XXXVI Congresso Brasileiro de Genética Médica — CBGM 2025 RECIFE — 24 a 27 de setembro de 2025

FIOCRUZ

Descriptive Analysis of Short Tandem Repeats in COVID-19 Patients

Matheus Azevedo Bomfim[1], Pablo Rafael Silveira Oliveira[2], Marcus Villander Barros de Oliveira Sá[3], Rodrigo Feliciano do Carmo[4], Túlio de Lima Campos[1], Luydson Richardson Silva Vasconcelos[1]

1- Instituto Aggeu Magalhães, Fiocruz Pernambuco, Fundação Oswaldo Cruz, Fiocruz, Recife, Pernambuco, Brazil.
2- Instituto de Biologia, Universidade Federal da Bahia (UFBA), Salvador, Bahia, Brazil.
3- Real Hospital Português, Recife, Pernambuco, Brazil.
4- Laboratório Avançado de Diagnóstico e Estudos em Saúde e Ambiente (LADESA), Colegiado de Medicina, Universidade Federal do Vale do São Francisco (UNIVASF), Petrolina, Pernambuco, Brazil.

Corresponding author: matheusazevedo98@gmail.com

INTRODUCTION

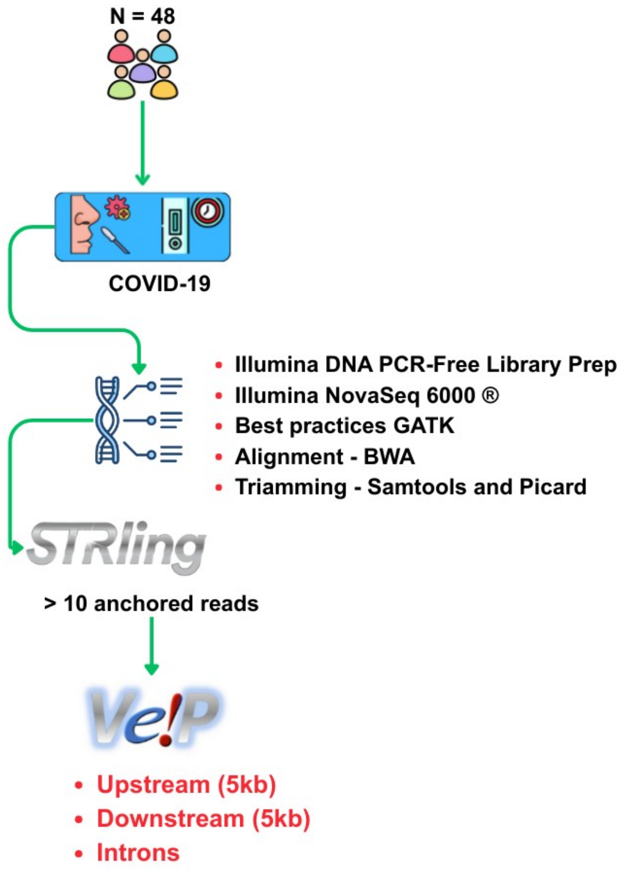
Short Tandem Repeats (STRs) are repetitive DNA sequences of 2–6 base pairs widely distributed throughout the human genome, involved in regulatory functions that influence susceptibility to viral infections, as well as immune responses. Recent studies during the COVID-19 pandemic, which has affected over 700 million people, have focused mainly on STRs in the viral genome, with only one examining host genome STRs' impact on disease response, revealing a key research gap.

OBJECTIVES

Here, we aimed to detect and describe the distribution of STRs across individuals with COVID-19.

