



ASHG 2014 HIGHLIGHTS

Frank Dudbridge

Gene-Environment Dependence Creates Spurious
Gene-Environment Interaction

Session: Statistical Methods for Multigene, Gene
Interaction and Pathway Analyses

Gene, Environment: Interaction and Dependence

- Typically we study markers thought to be in LD with causal variants (thus we are measuring causal variant with error)
- Relatively few gene-environment interactions reported to date
- Most research and simulation studies assume causal genetic variant and environmental exposure are independent
- If this is not true, a spurious Marker-by-Environment interaction may be observed

Gene-Environment Dependence Examples (Mediation)

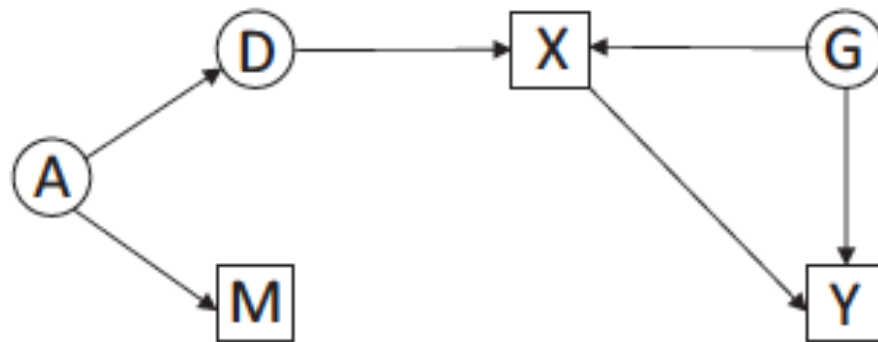


Figure 1. Directed Acyclic Graph Showing Gene-Environment Dependence by Mediation

Gene-Environment Dependence Examples (Pleiotropy)

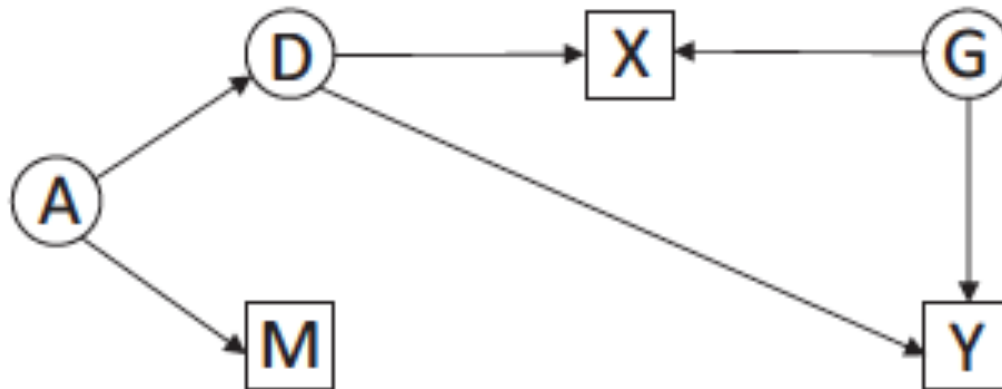


Figure 2. Directed Acyclic Graph Showing Gene-Environment Dependence by Pleiotropy

Gene-Environment Dependence Examples (Confounding)

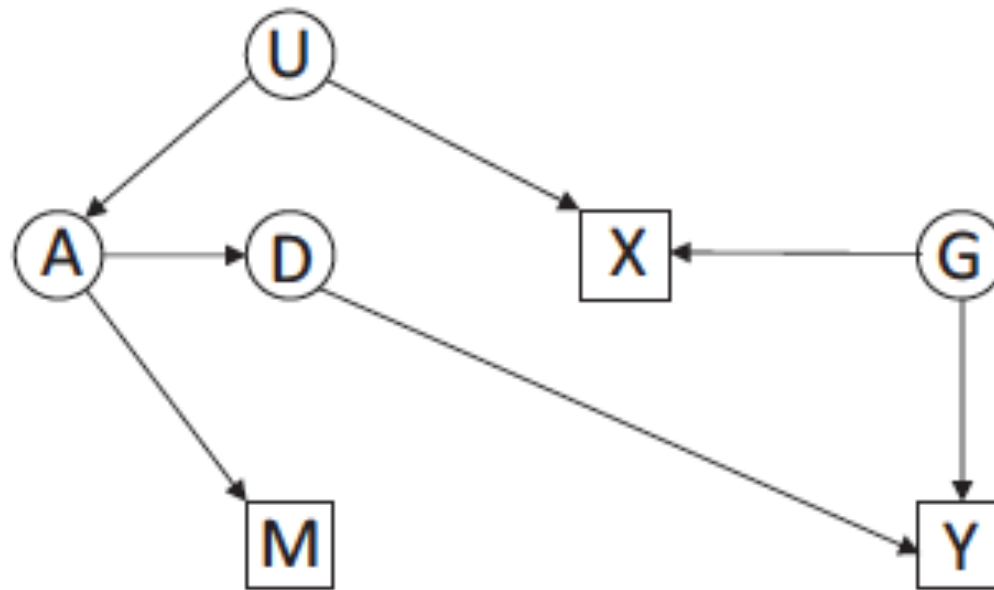


Figure 3. Directed Acyclic Graph Showing Gene-Environment Dependence by Confounding

Dudbridge, F. & Fletcher, O. Gene-environment dependence creates spurious gene-environment interaction. *American journal of human genetics* **95**, 301-307 (2014).

Lemma

If $\beta_{DX}=0$, then $\beta_{MX}=0$ if any of the following conditions hold

- (1) there is no main effect of the causal variant on the outcome, $\beta_D=0$
- (2) the marker is perfectly correlated with the causal variant, $D=M$
- (3) the causal variant is independent of the marker, conditional on the exposure, $\Pr(D|M,X) = \Pr(D|X)$

Furthermore, under linear or log-linear regression, $\beta_{MX}=0$ if

- (4) the causal variant is independent of the exposure, conditional on the marker, $\Pr(D|M,X) = \Pr(D|M)$

Gene x Exposure

- (In presence of causal variant/exposure dependence) Spurious interaction is due to imperfect LD between marker and causal variant
 - Dichotomous outcome (Y):
 - Misclassification error between cases and controls (due main effect of causal variant on outcome)
 - This misclassification is further modified (differently) for exposed and unexposed subjects (due to the causal variant-exposure association)
 - RR of marker may differ between exposed and unexposed samples (i.e interaction)
- Spurious ix Cannot be removed by transformation of variables
 - A direct result of measurement error of the causal variant
- Magnitude of bias depends on the causal variant –exposure association
 - Cannot be assessed from marker data (marker may not appear associated with exposure at all)

Gene x Gene

- GxG interaction has similar fate
 - Two causal variants in LD and at least one measured with error (i.e. a marker)
- A.R. Wood (*Analysis of variants obtained through whole-genome sequencing provides an alternative explanation to apparent epistasis*)
 - Numerous *cis* interactions can be explained by a single variant in LD with both of the interacting markers (here the two causal variants are actually one and the same thus not an interaction).