**Конец формы**

**Конец формы**

**Общие модели эволюции. Методы теоретической популяционной генетики. Теория нейтральности М.Кимуры**

1. Классическая популяционная генетика

В этой лекции мы рассмотрим модели, характеризующие общие свойства эволюции. Начнем с синтетической теории эволюции. Эта теория была развита в начале 20-го века. Она основана на учении Ч.Дарвина о естественном отборе и на представлениях Г.Менделя о генах - дискретных элементах передачи наследственных признаков. Большую роль в становлении синтетической теории эволюции сыграла маленькая плодовая мушка Drosophila. Именно эксперименты на этой мушке позволили примирить кажущиеся противоречия между Дарвиновским представлением о постепенном накоплении полезных изменений и наследовании этих изменений и дискретным характером Менделевской генетики. Эксперименты на дрозофиле показали, что мутационные изменения могут быть очень небольшими.

Математические модели синтетической теории эволюции были разработаны Р. Фишером, Дж. Холдейном и С. Райтом. В основном эта математическая теория классической популяционной генетики была завершена к началу 30-х годов.

Согласно синтетической теории эволюции, основным механизмом прогрессивной эволюции является отбор организмов, которые получают выгодные мутации.

2. Математические методы популяционной генетики

Математические модели популяционной генетики количественно характеризуют динамику распределения частот генов в эволюционирующей популяции [1-4,6,8]. Есть два основных типа моделей: 1) детерминистические модели и 2) стохастические модели.

Детерминистические модели предполагают, что численность популяции бесконечно велика, в этом случае  флуктуациями в распределении частот генов можно пренебречь, и динамику популяции можно описать в терминах средних частот генов.

Стохастические модели описывают вероятностные процессы в популяциях конечной численности.

Здесь мы кратко охарактеризуем основные уравнения и математические методы популяционной генетики. Наше изложение будет основываться на рассмотрении наиболее характерных примеров. Уравнения моделей мы будем приводить в основном в демонстрационных целях – без вывода, с пояснением смысла этих уравнений; тем не менее, мы будем приводить ссылки на литературу, в которой сделаны соответствующие математические выводы.

2.1. Детерминистические модели

Рассмотрим популяцию диплоидных1) организмов, которые могут иметь несколько аллелей2) *A1* , *A2* ,..., *AK* в некотором локусе3). Мы предполагаем, что приспособленности организмов определяются в основном рассматриваемым локусом. Обозначая число организмов и приспособленность генной пары *Ai* *Aj* через *nij* и *Wij* , соответственно, мы можем определить частоты генотипа и гена *Pij* и *Pi* , а также средние приспособленности генов *Wi* в соответствии с выражениями [1,2,4]:

*Pij* = *nij* /*n* , *Pi* = S *j Pij* , и *Wi* =*Pi*-1 S *j Wij Pij* ,                                              (1)

где *n* – численность популяции, индекс *i* относится к классу организмов {*Ai* *Aj*} ,  *j* = 1,2,..*., K*, которые содержат ген *Ai* . Популяция предполагается панмиктической4) : при скрещивании новые комбинации генов выбираются случайным образом из всей популяции.

Для панмиктической популяции приближенно справедлив принцип Харди-Вайнберга [1]:

*Pij* =*Pi Pj* , *i, j* = 1,..., *K*.                                                                             (2)

Уравнение (2) означает, что во время скрещивания генотипы формируются пропорционально частотам генов.

Эволюционная динамика популяции в терминах частот генов *Pi* может быть описана следующими дифференциальными уравнениями [1,2,4]:

*dPi /dt* = *Wi Pi* - <*W> Pi* - S *j uji Pi* + S *j uij Pj* , *i* = 1,..., *K*,               (3)

где *t* – время, <*W>* = S *ij Wij Pij* – средняя приспособленность в популяции; *uij* – параметры, характеризующие интенсивности мутационных переходов *Aj* --> *Ai* , *uii* =0 (*i, j* = 1,..., *K*). Первое слагаемое в правой части уравнения (3) характеризует отбор организмов в соответствии с их приспособленностями, второе слагаемое учитывает условие S *i Pi* = 1, третье и четвертое слагаемые описывают мутационные переходы.

