом и митохондриальной ДНК, транскрибируются в молекулы мРНК и транслируются в молекулы белков в цитоплазме [29, 30].

ФТ имеют модульную структуру и состоят из трех доменов: 1) ДНК-связывающего домена (ДСД/DBD), комплементарного специфичным последовательностям ДНК (мотивам) промоторов и энхансеров (выделяют димеризационный или олигомеризационный домен, который связывается с ДНК в виде димеров или с комплексами более высокого порядка); 2) транс-активирующего

или транс-репрессирующего домена (ТАД/ТАД), содержащего участки связывания с другими белками, например с ко-регуляторами – активаторами, репрессорами; 3) сигнал-распознающего или лиганд-связывающего домена (СПД/SSD) (например, домен, который связывает лиганд, чувствительный к воздействию внешних сигналов, и отвечает за их передачу другим компонентам транскрипционного комплекса, что проявляется повышением или понижением уровня экспрессии) [30, 31].

Семейства и классы ФТ. По структурной организации ФТ подразделяются на следующие семейства: а) спираль–поворот–спираль (helix-turn-helix – НТН), функционируют в форме димеров (гомеодомены-белки); б) цинковые пальцы (zink finger-ZnF), содержатся в клетке от 1 до 10 (например, Sp1, Elr-1, WT1); в) лейциновая молния (basic leucine zipper-bzip), функционируют в форме гомо- и гетеродимеров (C/EBPb – c-Fos, c-Jun), CREB; г) спираль–петля–спираль (basic helix – loop-helix- bHLH), функционируют в форме гомо- или гетеродимеров (MyoD, E-protein). Различают три класса ФТ: 1) главные, или общие (general transcription factors – GTFs), вовлечены в образование инициативного комплекса, присутствуют во всех клетках и взаимодействуют с кором промотора генов, транскрибируемых РНК-полимерозой II; 2) взаимодействующие с upstream участком ДНК, т. е. с областями, расположенными впереди промотора; 3) индуцируемые ФТ, сходны с предыдущими. Фоновая транскрипционная активность генома обеспечивается набором главных ФТ. Многие из них связываются с ДНК не прямо, а входят в состав комплекса инициации транскрипции, взаимодействующего с РНК-полимеразой. Наиболее распространенными ФТ общего свойства являются FT-IIA, FT-IIB, FT-IID (связываются с ТАТА-боксом), а также FT-IE, FT-IF и FT-IH. Они необходимы для инициации транскрипции всех промоторов, определяют участок инициации и образуют с РНК Pol-II комплекс инициации, называемый аппаратом базальной транскрипции и окружающий стартовую точку [32–35].

ФТ с upstream-участком ДНК – белки, распознающие короткие специфические сиквенсы (мотивы) консенсусных элементов ДНК генов, расположенных перед стартовой точкой транскрипции (например, Sp1, связывающиеся с GC-элементами). Они повышают эффективность функционирования состояния инициации. Индуцибельные факторы (inducible factors) функционально близки к последним, обладают регуляторными свойствами, синтезируются и активируются в строго определенное время в специфических тканях [28]. Дополнительные белки, такие как активаторы, ремодуляторы гистонов (гистонацетилазы, гистондеацетилазы), киназы, метилазы, белки цинкового пальца, ядерные факторы, STAT белки (signal transducers and activator transcription), белки семейства Мус и белки, регулирующие клеточный цикл семейства E2F и p53, также участвуют в регуляции активности генов [30, 31].

Механизм действия ФТ. Для проявления активности ФТ связываются с комплементарными им промоторными или энхансерными участками ДНК регулируемых генов. В зависимости от ФТ транскрипция ближайшего гена будет или усилена (up regulated), или снижена (down regulated) [29]. Важное значение в регуляции экспрессии генов играют и другие факторы. С этой целью ФТ используют разные механизмы: а) стабилизацию или блокирование связывания РНК-полимеразы с ДНК на ТАТА бокс-инициаторном элементе за счет взаимодействия их транс-активирующих доменов с компонентами базального транскрипционного комплекса; б) изменение структуры гистонов путем катализа реакций ацетилирования или деацетилирования. Многие ФТ используют один из двух противоположных механизмов регуляции транскрипции: 1) ацетилирование гистонов ацетилтрансферазой ослабляет ассоциацию ДНК с гистонами и повышает доступность генов для транскрипции (повышение транскрипции); 2) деацетилирование гистонов деацетилазой усиливает их ассоциацию с ДНК и, соответственно, приводит к утрате доступности генов к транскрипции (угнетение экспрессии). Кроме того, ФТ рекрутируют молекулы ко-активаторов и ко-репрессоров для образования комплексов ДНК–ФТ. Существуют также специфические факторы транскрипции, включающие/выключающие определенные гены в нужный момент. Одни ФТ участвуют в процессах онтогенеза человека (семейство гомеобоксных ФТ, ответственных за морфологию, Y-регион, определяющий пол (sex-determining region Y), другие отвечают на внеклеточные сигналы (эстрогеновый сигнальный путь, цитокиновые сигнальные пути), а также на изменения в окружающей среде (факторы теплового шока, индуцируемый ги-

поксией фактор и др., белки, связывающиеся со стерол-регуляторными элементами) и на факторы клеточного цикла (Мус-онкогенами) [27–31].

