Beyond Sliding and Jumping: Mutation Clustering, Intermutation Spacing, and VAF Concordance Suggest Cross-Strand Switching as a Third Mode of APOBEC Movement on DNA

SMTB

Sabancı Üniversitesi



Cancer Genomics Lab

Наша цель состояла в выявлении кластеров APOBEC-ассоциированных мутаций в геномах BLCA и анализу их свойств, включая:

- вклад APOBEC-индуцированного мутагенеза и изучение размеров APOBEC-ассоциированных кластеров;
- связь между размерами кластеров и долей мутаций с APOBEC-подписью;
- геномный контекст формирования кластеров, включая время репликации, транскрипцию и доступность хроматина;
- предпочтения к определенному нуклеотидному контексту, в частности разнообразие нуклеотида в третьей позиции мотива TC;
- распределение APOBEC-ассоциированных мутаций между цепями ДНК внутри кластеров.

Our aim was to detect APOBEC-associated mutation clusters in BLCA genomes and characterize their properties, including:

- the contribution of APOBEC mutagenesis and the size distribution of APOBEC-associated clusters;
- the relationship between cluster size and the fraction of APOBEC-signature mutations;
- the genomic context of cluster formation, including replication timing, transcription, and chromatin accessibility;
- sequence-context preferences within clusters, particularly nucleotide diversity at the third position of the TC motif;
- the strand distribution of APOBEC-associated mutations within clusters.

RESULTS

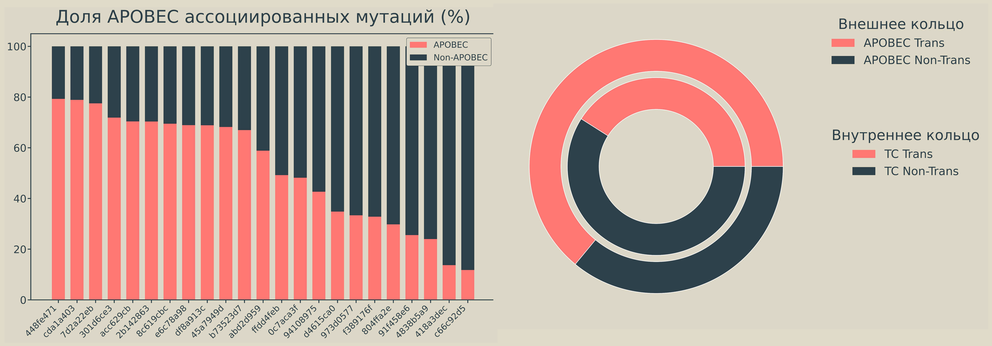
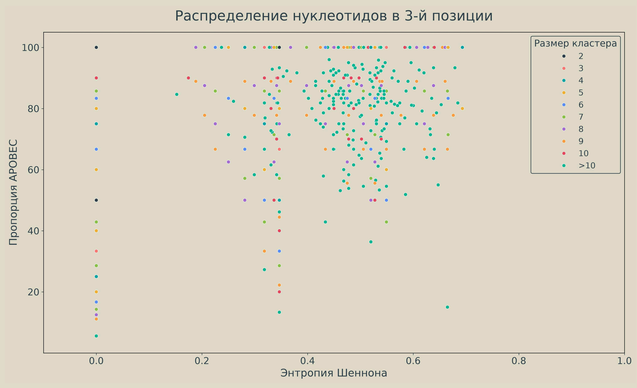


Рис. 4. (A) Доля APOBEC-ассоциированных мутаций в исследованных опухолевых образцах. Более чем в половине образцов APOBEC-ассоциированные мутации составляют большинство мутаций, что указывает на выраженный вклад APOBEC-индуцированного мутагенеза в формирование мутационного ландшафта этих опухолей. **(B)** APOBEC-ассоциированные мутационные кластеры обогащены в транскрибируемых областях по сравнению с распределением потенциальных APOBEC-мишеней (TC). Транскрибируемые области также преимущественно ассоциированы с открытым хроматином, что может способствовать доступности ДНК для APOBEC.

Рис. 5. Разнообразие нуклеотида в третьей позиции APOBEC-мотива TC в мутационных кластерах, оцененное с помощью энтропии Шеннона. Преобладание высоких значений энтропии указывает на разнообразие нуклеотидов, следующих за TC, и отсутствие выраженного предпочтения одного конкретного тринуклеотидного контекста.



