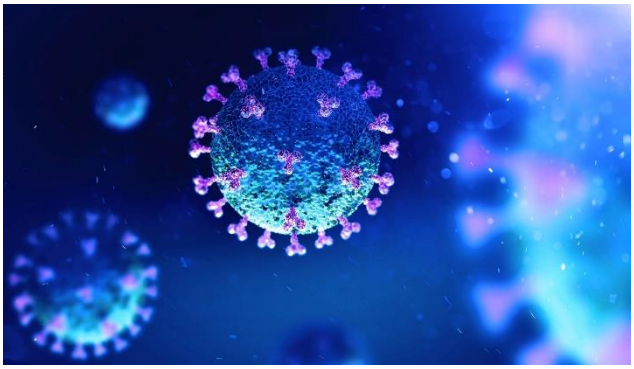
In Tracking SARS-CoV-2 Mutations New TARC-Seq Method Enhances Precision

Researchers developed a targeted accurate ribonucleic acid (RNA) consensus sequencing (tARC-seq) approach to precisely determine severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) mutation frequency and types in cell culture and clinical samples.



Introduction

SARS-CoV-2 replicates via RNA-dependent RNA polymerases (RdRp), which are prone to errors. Monitoring replication mistakes is critical to understanding the virus's development, but existing approaches are insufficient to identify infrequent de novo ribonucleic acid alterations.

During the coronavirus disease 2019 (COVID-19) pandemic, SARS-CoV-2 mutation rates ranged from 10−6 to 10−4 per base per cell. Exonuclease proofreading activity boosts mutation rates, leading to a mean of two mutations in each genome monthly.

Study

In the present study, researchers created tARC-seq to investigate the mechanisms of replication errors impacting the divergence of SARS-CoV-2.

The tARC-seq approach combines ARC-seq characteristics with hybrid capturing technology to enhance targets, allowing in-depth variant interrogation of these samples.

The researchers used tARC-seq to discover RNA variations in the original SARS-CoV-2 wild-type (WT) strain, SARS-CoV-2 Alpha and Omicron variants, and clinical and Omicron samples.

The researchers sequenced SARS-CoV-2 wild-type RNA following 4.0 infectious cycles, generating 9.0×10^5 plaque-forming units (pf.u.) of SARS-CoV-2 RNA. They added E. coli messenger RNA (mRNA) as an enzyme carrier to prepare libraries. Hybrid capture detected E. coli RNA in the genetic library, which the researchers examined individually and used as internal controls.

To further investigate selections in tARC sequencing data, the researchers mapped non-sense-type, synonymous, and non-synonymous variant frequencies identified by tARC sequencing across mon-structural protein 12 (nsp12), a critical gene that encodes SARS-CoV-2 RdRp.

They determined the evolutionary action (EA) scorings and variation frequencies for nonsense-type and non-synonymous single-nucleotide polymorphisms (SNPs) found in SARS-CoV-2 spike (S) and nsp12. They also computed the average mutational frequencies of open reading frames (ORFs) in the wild-type virus, broken down by mutational type and base alterations.

The researchers investigated the random distribution of RNA variants across the SARS-CoV-2 genome using location-based estimations and nucleotide identity analysis. They also used tARC-seq on two clinical samples to look for de novo mutations caused by spontaneous infection.

They matched the top ten most common C>TT and G>AA mutations to known A3A editing sites in the wild-type virus. The researchers examined all SID occurrences with ≥2 nucleotides of complementarity between donor and acceptor sites downstream in WT, Alpha, and Omicron. They investigated the genome-wide prevalence of TC>TT mutations in WT-Vero cells.

Findings and Conclusion

Researchers found 2.7 × 10−5 (mean) de novo mistakes per cycle in the SARS-CoV-2 virus, with C>T biases not primarily due to apolipoprotein B mRNA-editing enzyme, catalytic polypeptide (APOBEC) editing.

They identified cool and hot areas across the genome, according to low or high GC concentration, and highlighted transcription regulatory regions as sites more prone to mistakes. The tARC-seq approach enables the detection of template switches such as deletions, insertions, and complicated alterations.

The WT virus has 1.1 × 10−4 RNA variations per base, with base substitutions accounting for the majority (8.4 × 10−5), followed by insertions (2.5 × 10−6) and deletions (2.1 × 10−5). The G > A and C > T transitions dominate the viral mutation landscape, contributing 9.0% and 44% of all occurrences.

The mutational spectrum and frequency of wild-type SARS-CoV-2 off-target reads differ from those of E. coli, showing that these mutational events are genuine viral alterations rather than library preparation artifacts.

Random distributions and comparable rates of all three nsp12 mutation types suggest that most RNA variations found by tARC sequencing were de novo-type replication mistakes. The researchers found no differences in variant frequencies between the SNPs with low evolutionary action scores (estimated neutral effects) and those with high EA values (estimated harmful impacts) over the base substitution range, indicating that selection has a limited influence.

Variant rates vary considerably between locations, with 643 loci in WT viral duplicates showing considerably higher base substitution frequencies and 80 recurring throughout both WT replicates.

The researchers found no overlap between the highest-frequency tARC sequencing C>TT hotspots and A3A editing regions in the wild-type virus. The tARC sequencing C>TT frequencies at A3A editing regions were lower than the C>TT frequencies of the highest-frequency tARC sequencing C>TT hotspots by one to two orders of magnitude.

The study highlighted tARC-seq, a specialized sequencing approach, to investigate the replication mistakes that influence SARS-CoV-2 divergence. This approach selectively reads specific RNA molecules to generate a consensus sequence, allowing researchers to detect and evaluate minor differences and mistakes during viral replication.

It may also detect de novo insertions and deletions in SARS-CoV-2 resulting from cell culture infection, corroborating worldwide pandemic sequencing findings.

The study also discovered that SARS-CoV-2 possesses exonuclease proofreading capabilities, which may aid in understanding ExoN's critical function.

Source:
https://www.news-medical.net/news/20240424/New-tARC-seq-method-enhances-precision-in-tracking-SARS-CoV-2-mutations.aspx