

УДК 577.3:616-006:575.17:519.22/25

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ РАСПРЕДЕЛЕНИЯ В ЧЕЛОВЕЧЕСКОЙ ПОПУЛЯЦИИ ГЕНОВ, ПОВРЕЖДЕНИЕ КОТОРЫХ ОБУСЛОВЛИВАЕТ ВОЗНИКНОВЕНИЕ ОНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ

М.А. Бондаренко, В.Г. Книгавко, Н.А. Гордиенко, О.В. Зайцева, Т.Р. Кочарова

Харьковский национальный медицинский университет, пр. Науки, 4, Харьков, 61022, Украина

e-mail: vkni@mail.ru

Поступила в редакцию 1 марта 2016 года

Принята 17 марта 2016 года

Объектом исследования является канцерогенез. Целью исследования является создание математической модели канцерогенеза, позволяющей оценить распределения в человеческой популяции числа тех генов, повреждение которых является причиной образования онкологических заболеваний (а-генов). Основной задачей является построение такой вероятностной математической модели, которая описывает квазистационарное равновесие двух противоположных процессов, а именно: 1) процесса сокращения в популяции числа вышеуказанных генов вследствие их мутационного повреждения; 2) увеличение в популяции числа этих генов за счет того, что индивиды, имеющие малое число таких генов в их генотипе, с большей вероятностью приобретают онкологические заболевания на ранних стадиях жизни и не успевают до смерти осуществить репродукцию, и поэтому рост общей численности популяции идет больше за счет репродукции индивидов с высоким числом а-генов. Оценка распределения таких генов в популяции проводилась путем определения вероятности того, что наугад выбранный из популяции индивид будет иметь одно из возможных значений (по литературным данным, от 0 до 8) числа вышеуказанных генов.

КЛЮЧЕВЫЕ СЛОВА: канцерогенез, стволовые клетки, повреждения генов, математическое моделирование, распределение генов в популяции.

MATHEMATICAL MODELLING OF DISTRIBUTION OF GENES THE DAMAGE OF WHICH LEADS TO ONCOLOGIC DISEASES IN HUMAN POPULATION

M.A. Bondarenko, V.G. Knigavko, N.O. Gordiyenko, O.V. Zaytseva, T.R. Kocharova

Kharkiv National Medical University, Nauky ave., 4, 61022 Kharkiv, Ukraine

Carcinogenesis is subject of the research. The research aims at creating the mathematical model of carcinogenesis allowing assessing the distribution in human population of the genes which when damaged lead to oncology diseases. The main task is to build a probability mathematical model describing the quasistationary equilibrium of two contrary processes, and namely: 1) the process of reduction in population of the number of the aforesaid genes due to their mutative damage; 2) increase in population of the number of these genes due to the fact that persons with a few genes of the kind in their genotype acquire oncological diseases with higher probability at early stages of their lives and do not manage to reproduce themselves before they die, and so the growth of the total population size is more due to the reproduction of individuals with a high number of the a-genes. Assessment of the distribution of these genes in the population was carried out by determining the probability that a randomly selected individual from the population has one of the possible values (according to the literature, from 0 to 8) of the aforementioned genes.

KEY WORDS: carcinogenesis, stem cells, genes damage, mathematical modelling, distribution of genes in a population.

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ РОЗПОДІЛУ В ЛЮДСЬКІЙ ПОПУЛЯЦІЇ ГЕНІВ, ПОШКОДЖЕННЯ ЯКИХ ЗУМОВЛЮЄ ВИНИКНЕННЯ ОНКОЛОГІЧНИХ ЗАХВОРЮВАНЬ

М.А. Бондаренко, В.Г. Книгавко, Н.О. Гордієнко, О.В. Зайцева, Т.Р. Кочарова

Харківський національний медичний університет, пр. Науки, 4, Харків, 61022, Україна

Об'єктом дослідження є канцерогенез. Метою дослідження є створення математичної моделі канцерогенезу, що дозволить оцінити розподіл в людській популяції кількості тих генів, пошкодження яких є причиною виникнення онкологічних захворювань. Основним завданням є побудова такої імовірнісної математичної моделі, яка описує квазістаціонарну рівновагу двох протилежних процесів, а саме: 1) процесу скорочення в популяції кількості вищевказаних генів внаслідок їх мутационного пошкодження; 2) збільшення в популяції кількості цих генів за рахунок

