

А. Р. Каюмов, Е. О. Михайлова, М. И. Богачев

СТАТИСТИЧЕСКИЙ АНАЛИЗ ГЕНОМОВ БИОИНДИКАТОРНЫХ ОРГАНИЗМОВ

Ключевые слова: статистический анализ, геномы, биоиндикаторы.

Многие химические вещества и лекарственные средства проходят обязательный контроль на токсичность, мутагенность и канцерогенность. Эти анализы проводят как на культурах клеток, так и на тестерных организмах, от бактерий и дрожжей до рыб и млекопитающих. Основным вопросом в использовании биоиндикаторных организмов является достоверность оценки безопасности вещества, полученной в лабораторных условиях. Данная работа посвящена статистическому анализу нуклеотидных последовательностей эукариотических геномов для оценки правомочности тестирования токсичных и канцерогенных свойств веществ на модельных организмах. Полученные результаты выявили неравномерное изменение размеров структурных фрагментов ДНК эукариот с эволюцией. Этот факт позволил сделать предположение, что результаты тестов, проведенных на клетках дрожжей и простейших, требуют проверки на организмах более высокого эволюционного развития.

Keywords: statistical analysis, genomes, bioindicators.

Many chemical compounds are compulsorily tested for toxicity, mutagenicity and cancerogenicity. These tests are carried out on cell cultures and special test organism, from bacteria and yeasts to fishes and mammals. The main question concerning bioindicators is the relevance of laboratory estimation of compound safety. Here we investigated the statistical properties of eukaryotic genomes to estimate the relevance of testing of substances toxicity and cancerogenicity using model organisms. Our results have revealed uneven changes of DNA structural fragments sizes in eukaryotes with the evolution. This fact allowed us to propose that the tests results obtained on the yeast cells and protozoa require checking on higher organisms.

Введение

В связи с бурным развитием химической промышленности ежегодно образуются тысячи новых химических соединений – красителей, лекарственных средств, консервантов, строительных материалов. Каждое новое вещество, с которым предполагается контакт человека, должно проходить тесты на токсичность и мутагенность – способность повреждать ДНК клеток и в дальнейшем вызывать образование опухолей. В настоящее время для этих целей используют бактерий рода *Salmonella*, культуры клеток растений и животных (например, клетки HELA), различных членистоногих (плодовая мушка *Drosophila melanogaster*, рачки *Daphnia*) и животных (лабораторные мыши *Mus musculus* и крысы *Rattus norvegicus*).

Основной вопрос в использовании различных биоиндикаторных организмов – насколько точно соотносятся полученные результаты с безопасностью для человека? Или другими словами, какова достоверность оценки безопасности вещества, полученной в лабораторных условиях. Данная работа посвящена статистическому анализу нуклеотидных последовательностей эукариотических геномов для оценки правомочности тестирования токсичных и канцерогенных свойств веществ на модельных организмах.

Результаты и обсуждение

С использованием открытой базы данных ГенБанка (<ftp.ncbi.nlm.nih.gov/genomes>) была сформирована база данных геномов организмов, применяемых в биоиндикации и расположенных в порядке усложнения по эволюционной лестнице (табл. 1). В таблице указан размер генома, средние размеры гена, межгенных участков, а также интрона и экзона. Эти фрагменты мы обозначили как «структурные фрагменты ДНК».

Таблица 1 - Размеры генома, средних размеров гена, экзонов, интронов и межгенных последовательностей различных организмов

Организм	Геном, Млн.п.о.	Средние значения тыс.п.о.			
		Ген	Меж генн.	Экзон	Инtron
<i>Saccharomyces cerevisiae</i>	12	1,5	2,62	0,16	0,21
<i>Arabidopsis thaliana</i>	115	1,9	6,91	0,16	0,16
<i>Caenorhabditis elegans</i>	100	2,8	7,26	0,22	0,31
<i>Drosophila melanogaster</i>	130	4,4	14,1	0,37	0,94
<i>Danio rerio</i>	1412	23,4	62,6	0,15	2,79
<i>Mus musculus</i>	3400	36,7	188	0,15	4,52
<i>Rattus norvegicus</i>	3360	34,3	198	0,15	4,45
<i>Macaca mulatta</i>	3140	42,8	189	0,15	5,43
<i>Pan troglodytes</i>	3700	46,2	203	0,16	5,72
<i>Pongo abelii</i>	3600	36,8	120	0,16	4,93
<i>Homo sapiens</i>	3500	48,3	218	0,15	5,50

Как видно из таблицы, с усложнением организма, значительно возрастает размер его генома, который у низших эукариот (дрожжей *Saccharomyces cerevisiae*) и человека отличается в 100 раз. При этом увеличение размера генома происходит за счет значительного повышения доли некодирующей ДНК: длина межгенных участков и интронов также возрастает в 100 раз, тогда как длина гена увеличивается 30