

Systems Genetics

Spring 2025

Lecture 12

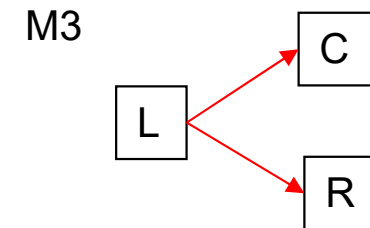
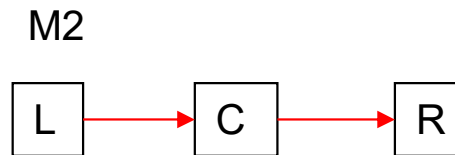
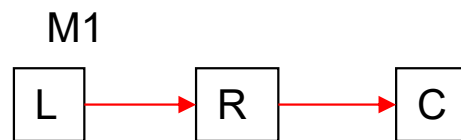
Inferring causal relations
(mediation test)

Irit Gat-Viks

Department of cell research and
Immunology

Tel-Aviv University

Inferring causal relations: mediation test



L - locus

R - gene

C – complex trait

Exploiting Structure in Probability Distributions:

Based on presentation and lecture notes of
Nir Friedman, Hebrew University

Basic Probability Definitions

- **Product Rule:** $P(A,B)=P(A | B)*P(B)= P(B | A)*P(A)$
- **Independence** between A and B: $P(A,B)=P(A)*P(B)$,
or alternatively: $P(A|B)=P(A)$, $P(B|A)=P(B)$.
- **Total probability theorem:** $\bigcup_{i=1}^n B_i = \Omega, \quad \forall i \neq j \quad B_i \cap B_j = \phi$

$$P(A) = \sum_{i=1}^n P(A, B_i) = \sum_{i=1}^n P(B_i) * P(A | B_i)$$

Basic Probability Definitions

□ Bayes Rule:

$$P(A | B) = \frac{P(B | A) \cdot P(A)}{P(B)}$$

$$P(A | B, C) = \frac{P(B | A, C) \cdot P(A | C)}{P(B | C)}$$

□ Chain Rule:

$$P(X_1, \dots, X_n) =$$

$$P(X_1 | X_2, \dots, X_n) \cdot P(X_2 | X_3, \dots, X_n) \cdot P(X_3 | X_4, \dots, X_n) \cdot \dots \cdot P(X_{n-1} | X_n) \cdot P(X_n)$$

Exploiting Independence Property

- G: whether the woman is pregnant
- D: whether the doctor's test is positive

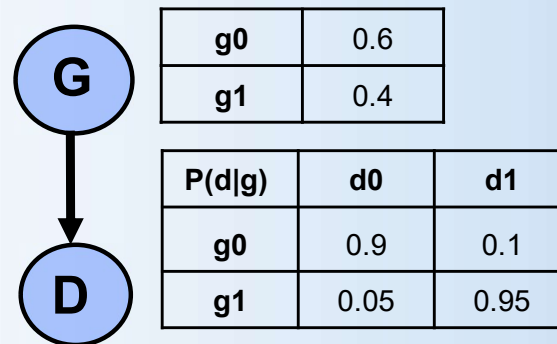
The joint distribution representation $P(g,d)$:

G	D	$P(G,D)$
0	0	0.54
0	1	0.06
1	0	0.02
1	1	0.38

Factorial representation

Using conditional probability: $P(g,d)=P(g)*P(d|g)$.

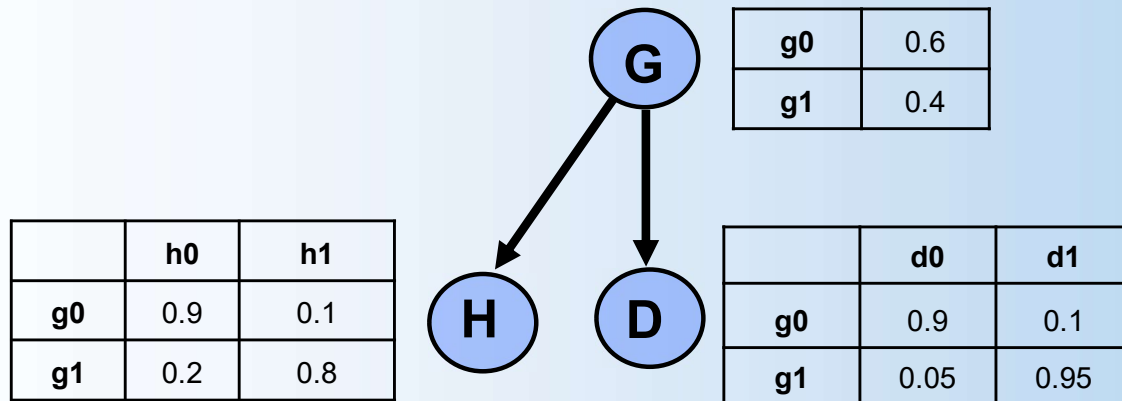
The distribution of $P(g)$, $P(d|g)$:



Example: $P(g0,d1)=0.06$ vs. $P(g0)*P(d1|g0)=0.6*0.1=0.06$

Exploiting Independence Property

- H: home test
- Independence assumption: $\text{Ind}(H;D|G)$ (i.e., given G, H is independent of D).



Joint
distribution

Factorial
representation

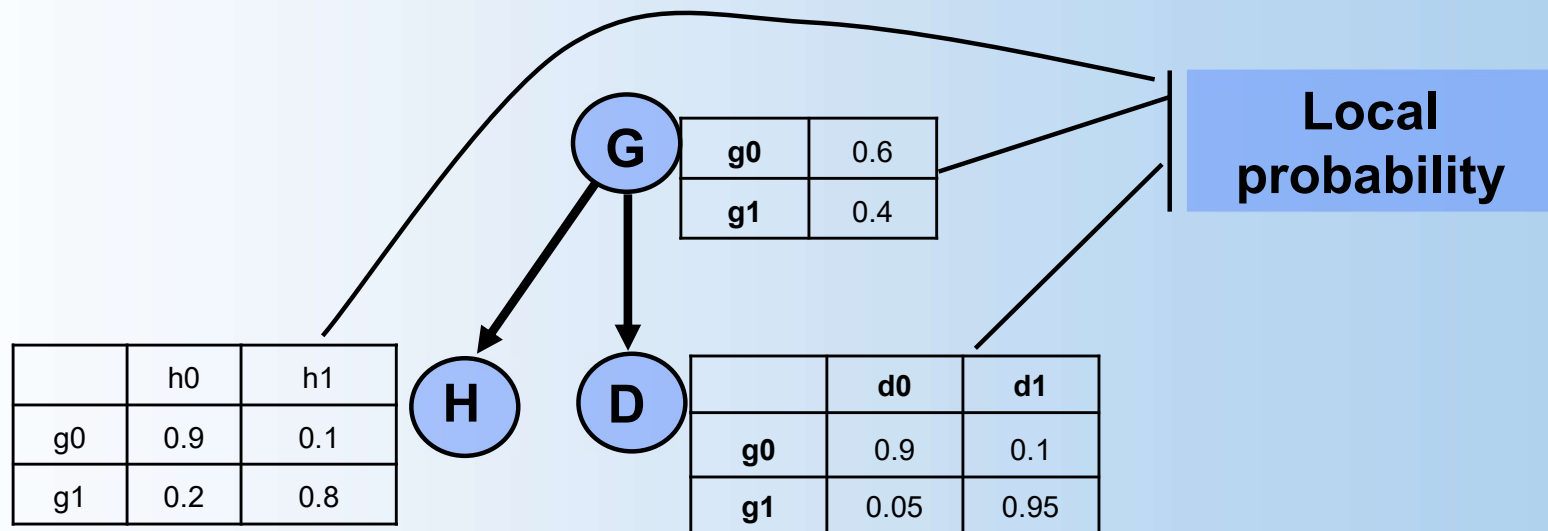
$$P(d,h,g)=P(d,h|g)*P(g)=P(d|g)* P(h|g)*P(g)$$

