

# Funktionelle Genomanalysen 2023 (09-CRT-A006)

## Übung 2: GWAS & Sekundäranalysen

Dr. Janne Pott

09.-11. Juni 2023

### Allgemeine Hinweise:

- Die Aufgaben werden in der Übung gemeinsam bearbeitet.
- Zur Lösung von manchen Aufgaben wird ein Taschenrechner o.ä. benötigt.
- Am Ende des Moduls wird eine Musterlösung bereitgestellt.

## Aufgabe 1: Regressionsmodelle

Nehmen Sie an, die autosomalen biallelischen SNPs **rs123456** und **rs127890** und ein normalverteilter Phänotyp  $X$  wurden in vier unabhängigen Studien gemessen ( $n_1 = 1911$ ,  $n_2 = 3350$ ,  $n_3 = 1933$ ,  $n_4 = 4964$ ). Sie sind verantwortlich für Studie 1, und erforschen den Zusammenhang zwischen Gen  $ABC$  und  $X$ , wobei die SNPs in der Promoterregion des Gens liegen.

- Betrachten Sie die unten gezeigten Boxplots der SNP Genotypen und  $X$  aus Studie 1 (Abb. 1). Überlegen Sie sich geeignete Regressionsmodelle, mit dem Sie testen können, ob rs123456 oder rs127890 einen Effekt auf  $X$  haben. Welche Hypothese wird dabei getestet? Kann damit zwischen dominant und rezessiv unterschieden werden?
- Um die Power zu maximieren sollen die Daten der vier Studien in einer Meta-Analyse kombiniert werden. Welche Annahme wird hier häufig getroffen und wie kann diese geprüft werden?
- Nachdem alle Ihre Analysen abgeschlossen sind, meldet sich ein Kollege von Studie 2 bei Ihnen. Er teilt Ihnen mit, dass bei der Analyse leider vergessen wurde auf die Populationsstruktur zu korrigieren. Was bedeutet das und welche Konsequenzen hat das für Ihre Analyse? (Stichwort **Stratifikationsbias**)
- In Ihrer finalen Analyse erklärt der SNP 4% der Varianz von  $X$ . Die Gesamt-Heritabilität von  $X$  liegt jedoch laut Literatur bei 40%. Definieren Sie den Begriff Heritabilität und erklären Sie den Unterschied zwischen den Werten!

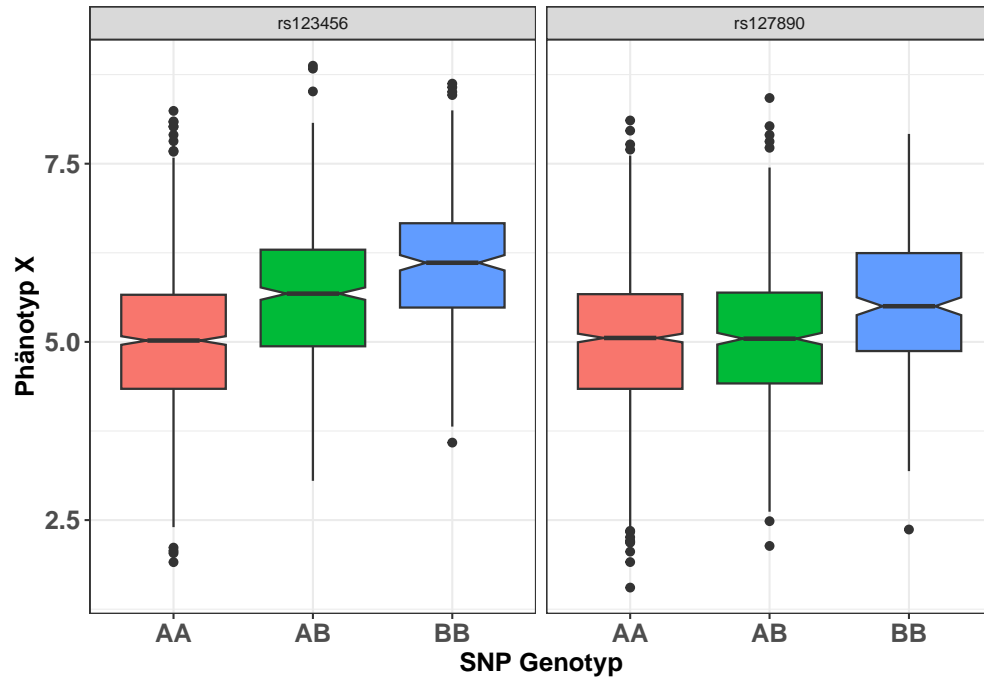


Figure 1: Boxplots per Genotypen pro SNP.

## Aufgabe 2: Colokalisierung und Mendelische Randomisierung

Ein Kollege forscht ebenfalls am Gen  $ABC$ , aber im Bezug auf die Krankheit  $Y$ , die ebenfalls mit dieser Region assoziiert ist. Sie möchten nun eine **Mendelische Randomisierung (MR)** planen um auf einen kausalen Effekt von  $X$  auf  $Y$  zu testen.

- Was ist die Idee der MR und welche drei Bedingungen müssen dafür gelten?
- Betrachten Sie die drei gerichteten azyklischen Graphen in Abbildung 2 (DAGs, *directed acyclic graphs*). Erläutern Sie das jeweilige Szenario!
- Ihr Kollege schlägt als Sensitivitätsanalyse der MR einen Test auf Colokalisierung vor. Was ist unter diesem Begriff zu verstehen und wie unterscheidet sich die Analyse von der MR?

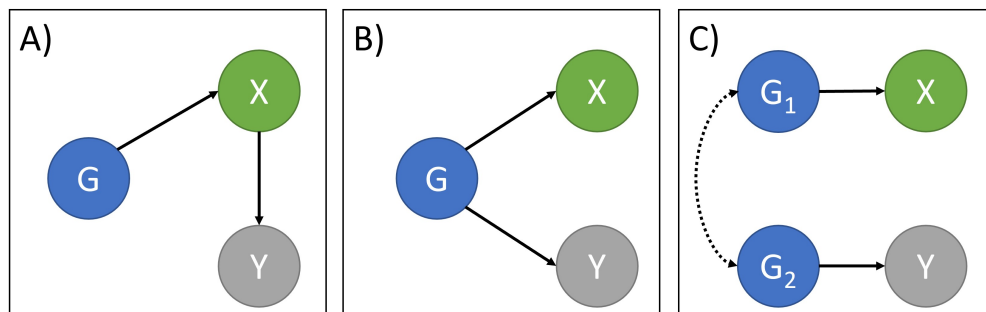


Figure 2: Drei DAGs