Отметим, что подобные уравнения используются в модели квазивидов [5], см Лекция 2

Пренебрегая мутациями, мы можем анализировать динамику генов в популяции посредством уравнений:

*dPi /dt* = *Wi Pi* - <*W> Pi* , *i* = 1,..., *K.*                                                     (4)

Используя (1), (2), (4), можно получить (при условии, что величины *Wij* постоянны), что

скорость роста средней приспособленности пропорциональна дисперсии приспособленности *V* = S *i* *Pi* (*Wi* - <*W>*)2 [1,3]:

*d*<*W>/dt* = 2 S *i* *Pi* (*Wi* - <*W>*)2*.*                                                                      (5)

Таким образом, средняя приспособленность – неубывающая величина. В соответствии с (4), (5), величина *L* = *Wmax* - <*W>* есть функция Ляпунова для рассматриваемой динамической системы (*Wmax* – локальный или глобальный максимум приспособленности, в окрестности которого рассматривается динамика популяции) [3]. Это означает, что величина *L* всегда уменьшается до тех пор, пока не будет достигнуто равновесное состояние (*dPi /dt* = 0).

Уравнение (5) характеризует фундаментальную теорему естественного отбора (Р.Фишер,1930), которая в нашем случае может быть сформулирована следующим образом [3]:

*"В достаточно большой панмиктической популяции, наследование в которой определяется одним n-аллельным геном, а давление отбора, задаваемое Wij , постоянно, средняя приспособленность популяции возрастает, достигая стационарного значения в одном из состояний генетического равновесия. Скорость изменения средней приспособленности пропорциональна аддитивной генной дисперсии и обращается в нуль при достижении генетического равновесия."*

Описанная модель – простой пример модели, использующей детерминистический подход. В рамках этого подхода был разработан широкий спектр аналогичных моделей, которые описывают различные особенности динамики генных распределений, например, учитывают несколько генных локусов, возраст особей, число мужских и женских особей, пространственную миграцию особей, подразделение популяции на субпопуляции и т.п. Многие из моделей и расчетов были предназначены для интерпретации конкретных генетических экспериментальных данных [1,3,4] .

2.2. Стохастические модели

Детерминистические модели позволяют эффективно описывать динамику распределения генов в эволюционирующих популяциях. Однако эти модели основаны на предположении бесконечного размера популяции, которое является слишком сильным для многих реальных случаев. Чтобы преодолеть это ограничение, были разработаны вероятностные методы теоретической популяционной генетики [1,3,4,6-8]. Эти методы включают анализ с помощью цепей Маркова (в частности, метод производящих функций) [4,7], и диффузионные [1,3,4,6,8] методы.

Ниже мы кратко рассмотрим основные уравнения и характерные примеры применения диффузионного метода. Этот метод достаточно нетривиален и его применение приводит к достаточно содержательным результатам.

2.2.1. Прямое и обратное уравнения Колмогорова

Рассмотрим популяцию диплоидных организмов с двумя аллелями *A*1 и *A*2 в некотором локусе. Численность популяции *n* предполагается конечной, но достаточно большой, так что частоты гена могут быть описаны непрерывными величинами. Мы также предполагаем, что численность популяции *n* постоянна.

Введем функцию j = j (*X,t*|*P,0*) , которая характеризует плотность вероятности того, что частота гена *A*1 равна *X* в момент времени *t* при условии, что начальная частота (в момент *t* = 0) была равна *P*. В предположении малого изменения частот генов за одно поколение, динамика популяции может быть описана приближенно следующими дифференциальными уравнениями в частных производных [1,3,4,8]:

*¶* j/¶ *t* = - ¶ (*Md X* j)/¶ *X* + (1/2)¶ 2(*VdX* j)/¶ *X*2 ,                                          (6)

*¶* j/¶ *t* = *Md P ¶* j/¶ *P* + (1/2)*Vd P ¶* 2j/¶ *P* 2 ,                                                  (7)

где *Md X* ,  *Md P* и *VdX* , *Vd P* – средние значения и дисперсии изменения частот *X*, *P* за одно поколение, соответственно; единица времени равна длительности одного поколения. Уравнение (6) есть прямое уравнение Колмогорова. (В физике это уравнение называют уравнением Фоккера-Планка), уравнение (7) – обратное уравнение Колмогорова.