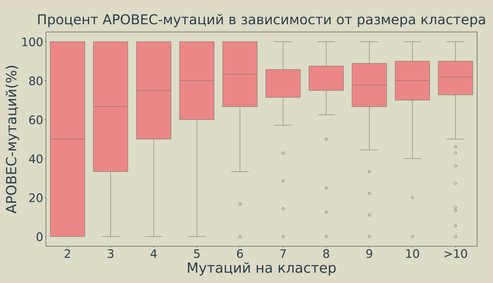


Рис. 6. Доля мутаций с APOBEC-подписью в зависимости от размера мутационного кластера. Доля APOBEC-мутаций увеличивается с ростом размера кластера.

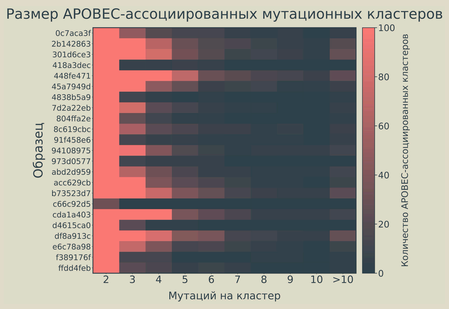
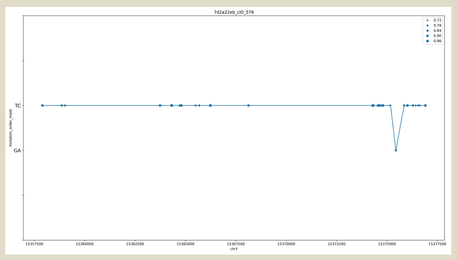
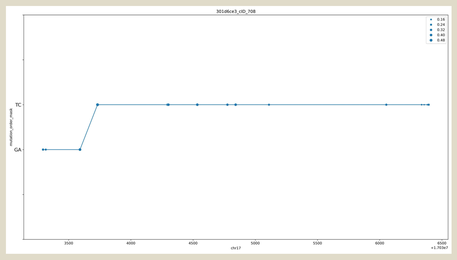


Рис. 7. Распределение APOBEC-индуцированных кластеров по размеру в отдельных образцах BLCA.





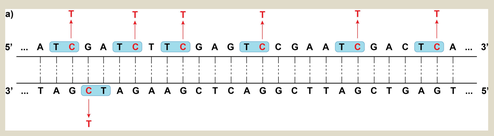
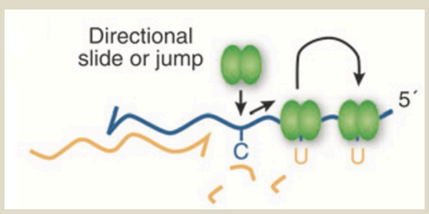
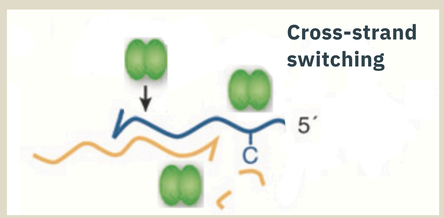


Рис. 11. Два часто наблюдаемых паттерна распределения APOBEC-ассоциированных мутаций между цепями ДНК. Первый паттерн (44%) согласуется с возможным переходом APOBEC на противоположную цепь с последующим возвратом на исходную, тогда как второй (56%) — с переходом на противоположную цепь и продолжением процессивного перемещения по ней. Схемы (a) и (b) представляют гипотетическую интерпретацию наблюдаемых мутационных паттернов.



*Chelico et al., Nat Struct Mol Biol 2006



*Model proposed in our study

BACKGROUND

Цитидиновые деаминазы APOBEC3A и APOBEC3B являются источником соматических мутаций в опухолях человека и преимущественно вызывают замены C→T и C→G в контексте TC.
APOBEC-индуцированные мутации часто формируют локальные кластеры (катаэгис, рис. 2), возникающие в результате эпизодов интенсивного мутагенеза на одноцепочечной ДНК (рис. 1).
При этом APOBEC-индуцированные кластеры значительно различаются по размеру, плотности мутаций и нуклеотидному составу, а причины этой гетерогенности остаются недостаточно изученными.