Product rule

$\text{Ind}(H;D|G)$

Exploiting Independence Property

	representation of $P(d,g,h)$	
	joint distribution	factored distribution
No. of parameters	7	5
Adding new variable H	changing the distribution entirely	Modularity: reuse the local probability model. (Only new local probability model for H.)

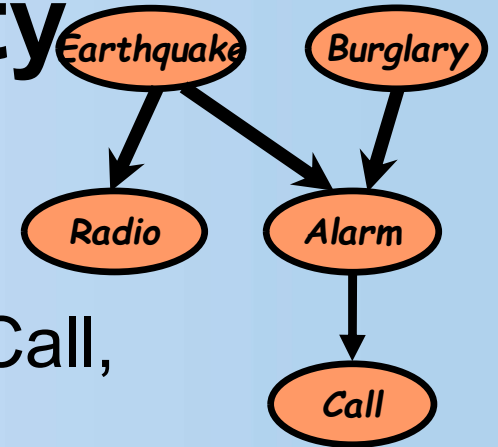


=> **Bayesian networks:** Exploiting independence properties of the distribution in order to allow a compact and natural representation.

Outline

- Introduction
- Bayesian Networks
 - » Representation & Semantics
 - Inference in Bayesian networks
 - Learning Bayesian networks
 - Conclusion

Representing the Uncertainty



- A story with five random variables:
 - Burglary, Earthquake, Alarm, Neighbor Call, Radio Announcement
 - Specify joint distribution with $2^5=32$ parameters

maybe...

- An expert system for monitoring intensive care patients
 - Specify joint distribution over 37 variables with (at least) 2^{37} parameters

no way!!!

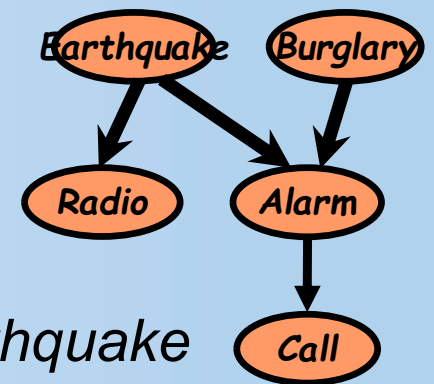
Probabilistic Independence: a Key for Representation and Reasoning

- Recall that if X and Y are **independent** given Z then

$$P(X | Z, Y) = P(X | Y)$$

- In our story...if

- *burglary* and *earthquake* are **independent**
- *alarm sound* and *radio* are **independent** given *earthquake*



- then instead of 15 parameters we need 8

$$P(A, R, E, B) = P(A | R, E, B) \cdot P(R | E, B) \cdot P(E | B) \cdot P(B)$$

versus

$$P(A, R, E, B) = P(A | E, B) \cdot P(R | E) \cdot P(E) \cdot P(B)$$

Need a language to represent independence statements

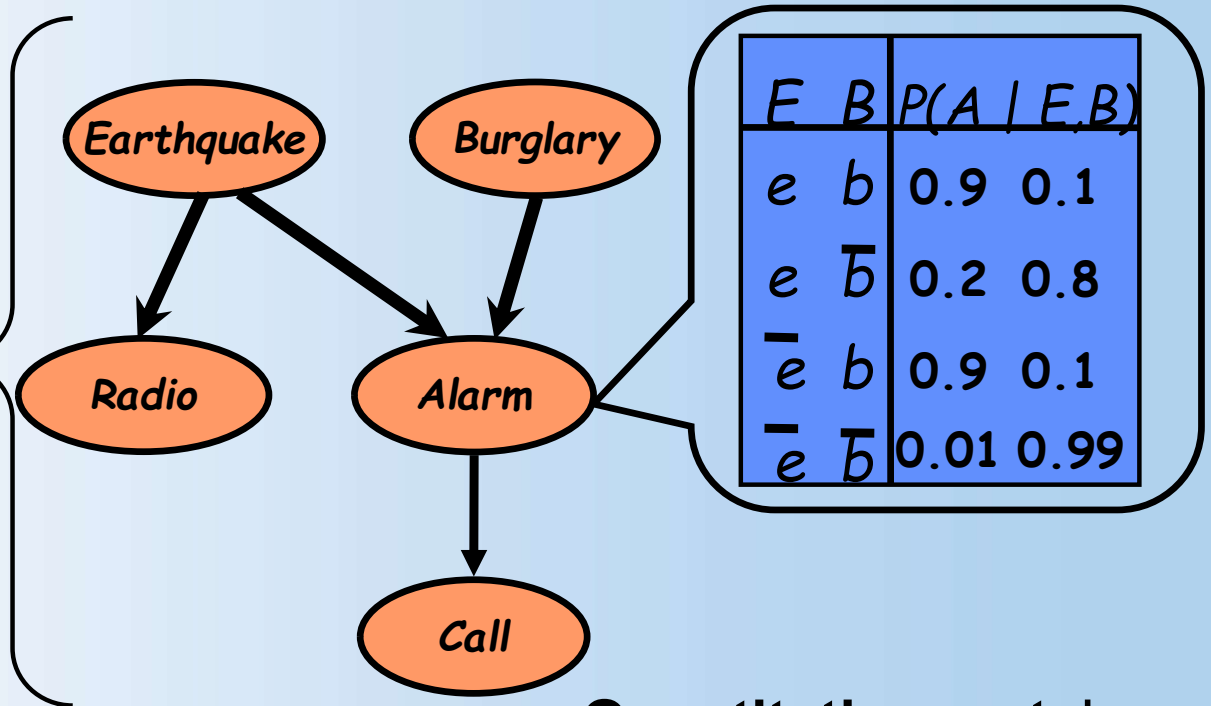
Bayesian networks

Efficient representation of probability distributions via conditional independence

Qualitative part: statistical independence statements

Directed acyclic graph (DAG)

- Nodes - random variables of interest (exhaustive and mutually exclusive states)
- Edges - direct influence



- **Quantitative part:** Local probability models. Set of conditional probability distributions.