Энхансеры – геномные цис-регуляторные сиквенсы, интегрирующие пространственно-временные сигналы, которые контролируют экспрессию генов. Активность энхансеров зависит от комбинации их связывания с ФТ. Это фрагменты ДНК размером 10–20 п. о., расположенные на удалении от гена или находящиеся внутри интронов. В геноме человека их насчитывается примерно 110 000. Многие из них отделены от их промоторов сотнями и тысячами оснований. Более 30 лет назад Vanerji и соавт. [36–55] обнаружили, что фрагмент ДНК Simian Virus 40 усиливает транскрипцию гена бетаглобина кроликов в 200 раз. Связывание белков-активаторов или репрессоров помогает прикреплению молекул РНК полимеразы, усилению или понижению уровня базальной транскрипции. Они присутствуют во всех клетках, но активными являются в определенных тканях [4, 24].

Локус-контрольные области (locus control regions – LCR, ЛКО) – крупные сиквенсы (много kb), расположенные в начале гена, имеют первостепенное значение для экспрессии генов. Основной их функцией является модулирование структуры хроматина. Один и тот же ЛКО может быть необходим для экспрессии многих генов. Кроме того, они, возможно, контролируют состояние крупных структурных единиц хроматина и посредством связывания с ними соответствующих белков регулируют активность генов [37].

Области прикрепления к матриксу (matrix attachment regions – MAR) – последовательности ДНК, ко-локализирующиеся с энхансерами, промоторами и началом репликации и обозначающие границы между доменами хроматина, являются АТ-обогащенными областями [37].

Сайленсеры – последовательности нуклеотидов ДНК, расположенные вблизи промотора или внутри интронов, подобны по функции энхансерам, способны снижать уровень транскрипции гена, при присоединении к ним белков-репрессоров происходит подавление транскрипции [37].

Инсуляторы – последовательности ДНК, блокирующие влияние энхансеров, сайленсеров или ЛКО на экспрессию генов. Основная функция инсуляторов – поддержание трехмерной организации генома. Обладают двумя видами активности: а) прерывают ассоциацию между промотором и энхансером; б) препятствуют распространению гетерохроматина (барьерная активность). Они располагаются между ЛКО и фрагментом ДНК, который они защищают и, подобно ЛКО, воздействуют на крупные фрагменты хроматина (петли) [38].

Элементы отклика (ответа) – короткие последовательности нуклеотидов (мотивы), находящиеся внутри промоторов, энхансеров или сайленсеров и повышающие их функциональные возможности, оказывают модулирующий эффект на уровень транскрипции генов при ответе на ряд стимулов, таких как воздействие циклического аденозинмонофосфата (сАМР), ростовых факторов, цитокинов (гамма-интерферона, интерлейкинов), тяжелых металлов, белков теплового шока, стероидов [39].

Множество ФТ участвует в реализации определенных фаз и этапов патогенеза заболеваний человека, но большинство из них еще неизвестны. Имеется несколько путей активации/деактивации ФТ: а) связывание лиганда, который необходим для функционирования субстанций, не входящих в состав полипептида (например, ионов цинка); б) фосфорилирование для обеспечения связывания с ДНК; в) взаимодействие с другими ФТ и/или ко-регулируемыми белками [39]. GATA-3 является высокоэкспрессируемым фактором транскрипции CD4⁺ клеток Th2 типа («мастер регуляции» функции), особенно Т-рег клеток, посредством изменения экспрессии Foxp3 при связывании с консервативными некодируемыми сиквенсами CNS2, что повышает активность Foxp3 цис-действующих элементов. Делеция Т-рег клеточно-специфичного GATA-3 ассоциируется с иммунным воспалением, обусловленным эффекторными функциями Т-клеток [30]. Кроме того, он проявляет себя и как опухолевый супрессорный фактор транскрипции онкогенов *HER2*, *Skp2*, *SATB1* и *Мус* и индуцирует экспрессию супрессорных генов *P21* и *LATS2* при раке молочной железы и раке простаты. Избыточная экспрессия его отмечена при меланоме, гемангиомах и раке яичников [30]. Транскрипционный фактор CREB регулирует различные типы ответной реакции клеток, включая дифференцировку, пролиферацию и выживание. Он индуцируется различными ростовыми факторами, воспалительными сигналами и регулирует транскрипцию ге-

нов, содержащих сАМР-зависимые элементы ответа, включая гены *ИЛ-2*, *ИЛ-6*, *ИЛ-10* и *ФНО-α*. Фосфорилированная его форма непосредственно ингибирует активацию ядерного фактора NF-κB, участвует в формировании и регуляции Treg-клеток [31]. Транскрипционный фактор NF-κB регулирует экспрессию генов цитокинов и участвует в патогенезе многих заболеваний, включая аллергические и аутоиммунные [32, 39].

