

## **Abstrakt**

Variant call format (VCF) je formát souborů používaný k reprezentaci a ukládání informací o variantách. Genetické varianty ve VCF mohou být reprezentovány více způsoby, protože specifikace VCF umožňuje nejednoznačnost, která může nastat kvůli různým variant call pipelineům nebo rozdílům v alignmentech sekvencí. Nejednoznačnosti narušují srovnávání souborů ve VCF a jejich variant, což vede ke komplikacím při další analýze variant.

Tato práce zkoumá rozdíly v reprezentaci genetických variant, které se mohou vyskytnout, a také jejich pravděpodobné příčiny a dopady na další analýzu. Dále je zkoumána normalizace souborů VCF a je uveden algoritmus pro atomizaci a deatomizaci souborů VCF.

**Klíčová slova:** VCF, variant call format, ambivalentní reprezentace variant, srovnání variantů, atomizace variantů, deatomizace variantů