



Contribution ID: 25

Type: **Invited talk**

## Site frequency spectrum and genomic mutations analysis as a possible tool in predicting the emergence of a new SARS-CoV-2 subvariants

*Monday, 15 September 2025 09:30 (30 minutes)*

The rapid succession of SARS-CoV-2 variants has underscored the importance of tracing the emergence of new subvariants with evolutionary advantages. Based on almost 15 million complete viral sequences from GISAID we investigated how the individual mutations that define Variants of Concern have emerged over time. We found that rather than accumulating mutations one at a time, key changes appeared in clusters, leading to the accelerated emergence of mature variant lineages. The timing and combinatorial nature of the mutations that define each variant reveal strong but hidden selective forces at play.

This observation and analysis of the site frequency spectrum (SFS) of the viral genomes led us to retrospective discovery of a characteristic pattern of the SFS cumulative tails, which includes a discontinuity, that can be traced down to a set of mutations that maintain identity over a certain time interval. Furthermore, the discontinuity shifts over time toward higher frequencies. Ultimately, subvariants with this cluster of mutations dominate over the parental variant. This observation leads to the potential use of prospective SFS tail analysis to identify emerging new viral substrains. In addition, we present a mechanistic model, which allows quantitative description of the dynamics of this transient process.

**Primary author:** KURPAS, Monika (Silesian University of Technology, Department of Systems Biology and Engineering)

**Co-authors:** KOSTKA, Daria (Silesian University of Technology, Department of Systems Biology and Engineering); PŁONKA, Wiktoria (Silesian University of Technology, Department of Systems Biology and Engineering); JAKSIK, Roman (Silesian University of Technology, Department of Systems Biology and Engineering); KIMMEL, Marek (Rice University)

**Presenter:** KURPAS, Monika (Silesian University of Technology, Department of Systems Biology and Engineering)

**Session Classification:** Session 3: Computational Models of Disease, Development and Evolution