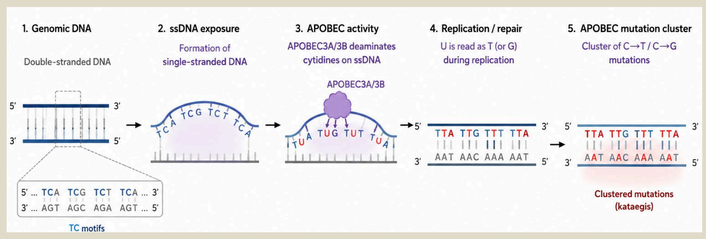


Figure 1. Schematic representation of APOBEC-induced clustered mutagenesis (kataegis) through cytidine deamination on exposed single-stranded DNA (ssDNA).

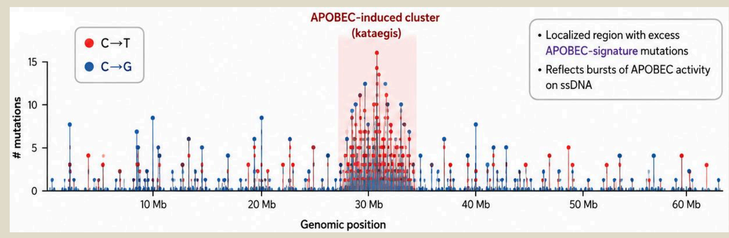


Figure 2. Example of an APOBEC-induced mutational cluster (kataegis), representing a localized burst of APOBEC mutagenesis along the genome.

APOBEC3A and APOBEC3B cytidine deaminases are important sources of somatic mutations in human cancers, preferentially inducing C→T and C→G mutations in TC contexts.
APOBEC-induced mutations frequently occur in spatially localized clusters (kataegis, Fig.2), reflecting episodes of intense mutagenesis on exposed single-stranded DNA (Fig.1).
However, APOBEC mutation clusters are highly heterogeneous in size, density, and sequence composition, and the mechanisms underlying this heterogeneity remain poorly understood.

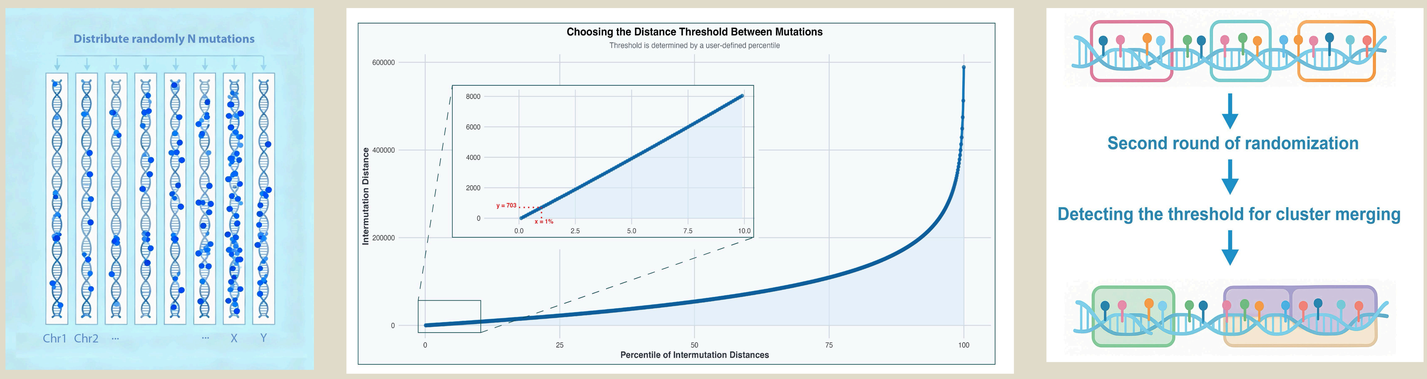


Figure 3. Overview of the GKclust workflow. (A) Chromosome-wise randomization of mutation positions. (B) Determination of the intermutation-distance threshold for initial cluster detection. (C) Second randomization step to determine the threshold for merging neighboring clusters and refine cluster boundaries.

Данные полногеномного секвенирования образцов рака мочевого пузыря (BLCA) были получены из проекта PCAWG. Для выявления мутационных кластеров использовали метод Gkclust (https://github.com/KazanovLab/GKclust), предназначенный для обнаружения кластеризации мутаций. На первом этапе рандомизации позиции мутаций случайным образом распределяются в пределах каждой хромосомы с сохранением общего числа мутаций, что позволяет определить статистически значимый порог расстояния между мутациями и выявить первичные кластеры. На втором этапе рандомизации анализируются расстояния между выявленными кластерами для определения порога их объединения, что позволяет объединять соседние кластеры, когда такое объединение статистически обосновано.