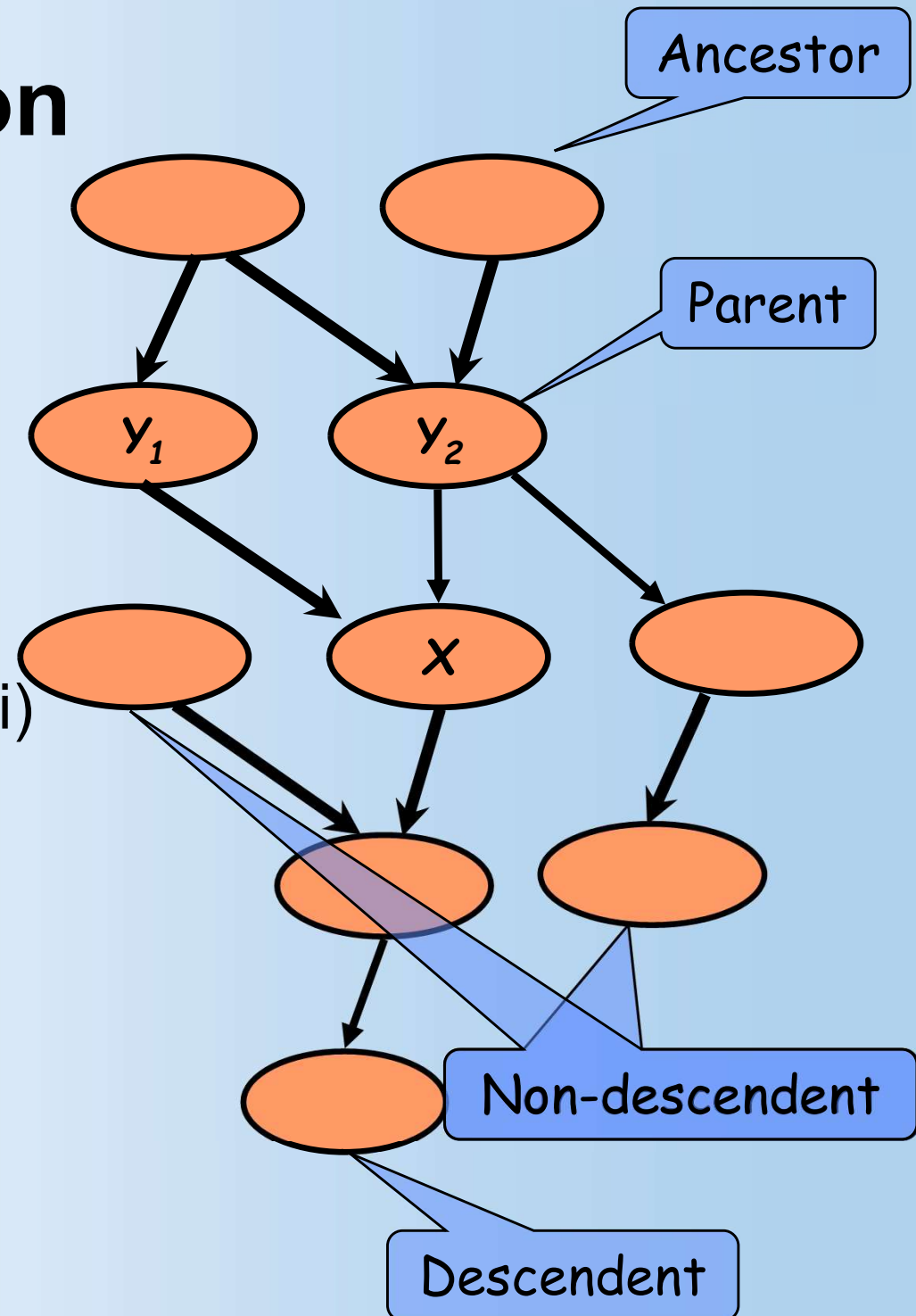
Markov Assumption

Generalizing:

- A child is **conditionally independent** from its non-descendants, given the value of its parents.

$\text{Ind}(X_i ; \text{NonDescendant}_i \mid \text{Pa}_i)$

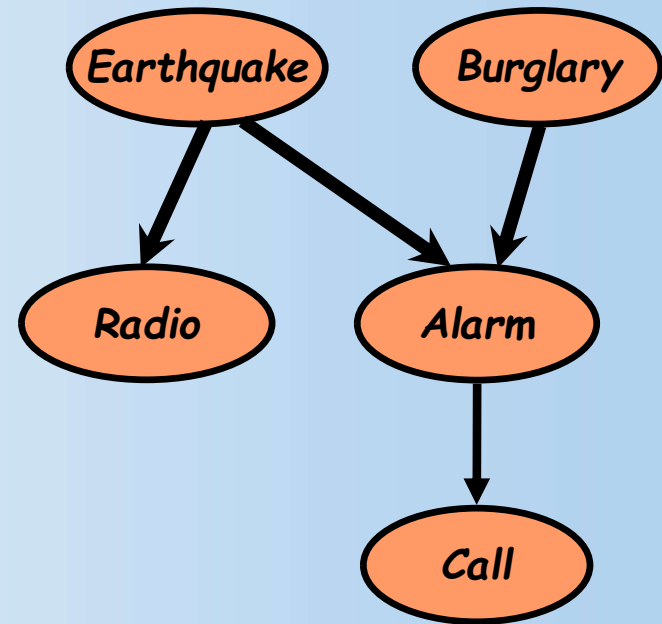
- It is a natural assumption for many **causal** processes



Markov Assumption (cont.)

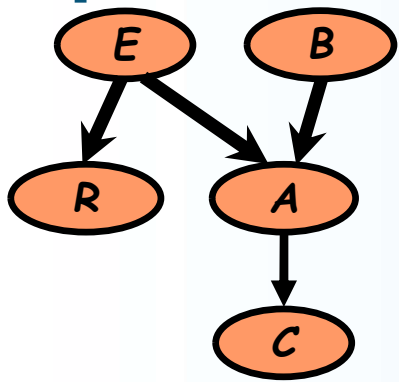
□ In this example:

- R is independent of A, B, C , given E
- A is independent of R , given B and E
- C is independent of B, E, R , given A



- Are other independencies implied by these ones?
- A graph-theoretical criteria identifies all such independencies

Bayesian Network Semantics



Qualitative part
conditional
independence
statements
in BN structure

Quantitative part
local
probability
Models
+ (e.g., multinomial,
linear Gaussian) =

Unique joint
distribution
over domain

□ Compact & efficient representation:

- nodes have $\leq k$ parents $\Rightarrow O(2^k n)$ vs. $O(2^n)$ params
- parameters pertain to local interactions

$$P(C, A, R, E, B) = P(B) * P(E|B) * P(R|E, B) * P(A|R, B, E) * P(C|A, R, B, E)$$

versus

$$P(C, A, R, E, B) = P(B) * P(E) * P(R|E) * P(A|B, E) * P(C|A)$$

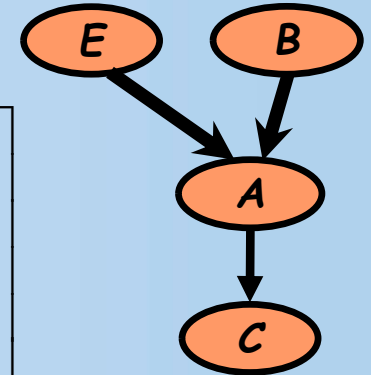
→ In general:
$$P(x_1, \dots, x_n) = \prod_{i=1, \dots, n} P(x_i | Pa_{x_i})$$

