

Коанцестри и уставное число популяций применительно к оценке генетического разнообразия лесных древесных пород

А. П. Царёв¹

Петрозаводский государственный университет

АННОТАЦИЯ

Изложены содержание понятия коанцестри и концепция шведских исследователей по использованию коанцестри и уставного числа популяций для оценки генетического разнообразия лесных древесных пород.

Ключевые слова: коанцестри, инбридинг, идентичность по происхождению, уставное число, генетическое разнообразие, лесные древесные породы, популяция.

SUMMARY

The paper considers the ideas of Sweden researchers about coancestry and status number of population. The opportunity to use these figures for assessment of forest tree species genetic diversity is described.

Keywords: coancestry, inbreeding, identical by descens, status number, genetic diversity, forest tree species, population.

ВВЕДЕНИЕ

Оценка генетического разнообразия лесных древесных пород в настоящее время представляет значительные трудности. Для получения репрезентативных данных требуются трудоемкие и дорогие исследования, поэтому ученые ищут подходы для удешевления и снижения трудоемкости работ. Одним из таких подходов является концепция шведского исследователя Дага Линдгрена и его школы по использованию показателя коанцестри и уставного числа популяции для оценки генетического разнообразия лесных древесных пород. Поскольку в отечественной практике эти понятия ещё не нашли своего применения, целесообразно более подробно остановиться на их характеристиках.

В любой отдельно взятой популяции древесных растений могут находиться особи, связанные различной степенью родства. Требуется какая-то мера, выражающая степень родства особей в популяции. Для этого предлагается показатель, который называется коэффициент коанцестри, или коэффициент родства. Этот коэффициент отражает вероятность происхождения исследуемых индивидуумов от общего предка.

В иностранной литературе он имеет разные названия: coefficient of coancestry, coefficient of consanguinity,

coefficient of kinship (англ.), что означает коэффициент родства (кровного), близости, сходства. Во французской литературе этот коэффициент родства обозначается как coefficient de parent'e.

ИЗЛОЖЕНИЕ ПРОБЛЕМЫ

Определение этого показателя может быть сформулировано следующим образом. **Коэффициент родства (коанцестри)** - это вероятность того, что отдельные гены, взятые случайно от двух индивидуумов x и y , являются идентичными по происхождению (Lynch M., Walsh B., 1997; Lindgren D. et al., 1997, a, b, c; Lindgren D., 1999, и др.).

Из определения коанцестри видно, что в то время как коэффициент инбридинга (F) описывает родство между гомологичными генами индивидуума, коэффициент коанцестри (f) есть вероятность того, что ген, взятый случайно от одного индивидуума, будет идентичным по происхождению случайному гомологичному гену от другого индивидуума (Mullin T. J., Bertrand S., 1998).

Эти подходы могут оказаться очень полезными как для программ селекции и семеноводства отдельных лесных древесных пород, так и лесовосстановления в целом. Для дальнейшего изучения предмета необходимо дать или напомнить определения еще нескольких показателей.

Идентичный по происхождению (Identical by descens - IBD) означает, что гены одного и того же локуса являются копиями тех же самых первоначальных (оригинальных) генов какого-то предка, или анцестора (ancestor - англ.). Например, Случай, когда оба гомологичных гена в той же самой зиготе являются идентичными по происхождению, называется инбридингом (F), или коэффициентом инбридинга. Если две особи спариваются, их коанцестри становится инбридингом в их потомстве.

Основатель популяции является отправной точкой отсчета или вычислений. Если весь инбридинг и коанцестри основателя популяции известны, то инбридинг и коанцестри могут быть рассчитаны по родословной (педигри). Практично и удобно полагать инбридинг и коанцестри в «диком лесу» нулевыми и рассматривать основателей (плюсовые деревья) как образцы из дикого леса.

Самокоанцестри - это коанцестри индивидуума с самим собой. Оно равно $f = 0,5$. Его можно представить, например, при допущении, что коанцестри в предыдущем поколении становится инбридингом в последующем при самоопылении.

Генный пул означает совокупность всех генов в популяции. Удобно рассматривать гены одного локуса. Пул генов не зависит от того, как популяция организована в зиготы.