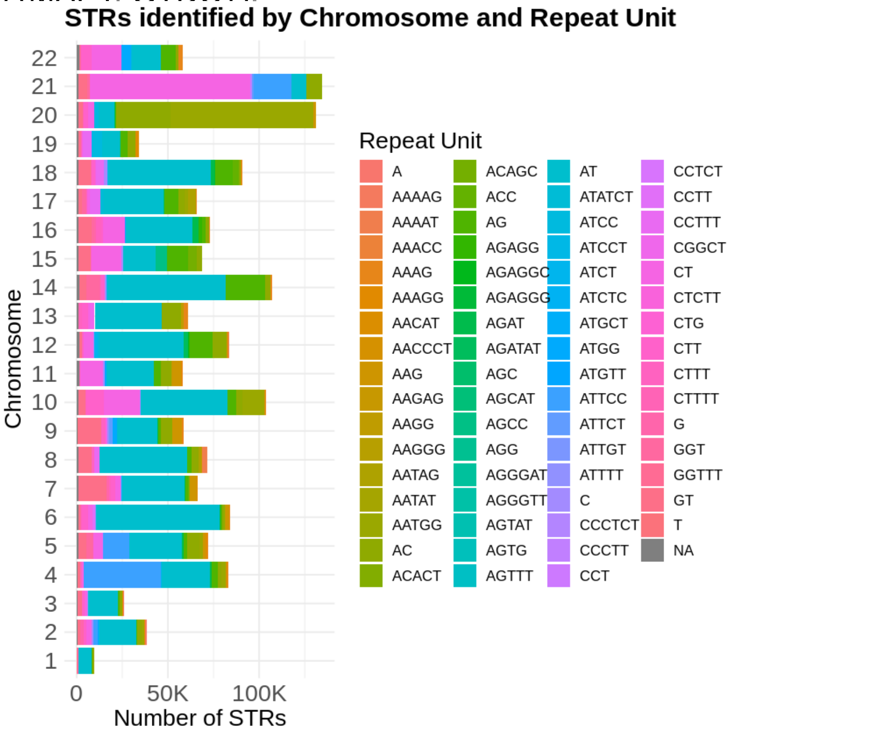
MATERIALS AND METHODS

This cross-sectional study included 48 patients without prior chronic diseases admitted to intensive care units in Brazil between 2020 and 2021. SARS-CoV-2 infection was confirmed by RT-qPCR or serological assays. Genomic DNA was extracted from whole blood, with sequencing libraries prepared following the Illumina DNA PCR-Free Library Prep protocol. Whole-genome sequencing was performed on an Illumina NovaSeq 6000® system in paired-end mode. Data processing followed GATK best practices, with trimming by Trimmomatic, alignment by BWA, and duplicate removal using samtools and Picard. STRs were identified using STRling and annotated with Variant Effect Predictor (VEP), including variations with at least ten anchored reads. Following descriptive analysis, the top five STRs with the highest incidence in regulatory regions (upstream, downstream and introns) were evaluated at NCBI.

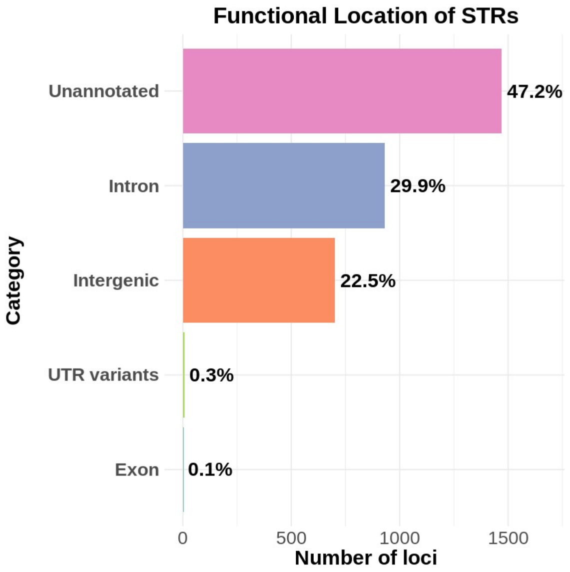


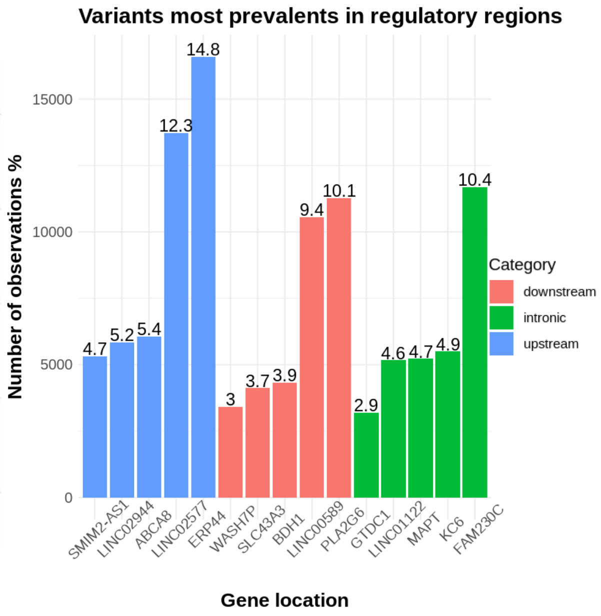
RESULTS AND DISCUSSION

A broad distribution of STRs was identified, and the most prevalent repeat motifs were AT (n = 46,632) with an average of 14.5 repeats units (RU), followed by ATTCC (n = 10,652; RU = 8) and CT (n = 8,932; RU = 21.6). VEP annotated 52.8% of the STRs, 29.9% mapped to intronic regions across 538 loci, 22.5% along intergenic regions (n = 94,653), 505 at 3′ or 5′ untranslated regions, and 152 in exonic regions. mong prevalent regulatory variants, eight are in non-coding RNA regions (FAM230C, KC6, LINC01122, WASH7P, FAM230C, SMIM2-AS1, LINC02944, LINC02577), one in pseudogene LINC02982, and six near protein-coding genes, linked to lipid metabolism (ABCA8, PLA2G6, BDH1, SLC43A3) and neuronal functions (MAPT, GTDC1).



The most prevalent alteration in downstream regions is associated with innate immune pathways, suggesting a potential role in the response to SARS-CoV-2. Given that STRs are known to influence neurological and oncological diseases, it is plausible that they may also modulate outcomes in infectious diseases. These findings highlight the importance of further research into their functional impact.





CONCLUSION

Given that, to advance our understanding of the role of STRs in disease, there is an urgent need for standardization in STR analysis and the development of specific tools for phenotypic annotation. Such efforts will be crucial to uncover the full extent of STRs functional significance, particularly in the context of infectious diseases like COVID-19.

AGRADECIMENTOS

UNIVASF — Universidade Federal do Vale do São Francisco · Real Hospital Português · UFBA · FACEPE — Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco · CNPq — Conselho Nacional de Desenvolvimento Científico e Tecnológico · CAPES