MLE principle for Bayesian networks

- Training data has the form:

$$D = \begin{bmatrix} E[1] & B[1] & A[1] & C[1] \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ E[M] & B[M] & A[M] & C[M] \end{bmatrix}$$

- Assume i.i.d. samples

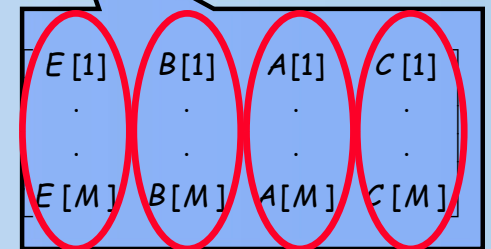
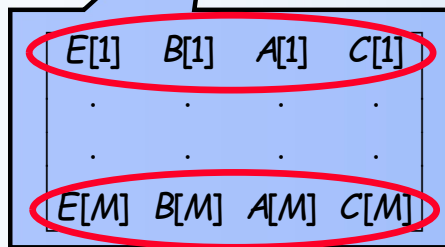


$$L(\Theta : D) = \prod_m P(E[m], B[m], A[m], C[m] : \Theta)$$

$$= \prod_m \begin{pmatrix} P(E[m] : \Theta) \\ P(B[m] : \Theta) \\ P(A[m] | B[m], E[m] : \Theta) \\ P(C[m] | A[m] : \Theta) \end{pmatrix}$$

By
definition
of network

$$= \prod_m P(E[m] : \Theta) \prod_m P(B[m] : \Theta) \prod_m P(A[m] | B[m], E[m] : \Theta) \prod_m P(C[m] | A[m] : \Theta)$$



MLE principle for Bayesian networks

- Generalizing for any Bayesian network:

$$L(\Theta : D) = \prod_i \prod_m P(x_i[m] | Pa_i[m] : \Theta_i) = \prod_i L_i(\Theta_i : D)$$

$$\begin{aligned} L_i(\theta_i : D) &= \prod_m P(x_i[m] | Pa_i[m] : \theta_i) = \prod_{pa_i} \prod_{m: Pa_i[m]=pa_i} P(x_i[m] | pa_i[m]; \Theta_{x_i|Pa_i}) \\ &= \prod_{pa_i} \prod_{x_i} P(x_i | pa_i : \theta_i)^{N(x_i, pa_i)} = \prod_{pa_i} \prod_{x_i} \theta_{x_i|pa_i}^{N(x_i, pa_i)} \end{aligned}$$

- The likelihood decomposes according to the network structure.
- **Decomposition \Rightarrow Independent estimation problems**
(If the parameters for each family are not related)
- For each value pa_i of the parent of X_i we get independent multinomial problem.

- The **MLE** is $\hat{\theta}_{x_i|pa_i} = \frac{N(x_i, pa_i)}{N(pa_i)}$

Continuous (Gaussian) variables



$$L(\Theta : D) = \prod_i \prod_m P(x_i[m] | Pa_i[m] : \Theta_i) = \prod_i L_i(\Theta_i : D)$$

$$L_i(\theta_i : D) = \prod_m P(x_i[m] | Pa_i[m] : \theta_i)$$

$$X_i | Pa_i \sim N(\mu_{i, Pa_i}, \sigma_{i, Pa_i}^2)$$

- The likelihood decomposes according to the network structure.
- **Decomposition \Rightarrow Independent estimation problems**
(If the parameters for each family are not related)
- For each value pa_i of the parent of X_i we get independent maximization problem.

- The **MLE** is

$$\hat{\mu}_{i, Pa_i} = \frac{\sum_{m: Pa_i[m] = pa_i} x_i[m]}{N(pa_i)}$$
$$\hat{\sigma}_{i, Pa_i}^2 = \frac{\sum_{m: Pa_i[m] = pa_i} (x_i[m] - \hat{\mu}_{i, Pa_i})^2}{N(pa_i)}$$

Continuous (Gaussian) variables

$$X \sim N(\mu, \sigma_X^2) \quad \text{X} \rightarrow \text{Y} \quad Y | X \sim N(ax + b, \sigma^2)$$

$$Pa_i \sim N(\mu_{Pa_i}, \sigma_{Pa_i}^2) \quad Pa_i \rightarrow X_i \quad X_i | Pa_i \sim N(a \cdot pa_i + b, \sigma^2)$$

$$L(\Theta : D) = \prod_i \prod_m P(x_i[m] | Pa_i[m] : \Theta_i) = \prod_i L_i(\Theta_i : D)$$

$$L_i(\theta_i : D) = \prod_m P(x_i[m] | Pa_i[m] : \theta_i)$$

- The likelihood decomposes \Rightarrow **Independent estimation problems**
The **MLE** is

$$X_i | Pa_i \sim N(a \cdot pa_i + b, \sigma^2)$$

Statistical background - regression

Assume $Y = aX + b + Z$, where $Z \sim N(0, \sigma_z^2)$, $\text{Ind}(X, Z)$.

σ_x^2, μ_x are population variance and mean of X .

Thus :

1. $\mu_y = a\mu_x + b$

2. $Y | X \sim N(aX + b, \sigma_z^2)$.

Using least - squares estimation of a and b :

$$\hat{a}, \hat{b} = \arg \min \sum_i (y_i - (ax_i + b))^2$$

Solving max likelihood estimation of $Y | X$:

$$\hat{a}, \hat{b} = \arg \max \log L(X | Y | a, b)$$

$$= \arg \max \frac{n}{2} \ln \frac{1}{2\pi\sigma^2} - \sum_i (y_i - (ax_i + b))^2$$

$$= \arg \min \sum_i (y_i - (ax_i + b))^2$$

Statistical background - regression

$$\hat{a}, \hat{b} = \arg \min \sum_i (y_i - (ax_i + b))^2$$

$$\hat{a} = \frac{\text{cov}(X, Y)}{\sigma_x^2} = \frac{\rho_{xy} \sigma_x \sigma_y}{\sigma_x^2} = \frac{\rho_{xy} \sigma_y}{\sigma_x}$$

$$\hat{b} = E(Y) - \hat{a} E(X) = \mu_y - \hat{a} \mu_x$$

Express $Y | X$ using population parameters $\mu_y, \mu_x, \sigma_y^2, \sigma_x^2, \rho_{xy}$

$$E(Y | X) = \hat{a} X + \hat{b} = \hat{a} X + \mu_y - \hat{a} \mu_x = \mu_y + \hat{a} (X - \mu_x) = \mu_y + \frac{\rho_{xy} \sigma_y}{\sigma_x} (X - \mu_x)$$

$$\text{Var}(Y | X) = \dots = \sigma_y^2 (1 - \rho_{xy}^2)$$

Continuous (Gaussian) variables

$$X \sim N(\mu, \sigma_X^2) \quad \text{X} \rightarrow \text{Y} \quad Y | X \sim N(ax + b, \sigma^2)$$

$$E(Y | X) = \mu_y + \frac{\rho_{xy} \sigma_y}{\sigma_x} (X - \mu_x)$$

$$\text{Var}(Y | X) = \sigma_y^2 (1 - \rho_{xy}^2)$$

$$Pa_i \sim N(\mu_{Pa_i}, \sigma_{Pa_i}^2) \quad \text{Pa}_i \rightarrow \text{X}_i \quad X_i | Pa_i \sim N(a \cdot pa_i + b, \sigma^2)$$