Известно, что в процесс активации иммунокомпетентных клеток вовлекаются протеинкиназы и кальций-зависимые сигнальные пути, включая тирозин и серин/треониновые киназы, фосфатазы, протеинкиназу C и кальциневрин. Образование сигналов и их передача в ядро включает несколько этапов: 1) связывание активаторов с рецепторами клеток; 2) активацию (фосфорилирование) Jak-киназ и цитокиновых рецепторов; 3) связывание фосфорилированными участками протеинкиназ и рецепторов цитокинов STAT-белков посредством SH2 доменов и перенос димеров STAT-белков в ядро; 4) взаимодействие димерных белков с ДНК. Следствием этих взаимодействий является вовлечение в ответную реакцию семейств ядерных факторов транскрипции – NF-κB, NF-AT, AP-1 и др. [3, 26], которые участвуют в активации транскрипции генов цитокинов (ИЛ-1, ИЛ-2, ИЛ-6, ФНО, КСФ и др.), ускоряют или ингибируют синтез мРНК. При активации лимфоцитов синтез РНК увеличивается уже через 6 ч. Через 24 ч ее количество удваивается, а спустя 96 ч ее концентрация возрастает еще в 2 раза. Ответ клеток на стимуляцию антигенами и митогенами характеризуется разной временной и клеточно-типовой динамикой и специфичностью синтеза мРНК и изменяется в значительных пределах (в 10–100 раз и более) [33, 39]. Этапы взаимодействия клетки с факторами внешней среды, пути реализации генетической информации и регуляции экспрессии генов представлены на рисунке.

Посттранскрипционная регуляция экспрессии генов. Экспрессия генов является сложным процессом, регулируемым на различных этапах. В случаях стабильности окружающей клетку среды скорость синтеза мРНК и их деградация равны. Варьирование количественного содержания мРНК в клетке может быть следствием как изменения интенсивности транскрипции, так и разрушения их молекул. В обеспечении разнообразия жизненно важных белков ведущую роль играет посттранскрипционный процессинг мРНК или ее созревание. Основным механизмом данного типа регуляции являясь альтернативный сплайсинг, редактирование мРНК и изменение стабильности молекул [4].

Альтернативный сплайсинг (от англ. *splice* – сращивать или склеивать концы чего-либо). Большинство генов иммунной системы имеют полиэкзонное и полиинтронное строение. В процессе созревания из пре-мРНК вырезаются интроны, а экзоны сшиваются, при этом молекула мРНК существенно укорачивается. Длина зрелой молекулы мРНК может колебаться от нескольких сотен до нескольких тысяч нуклеотидов. Более того, при сплайсинге из первичного транскрипта мРНК (пре-мРНК) образуется множество разных молекул в зависимости того, какие комбинации экзонов сшиваются в зрелую молекулу. В результате из кодируемой одним геном генетической информации синтезируется некая совокупность структурно и функционально близких, но различающихся полипептидов. Сплайсингу подвержено до 94 % генов человека. Таким образом, при наличии 22 000–25 000 генов в одной клетке может синтезироваться не менее 100 000 белков (от 100 000 до 300 000) [4, 34].

«Редактирование РНК» – молекулярно-биологический процесс, при котором в молекуле РНК путем химических модификаций заменяются основания. Редактирование РНК может происходить в ядре, цитозоле и митохондриях. Разнообразие механизмов редактирования мРНК включает модификацию нуклеозидов, вставки и выпадения нуклеотидов, в результате чего может меняться структура и функция конечного белкового продукта [4, 6, 12].

Стабильность мРНК. Молекулы мРНК после их синтеза в ядре через поры транспортируются в цитоплазму клетки. Качественный и количественный набор мРНК в ядре (транскриптом) разнообразнее такового в цитоплазме, что связано с частичным разрушением нуклеазами. Функционально способными в цитоплазме являются мРНК, защищенные белками. Время жизни мРНК в цитоплазме – критический фактор биосинтеза белков. Последовательности нуклеотидов, влияющие на время полужизни, локализуются в нетранслируемой области 3' конца молекулы (UTR). Стабильность мРНК сказывается на трансляции и конечном содержании белкового

продукта. Продолжительность жизни мРНК варьируется в широких пределах. Некоторые мРНК кодируют продукт с большой продолжительностью жизни (10 ч и более), а другие (например, мРНК факторов роста) – с короткой [34, 35].

Роль микро-РНК в регуляции транскрипции. Важнейшим механизмом посттранскрипционной регуляции экспрессии генов являются молекулы микро-РНК. Они представляют собой не кодирующие РНК длиной 21–22 нуклеотида. На их долю приходится только 1 % транскриптов клетки, но они регулируют экспрессию более 60 % генов [40], связываясь со специфическими последовательностями мРНК 3' нетранслируемой области, ингибируют трансляцию и удаляют поли-А хвост [41]. Из-за неполного спаривания с молекулами мРНК-мишенями они способны ингибировать трансляцию широкого спектра мРНК со сходными последовательностями, т. е. изменяют экспрессию генов без изменений структуры ДНК и мРНК или, иными словами, осуществляют эпигенетический контроль реализации генетической информации [40, 41]. Данное явление получило название РНК-интерференции, или подавления, экспрессии гена на стадии транскрипции, трансляции, деаденилирования или дегградации мРНК. Сайленсинг (от англ. *silencer* – глушитель, от лат. *silentum* – молчание) – процесс понижения (подавления, исключения) транскрипции генов, индуцированный микро-РНК (induced transcriptional silencing, RITS), химическими веществами, вирусами, регулируется комплексом белков и представляется важнейшим механизмом биологической регуляции функции генов [45, 46]. При этом для сохранения существующих районов гетерохроматина RITS образует комплексы с малыми, интерферирующими РНК, комплементарными соответствующим генам. Именно этот комплекс вызывает дегградацию молекул пре-мРНК, синтезируемых РНК-полимеразой, и блокирует транскрипцию [41, 42]. В регуляторных механизмах иммунной системы участвует более 100 разных типов микро-РНК, регулирующих дифференцировку и активацию Т- и В- лимфоцитов, моноцитов, макрофагов, дендритных клеток, а также полинуклеарных лейкоцитов. Наиболее изученными являются MiR-17-92 (IFN-gamma, IL-4, IL-5, IL-13), miR-155 (IL-10, IL-17, IL-22), let-7 (IL-10, IL-13, IL-17) [41–43]. Установление роли эпигенетических событий, влияющих на устойчивость/восприимчивость индивидуума к возбудителям инфекций, иммунопатологическим процессам и онкопатологии, является важнейшим направлением персонализированной медицины [44–47].