Рассмотрим более подробно взаимоотношения между генами в популяции. Популяция с N зиготами имеет

¹ Автор – профессор кафедры лесного хозяйства
© А. П. Царёв, 2001

$2N$ генов в генном пуле. Каждый ген имеет частоту $1/2N$. Вероятность выбрать тот же самый ген дважды равна $1/2N$. Вероятность, что будут отобраны различные гены, равна $(1-0,5/N)$. Гены могут быть идентичные по происхождению. Вероятность этого есть коанцестри.

Генный пул структурирован по индивидуумам. Вероятность того, что гены в двух индивидуумах идентичны по происхождению, есть коанцестри между этими двумя индивидуумами.

Вероятность того, что различные гены одной и той же зиготы являются идентичными по происхождению, есть коэффициент инбридинга.

Для самокоанцестри гены не нуждаются в различии. Если они различны, $f=F$, если те же самые, $f=I$, в среднем $f=(I+F)/2$.

В исследуемой популяции возможны три различных механизма определения генов, идентичных по происхождению:

1. Тот же самый ген отбирается дважды (дрейф);
2. Гены – это гомологичные гены от того же самого индивидуума (инбридинг),
3. Гены происходят от различных индивидуумов (родство).

Оценки значений коанцестри спариваемых индивидуумов могут быть представлены в виде матрицы коанцестри (табл. 1).

Таблица 1

Значения коанцестри спариваемых индивидуумов

Индивидуумы	1	2	3
1	0,5	0,25	0
2	0,25	0,5	0
3	0	0	1

Обозначим оценку некоторого спаривания $f_{2,1} = 0,25$. Поскольку ковариационная матрица симметричная, то $f_{2,1} = f_{1,2}$. Значения показателей по диагонали (самокоанцестри) появляются только однажды. Коэффициент родства представляется в такой матрице числителем. В отсутствие инбридинга эти оценки в два раза больше. Следовательно, коанцестри – это вероятности, и, таким образом, $0 \leq f \leq 1$.

Некоторые примеры значений коанцестри:

Родство особей	Коанцестри
Отсутствует	0
Полусибсы	0,125
Полные сибсы	0,25
Родители-потомки	0,25

Групповое коанцестри. Если все гомологичные гены поместить в один большой пул и выбрать наугад два из них (с замещением), то вероятность того, что они являются идентичными по происхождению, определяется как *групповое коанцестри* (Θ). Этот термин был предложен Кокерхэмом (Lindgren D., 1999).

Групповое коанцестри равняется среднему числу всех N^2 значений коанцестри всех комбинаций N индивидуумов в популяции (или среднее число всех $4 N^2$ комбинаций индивидуальных генов). Возможно также определять групповое коанцестри как среднее число. Преимущество вероятностного определения проявляется в более сложных ситуациях.

Сформируем, или получим, среднее число из матрицы коанцестри (табл. 1). Сумма из 9 оценок в матрице = 2,5. Среднее число = групповому коанцестри = $2,5/9 = 0,278$. То есть групповое коанцестри здесь – это среднее число.

Следует обратить внимание на то, что показатели самокоанцестри появляются однажды, в то время как другие показатели коанцестри появляются дважды (реципрокно). Если все индивидуумы в популяции имеют один и тот же тип, достаточно вычислить N показателей коанцестри для одного отдельного индивидуума. Самокоанцестри – это групповое коанцестри для популяции с одним единственным членом.

Все члены в семействе полных сибсов имеют одинаковый показатель коанцестри. Таким образом, достаточно построить матрицу коанцестри для семейства полных сибсов, чтобы иметь представление о некоторых генетических параметрах исследуемой совокупности деревьев. Групповое коанцестри зависит от родства, независимо от того, как гены объединены в гаметы. Брат одинаково связан со своим братом, как и со своей сестрой, несмотря на то, что его гаметы способны объединиться только с таковыми его сестры.

Если в популяции известны инбридинг и коанцестри, групповое коанцестри может быть рассчитано. Формула для группового коанцестри в этом случае может быть следующей:

$$\Theta = \frac{1 + \overline{F} + (2N - 2) \overline{f}}{2N}, \quad (1)$$

$$\overline{f} = \frac{N \Theta - 0,5(1 + \overline{F})}{N - 1}. \quad (2)$$

Здесь Θ – групповое коанцестри; N – индивидуумы;

\overline{f} – среднее значение коанцестри для различных индивидуумов; \overline{F} – среднее значение инбридинга.