того, що індивіди, які мають малу кількість таких генів в їхньому генотипі, з більшою імовірністю набувають онкологічних захворювань на ранніх стадіях життя і не встигають до смерті здійснити репродукцію, і тому зростання загальної чисельності популяції йде більше за рахунок репродукції індивідів з високим числом а-генів. Оцінка розподілу таких генів в популяції проводилася шляхом визначення ймовірності того, що навмання вибраний з популяції індивід має одне з можливих значень (за літературними даними, від 0 до 8) числа вищевказаних генів.

КЛЮЧОВІ СЛОВА: канцерогенез, стовбурові клітини, пошкодження генів, математичне моделювання, розподіл генів в популяції.

ПОСТАНОВКА ЗАДАЧИ

В литературе [1] уже достаточно давно высказывается мнение о том, что необходимым условием возникновения онкологического заболевания индивида является образование стволовой клетки с нестабильным генотипом, что, в свою очередь связано с повреждением в этой клетке некоторых определенных генов. Предполагается также, что указанные гены – это гены, участвующие в репарации поврежденных ДНК клеток и число этих генов не более 8. Далее эти гены будем называть а-генами.

В работе [2], исходя из вышеуказанных соображений, была построена математическая модель процесса канцерогенеза, позволившая получить математические выражения для таких вероятностных характеристик этого процесса, как функция распределения времени образования опухоли, плотность вероятности времени образования опухоли, среднее и наиболее вероятное значения времени образования опухоли.

Вместе с тем, очевидно, что вследствие мутационного повреждения указанных генов, имеющего случайный характер, число а-генов в генотипах разных индивидов является разным, причем 8 – это максимально возможное значение числа а-генов. Очевидно также, что в процессе жизни индивида число его а-генов будет или уменьшаться, или (что менее вероятно) оставаться постоянным.

Учитывая, что в настоящее время наблюдается возрастание онкологической заболеваемости, весьма актуальной становится проблема оценки распределения в человеческой популяции индивидов с разными значениями числа а-генов. Говоря об оценке этого распределения, мы имеем в виду определение вероятностей того, что наугад выбранный из популяции индивид имеет в генотипе определенное количество (от нуля до восьми) а-генов. Это распределение, разумеется, изменяется со временем, но изменяется очень медленно, то есть является квазистационарным. Поэтому можно с достаточной точностью считать его стационарным, то есть считать, что вероятность обнаружения генотипа с определенным количеством а-генов является величиной постоянной.

Стационарность распределения определяется тем, что в популяции происходят одновременно два противоположно направленных процесса: 1) мутации уменьшают число а-генов в генотипе индивида; 2) у индивидов с исходно малым количеством а-генов в какой-то стволовой клетке достаточно рано повреждаются все а-гены, происходит малигнизация клетки, возникает онкологическое заболевание до репродукции (до рождения ребенка) и этот индивид элиминируется из популяции.

Дальнейшее изложение целесообразно проводить, анализируя изменения со временем числа а-генов в геноме индивида.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Поскольку максимальное количество а-генов у человека равно 8, то количество таких генов в гамете может принимать значения от 0 до 4.

Поскольку гаметы сохраняют свою индивидуальность в составе диплоидной клетки, в дальнейшем изложении можно, главным образом, анализировать процессы,

происходящие с гаметой в течение времени от предыдущей репродукции до следующей. Будем называть это время репродуктивным периодом.

Вероятность сохранения гаметы зависит как от времени репродукции (возраста родителей), так и от длительности жизни индивида. Факторы, влияющие на продолжительность жизни, можно подразделить на те, что связаны с онкологическими заболеваниями, и те, что с этими заболеваниями не связаны. В свою очередь, длительность жизни онкологического больного и возможность репродукции зависят от эффективности лечения. В дальнейшем изложении необходимо учитывать все эти факторы. Для построения вероятностной математической модели распределения числа а-генов в человеческой популяции введем следующие обозначения.