Белковые молекулы по завершении трансляции накапливаются в цитоплазме, секретируются в межклеточное пространство, транспортируются в кровь и органы или подвергаются дегградации. Процесс их дегградации происходит в протеасомах цитоплазмы (гигантских белковых комплексах) с участием белка убиквитина. Важно обеспечение слаженной работы всех звеньев генома с целью достижения полноценного функционирования всех типов клеток и систем организма, своевременной коррекции нуклеотидных последовательностей генов, молекул мРНК, сбалансированной продукции белков с регуляторными и эффекторными свойствами, дифференцировки клеток и элиминации генетически измененных молекул и клеток [44].

Вариабельность генома. Геном современного человека, как результат эволюционного процесса, подобен геномам других организмов. Перспективы развития персонализированной терапии базируются на ожиданиях расшифровки и анализа полной информации об определенных генах и геноме индивидуума в целом (3 млрд оснований), а также на ограниченных знаниях биологической роли содержащихся в них генетических вариаций. Эти вариации являются генетической основой индивидуальности каждого члена нашего вида и, соответственно, основой эволюционного процесса. Вариабельность генома индивидуума определяет фенотип заболеваний или его предрасположенность к воздействию факторов окружающей среды [6, 7]. В результате расшифровки двух первых геномов – Крейга Вентера (Creig Wenter, 2007) и Джеймса Уотсона (James Watson, 2008) и установлению их вариабельности были развернуты обширные исследования по ПГС, изучению частоты генетических вариаций, установлению основных типов и механизмов их формирования [48–53].

В настоящее время в структуре генома человека выявляют несколько типов генетических вариаций.

Полиморфизм одиночных нуклеотидов (Single nucleotide polymorphism – SNPs-снупсы), которые представляют собой класс хорошо изученных вариаций и встречаются в определенных

участках генов. Большинство вариантов ДНК представлено в виде мутаций – замен одиночных нуклеотидов в кодонах или полиморфизмов. Первые SNPs были идентифицированы в 1978 г. в лаборатории Y. W. Кан в 3' области гена бета-глобина [56]. Полиморфные участки имели две альтернативные аллели. Частота каждой аллели в разных популяциях может значительно варьироваться. На 1000 нуклеотидов хромосомы в среднем приходится один SNPs. Многие снипсы встречаются весьма часто, другие – редко. База данных содержит 25 млн общих и редких снипсов (<http://www.ncbi.nlm.nih.gov/SNP/>), из которых 301 000 локализована в генах, кодирующих белки. Из них 188 000 сопровождаются заменами аминокислот в полипептиде, т. е. являются несинонимичными. В геноме человека выявляется примерно 11 млн снипсов [16]. Посредством взаимосвязанных исследований и ПГС множественные SNPs ассоциированы с широким спектром болезней – диабетом II типа, раком молочной железы, рассеянным склерозом, сердечно-сосудистыми заболеваниями, а также с ответом на терапию лекарственными препаратами. Информация об ассоциации полиморфизмов с заболеваниями имеется в базах данных Gynotypes and Phenotypes (dbGaP) online Mendelian Inheritance in Man (OMIM) [17, 50].

Короткие повторы сиквенсов (short sequence repeats – SSR), или структурные вариации. Наиболее частыми из них являются динуклеотидные повторы, но встречаются три-, тетра- и пента-повторы, называемые микросателлитами. Короткие повторяющиеся единицы содержат до 15 нуклеотидов, а более длинные – 15–500 нуклеотидов и называются минисателлитами. Они занимают до 5 % генома, на 1 SSR приходится до 5 kb информации. Наиболее частыми короткими повторами являются GT-элементы, встречающиеся в геноме 28 раз на мегаоснование, затем AT-элементы с частотой 19 раз на мегаоснование. Более десятка неврологических заболеваний обусловлены повторами определенных триплетов. При этом отмечается накопление (экспансия) определенных триплетов внутри или на концах отдельных генов. Наиболее частым триплетом является TAA с частотой 4 раза на мегаоснование. Например, повторы триплета CGG в первом экзоне гена *FMR1* у здоровых лиц встречаются от 6 до 20 раз, а при патологии (ломкость X-хромосомы) достигает 230 [18, 51]. Показано, что разные мутации в одном и том же гене могут определять один и тот же или разные типы патологии.