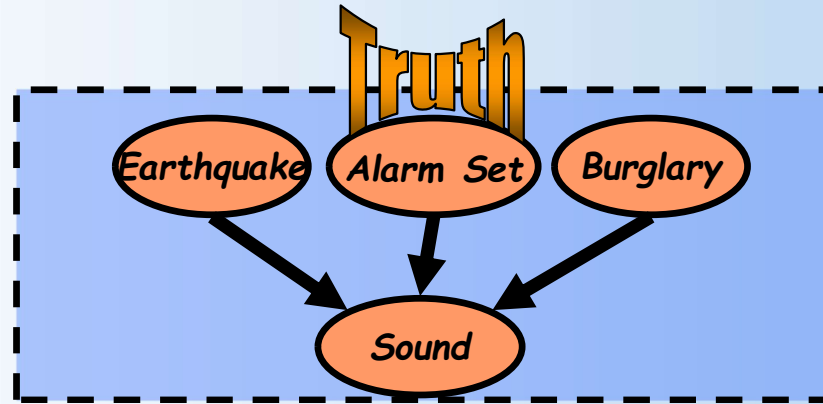
$$E(X_i | Pa_i) = \mu_{X_i} + \frac{\rho_{Pa_i X_i} \sigma_{X_i}}{\sigma_{Pa_i}} (Pa_i - \mu_{Pa_i})$$

$$\text{Var}(X_i | Pa_i) = \sigma_{X_i}^2 (1 - \rho_{Pa_i X_i}^2)$$

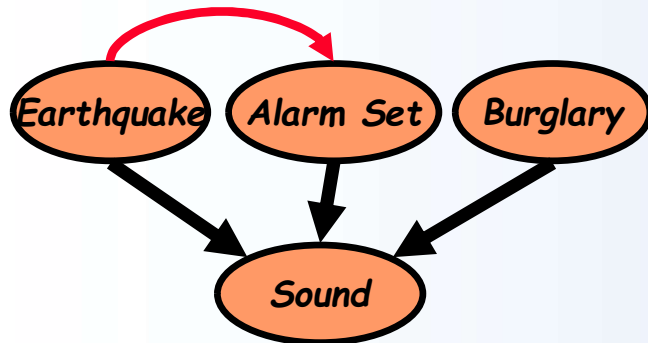
Outline

- Introduction
- Bayesian Networks
 - Representation & Semantics
 - Inference in Bayesian networks
 - Learning Bayesian networks
 - Parameter Learning
 - » Structure Learning

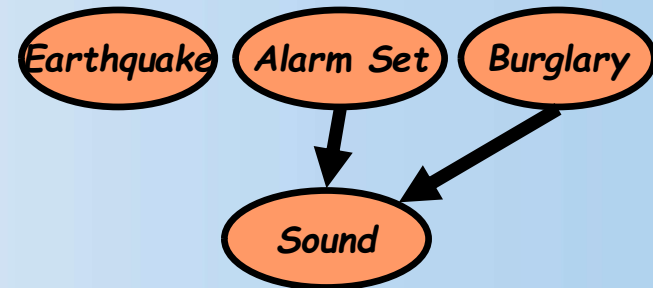
Learning Structure: Motivation



Adding an arc



Missing an arc



Optimization Problem

Input:

- Training data
- Scoring function (including priors)
- Set of possible structures

Output:

- A network (or networks) that maximize the score

Key Property:

- **Decomposability:** the score of a network is a sum of terms.

Scores

For example. The BDE score:

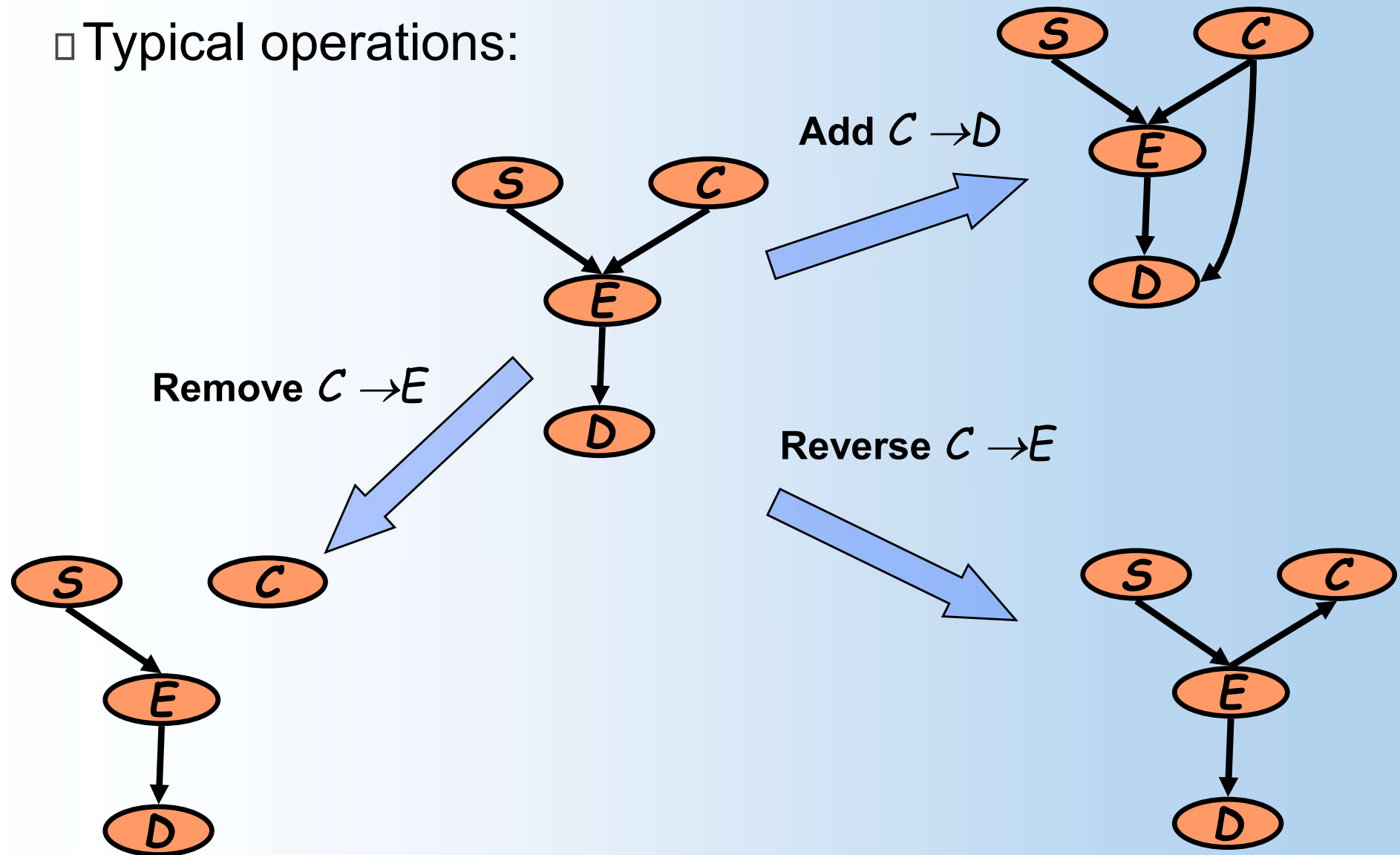
$$\begin{aligned} \text{Score}(G : D) &= P(G \mid D) \propto P(D \mid G)P(G) \\ &= \int P(D \mid G, \theta)P(\theta \mid G)d\theta P(G) \end{aligned}$$

