

Смирнов В.В.

Соотнесение функций аминокислот, кодируемых кодонами, с порядком расположения гексаграмм по Вэнь-вану

Аннотация

В процессе исследования многоуровневой системы кодирования генетической информации предлагается использовать сопоставление кодонного алфавита с 64-символьным алфавитом гексаграмм, описанных в классической китайской "Книге перемен" ("И цзин"). Применяемый в данной работе авторский способ такого сопоставления позволяет выделить биологически значимые варианты соотнесения единиц генетического кода с гексаграммами. Этот способ учитывает следующее: порядок расположения гексаграмм по Вэнь-вану; смысловые значения, приписываемые гексаграммам; возможность того, что скорость синтеза белка может отличаться при участии в этом процессе различных кодонов, кодирующих одни и те же аминокислоты; информацию о биологических функциях аминокислот; авторское предположение, что одни и те же позиции нуклеотидов в кодоне могут по-разному интерпретироваться рибосомами.

Оглавление

1. Введение
 2. Сопоставление кодонов с гексаграммами
 3. Заключение
- Литература

1. Введение

Начиная с 1953 года, когда Фрэнсис Крик, Джеймс Уотсон и Морис Уилкинс открыли структуру молекулы ДНК, исследователями генома пройден большой путь, результатом которого в настоящее время является возможность использования генома человека в качестве индивидуального информационного продукта. Тем не менее ученые признают, что геном расшифрован,

но не понят [1]. В частности, все еще выполняются исследования с целью получения обоснований существования многоуровневой системы генетических алфавитов, включающих 4-символьный алфавит азотистых оснований (нуклеотидов), 64-символьный алфавит кодонов (кодирующих тринуклеотидов), 20-символьный алфавит аминокислот и другие алфавиты [2]. Например, обоснования могут базироваться на теории помехоустойчивого кодирования [3,4], на принципе экономии свободной энергии цепочек нуклеотидов [5], на принципе комплементарности [6], на тенденции генома к повышению информационной емкости, проявлением которой являются механизмы альтернативного сплайсинга [7], и др. Результаты недавних исследований генома показали наличие новых еще более глубоких уровней кодирования информации, связанных со скоростью синтеза белков [8], что требует формирования новых и пересмотра ранее предложенных теоретических обоснований, касающихся, например, предполагаемой избыточности 64-символьного алфавита кодонов.

Для решения этих задач оказывается полезным сопоставление кодонного алфавита с 64-символьным алфавитом гексаграмм, описанных в классической китайской "Книге перемен" ("И цзин"). Алфавит гексаграмм можно рассматривать как двухуровневый, поскольку гексаграммы состоят из шести линий, каждая из которых может быть прерывистой или сплошной, образуя 2-символьный алфавит линий. То есть способ кодирования гексаграмм похож на способ кодирования кодонов, отличаясь, по существу, лишь системами счисления. Порядок расположения линий на гексаграммах, группы линий гексаграмм, гексаграммы целиком и порядки расположения гексаграмм могут ассоциироваться с различными аспектами жизни людей, в том числе со структурой и функциями человеческого организма [9]. Вместе с тем, некоторые принципы, положенные в основу "Книги перемен", в частности, связанные с порядком расположения гексаграмм по Вэнь-вану, пока еще остаются не известными [10]. Поэтому нахождение связей между этими двумя системами кодирования информации позволяет лучше понять как смысл, заложенный в "Книге перемен", так и принципы, на которых построен генетический код.

Результаты некоторых исследований, проведенных в данном направлении, представлены в работах [3,4,6]. Пожалуй, наиболее интересной является работа С.В. Петухова [3], в которой предложен подход к сопоставлению кодонов и гексаграмм, основанный на "алгебре бисексов". К сожалению, в этом подходе не принимаются

во внимание смысловые значения, приписываемые гексаграммам. Наиболее показательной в этом отношении являются гексаграмма "Уже конец", имеющая номер 63 по Вэнь-вану. Логично предположить, что существуют такие соотношения кодонов с гексаграммами, в которых один из стоп-кодонов, т.е. кодонов, кодирующих прекращение синтеза белка, сопоставляется с гексаграммой "Уже конец". Несмотря на то, что С.В. Петухов продолжил развивать свой подход, в частности, рассмотрев в работе [4] стоп-кодоны для различных вариантов (диалектов) генетического кода, но даже в рамках стандартного диалекта генетического кода (соответствующие кодоны РНК представлены в таблице 1) он не исследовал такие решения, в которых бы стоп-кодоны совпадали бы с 63-й гексаграммой. Рассмотренные в других работах закономерности порядка расположения гексаграмм, например, представленные в работах [10, 11], не позволяют однозначно сопоставить кодоны ДНК или РНК с гексаграммами "Книги перемен" таким образом, чтобы обеспечивалось совпадение одного из стоп-кодонов с гексаграммой № 63 по Вэнь-вану.