Вставки/делеции – инделы. Эти вариации обусловлены наличием или отсутствием в ДНК генома определенных последовательностей – сиквенсов. Такие сиквенсы могут состоять из нескольких нуклеотидов, но могут быть и в виде транспозонов или внутренних повторов, таких как длинные (LINE) и короткие (SINE) элементы, альтернативно это могут быть псевдогены или другие элементы. Малые вставки и делеции (инделы) имеют размер от 1 до 10 000 нуклеотидов и их число достигает 1 млн на геном [37]. Инделы локализуются в кодируемой области и могут вызывать сдвиг рамки считывания или инициировать блок синтеза белка, иногда с нарушением структуры белкового продукта. Такие инделы обозначают как бессмысленные мутации или нон-сенс-мутации [19, 52, 53].

Копийные вариации (copy number variation – CNV). Они включают наличие в геноме перестроек более крупных областей, которые удвоены в определенных положениях генов хромосом. Это так называемые копийные варианты, или CNV, или инвертированные повторы, т. е. инверсии. В противоположность инделам CNVs наиболее хорошо изучены с точки зрения их связи с патологией и часто ассоциируются с аутоиммунными заболеваниями (системной красной волчанкой), аутизмом, шизофренией, онкозаболеваниями [19, 54, 55].

Мультифакториальные заболевания, такие как рак, диабет, сердечно-сосудистые и психоневрологические заболевания, определяются как состояния с множественными причинными факторами генетического (мутации, перестройки) и негенетического (влияние окружающей среды) генеза. К настоящему времени идентифицировано множество вариантов генов (генотипов), присутствие которых у индивидуума повышает риск развития онкологических, сердечно-сосудистых, аутоиммунных, аллергических, иммунодефицитных заболеваний, болезни Альцгеймера и др. Несмотря на то что информация по медицинской геномике накапливается быстрыми темпами, мы еще находимся в начале познания типов и механизмов вариабельности генома (вставка/делеций), элементов копийности, ретроэлементов, полиморфизма одиночных нуклеотидов, крупных перестроек и естественной селекции [51–55].

Врожденные генные болезни являются преимущественно моногенными [6, 7]. Общая частота новорожденных с генными болезнями в популяции составляет примерно 1 %, из них с ауто-сомно-доминантным типом – 0,5 %, с ауто-сомно-рецессивным – 0,25, с X-сцепленным – 0,25, с Y-сцепленным – 0,25 %. Распространенность генных мутаций можно считать высокой, если 1 пациент встречается на 10 000 новорожденных и чаще, средней – если это соотношение составляет 1:10 000–1:40 000, низкой – в очень редких случаях. Например, тяжелый комбинированный иммунодефицит (ТКИД) – генетическое заболевание, при котором в результате дефекта одного из генов нарушается работа каскада адаптивной иммунной системы В- и Т-лимфоцитов. Результат этого дефекта является настолько сильным повреждением иммунной системы, что ее можно считать практически отсутствующей. Частота встречаемости этого дефекта в популяции существенно варьируется и в среднем составляет 1:100 000, в Австралии – 1:65 000, а в популяции индейцев Навахо – 1:2500, что является основной причиной заболеваемости и смертности среди детей данной народности [53]. К группе распространенных генных заболеваний относятся не более 15 болезней, которые суммарно обуславливают наличие почти 50 % больных с наследственной патологией. Важно установить характер взаимосвязей между ними, оценить, как работают гены в конкретных условиях – внутренних и внешних, так как многие болезни человека обусловлены не дефектами генов, а нарушением их скоординированных действий и регуляторных механизмов.

Закключение. Успешное развитие медицинской геномики позволит установить биологическое и медицинское значение мутаций, генетических перестроек, их связей с состоянием здоровья, предболезни и разнообразной патологии, разработать новые подходы к диагностике, профилактике и терапии. Разработка и внедрение в медицинскую практику новых методов быстрого секвенирования молекул ДНК и РНК (Ion torrent, Illumina, RNA-seg), микроэрепей технологии и биоинформатики обеспечат прогресс фактически во всех областях медицины. В настоящее время в США и Европе функционируют крупные биомедицинские компании, которые быстро и качественно выполняют молекулярно-генетические исследования, включая полное секвенирование генома человека (например, 23 and Me, США) [49, 55]. Одновременно разрабатываются требования к проведению подобных исследований, оценке и интерпретации результатов [53, 55]. В этом направлении начинают работать ученые и практики России, Украины и Беларуси. В связи с этим потребители данных технологий ожидают постепенного снижения стоимости ПГС индивидуума (скоро – 1000 долларов, а затем 300 и даже 100), что сопоставимо со стоимостью многих широко выполняемых рутинных лабораторных анализов. Коммерциализация исследований в области геномики создает рынок нового типа услуг, новые права собственности (на первичную структуру генов, новые тесты), ускоряет внедрение результатов фундаментальных исследований в практику здравоохранения [57–59]. Постепенное накопление результатов ПГС здоровых лиц и пациентов с разнообразными заболеваниями позволит сформировать национальные базы данных, анализировать и выявлять специфичные для определенной патологии генетические вариации и использовать эти результаты в профилактических и клинических целях. Новые популяционные данные о геноме человека будут полезны для создания вакцин и принципов геном-адресованных или персонализированных подходов в иммунопрофилактике. Лабораторная диагностика на уровне генома детей и взрослых с нарушениями интеллекта, психическими расстройствами, аутизмом, врожденными дефектами, установление молекулярно-генетических механизмов и их причин станут более доступными и рутинными [58, 59].