Пусть N – число стволовых клеток организма, t – время, прошедшее после образования гаметы, α – вероятность повреждения какого-либо а-гена в одном цикле деления; $\beta = 1 - \alpha$; n – число циклов деления от момента образования гаметы; k – начальное число а-генов в клетке; T – средняя продолжительность клеточного цикла. Пусть также m – начальное число а-генов в гамете, $\phi(m)$ – вероятность того, что при образовании гаметы в ней содержалось m а-генов; $\psi(m)$ – вероятность того, что к моменту следующей для данной гаметы репродукции в гамете содержатся m а-генов; $s(t)$ – функция распределения длительности жизни индивидов, не имеющих онкологического заболевания; $v(\tau)$ – плотность вероятности времени репродукции; $\Phi(m, t)$ – функция распределения времени появления злокачественного заболевания при изначальной наявности в гамете m а-генов; $F(k, t)$ – функция распределения времени появления злокачественного заболевания при исходном наличии в генотипе индивида k а-генов; L – вероятность излечения онкологического больного.

Проанализируем процессы сохранения числа а-генов и изменения этого числа в гаметах со временем.

Рассчитаем сначала вероятность того, что гамета, изначальное содержащая 4 а-гена, сохранила до следующей репродукции все эти 4 а-гена. Обозначим эту вероятность как $P_{4 \rightarrow 4}$. Для сохранности гаметы необходимо, чтобы в течение времени до следующей репродукции выполнялись такие условия: мутации в течение этого времени не повредили ни один из а-генов в гамете, продолжительность жизни индивида была больше этого же времени, индивид за это время не приобрел онкологическое заболевание, либо приобрел, но вылечился.

Вероятность того, что репродукция произойдет в момент времени t (точнее в интервале времени от t до $t + dt$), равна $v(t)dt$. Вероятность того, что индивид еще жив к этому моменту, равна $1 - s(t)$. Вероятность того, что за это время ни один из а-генов в гамете не будет поврежден, равна, как это показано в работе [2], такому выражению:

$\beta^{\frac{4t}{T}}$. Вероятность того, что индивид, имеющий эту гамету, к моменту t не приобретет онкологическое заболевание, равна $1 - \Phi(4, t)$. Вероятность же того, что к этому же моменту индивид приобретет онкологическое заболевание, но затем выздоровеет, равна $\Phi(4, t) \cdot L$. Отсюда:

$$\psi(4) = \phi(4)P_{4 \rightarrow 4} = \phi(4) \int_0^{\infty} \beta^{\frac{4t}{T}} v(t) \cdot (1 - s(t)) \cdot (1 - \Phi(4, t)) \cdot (1 - L) dt, \quad (1)$$

где

$$\Phi(m, t) = \sum_{i=0}^4 (\phi(i)F((i + m), t)), \quad (2)$$

а

$$F(t) = 1 - \left(1 - \left(1 - \beta^{\frac{t}{T}} \right)^k \right)^N. \quad (3)$$

Пусть

$$f(m, t) = v(t) \cdot (1 - s(t)) \cdot (1 - \Phi(m, t)) \cdot (1 - L). \quad (4)$$

Тогда формулу (1) можно записать в следующем виде:

$$\psi(4) = \phi(4)P_{4 \rightarrow 4} = \phi(4) \int_0^{\infty} \beta^{\frac{4t}{T}} f(4, t) dt. \quad (5)$$

Теперь определим вероятность того, что к моменту репродукции гамета содержит 3 а-гена. Такое случайное событие может быть следствием либо процесса $4 \rightarrow 3$, либо процесса $3 \rightarrow 3$.

Процесс $3 \rightarrow 3$ описывается формулой, подобной формуле (5):

$$P_{3 \rightarrow 3} = \int_0^{\infty} \beta^{\frac{3t}{T}} f(3, t) dt. \quad (6)$$

Процесс $4 \rightarrow 3$ описывается следующей формулой:

$$P_{4 \rightarrow 3} = \int_0^{\infty} 4\beta^{\frac{3t}{T}} \left(1 - \beta^{\frac{t}{T}} \right) f(3, t) dt. \quad (7)$$

Таким образом, вероятность обнаружить к началу следующей репродукции гамету, содержащую 3 а-гена, равна:

$$\psi(3) = \phi(4)P_{4 \rightarrow 3} + \phi(3)P_{3 \rightarrow 3} = \int_0^{\infty} \beta^{\frac{3t}{T}} \left(\phi(3) + 4\phi(4) \left(1 - \beta^{\frac{t}{T}} \right) \right) f(3, t) dt. \quad (8)$$

Аналогично можно получить выражения также для $\psi(2)$, $\psi(1)$ и $\psi(0)$.