Таблица 1. Таблица кодонов РНК.

2-е азотистое основание	1-е азотистое основание			
	С	U	G	A
U	CUU Лейцин CUG Лейцин CUA Лейцин CUC Лейцин	UUU Фенилаланин UUG Лейцин UUA Лейцин UUC Фенилаланин	GUU Валин GUG Валин GUA Валин GUC Валин	AUU Изолейцин AUG Метгониин AUA Изолейцин AUC Изолейцин
C	CCU Пролин CCG Пролин CCA Пролин CCC Пролин	UCU Серин UCG Серин UCA Серин UCC Серин	GCU Аланин GCG Аланин GCA Аланин GCC Аланин	ACU Треонин ACG Треонин ACA Треонин ACC Треонин
A	CAU Гистидин CAG Глутамин CAA Глутамин CAC Гистидин	UAU Тирозин UAG Стоп-кодон UAA Стоп-кодон UAC Тирозин	GAU Аспарагиновая кислота GAG Глутаминовая кислота GAA Глутаминовая кислота GAC Аспарагиновая кислота	AAU Аспарагин AAG Лизин AAA Лизин AAC Аспарагин
G	CGU Аргинин CGG Аргинин CGA Аргинин CGC Аргинин	UGU Цистеин UGG Триптофан UGA Стоп-кодон UGC Цистеин	GGU Глицин GGG Глицин GGA Глицин GGC Глицин	AGU Серин AGG Аргинин AGA Аргинин AGC Серин

Применяемый в данной работе оригинальный способ позволяет выделить биологически значимые варианты сопоставления единиц генетического кода с гексаграммами. Способ учитывает следующее: порядок расположения гексаграмм по Вэнь-вану; смысловые значения, приписываемые гексаграммам; результаты исследований, показывающие, что скорость синтеза белка зависит от участвующих в этом процессе различных кодонов, кодирующих одни и те же аминокислоты [8]; информацию о биологических функциях

аминокислот; предположение о том, что одни и те же позиции нуклеотидов (азотистых оснований) в кодоне могут по-разному интерпретироваться рибосомами, что, вероятно, связано с механизмами альтернативного сплайсинга [7].

2. Сопоставление кодонов с гексаграммами

Рассмотрим ограничения, учитываемые при выполнении процесса соотнесения кодонов с гексаграммами. Первое из таких ограничений происходит из предания о том, что гексаграммы получились из 8-ми триграмм путем их удвоения Вэнь-ваном (XII в. до н.э.), отцом У-вана, ставшего основателем династии Чжоу (прав. в 1121—1115 гг. до н.э.) [9]. В свою очередь, триграммы, описанные в "Книге перемен", получены путем последовательного дихотомического деления Великого предела, основанного на принципах *инь* и *ян*. В качестве графических символов, обозначающих *инь* и *ян*, в китайской *арифмосемиотике* используются, соответственно, прерывистая и сплошная линии. Важное значение для соотнесения кодонов с гексаграммами имеет также учение о пяти "стихиях". Под "стихиями" подразумеваются *шуй* (Вода), *хо* (Огонь), *му* (Дерево), *цзинь* (Металл), *ту* (Земля). Существуют различные варианты соответствий между триграммами и "стихиями" [12,13]. Последовательность формирования триграмм из Великого предела представлена на рис.1. Также на рис. 1 показано такое соответствие триграмм и "стихий", которого мы придерживаемся в данной статье при соотнесении кодонов с гексаграммами.