Особый прогресс ожидается в области фармакогеномики, создании новых лекарств и способов применения. Ускоренное развитие ПГС геномов пациентов, не отвечающих на терапию определенными лекарствами, позволит установить причины и пути преодоления этого явления. Эффективность применения большинства лекарственных препаратов сейчас оценивается примерно в 30–60 %. Интересным представляется применение результатов ПГС пациентов с избыточной реакцией на лекарства – лекарственной аллергией [60, 61]. Сравнительный анализ генетических вариаций геномов детей с геномами их родителей позволит выявить спектр родительских мутаций, а также обнаружить новые индивидуальные вариации, что поможет лучше понять роль генотипа в генезе наследственных и приобретенных заболеваний. Молодые пары смо-

гут использовать результаты ПСГ для планирования семьи. Дальнейшее развитие медицинской геномики будет наблюдаться и в криминалистике, что позволит осуществлять генетическую идентификацию, устанавливать наследственность (отцовство, материнство) [61, 62]. Прогрессивное внедрение в практику достижений этого направления тесно связано с подготовкой новой генерации специалистов по медицинской геномике, разработкой специальных обучающих курсов для врачей-лаборантов и врачей-клиницистов [63].

Уровень генетических вариаций в популяции человека относительно невелик – примерно 1 %, но важно определить доминирующие и редкие мутации, оценить их роль и дополнительный риск для здоровья человека [64]. В общебиологическом плане важным является реконструкция истории эволюции человека – генетической структуры предковой популяции, расселения по планете, происхождения народов, связей с гоминидами, чувствительности к инфекциям, радиации и др. [62, 63, 65]. Развитие этих направлений будет сопровождаться обширным и глубоким научным поиском, совершенствованием концепций и методов – от анализа больших объемов информации к персонализированной геномике для всех.

Список использованной литературы

1. *Титов, Л. П.* Вирусы и эукариотические клетки: стадии взаимодействия, стратегия геномов, репродукция и исходы инфекции / Л. П. Титов // Мед. журн. – 2008. – № 1. – С. 10–16.
2. *Braga-Neto, U. M.* From Genomics to Functional Immunomics: New challenges, Old problems, Big Rewards / U. M. Braga-Neto, E. T. A. Marques // Plos comput. biol. – 2006. – Vol. 2. – Iss. 7. – eB1.P.0651–0662.
3. Microbiome Composition by Pyrosequencing in Mesenteric Lymph Nodes of Rats with CCL4-induced Cirrhosis / S. Cuenca [et al.] // J. Innate Immunol. – 2014. – Vol. 6. – P. 263–271.
4. An Immunomics Approach to Schistosome antigen discovery: antibody signatures of naturally resistant and chronically infected individuals from endemic areas / S. Gaze [et al.] // PLOS Pathogens. – 2014. – Vol. 10. – Iss. 3. – P. 1–16. – E1004033.
5. *Титов, Л. П.* Иммунология: терминологический словарь / Л. П. Титов. – М.: МИА, 2008. – 521 с.
6. *Хромов-Борисов, Н. Н.* Эволюционная медицинская геномика / Н. Н. Хромов-Борисов, А. В. Рубанович // Молек. медицина. – 2014. – № 2. – С. 13–17.
7. Структурная геномика и медицина / И. Е. Ясный [и др.] // Молек. медицина. – 2009. – № 6. – С. 15–20.
8. *Groot, A. S. D.* Immunomics discovering new targets for vaccine and therapeutics / A. S. D. Groot // Drug Discov. Today. – 2006. – Vol. 11. – P. 203–209.
9. *Miller, M. B.* Basic concept of microarrays and potential application in clinical microbiology / M. B. Miller, Y. W. Tang // Clin. Microbiol. Rev. – 2009. – Vol. 22 (4). – P. 611–633.
10. *Пинегин, Б. В.* Роль клеток иммунной системы и цитокинов в развитии псориаза / Б. В. Пинегин, О. Л. Иванов, В. Б. Пинегин // Иммунология. – 2012. – Т. 33, № 4. – С. 213–219.
11. *Титов, Л. П.* Противовирусный иммунитет: молекулярно-клеточные механизмы, закономерности развития и иммунопатология / Л. П. Титов, И. А. Карпов // Мед. журн. – 2007. – № 1. – С. 4–14.
12. *Бокуть, С. Б.* Молекулярная биология / С. Б. Бокуть, Н. В. Герасимович, А. А. Милютин. – Минск: Вышэйш. школа, 2005. – 463 с.
13. The sequence of human genome / J. C. Venter [et al.] // Science. – 2001. – Vol. 291. – P. 1304–1309.
14. *Read, A.* The human Genome: structure and organization / A. Read // Genetics and Medicine / ed. D. Kumar, D. Weatherall. – Oxford press, 2008. – P. 17–29.
15. *Antonarakis, S. E.* Human Genome Sequence and Variation / S. E. Antonarakis // Vogel and Matulsky's Human Genetics, Problems and Approaches / eds. M. R. Speicher et al. – 2010. – Berlin; Heidelberg: Springer-Verlag. – P. 31–53.
16. Global gene expression analysis reveals evidence for decreased lipid biosynthesis and increased innate immunity in uninvolved psoriatic skin / J. E. Gudjonsson [et al.] // J. Invest. Dermatol. – 2009. – Vol. 129. – P. 2795–2804.
17. Collaborative association study of genome-wide scan reveals association of psoriasis with IL-23 and NF-kappa B pathways / R. P. Nair [et al.] // Nat. genetics. – 2009. – Vol. 41. – P. 199–204.
18. *Garaev, P. P.* Another Understanding of the Model of Genetic Code Theoretical Analysis / P. P. Garaev // Open J. of Genetics. – 2015. – Vol. 5, N 1. – P. 92–109.
19. *Mardis, E. R.* Cancer genome sequencing / E. R. Mardis, R. K. Wilson // Human mol. genetics. – 2009. – Vol. 18. – Iss. 2. – Doi:10.1093/ddp396.
20. Consortium biology in immunology: the perspective from the Immunological Genome Project / C. Benoist [et al.] // Nature Rev. Immunol. – 2012. – P. 1–6.
21. *Heng, T. S. P.* The immunological Genome Project: network of gene expression in immune cells / T. S. P. Heng, M. W. Painter // Immunology. – 2008. – Vol. 4, N 10. – P. 1091–1094.
22. *Grander, D. I.* Immunomics: principles and practis / D. I. Grander // IRTI. – 2004. – Vol. 2. – P. 1–6.
23. *Титов, Л. П.* Геномико-протеомические основы эволюции и молекулярной эпидемиологии вирусов / Л. П. Титов, В. И. Вотяков // Вес. НАН Беларусі. Сер. мед. навук. – 2011. – № 1. – С. 109–124.