В частности, $\psi(2) = \phi(2)P_{2 \rightarrow 2} + \phi(3)P_{3 \rightarrow 2} + \phi(4)P_{4 \rightarrow 2}$.

В общем случае можно записать:

$$\psi(m) = \int_0^{\infty} \beta^{\frac{mt}{T}} \chi(m, t) f(m, t) dt,$$

где

$$\chi(m, t) = \left(\sum_{i=0}^{4-m} \left(\phi(4-i) \frac{(4-i)!}{m!(4-m-i)!} \left(1 - \beta^{\frac{t}{T}} \right)^{4-m-i} \right) \right).$$

Очевидно, что $\sum_{m=0}^4 \phi(m) = 1$. Также очевидно, что $\sum_{m=0}^4 \psi(m) < 1$, поскольку в последней сумме не содержатся вероятности событий, приводящих к элиминации гамет. Вместе с тем, выражение $\frac{\phi(m)}{\psi(m)}$ должно быть одинаковым для всех возможных значений m . Поэтому связь между величинами $\phi(m)$ и $\psi(m)$ имеет вид:

$$\phi(m) = \frac{\psi(m)}{\sum_{m=0}^4 \psi(m)}.$$

Если известны значения величин $s(t)$, $v(t)$, $\Phi(m, t)$ и величины L , то последняя формула представляет собой систему уравнений, позволяющую рассчитать численные значения величин $\phi(m)$, что в свою очередь позволяет рассчитать искомые значения

величин $p(k)$, которые и являются вероятностями, определяющими распределение людей по количеству а-генов в их генотипах. При этом используются следующие формулы:

$$\begin{aligned} p_0 &= \phi(0)^2; & p_1 &= 2\phi(0)\phi(1); & p_2 &= 2\phi(0)\phi(2) + \phi(1)^2; \\ p_3 &= 2(\phi(0)\phi(3) + \phi(1)\phi(2)); & p_4 &= 2(\phi(0)\phi(4) + \phi(1)\phi(3)) + \phi(2)^2; \\ p_5 &= 2(\phi(1)\phi(4) + \phi(2)\phi(3)); & p_6 &= 2\phi(2)\phi(4) + \phi(3)^2; \\ p_7 &= 2\phi(3)\phi(4); & p_8 &= \phi(4)^2, \end{aligned}$$

где p_i – вероятность наличия i генов в генотипе индивида при рождении, а $\phi(j)$ – вероятность наличия j генов в гамете этого индивида при рождении.

Подводя итоги, можно сделать вывод о том, что вышеприведенная математическая модель может быть использована для определения распределения а-генов в человеческой популяции. Однако определение численного значения вероятности обнаружить индивида с определенным значением числа а-генов в генотипе требует знания численных значений ряда параметров модели, в частности, α , T , $s(t)$, $v(\tau)$, L , N . И, если значения параметров T , $s(t)$, $v(\tau)$ и L можно с неплохой точностью оценить, то значения остальных параметров в настоящее время оценить трудно, и для их оценки необходимы дополнительные исследования.

СПИСОК ЛИТЕРАТУРЫ

1. Гродзинський Д.М. Радіобіологія / Д.М. Гродзинський. - К.: Либідь, 2000. – 448 с. / Grodzins'kij D.M. Radiobiologija / D.M. Grodzins'kij. - K.: Libid', 2000. – 448 s.
2. Книгавко В.Г. Математическое моделирование канцерогенеза / В.Г. Книгавко, Е.Б. Радзишевская, М.А. Бондаренко // Біофізичний вісник. – 2010, № 25(2). – С. 93 -100. / Knigavko V.G. Matematicheskoe modelirovanie kancerogeneza / V.G. Knigavko, E.B. Radzishevskaja, M.A. Bondarenko // Biofizichnij visnik. – 2010, N 25(2). – S. 93 -100.