Whole-genome sequencing (WGS) data from bladder cancer (BLCA) samples were obtained from the PCAWG dataset. Mutational clusters were identified using GKclust (https://github.com/KazanovLab/GKclust), a randomization-based method for detecting non-random mutation clustering. In the first randomization step, mutation positions are randomized chromosome-wise while preserving mutation counts to determine a statistically significant intermutation-distance threshold and identify initial clusters. In the second randomization step, distances between the detected clusters are analyzed to determine a cluster-merging threshold, allowing statistically justified neighboring clusters to be merged.

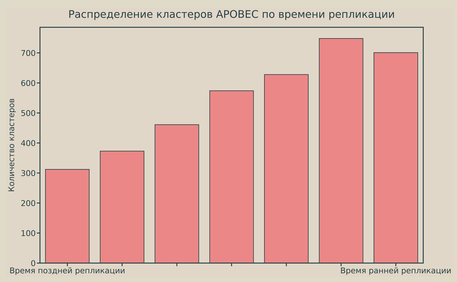


Рис. 8. Распределение APOBEC-ассоциированных мутационных кластеров по времени репликации. Кластеры преимущественно локализованы в областях ранней репликации, которые, как правило, характеризуются более открытым состоянием хроматина.

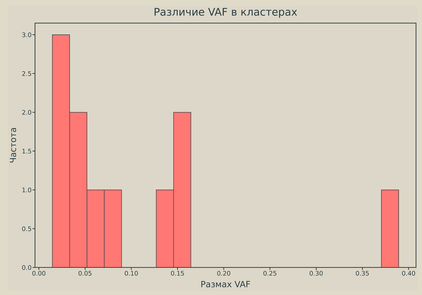


Рис. 9. Для кластеров, в которых APOBEC-ассоциированные мутации обнаружены на обеих цепях ДНК, было проанализировано распределение разницы между средними VAF мутаций на противоположных цепях. В большинстве случаев эта разница невелика, что согласуется с возникновением мутаций на обеих цепях в рамках одного и того же мутационного эпизода.

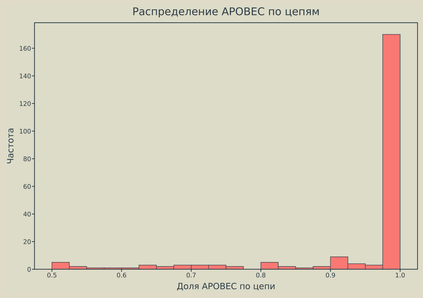


Рис. 10. Доля APOBEC-ассоциированных мутаций, локализованных на одной цепи ДНК, в крупных мутационных кластерах. В большинстве кластеров APOBEC-мутации преимущественно возникают на одной и той же цепи ДНК, однако примерно в 23% случаев мутации распределены между обеими цепями.

CONCLUSIONS

APOBEC-ассоциированные мутации в кластерах демонстрируют выраженную цепеспецифичность, что согласуется с процессивным перемещением APOBEC вдоль одноцепочечной ДНК.

Мы обнаружили, что в наших данных около 23% кластеров содержат мутации с APOBEC-подписью на обеих цепях ДНК, что указывает на возможные эпизоды перехода между цепями.

Сходные значения VAF мутаций на противоположных цепях позволяют предположить, что данные кластеры, как правило, возникают в рамках одного мутационного эпизода.

Два регулярно наблюдаемых паттерна согласуются с возможным переходом APOBEC на противоположную цепь с последующим возвращением на исходную либо продолжением перемещения по новой цепи.

В совокупности эти наблюдения позволяют предположить существование третьего возможного механизма перемещения APOBEC по ДНК — смены цепи, наряду со скольжением и прыжками.

APOBEC-associated mutations are strongly strand-coordinated within clusters, consistent with processive movement along single-stranded DNA.

We found that ~23% of clusters contain APOBEC-signature mutations on both DNA strands, indicating occasional cross-strand events.

Similar VAFs of mutations on opposite strands suggest that these mutations generally arise during the same mutational episode.

Two recurrent patterns are consistent with APOBEC switching to the opposite strand and either returning to the original strand or continuing along the new strand.

Together, these observations suggest cross-strand switching as a putative third mode of APOBEC movement on DNA, in addition to sliding and jumping.