

# Allelyser – tool for analysing genotype data in

E4013

Jan Böhm

RECETOX

6. března 2024

Diagnosis	SNP <sub>1</sub>	SNP <sub>2</sub>	SNP <sub>3</sub>
Disease A	AG	GG	AA
Disease B	GG	AA	GG
Control	AA	AG	AG
Disease C	GG	GG	AA
Control	AG	AA	GG
⋮	⋮	⋮	⋮

## Motivace

Chceme nástroj, který umožní provést rychlou sadu analýz pro tento typ dat. Tento nástroj by měl být easy-to-use pro biology a pomoci jim se správnou interpretací výsledků, které umožňuje exportovat ve formě vhodné pro další prezentaci.

## Oficiální zadání projektu

Vaším úkolem je vytvořit aplikaci v RShiny, která načte vstupní soubor (.xlsx) obsahující aspoň jednu proměnnou která popisuje skupiny a několik (až mnoho) SNP.

Nad těmito daty pak probíhají vhodné analýzy (HW rovnováha,  $\chi^2$  test, korelační analýza, porovnání šancí, LD graf. . .) na základě požadavků uživatele (volba SNP, volba skupin, filtrování, analýza více SNP současně. . .)

Výsledky obsahují komentář s interpretací a jsou snadno exportovatelné (png/pdf pro obrázky, excel pro tabulky).



## Biologie

- práce s genotypickými daty
- porozumění klíčovým pojmům z oblasti genetiky, např. genotyp, haplotyp, haplogenotyp, alela
- interpretace výsledků statistických nástrojů v biologickém kontextu

## Statistika

- analýza kategoriálních dat – kontingenční tabulky
- vhodné testy a vizualizace

## Programování

- tvorba aplikace pomocí RShiny
- efektivní prototypování s jazykovými modely
- tvorba UX pro nestatistiky