Первые слагаемые справа в уравнениях (6), (7) описывают давление отбора, которое обусловлено разностью приспособленностей генов *A*1 и *A*2. Вторые слагаемые характеризуют случайный дрейф частот, который обусловлен флуктуациями в популяции конечной численности.

Используя уравнение (6), можно определять динамику частот генов во времени. Уравнение (7) позволяет оценивать вероятности фиксации генов.

Предполагая, что 1) приспособленности генов *A*1 и *A*2 равны 1 и 1 - *s* , соответственно и 2) вклады генов в приспособленности генных пар *A*1 *A*1, *A*1 *A*2 и *A*2 *A*2 аддитивны, можно получить, что величины *Md X* ,  *Md P* и *VdX* , *Vd P*  определяются следующими выражениями [1,3,4,8]:

*Md  X* = *sX*(1-*X*), *Md  P* = *sP*(1-*P*), *Vd  X* = *X*(1-*X*)/(2*n*), *Vd P* = *P*(1-*P*)/(2*n*) .            (8)

2.2.2. Случай чисто нейтральной эволюции

Если эволюция чисто нейтральная (*s* = 0), то уравнение (6) принимает вид:

*¶* j/¶ *t* = (1/4*n*)¶ 2[*X*(1-*X*)j]/¶ *X* 2 .                                                                  (9)

Это уравнение было решено аналитически М. Кимурой [1,6]. Само решение имеет сложный вид, основные результаты этого решения сводятся к следующему: 1) в конечной популяции фиксируется только один ген (*A*1либо *A*2); 2) типичное время перехода от начального распределения к конечному составляет величину порядка 2*n* поколений. Отметим, что этот результат согласуется с оценками лекции 4  , где была рассмотрена несколько иная модель "чисто нейтральной" эволюции.

2.2.3. Вероятность фиксации гена

Используя уравнение (7), мы можем оценить вероятность фиксации гена *A*1 в конечной популяции. Действительно, рассматривая асимптотику при времени, стремящемся к бесконечности( *t* --> inf ), мы можем положить ¶ j /¶ *t* = 0 и *X* = 1 ; тогда аппроксимируя вероятность *u*(*P*) , которую нужно найти, величиной *u*(*P*) = j (1, inf |*P,0*)/(2*n*) (здесь *u*(*P*) = j(1, inf |*P,0*)*DX* , где *DX* = 1/2*n* – минимальный шаг изменения частоты в популяции, см. также [3] для более строгого рассмотрения) и комбинируя (7), (8), мы получаем:

*sdu* /*dP* + (1/4*n*) *d* 2*u* /*dP* 2 = 0 .                                                                      (10)

Решая это простое уравнение при естественных граничных условиях: *u* (1) = 1, *u* (0) = 0 , мы получим вероятность фиксации гена *A*1 в конечной популяции [1,3,6]:

*u*(*P*) = [1 - *exp* (- 4*nsP*)] [1 - *exp* (- 4*ns*)]-1 .                                                       (11)

Выражение (11) показывает, что если 4*ns* < < 1 , то имеет место нейтральная фиксация гена: *u*(*P*) » *P* , если 4*ns* > > 1, то отбирается предпочтительный ген *A*1: *u*(*P*) » 1; размер популяции *nc* ~ (4*s*)-1 есть граничное значение, разделяющее области "нейтрального" и "селективного" отбора.

Итак, математические методы популяционной генетики описывают динамику частот генов в эволюционирующих популяциях. Детерминистические методы используются при описании динамики частот в среднем; стохастические методы учитывают флуктуации в популяциях конечной численности.