Связь между поколениями. Изменения группового коанцестри при изменении поколения можно вычислить ретроспективно, если известна родословная, связывающая данную популяцию с основателями.

Будущее групповое коанцестри может быть рассчитано, если имеется представление о будущих родословных. Для других случаев прогноз также возможен, однако это более сложный процесс. Связь между поколениями осуществляется через гаметы. Генный пул потомства идентичен генному пулу успешных гамет родителей.

Рассмотрим пару генов, которые могут быть эквивалентно расценены как в зиготах потомства, так и в успешных родительских гаметах. Пара генов может быть идентична по происхождению, поскольку они - копии того же самого гена в родительской популяции. Это может случаться, если родитель имеет больше чем одного потомка.

Пара генов может происходить из однородных генов той же самой родительской зиготы в родительской популяции. Если был инбридинг, рассматриваемые гены могут быть идентичны по происхождению.

Сибсы - это те особи, которые делят одного родителя и получают коанцестри, равное $(1 + F_p) / 2$. Если рассматриваемая пара генов происходит от различных родителей, коанцестри будет равным f_p . Индекс p при показателях инбридинга (F) и коанцестри (f) означает, что данные показатели относятся к родительским поколениям.

Идентичность по происхождению может обеспечиваться следующими механизмами:

1. Тот же самый ген в текущем поколении отбирается дважды;
2. Гены являются копиями того же самого гена в родительском поколении;
3. Гены происходят от гомологичных генов в одном и том же инбредном родителе;
4. Гены приходят от различных, но связанных родством, родителей.

Связь между групповым коанцестри и разнообразием генов. Групповое коанцестри обеспечивает вероятность того, что два гена в популяции являются идентичными по происхождению.

Разнообразие означает, что мы имеем дело с различными объектами. Разнообразие генов означает, что гены различны.

Очевидно, что групповое коанцестри может быть оценено как 1 (единица), если существует полная вероятность того, что гены не являются идентичными, то есть они разнообразны. Для оценки разнообразия генов в какой-либо популяции может быть использована формула:

$$GD = 1 - \Theta, \quad (3)$$

где GD - мера разнообразия генов, а групповое коанцестри Θ - мера потерянного разнообразия генов в рассматриваемой популяции.

Этот подход предполагает, что все гены в индивидуумах рассматриваемых популяций уникальны.

Мера, основанная на групповом коанцестри, относится к рассматриваемой популяции. Генное разнообразие GD является ожидаемым средним значением гетерозиготности.

Для селекции лесных древесных пород дикий лес представляет хорошую основу. Разнообразие генов дикого леса равно 1, и групповое коанцестри может дать представление о доле или проценте потерянного разнообразия генов.

Если контролировать групповое коанцестри при различных приемах селекции лесных древесных пород, то можно отметить, как много было потеряно разнообразия генов по сравнению с диким лесом.

Уставное число (*status number, N_s*) - это половина обратной величины группового коанцестри. Или, что является эквивалентным, это половина обратной величины вероятности того, что два гена, взятые случайно, являются идентичными по происхождению. Этот показатель ввел в научный обиход шведский ученый Даг Линдгрэн.

Привлекательная особенность уставного числа заключается в том, что оно отражает *списочное число* (перечень) N неродственных и неинбредных индивидуумов в популяции. Уставное число свидетельствует о том, что вероятность отобрать два идентичных по происхождению гена такая же, как если бы это было число неродственных неинбредных индивидуумов, составляющих генный пул. Поэтому его можно называть *эффективным числом* (Lindgren D., 1999).

Уставное число - интуитивно более привлекательный путь представления группового коанцестри, поскольку это связано с известной концепцией размера популяции.