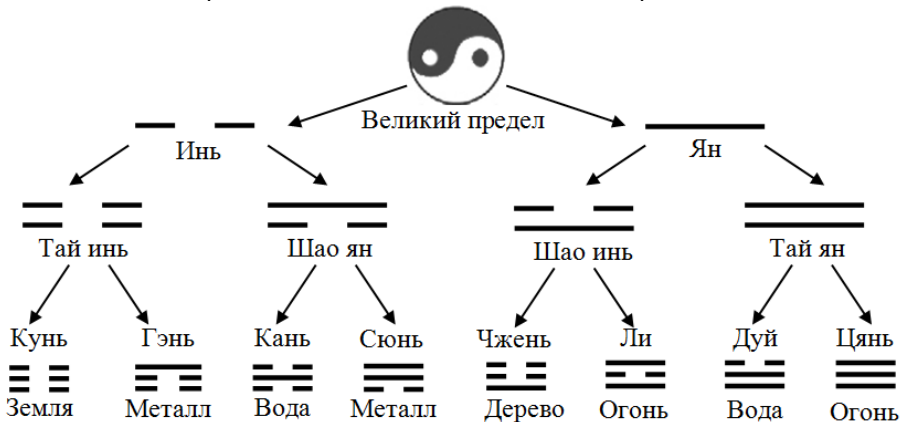


Рис. 1. Происхождения триграмм из Великого предела.

На рис. 1 представлена схема формирования гексаграмм из триграмм. Под каждой гексаграммой представлены ее код в в

двоичной системе (0 для прерывистой линии, 1 - для сплошной линии) и номер по Вэнь-вану в десятичной системе. В соответствии с тем, как образованы гексаграммы, в них можно выделить нижнюю ("земля", *ди*) и верхнюю ("небо", *тянь*) триграммы. Также линии могут быть разбиты на нижнюю, среднюю и верхнюю пары, называемые соответственно "земля" (*ди*), "человек" (*жэнь*) и "небо" (*тянь*). Существуют и другие группировки линий гексаграмм [10,14]. Каждая отдельная линия в гексаграмме также может иметь свое значение. Например, линии могут ассоциироваться с частями тела человека или животного [9].

Рассмотрим второе и третье ограничения процесса установления соответствия между кодонами и гексаграммами. Второе ограничение состоит в использовании разбиения линий гексаграмм на верхнюю, среднюю и нижнюю пары, и каждая пара ассоциируется, соответственно, с первой, второй и третьей позициями нуклеотидов в кодоне. Если предположить, что все три позиции нуклеотидов интерпретируется рибосомами одинаково, то каждому из четырех различных нуклеотидов, которые могут оказаться в любой из трех позиций, следует однозначно сопоставить одно из четырех возможных сочетаний прерывистых и сплошных линий в паре.

Пример одного из вариантов такого кодирования нуклеотидов представлен в таблице 2. В соответствии с формулой перестановки без повторений общее число вариантов подобного кодирования оказывается: $P_n = n! = 4! = 24$.

Таблица 2. Вариант кодирования нуклеотидов.

	1-й нуклеотид	2-й нуклеотид	3-й нуклеотид
U			
C			
A			
G			

Третье ограничение состоит в том, что один из стоп кодонов должен совпасть с гексаграммой 63. Перебрав все 24 комбинации можно убедиться в невозможности такого совпадения. Поэтому целесообразно изменить сделанное ранее предположение об одинаковой интерпретации позиций нуклеотидов рибосомами и принять, что одни и те же позиции нуклеотидов в кодоне могут интерпретироваться по-разному. Пример одного из вариантов такого кодирования нуклеотидов представлен в таблице 3. В соответствии с формулами размещения с повторениями и

перестановки без повторов общее число комбинаций оказывается:

$$A_m^k = A_{Pn}^k = A_n^3 = 24^3 = 13824$$

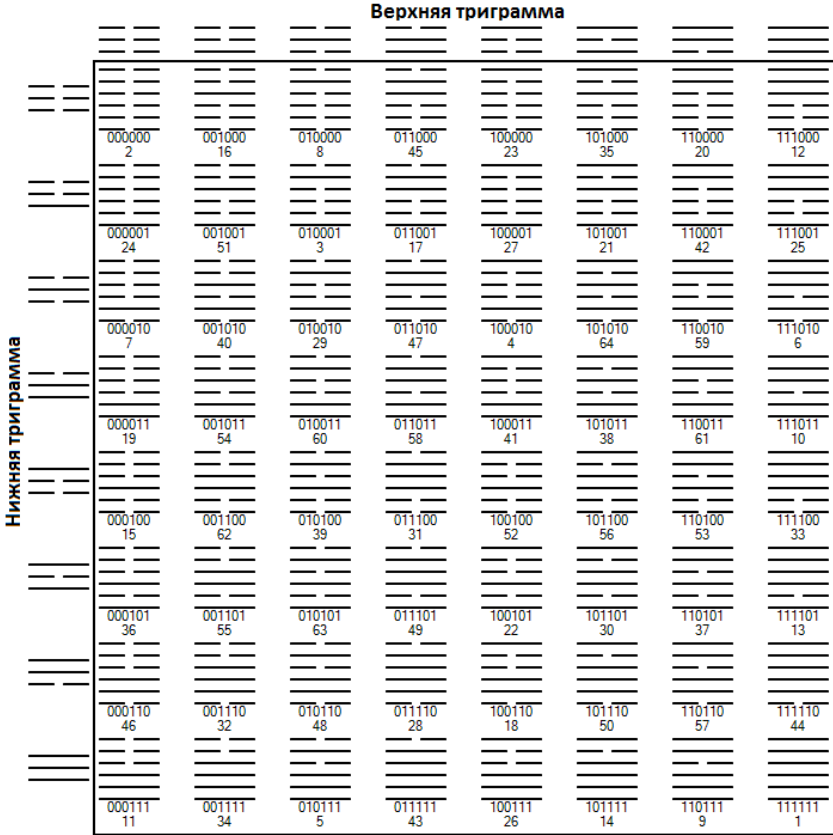


Рис. 2. Формирование гексаграмм из триграмм с двоичным представлением и номером по Вэнь-вану.

Таблица 3. Вариант кодирования нуклеотидов.

	1-й нуклеотид	2-й нуклеотид	3-й нуклеотид
U			
C			
A			
G			

Среди этих 13824 комбинаций присутствуют 648 таких, для которых стоп кодон совпадает с 63-й гексаграммой. Зная информацию о биологических функциях аминокислот среди

оставшихся 648 комбинаций можно отобрать биологически значимые. Рассмотрим эти комбинации с учетом функций аминокислот, выполняемых в человеческом организме. Для упрощения задачи можно использовать информацию о том, что триграммы ассоциируются со "стихиями". Очевидно, что гексаграммы, у которых составляющие их нижняя и верхняя триграммы совпадают, ассоциируются только с одной "стихией", с той, к которой относятся эти триграммы. Далее по тексту для обозначения таких гексаграмм будет использован термин "моностихийные". Названия этих гексаграмм совпадают с названиями соответствующих им триграмм. Выделив моностихийные гексаграммы (они располагаются на рис. 2 по диагонали от левого верхнего до правого нижнего углов), можно сфокусировать внимание на функциях аминокислот, характерных только для отдельно взятых органов человеческого организма, что оказывается удобным, если учитывать ассоциацию органов с пятью "стихиями", принятую в китайской медицине [15].

Рассмотрим некоторые моностихийные гексаграммы, расположенные на диагонали рис. 2. Начнем с гексаграммы Цянь (№ 1 по Вэнь-вану), иероглиф которой имеет значение Творчество. Поскольку творческую деятельность принято ассоциировать с функциями головного мозга, то среди 20 аминокислот логично выбрать такую, которая выполняет в головном мозге уникальное сочетание функций и тем самым может быть соотнесена с гексаграммой Цянь. По нашему мнению, уникальное сочетание функций серина в головном мозге проявляется в том, что он не только используется в синтезе глицина и протеинов, но также в синтезе липидов [16]. Кроме того D-серин, образующийся из L-серина при помощи фермента серин-рацемазы, является нейромодулятором [17]. Следовательно, именно серин логично соотнести с гексаграммой Цянь.

Перейдем к гексаграмме Кунь (№ 2 по Вэнь-вану), которая относится к "стихии" Земли. В китайской медицине к "стихии" Земли принято относить желудок, поджелудочную железу и селезенку. Поэтому, обратив внимание на то, что производство инсулина поджелудочной железой стимулируется лейцином, именно эту аминокислоту целесообразно ассоциировать с гексаграммой Кунь.