3. Молекулярная эволюция: теория нейтральности

Классическая теория популяционной генетики, содержательно основанная на синтетической концепции эволюции, интенсивно развивалась до 1960-х годов, до тех пор, пока не возникли трудности интерпретации экспериментальных данных молекулярной биологии. В  лекции 1  я уже отмечал, в 1950-1960-х годах произошла революция в молекулярной биологии. Была определена структура ДНК, расшифрован генетический код, ученые установили общие принципы работы молекулярно-генетической системы живой клетки.

Интенсивные исследования молекулярной биологии привели к серьезным результатам, касающимся биологической эволюции: была оценена скорость аминокислотных замен в белках, а также получены оценки, характеризующие полиморфизм белков.

Анализируя экспериментальные данные, М.Кимура обнаружил, что когда он пытался объяснить эти эксперименты на основе селекции благоприятных мутаций путем Дарвиновского отбора, то возникли серьезные затруднения. В своей книге [6] Кимура подробно описывает идеи, послужившие основанием для изобретения теории нейтральности. Например, в некоторых своих оценках, основанных на Дарвинском отборе, он получил, что для объяснения экспериментальных данных нужно потребовать, чтобы каждая особь в процессе эволюции давала 22 000 потомков. И для того, чтобы проинтерпретировать данные по молекулярной эволюции белков, Кимура предложил теорию нейтральности [6,9].

Основное предположение этой теории состоит в следующем: *на молекулярном уровне мутации (замены аминокислот или нуклеотидов) преимущественно нейтральны или слабо вредны (существенно вредные мутации также возможны, но они элиминируются из популяции селекцией).* Это предположение согласуется с экспериментально наблюдаемой скоростью аминокислотных замен и с тем фактом, что скорость замен в менее важных частях белков значительно больше, чем для активных центров макромолекул.

Используя математические методы популяционной генетики, Кимура получил ряд следствий теории, которые находятся в довольно хорошем согласии с данными молекулярной генетики [6].

Математические модели теории нейтральности существенно стохастические, т.е. относительно малая численность популяции играет важную роль в фиксации нейтральных мутаций. См. примеры расчетов, приведенных выше.

Но если молекулярные замены преимущественно нейтральны, как возможна прогрессивная эволюция? Чтобы ответить на этот вопрос, Кимура использует концепцию дупликации генов, развитую С.Оно [10]. Согласно теории Кимуры, дупликация генных участков создает дополнительные, избыточные ДНК-последовательности, которые в свою очередь дрейфуют далее за счет случайных мутаций, предоставляя тем самым сырой материал, из которого могут возникать новые, биологически значимые гены (Рис.1).

Рис. 1. Иллюстрация к механизму прогрессивной эволюции в теории нейтральности. Схема появления нового биологически значимого белка. Показаны участки ДНК ( *Ii* ) и кодируемые ими белки ( *Ei* ). a) ген *I*1кодирует белок *E*1, b) дупликация гена *I*1, новый участок (справа) кодирует тот же белок *E*1, c) случайный дрейф правого участка, d) возникновение нового биологически значимого белка *E*2кодируемого участком ДНК *I*2.

Заключая наш сжатый обзор теории нейтральности, процитируем пять принципов этой теории [6]. Первые четыре из них – эмпирические, а пятый установлен теоретическим путем.

1. *Скорость эволюции любого белка, выраженная через число аминокислотных замен на сайт в год, приблизительно постоянна и одинакова в разных филогенетических линиях, если только функция и третичная структура этого белка остаются в основном неизменными.*
2. *Функционально менее важные молекулы и их части эволюционируют* (*накапливая мутационные замены*) *быстрее, чем более важные.*
3. *Мутационные замены, приводящие к меньшим нарушениям структуры и функции молекулы* (*консервативные замены*)*, в ходе эволюции происходят чаще тех, которые вызывают более существенное нарушение структуры и функции этой молекулы.*
4. *Появлению нового в функциональном отношении гена всегда должна предшествовать дупликация гена.*
5. *Селективная элиминация вредных мутаций и случайная фиксация селективно нейтральных или очень слабо вредных мутаций происходят в ходе эволюции гораздо чаще, чем положительный дарвиновский отбор благоприятных мутаций.*