24. Epigenetic regulation of inducible gene expression in the immune system / P. S. Lim [et al.] // *Immunology*. – 2013. – Vol. 139. – Iss. 3. – P. 285–293.
25. *Josefowicz, S. Z.* Regulatory of chromatin state and transcription in CD4 T-cells polarization / S. Z. Josefowicz // *Immunology*. – 2013. – Vol. 139. – Iss. 1. – P. 299–308.
26. *Титов, Л. П.* Особенности строения, развития и функционирования иммунной системы детского организма / Л. П. Титов, Е. Ю. Кирильчик, Т. А. Канашкова // *Мед. новости*. – 2009. – № 5. – С. 7–16.
27. Transcription factors and CD4 T-cell seeking identity: masters, minions, setters and spikers / G. Vahedi [et al.] // *Immunology*. – 2013. – Vol. 139. – Iss. 3. – P. 294–298.
28. *Chen, W.* Development of thymic Foxp3 regulatory T-cells: TGF-b matters / W. Chen, J. E. Konkel // *Eur. J. of Immunol.* – 2015. – Vol. 45. – P. 958–965.
29. *Lee, G. R.* Transcriptional regulation T-helper type 2 / G. R. Lee // *Immunology*. – 2014. – Vol. 141. – Iss. 4. – P. 498–505.
30. *Wang, Y.* An essential role of the transcriptional factor GATA-3 for the function of regulatory T-cells / Y. Wang, M. A. Su, Y. Y. Wan // *Immunity*. – 2011. – Vol. 35. – P. 337–348.
31. *Wen, A. Y.* The role of the transcriptional factor CREB in immune function / A. Y. Wen, K. M. Sakamoto, L. S. Miller // *J. Immunol.* – 2010. – Vol. 185. – P. 6413–6419.
32. *Rico-Rosillo, G.* The involvement of NF-kB transcription factor in asthma / G. Rico-Rosillo, G. B. Vega-Robledo // *Rev. Alergia Mex.* – 2011. – Vol. 58, N 2. – P. 107–111.
33. *Титов, Л. П.* Молекулярные механизмы активации Т- и В-лимфоцитов / Л. П. Титов // *Современные проблемы инфекционной патологии человека: материалы 2-й науч.-практ. конф.* – Минск, 2001. – С. 287–317.
34. From single-cell to cell pool transcriptomes: stochasticity in gene expression and RNA splicing / G. K. Maurinov [et al.] // *Genome Res.* – 2011. – Vol. 24. – P. 496–510.
35. *Youngblood, B.* T-cell memory differential signatures and epigenetics / B. Youngblood, J. S. Hall, R. Ahmed // *Immunology*. – 2013. – Vol. 139. – Iss. 3. – P. 325–335.
36. *Banerji, J.* Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences / J. Banerji, R. Sandro, M. Schaffner // *Cell*. – 1981. – Vol. 27. – P. 299–308.
37. *Ogbourne, S.* Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes / S. Ogbourne, T. M. Antalıs // *Biochemistry*. – 1998. – Vol. 231. – P. 1–14.
38. Enhancers and silencers: an integrated and simple model for their function / P. Kolovos [et al.] // *Epigenetics & Chromatin*. – 2012. – Vol. 5, N 1. – P. 2–8.
39. *Разин, С. В.* Регуляторные элементы эукариотического генома, контролирующие транскрипцию / С. В. Разин, А. А. Говрилов, С. В. Ульянов // *Молек. биология*. – 2015. – Т. 49, № 2. – С. 212–223.
40. *Гонзвгло, М.* МикроРНК как биомаркер для диагностики рака мочевого пузыря / М. Гонзвгло, Г. Ишханова // *Молек. медицина*. – 2014. – № 2. – С. 18–24.
41. *Титов, Л. П.* Микро-РНК: новый класс регуляторных молекул иммунного ответа и инфекционного процесса / Л. П. Титов // *Современные проблемы инфекционной патологии человека: сб. науч. тр.* – 2012. – В. 5. – С. 256–261.
42. *Perez-Ortin, J. E.* Genomics of mRNA turnover / J. E. Perez-Ortin // *Briefing in functional genomics and proteomics*. – 2007. – Vol. 6, N 4. – P. 282–291.
43. Cross-regulation between cytokine and microRNA pathways in T-cells / T. Amado [et al.] // *Eur. J. of Immunol.* – 2015. – Vol. 45. – Iss. 6. – P. 1584–1595.
44. Physiological and pathological roles for microRNAs in the immune system / R. M. O'Connel [et al.] // *Nat. Rev. Immunol.* – 2010. – P. 11–122.
45. A role for Dicer in immune regulation / B. S. Cobb [et al.] // *J. Exp. Med.* – 2006. – Vol. 203. – P. 2519–2527.
46. The RNaseIII enzyme Drosha is critical in T cells for preventing lethal inflammatory disease / M. M. Chang [et al.] // *J. Exp. Med.* – 2008. – Vol. 205. – P. 2005–2017.
47. Role of mir-155 in the regulatory of lymphocytes immune function and disease / N. Seddiki [et al.] // *Immunology*. – 2014. – Vol. 142. – Iss. 1. – P. 32–38.
48. *Ortutay, C.* Immunomic Knowledge base (IKB): an integrated service for immune research / C. Ortutay, M. Vikinen // *BMC Immunol.* – 2009. – Vol. 10. – doi: 101186/1471–2172–10–3.
49. *Kan, Y. W.* Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation / Y. W. Kan // *Proc. Natl. Sci.* – 1978. – Vol. 75, N 11. – P. 5631–5635.
50. Increased expression of Wnt5a in psoriatic plaques / J. Reischl [et al.] // *J. Invest. Dermatol.* – 2007. – Vol. 127. – P. 163–169.
51. Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infection disease / C. Skevaki [et al.] // *Clin. and Exp. Immunol.* – 2015. – Vol. 180. – P. 165–177.
52. *Marietta, E.* Immunogenetic control of the intestinal microbiota / E. Marietta, A. Rishi, V. Taneja // *Immunology*. – 2015. – Vol. 145. – P. 313–322.
53. Severe combined immunodeficiency. A national surveillance study / A. Yee [et al.] // *Pediatric Allergy and Immunol.* – 2008. – Vol. 19 (4). – P. 298–302.
54. Quantitative monitoring of gene expression patterns with a complementary DNA microarray / M. Schena [et al.] // *Science*. – 1995. – Vol. 270. – P. 467–470.
55. Clinical interpretation and implications of whole genome sequencing / F. E. Dewey [et al.] // *JAMA*. – 2014. – Vol. 31, N 10. – P. 1035–1041.
56. *Cordero, P.* Whole genome sequencing in personalized therapeutics / P. Cordero, E. A. Ashely // *Clin. Pharmacol. and Ther.* – 2012. – Vol. 91, N 6. – P. 1001–1009.