В результате соотнесения серина с гексаграммой Цянь и лейцина с гексаграммой Кунь из 648 различных соответствий между кодонами и гексаграммами остается только 8, которые образуют два

различных набора аминокислот, ассоциируемых с моностихийными гексаграммами. В первом наборе гексаграмме Кань (№ 29 по Вэнь-вану) соответствует лейцин, а гексаграмме Гэнь (№ 52 по Вэнь-вану) соответствует аспарагиновая кислота. Во втором наборе гексаграмме Кань соответствует фениланин, а гексаграмме Гэнь соответствует глутаминовая кислота. Можно заметить, что гексаграмма Кань относится к "стихии" Воды. Также к "стихии" Воды в китайской медицине принято относить почки и желчный пузырь. Кроме того, почки контролируют костный мозг, который также относится к "стихии" Воды, что проявляется, например, в виде анемии как осложнения почечной недостаточности. Анемия также может вызываться недостатком определенных веществ, например, аминокислот. По сравнению с недостатком фениланина, недостаток лейцина с большей вероятностью может привести к почечной недостаточности и анемии, поэтому, по нашему мнению, первый набор аминокислот, ассоциируемых с моностихийными гексаграммами, более предпочтителен для человеческого организма, чем второй.

Leu, CUU	Pro, CCU	Phe, UUU	Ser, UCU	Val, GUU	Ala, GCU	Ile, AUU	Thr, ACU
Leu, CUG	Pro, CCG	Leu, UUG	Ser, UCG	Val, GUG	Ala, GCG	Met, AUG	Thr, ACG
Leu, CUA	Pro, CCA	Leu, UUA	Ser, UCA	Val, GUA	Ala, GCA	Ile, AUA	Thr, ACA
Leu, CUC	Pro, CCC	Phe, UUC	Ser, UCC	Val, GUC	Ala, GCC	Ile, AUC	Thr, ACC
His, CAU	Arg, CGU	Tyr, UAU	Cys, UGU	Asp, GAU	Gly, GGU	Asn, AAU	Ser, AGU
Gln, CAG	Arg, CGG	stop, UAG	Trp, UGG	Glu, GAG	Gly, GGG	Lys, AAG	Arg, AGG
Gln, CAA	Arg, CGA	stop, UAA	stop_Sec, UGA	Glu, GAA	Gly, GGA	Lys, AAA	Arg, AGA
His, CAC	Arg, CGC	Tyr, UAC	Cys, UGC	Asp, GAC	Gly, GGC	Asn, AAC	Ser, AGC

Рис. 3. Распределение кодонов по гексаграммам, учитывающее всасывание триптофана в тонком кишечнике. Процессы возбуждения нервной системы ускоренные.

Таким образом, если ассоциировать лейцин с гексаграммой Кань, то останется всего 4 комбинации, связывающие кодоны и гексаграммы, которые представлены на рис. 3, 4, 5, 6. Полученные комбинации имеют одинаковое распределение аминокислот по гексаграммам, за исключением позиции триптофана. Следует отметить, что представленные комбинации относятся к организму человека и связаны со стандартным диалектом генетического кода (таблица 1), а распределения кодонов по гексаграммам для других живых организмов и других диалектов генетического кода в данной статье не рассматриваются. Несмотря на то, что мы считаем, что эти 4 варианта в наибольшей степени проявляются при функционировании человеческого организма, мы допускаем возможность и других распределений кодонов, входящих в набор из 13824 вариантов.

Leu, CUC	Pro, CCC	Phe, UUC	Ser, UCC	Val, GUC	Ala, GCC	Ile, AUC	Thr, ACC
Leu, CUG	Pro, CCG	Leu, UUG	Ser, UCG	Val, GUG	Ala, GCG	Met, AUG	Thr, ACG
Leu, CUA	Pro, CCA	Leu, UUA	Ser, UCA	Val, GUA	Ala, GCA	Ile, AUA	Thr, ACA
Leu, CUU	Pro, CCU	Phe, UUU	Ser, UCU	Val, GUU	Ala, GCU	Ile, AUU	Thr, ACU
His, CAC	Arg, CGC	Tyr, UAC	Cys, UGC	Asp, GAC	Gly, GGC	Asn, AAC	Ser, AGC
Gln, CAG	Arg, CGG	stop, UAG	Trp, UGG	Glu, GAG	Gly, GGG	Lys, AAG	Arg, AGG
Gln, CAA	Arg, CGA	stop, UAA	stop, Sec, UGA	Glu, GAA	Gly, GGA	Lys, AAA	Arg, AGA
His, CAU	Arg, CGU	Tyr, UAU	Cys, UGU	Asp, GAU	Gly, GGU	Asn, AAU	Ser, AGU

Рис. 4. Распределения кодонов по гексаграммам, учитывающее всасывание триптофана в тонком кишечнике. Процессы возбуждения нервной системы замедленные.