When the data is complete, the score is **decomposable**:

$$\text{Score}(G : D) = \sum_i \text{Score}(X_i \mid Pa_i^G : D)$$

Heuristic Search (cont.)

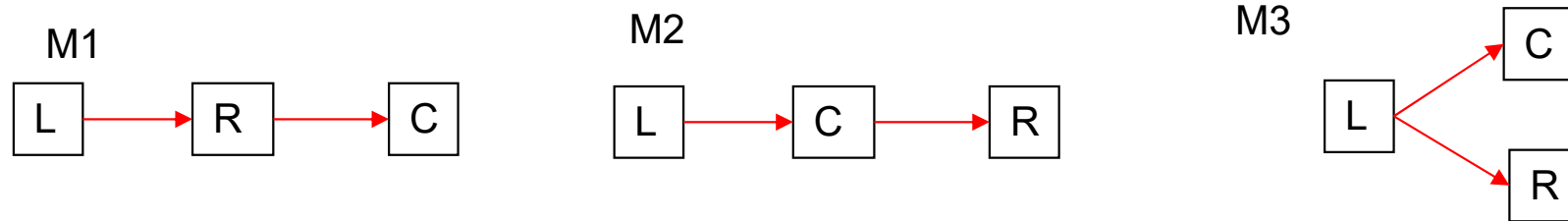
□ Typical operations:



Heuristic Search

- We address the problem by using heuristic search
- Traverse the space of possible networks, looking for high-scoring structures
- Search techniques:
 - Greedy hill-climbing
 - Simulated Annealing
 - ...
- Need to consider the number of parameters
 - ...

Inferring causal relations – mediation test



L - locus

R - gene

C – complex trait

$$M1. P(L, R, C) = P(L) P(R|L) P(C|R)$$

$$M2. P(L, R, C) = P(L) P(C|L) P(R|C)$$

$$M3. P(L, R, C) = P(L) P(C|L) P(R|C, L)$$

The likelihood for each model over all individuals in the population of interest are given by:

$$\text{likelihood function} = L(\theta; M) = p(\text{data} \mid \theta_M)$$

$$P(L, R, C \mid \theta_{M1}) = \prod_{i=1}^n P(L_i) \cdot P(R_i \mid L_i) \cdot P(C_i \mid R_i)$$

$$P(L, R, C \mid \theta_{M2}) = \prod_{i=1}^n P(L_i) \cdot P(C_i \mid L_i) \cdot P(R_i \mid C_i)$$

$$P(L, R, C \mid \theta_{M1}) = \prod_{i=1}^n P(L_i) \cdot P(R_i \mid L_i) \cdot P(C_i \mid L_i)$$

Joint probability for R|L and C|L

$$R | L \sim N(\mu_{R_L}, \sigma_{R_L}^2) \quad C | L \sim N(\mu_{C_L}, \sigma_{C_L}^2)$$

likelihoods for an individual given the joint probabilities given above are:

$$l(\theta_{r_i|L}; r_i | L) = \frac{1}{\sqrt{2\pi}\sigma_{R_L}} \exp\left(-\frac{(r_i - \mu_{R_L})^2}{2\sigma_{R_L}^2}\right), \text{ with } \theta_{r_i|L} = (\mu_{R_L}, \sigma_{R_L}),$$

$$l(\theta_{c_i|L}; c_i | L) = \frac{1}{\sqrt{2\pi}\sigma_{C_L}} \exp\left(-\frac{(c_i - \mu_{C_L})^2}{2\sigma_{C_L}^2}\right), \text{ with } \theta_{c_i|L} = (\mu_{C_L}, \sigma_{C_L}),$$

Joint probability for R|C

Assume $R = aC + b + Z$, where $Z \sim N(0, \sigma_z^2)$, $\text{Ind}(C, Z)$.

σ_C^2, μ_C are population variance and mean of C.

$$E(R | C) = \mu_R + \frac{\rho \sigma_R}{\sigma_C} (C - \mu_C), \quad \text{Var}(R | C) = \sigma_R^2 (1 - \rho^2)$$

$$R | C \sim N\left(\mu_R + \frac{\rho \sigma_R}{\sigma_C} (C - \mu_C), \sigma_R^2 (1 - \rho^2)\right)$$

Statistical background - regression

Assume $Y = aX + b + Z$, where $Z \sim N(0, \sigma_z^2)$, $\text{Ind}(X, Z)$.

σ_x^2, μ_x are population variance and mean of X .

Thus :

1. $\mu_y = a\mu_x + b$

2. $Y | X \sim N(aX + b, \sigma_z^2)$.

Using least - squares estimation of a and b :

$$\hat{a}, \hat{b} = \arg \min \sum_i (y_i - (ax_i + b))^2$$

Solving max likelihood estimation of $Y | X$:

$$\hat{a}, \hat{b} = \arg \max \log L(X | Y | a, b)$$

$$= \arg \max \frac{n}{2} \ln \frac{1}{2\pi\sigma^2} - \sum_i (y_i - (ax_i + b))^2$$

$$= \arg \min \sum_i (y_i - (ax_i + b))^2$$

Statistical background - regression

$$\hat{a}, \hat{b} = \arg \min \sum_i (y_i - (ax_i + b))^2$$

$$\hat{a} = \frac{\text{cov}(X, Y)}{\sigma_x^2} = \frac{\rho_{xy} \sigma_x \sigma_y}{\sigma_x^2} = \frac{\rho_{xy} \sigma_y}{\sigma_x}$$

$$\hat{b} = E(Y) - \hat{a} E(X) = \mu_y - \hat{a} \mu_x$$

Express $Y | X$ using population parameters $\mu_y, \mu_x, \sigma_y^2, \sigma_x^2, \rho_{xy}$

$$E(Y | X) = \hat{a} X + \hat{b} = \hat{a} X + \mu_y - \hat{a} \mu_x = \mu_y + \hat{a} (X - \mu_x) = \mu_y + \frac{\rho_{xy} \sigma_y}{\sigma_x} (X - \mu_x)$$

$$\text{Var}(Y | X) = \dots = \sigma_y^2 (1 - \rho_{xy}^2)$$

likelihoods for an individual given the probability distribution of R|C and C|R:

$$l(\theta_{c_i|r_i}; c_i | r_i) = \frac{1}{\sqrt{2\pi\sigma_C^2(1-\rho^2)}} \exp \left[-\frac{\left(c_i - \mu_C - \rho \frac{\sigma_C}{\sigma_R} (r_i - \mu_R) \right)^2}{2\sigma_C^2(1-\rho^2)} \right],$$