4. Другие модели, характеризующие общие закономерности эволюции

Теория нейтральности – одна из наиболее разработанных общих теорий эволюции. Однако есть ряд моделей и концепций, также характеризующих эволюцию на молекулярном уровне, которые в основном дополняют теорию нейтральности. Отметим наиболее известные из них.

В работах Д.С.Чернавского и Н.М.Чернавской [11,12] сделана оценка вероятности случайного формирования нового биологически значимого белка с учетом того, что в белке есть активный центр, в котором замены аминокислот практически недопустимы, и участки, свойства которых не сильно меняются при многих аминокислотных заменах. Там же сделана оценка количества возникающей в геноме информации при появлении нового белка. Полученная оценка указывает на то, что случайное формирование белка было вполне вероятно в процессе эволюции.

Интересна, хотя, по-видимому, не бесспорна, модель блочно-иерархического эволюционного отбора [13,14], согласно которой новые генетические тексты большой длины сначала случайно составляются из коротких текстов, оптимизированных в предыдущие эволюционные эпохи, а после составления оптимизируются. Модель блочно-иерархического эволюционного отбора критически проанализирована в [15].

Блочно-модульный принцип организации и эволюции молекулярно-генетических систем управления обосновывается В.А.Ратнером [16]. Согласно этому принципу эволюция генов, РНК, белков, геномов и молекулярных систем управления на их основе шла путем комбинирования блоков (модулей) снизу доверху, причем модулями, из которых составлялись вновь возникающие молекулярно-генетические системы, служили уже функционирующие макромолекулярные компоненты. По сравнению с моделью блочно-иерархического отбора блочно-модульный принцип более гибок и более реалистичен.

В модели "генов-мутаторов" [17] предполагается, что уровень мутаций может меняться и наследоваться, в результате чего при попадании популяции в новую среду, когда выгоден активный поиск новых свойств, уровень мутаций возрастает, а при длительном нахождении в постоянной среде, где важно сохранение уже найденных свойств, интенсивность мутаций падает.

Интересно проанализировать, как могли возникать достаточно нетривиальные системы обработки информации. Для простейших организмов (вирусов и бактерий) в качестве таковых можно рассматривать регулирование синтеза белков (функциональных и структурных элементов организма) в соответствии с определенной "программой". Например, в процессе онтогенеза фага Т4 происходит образование сложной пространственной структуры, в формировании которой участвуют несколько десятков белков, синтезируемых в соответствии с программой, закодированной геномом фага [18]. Иллюстративная модель эволюционного возникновения подобных "программ жизнедеятельности" предложена [19]. Согласно модели в процессе эволюционного формирования этих программ в генотип закладывается некоторая избыточность, которая приводит к тому, что при небольшой модификации генома часть блоков программ сохраняется. При введении в модель представлений о "генах-мутаторах" наблюдалось поведение, качественно сходное с явлением каскадного мутагенеза [20] – резким возрастанием интенсивности мутаций после дестабилизации генома.

В чрезвычайно интересном цикле работ С.Кауффмана с сотрудниками [21,22] исследуется эволюция автоматов, состоящих из соединенных между собой логических элементов. Отдельный автомат можно рассматривать как модель молекулярно-генетической системы управления живой клетки, при этом каждый логический элемент интерпретируется как регулятор синтеза определенного фермента. Модели Кауффмана позволяют сделать ряд предсказаний относительно "программ" жизнедеятельности клеток. В частности, продемонстрировано, что для одновременного обеспечения устойчивости и гибкости программы число входов логических элементов должно быть ограничено определенным интервалом, а именно составлять величину примерно равную 2-3. Для моделей Кауффмана разработаны эффективные методы анализа на базе статистической физики, эти модели получили широкую известность и исследовались рядом ученых. Подробнее основные результаты этой модели мы обсудим в следующей лекции.