Отношение уставного числа и списочного числа называется *относительным уставным числом* (N_r):

$$N_r = N_s / N. \quad (4)$$

Основные свойства уставного числа следующие:

- N_s никогда не может быть выше, чем списочное число (N);
- N_s никогда не может быть более низким, чем 0,5 (оценка гаметы);
- N_s рассматривает родство и инбридинг;
- не имеет значения, если члены «популяции» принадлежат тому же самому поколению или той же самой «подпопуляции». Уставное число может быть рассчитано для любой гипотетической популяции (с известной родословной);
- это - не то же самое, что инбридинг, но оно тесно связано с инбридингом, поскольку инбридинг, наступающий после равного количества всех возможных спариваний, будет равен $0,5 / N_s$. После случайного спаривания это является ожидаемой величиной. Но инбридинг может быть равным 1, в то время как уставное число - $0,5 \# N_s \# N/2$. Инбридинг также может

быть равным 0, в то время как уставное число - $1 \# N_s$ $\# N$:

- если имеется ряд дискретных поколений с различным списочным числом, N_s не может превышать минимум N в предшествующих поколениях;

- для уставного числа не имеет значения род членов популяции;

- уставное число после смены поколения зависит только от численности потомства для каждого родителя. Оно не зависит от того, как родители спаривались. Уставное число описывает генный пул, а не то, как он образован;

- уставное число обычно снижается при смене поколений, но оно может и повышаться, поскольку первоначальные геномы могут иметь более равное представительство после изменения поколения, чем прежде;

- уставное число может быть рассчитано для любой популяции с известной родословной, связывающей ее с основателями с известным коанцестри и инбридингом (подобно вычислениям инбридинга). Для этого Дагом Линдгреном с сотрудниками разработаны специальные алгоритмы (Lindgren D. et al., 1997; Lindgren D., 1999).

С помощью уставного числа и группового коанцестри можно измерять дисперсию гена. Согласно Кокерхэму (1969, по Lindgren D., 1999). варианса частоты гена (p), или среднего значения встречаемости гена, следующая:

$$\sigma_p^2 = \Theta p(1-p), \quad (5)$$

или

$$\sigma_p^2 = \frac{p(1-p)}{2N_s}. \quad (6)$$

Это - биномиальное выражение вариансы для частоты гена в популяции с N_s неинбредными и неродственными членами.

Таким образом, уставное число - это количество неродственных и неинбредных деревьев, отобранных из исследуемой популяции, которые испытали тот же самый дрейф, как и накопленный дрейф данной популяции.

Концепция уставного числа для лесных селекционеров более полезна, чем для генетиков и селекционеров, работающих с другими объектами. Это связано с тем, что лесные селекционеры ближе к основателям популяций и вообще имеют дело во многих случаях с диким лесом. Они также имеют дело с небольшим числом поколений, имеющих специфические смены между собой. Кроме того, лесные селекционеры могут осуществлять контроль над популяциями в течение длительного времени.

В заключение необходимо отметить, что то краткое изложение понятий и терминов, которые приводи-

лись в настоящей работе, дает только первое представление о рассматриваемом вопросе. Более полное знание можно почерпнуть из упоминавшихся выше публикаций Дага Линдгрена и других зарубежных исследователей.

Рассмотренные подходы следует апробировать при создании лесосеменной базы лесных древесных пород. Они позволяют вести мониторинг генетического разнообразия лесных древесных пород при внедрении их в лесокультурное производство.

ЛИТЕРАТУРА

1. Lindgren D. Doctorand course OH A2; Kurs 99 A. Doc / Dag Lindgren, gen/fys SLU; file: KURSA. DO6 // <http://linne.genfys.slu.se/bred/bred.htm> - 1999. 54 p.
2. Lindgren D., Gea L. D., Jefferson P. A. Status number for measuring genetic diversity // *Forest Genetics*. 1997. No 4 (2). P. 69 - 76 (a).
3. Lindgren D., Kang Kyu Suk. Status number - a useful tool for tree breeding // *The Research Report of the Institute of Forest Genetics*. 1997. No 33. P. 154 - 165 (b).
4. Lindgren D., Wei Run-Peng, Lee St. J. How to calculate Optimum Family Number When Starting a Breeding Program // *Forest Science*. 1997. Vol. 43. No 2. P. 206 - 212 (c).
5. Lynch M., Walsh B. *Genetics and Analysis of Quantitative Traits* / Sinauer Associates, Inc. Publishers Sunderland. Massachusetts, 01375 USA, 1997. 980 p.
6. Mulin T. J., Bertrand S. *Forest management on genetics of forest tree species: Report under Order Number 122994*. Ontario Ministry of Natural Resources / Genesis Forest Science Canada Inc. C. P. 64. Succursale Haute-Ville, Quebec, Quebec G1R 4M8. 1998. 36 p.