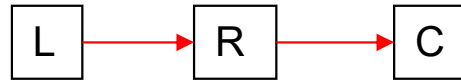
with $\theta_{c_i|r_i} = (\mu_R, \mu_C, \sigma_R, \sigma_C, \rho)$,

$$l(\theta_{r_i|c_i}; r_i | c_i) = \frac{1}{\sqrt{2\pi\sigma_R^2(1-\rho^2)}} \exp \left[-\frac{\left(r_i - \mu_R - \rho \frac{\sigma_R}{\sigma_C} (c_i - \mu_C) \right)^2}{2\sigma_R^2(1-\rho^2)} \right],$$

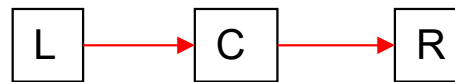
with $\theta_{r_i|c_i} = (\mu_R, \mu_C, \sigma_R, \sigma_C, \rho)$,

Comparing models with a different number of parameters

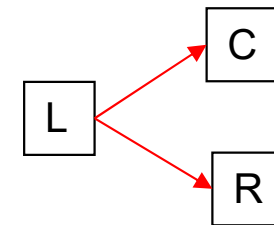
M1



M2



M3



L - locus

R - gene

C – complex trait

$$1. \quad L(\theta_{M_1}; M_1) = \prod_{i=1}^N \sum_{j=1}^3 P(L_j) l(\theta_{r_i|L_j}; r_i | L_j) l(\theta_{c_i|r_i}; c_i | r_i)$$

$$2. \quad L(\theta_{M_2}; M_2) = \prod_{i=1}^N \sum_{j=1}^3 P(L_j) l(\theta_{c_i|L_j}; c_i | L_j) l(\theta_{r_i|c_i}; r_i | c_i)$$

$$3. \quad L(\theta_{M_3}; M_3) = \prod_{i=1}^N \sum_{j=1}^3 P(L_j) l(\theta_{r_i|L_j}; r_i | L_j) l(\theta_{c_i|r_i, L_j}; c_i | L_j),$$

Model selection

- The **Akaike information criterion** is a measure of the relative goodness of fit of a statistical model.
- AIC values provide a means for model selection.

$$AIC = 2k - 2 \ln(L)$$

where k is the number of parameters in the statistical model, and L is the maximized value of the likelihood function for the estimated model.

- Given a set of candidate models for the data, *the preferred model is the one with the minimum AIC value.*

The likelihood for each model over all individuals in the population of interest are given by:

$$\text{likelihood function} = L(\theta; M) = p(\text{data} \mid \theta_M)$$

$$P(L, R, C \mid \theta_{M1}) = \prod_{i=1}^n P(L_i) \cdot P(R_i \mid L_i) \cdot P(C_i \mid R_i)$$

$$P(L, R, C \mid \theta_{M2}) = \prod_{i=1}^n P(L_i) \cdot P(C_i \mid L_i) \cdot P(R_i \mid C_i)$$

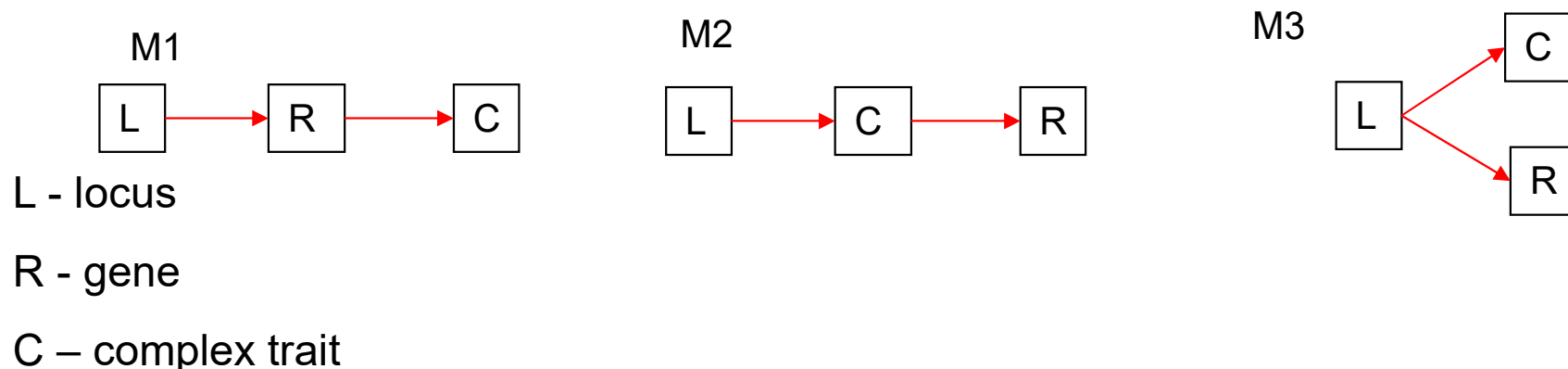
$$P(L, R, C \mid \theta_{M1}) = \prod_{i=1}^n P(L_i) \cdot P(R_i \mid L_i) \cdot P(C_i \mid L_i)$$

Model selection test

- H0: M1 and M2 are equally close to the true model
- H1: M1 is better than M2

$$LR_{12} = \ln \frac{\max_{\theta_{M1}} P(L, R, C | \theta_{M1})}{\max_{\theta_{M2}} P(L, R, C | \theta_{M2})} = \ln \frac{P(L, R, C | \hat{\theta}_{M1})}{P(L, R, C | \hat{\theta}_{M2})}$$

- Option 1: permutation test
- Option 2: approximated distribution
- The same rationale can be applied to test M1 against M3.
- It is also possible to have a joint test of M1 against Max(M2, M3):
 - H0: M1 no closer to the true model than M2 and M3
 - H1: M1 is closer to the true model than M2 and M3



Final Project - Causality analysis

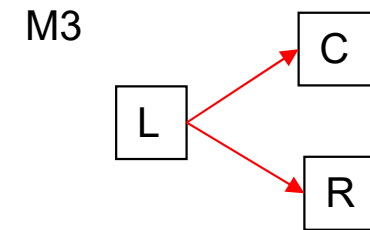
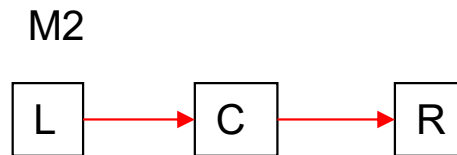
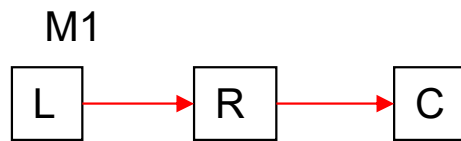
Please note: This assignment should be submitted in **pairs**. Each pair should solve the assignment independently, although general discussion between students is allowed. Copying code and solutions is forbidden.