Специальные термины:

1) Диплоидный организм: особь, имеющая два набора хромосом в каждой из ее клеток.

2) Аллель: Одна из различных форм гена, который может быть в заданном локусе.

3) Локус: участок хромосомы, в котором локализован ген.

4) Панмиксия: полностью случайное скрещивание.

Литература

1. *J.F. Crow, M. Kimura.* "An introduction to population genetics theory". New York etc, Harper & Row. 1970.
2. *T. Nagylaki.* "Introduction to theoretical population genetics ". Berlin etc, Springer Verlag. 1992.
3. Свирежев Ю.М., Пасеков В.П. Основы математической генетики. М.: Наука, 1982, 511 с.
4. *P.A.P. Moran.* “The statistical processes of evolutionary theory”, Oxford, Clarendon Press, 1962. Имеется перевод: П. Моран. Статистические процессы эволюционной теории. М.: Наука, 1973. 288 с.
5. Эйген М., Шустер П. Гиперцикл. Принципы самоорганизации макромолекул. М.: Мир, 1982. 270 с.
6. Кимура М. Молекулярная эволюция: теория нейтральности. М.: Мир, 1985, 400 с.
7. *S. Karlin.* "A first course in stochastic processes". Academic Press. New York, London, 1968. Имеется перевод: С.Карлин. Основы теории случайных процессов. М.: Мир, 1975.
8. Ратнер В.А. Математическая популяционная генетика. Новосибирск: Наука, 1976. 128 с.
9. *M. Kimura.* Evolutionary rate at the molecular level // Nature. London, 1968.V.217. PP.624-626.
10. Оно С. Генетические механизмы прогрессивной эволюции. М.: Мир, 1973, 228 с
11. Чернавская Н.М., Чернавский Д.С. Проблема возникновения новой информации в эволюции // Термодинамика и регуляция биологических процессов. Теория информации, управление в живых системах, проблема самоорганизации, эволюция и онтогенез. М.: Наука. 1984. С.247-255
12. Романовский Ю.М., Степанова Н.В., Чернавский Д.С. Математическая биофизика. М.: Наука, 1984. 304 с.
13. Иваницкий Г.Р., Есипова Н.Г., Абагян Р.А., Шноль С.Э. Блочное совершенствование генетического текста как фактор ускорения биологической эволюции // Биофизика. 1985. Т.30. N.3. С.418-421.
14. Шноль С.Э. Хватает ли времени для дарвиновской эволюции? // Природа, 1990. N.11. С. 23-26.
15. Моносов Я.А. О факторах ускорения биологической эволюции // Биофизика. 1991. Т.36. N.5. С.920-922.
16. Ратнер В.А. Блочно-модульный принцип организации и эволюции молекулярно-генетических систем управления (МГСУ) // Генетика. 1992. Т.28. N.2. С.5-23.
17. Семенов М.А., Теркель Д.А. Об эволюции механизмов изменчивости посредством косвенного отбора // Журн. общ. биологии. 1985.Т. 46. N.2. С. 271 -277.
18. Ратнер В.А. Молекулярно-генетические системы управления. Новосибирск: Наука, 1975. 288 с.
19. Редько В.Г. К теории эволюции. Модель происхождения "программ жизнедеятельности" // Журн. общ. биологии. 1991.Т.52. N.3. С. 334-342.
20. Корогодин В.И. Кариотаксоны, надежность генома и прогрессивная эволюция // Природа. 1984. N.2. С.3-14.
21. Kauffman S.A., Smith R.G. Adaptive automata based on Darwinian selection // Physica D. 1986. V.22. N.1-3. P.68-82.
22. Кауффман С. Антихаос и приспособление // В мире науки. 1991. № 10. С. 58.

Copyright © Vladimir Red'ko,  Oct 25, 1999 ( redko@keldysh.ru )

Last modified: Oct 25, 1999