57. Use of whole genome sequencing for diagnosis and discovery in the cancer genetics clinic / S. B. Foley [et al.] // *Ebiomedicine*. – 2015. – Vol. 2. – P. 74–81.
58. Whole genome sequencing in health care / C. G. van El [et al.] // *Eur. J. of Human Genetics*. – 2013. – Vol. 21. – P. 580–584.
59. *Swanson, C. I.* Rapid evolutionary of a structurally constrained eye enhancer / C. I. Swanson, D. B. Schwimmer, S. Barolo // *Curr. Biol.* – 2011. – Vol. 21, N 14. – P. 1186–1196.
60. Preparing the next generation of genomicists: a laboratory-style course in medical genomics / M. D. Lindemann [et al.] // *BMC med. genomics*. – 2015. – Vol. 8, N 47. – P. 1–3.
61. *Ngo, D. A.* Genomic approaches to identifying targets for treating bêta-hemoglobinopathies / D. A. Ngo, M. H. Steinberg // *BMC med. genomics*. – 2015. – Vol. 8, N 44. – P. 1–13.
62. *Karki, R.* Defining “mutation” and “polymorphism” in the era of personal genomics / R. Karki, D. Pandya, R. C. Elston // *BMC med. genomics*. – 2015. – Vol. 8, N 37. – P. 1–8.
63. *Alyass, A.* From big data analysis to personalized medicine for all: challenges and opportunities / A. Alyass, M. Turlotbe, D. Meyre // *BMC med. genomics*. – 2015. – Vol. 8, N 33. – P. 1–12.
64. *Тумов, Л. П.* Компьютерная иммунология: сравнительный анализ нуклеотидных замен в CDR и FR фрагментах генов иммуноглобулинов при гепатите С, криоглобулинемии и лейкозах / Л. П. Титов, Т. А. Столярова, Е. А. Столярова // *Вест. НАН Беларусі. Сер. мед. навук.* – 2010. – № 3. – С. 10–18.
65. Medical genomics: The intricate path from genetic variant identification to clinical interpretation / B. Quintas [et al.] // *Appl. & Transl. Genomics*. – 2014. – Vol. 3. – Iss. 3. – P. 60–67.

Поступила в редакцию 04.09.2015