

О РОЛИ G/C ТРЕКОВ В СТРУКТУРНОЙ ОРГАНИЗАЦИИ ГЕНОМА ЧЕЛОВЕКА

Комаров В.М., Самченко А.А., Кондратьев М.С.

Институт биофизики клетки РАН ФИЦ ПНЦБИ РАН,
142290, Пущино, Институтская ул. д. 3;
Тел.: +7(4967)73-06-81; факс: +7(4967)33-05-09; e-mail: komarov@icb.psn.ru

Установление общих закономерностей появления повторяющихся и уникальных нуклеотидных последовательностей (треков) в структурно-функциональной организации геномов различных групп организмов остается важнейшей задачей современной геномики. Устоялось положение о ключевой роли GC пар (как пар с наибольшей термодинамической устойчивостью) в формировании термоадаптации двойной спирали ДНК, в определении размеров кодирующих последовательностей геномов живых организмов, а также в иницировании механизма спонтанных мутаций. Нами недавно, исходя из данных секвенирования геномов широкого ряда эукариот было выявлено, однако, общее доминирование встречаемости моно poly(dA)_n, poly(dT)_n и смешанных W(A/T) треков над подобными треками из GC-пар в структуре ДНК. Природа наблюдаемого явления приоритетности AT-пар на наш взгляд нетривиальна. Вполне вероятно, что ключевую роль здесь играет установленная нами ранее скрытая, разная размерность исходного полиморфизма геометрии водородного связывания уотсон-криковских AT- и GC-пар. Полиморфизм обусловлен наличием в их структуре экзоциклических амино групп пирамидального строения. Квантово-химические оценки показали, что AT-пары обладают двукратным, а GC-пары – четырехкратным вырождением геометрии своего комплементарного Н-спаривания. Именно пониженный структурный полиморфизм AT-пар по сравнению с 4-кратным полиморфизмом комплементарных GC-пар выступает, по-видимому, главным фактором «предпочтительного использования» природой аденин-тиминных пар для надежности протекания процессов сохранения и передачи генетической информации в ДНК-системах.

Остается, однако, главный вопрос: где все же и каким образом в таком случае реализуется геномный GC состав? В данной работе на примере генома человека детально проанализирована с использованием методов сравнительной геномики встречаемость A/T и G/C треков различной длины в структуре экзонов и интронов отдельных хромосом, а также в их негенных областях. Было получено, что при общем доминировании в хромосомах A/T треков в структуре их экзонов реализуется необычно высокая концентрация треков S(G/C), как правило длиной ~ до 30-35 пар оснований. По виду асимптотики доминирования этих треков все хромосомы могут быть подразделены на 3 типа: **I-тип** (хромосомы 1, 6, 7, 8, 9, 11, 15, 16, 17, 19, 20, 21, 22 с самым высоким GC составом в экзонах, от 51 до 58%); **II-тип** (хромосомы 2, 3, 4, 5, 10, 12, 13, 14, 18, X с GC-составом экзонах, от 47-51%); **III-тип** хромосом где в экзонах, однако, нет явного доминирования S треков (это Y-хромосома, но GC состав здесь ~ 50% и мтДНК, GC=45%). Сделан анализ связи полученной систематики хромосом со спецификой протекания в них процессов точечных спонтанных мутаций.