Рис. 5. Распределение кодонов по гексаграммам, учитывающее всасывание триптофана в толстом кишечнике. Процессы торможения нервной системы ускоренные.



Рис. 6. Распределение кодонов по гексаграммам, учитывающее всасывание триптофана в толстом кишечнике. Процессы торможения нервной системы замедленные.

3. Заключение

В работе показано, каким образом авторский способ соотнесения функций аминокислот, кодируемых кодонами, с порядком расположения гексаграмм по Вэнь-вану позволил выделить 4 наиболее биологически значимых варианта соотнесения, характерных для человеческого организма. Биологическая значимость этих вариантов подтверждается некоторыми функциями аминокислот.

Литература

1. Коган Р.П. 4 запретных эксперимента. Региональная газета новостей, № 23 (23) 2012, с. 3.
2. Berger J.A., Mitra S.K., Carli M., Neri A. Visualization and analysis of DNA sequences using DNA walks. Journal of The Franklin Institute-engineering and Applied Mathematics, 2004, v. 341, No. 1, p. 37-53.
3. Петухов С.В. Матричная генетика, алгебры генетического кода, помехоустойчивость : монография / С. В. Петухов ; Рос. акад. наук. - М.: Регуляр. и хаотич. динамика, 2008. – 316 с.
4. Petoukhov S.V. The genetic code, 8-dimensional hypercomplex numbers and dyadic shifts. 2011, p. 1-31, from <http://arxiv.org/abs/1102.3596>.
5. Stewart I. The Mathematics of Life. New York: Basic books, 2011, 358p.
6. Беляев М.И. Новое мышление: Генетика Мироздания. М.: Милология, 2010. – 270 с.
7. Свердлов Е.Д. Взгляд на жизнь через окно генома: в 3 т. - М.: Наука. Т.1: Очерки структурной молекулярной генетики. - 2009. - 525 с.
8. Li G.-W., Oh E., Weissman J.S. The anti-Shine–Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature, V. 484(7395), 2012, p. 538–541.
9. Еремеев В.Е. Символы и числа «Книги перемен». М.: АСМ, 2002. – 400 с.
10. Нестеров А. В. Парадоксальная логика Книги Перемен. – Саарбрюккен: LAP LAMBERT Academic Publishing, 2011. – 204 с.
11. Разумов И.К., Еременко Г.А. Числовые закономерности порядка гексаграмм Вэнь Вана. – Доклады независимых авторов, Изд. «DNA», Россия - Израиль, 2008, Вып. 9, с. 95-105.

12. Лу Куань Юй. Даосская йога, алхимия и бессмертие. Пер.с англ. Торчинов Е., общ. ред. Добровин Д. СПб: Орис, 1993. – 368 с.
13. Огудин В.А. Золотые правила фэншуй. 10 простых шагов к успеху, благополучию и долголетию. М., Центрполиграф, 2010. – 512 с.
14. Храмов А., Цуманов А. И-Цзин. Алфавит судьбы М.: Открытый Мир, 2008. – 224 с.
15. Хоушен Л., Пэйной Л. Секреты китайской медицины: 300 вопросов о цигуне. - 2-е изд., доп. и перераб. – Новосибирск: Наука, 1995. – 410 с.
16. Jan Pawd Woronczakl, Ewa Siucińska, Matgorzata Kosst and Jolanta Barańska Temporal dynamics and regional distribution of [¹⁴C]serine uptake into mouse brain. Acta Neurobiol, 1995, p. 233-241.
17. Schell MJ, Brady RO Jr, Molliver ME & Snyder SH (1997). D-Serine as a neuromodulator: regional and developmental localization in rat brain glia resemble NMDA receptors. J Neurosci, 17, p. 1604–1615.