Causality analysis:

- Data: mention the data you're using, and write a short summary of the previous assignments HW2 and HW3 (QTL and eQTL analysis, with combined results).
- Apply the causality test on the results from previous parts. Run causality test on each pair of gene and phenotype where both the QTL and the eQTL are located in a nearby genomic position (or, of course, have the same position).
- Report the predicted relations among the QTL, associated gene, and phenotype.
- Apply permutation test to get statistical significance for one or a few specific causality hypotheses.
- Run the test on 10 triplets and explain the choice you made and the design of permutation test in detail.

Using mediation analysis to identify causal genes

Inferring causal relations

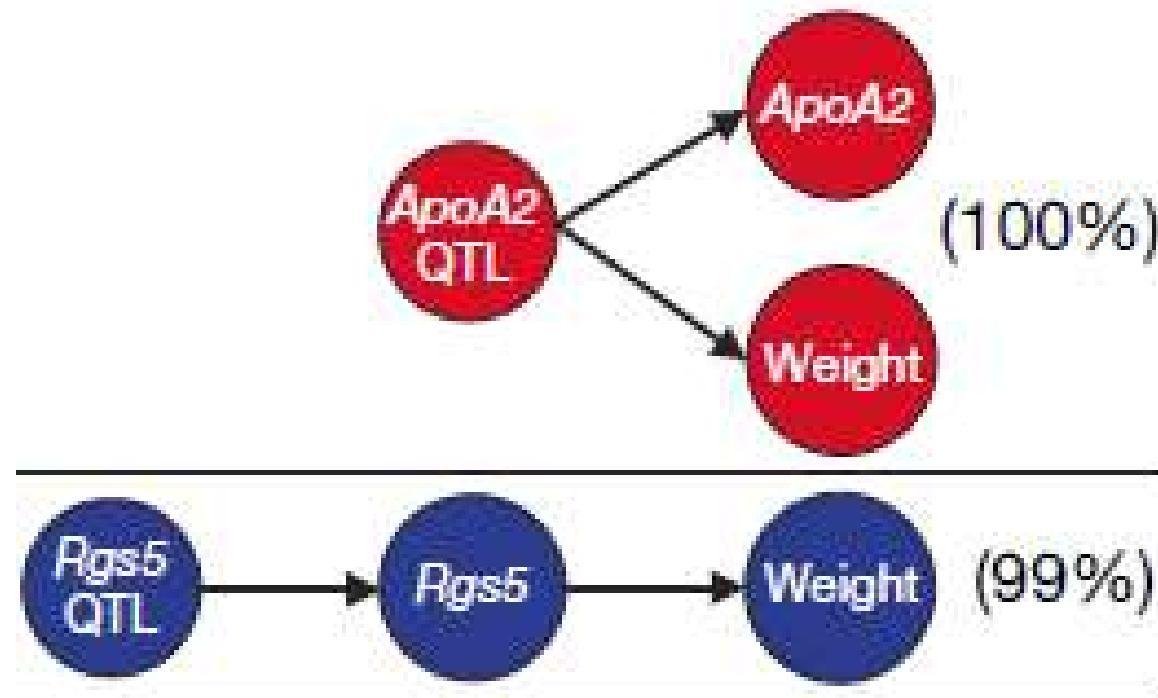


L - locus

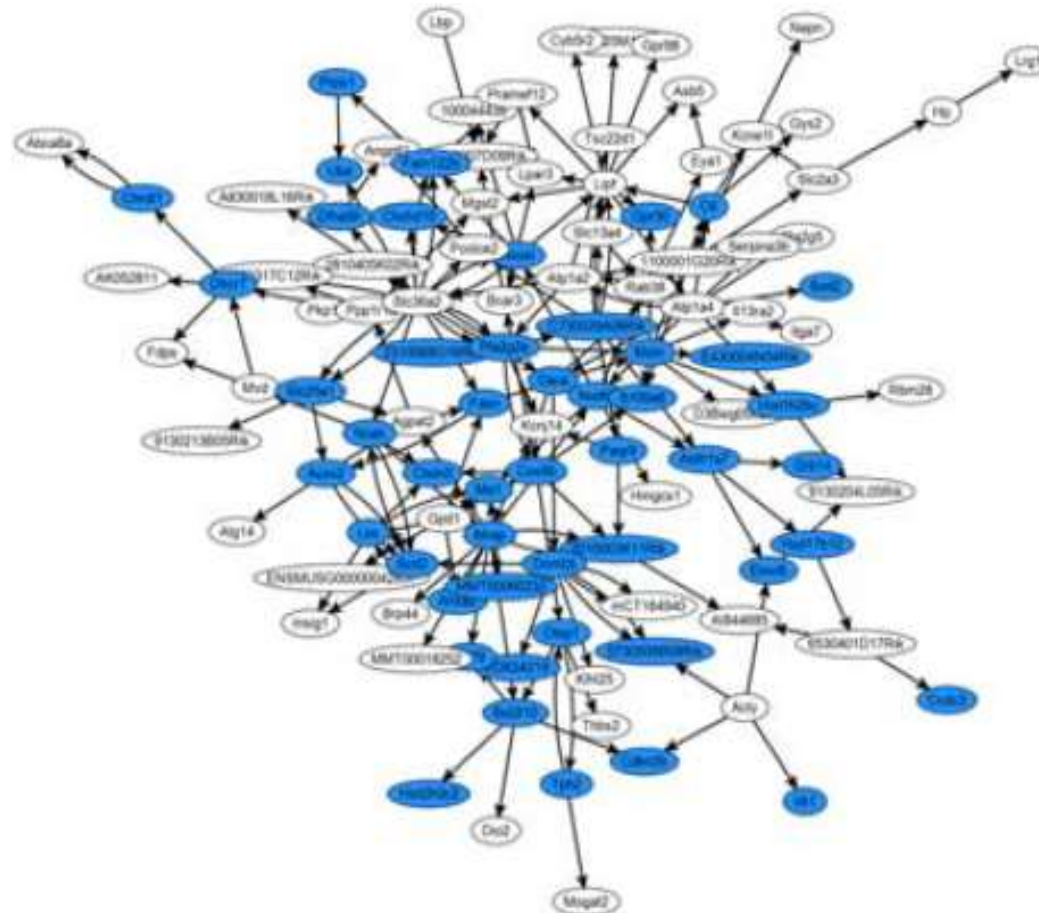
R - gene

C – complex trait

Example



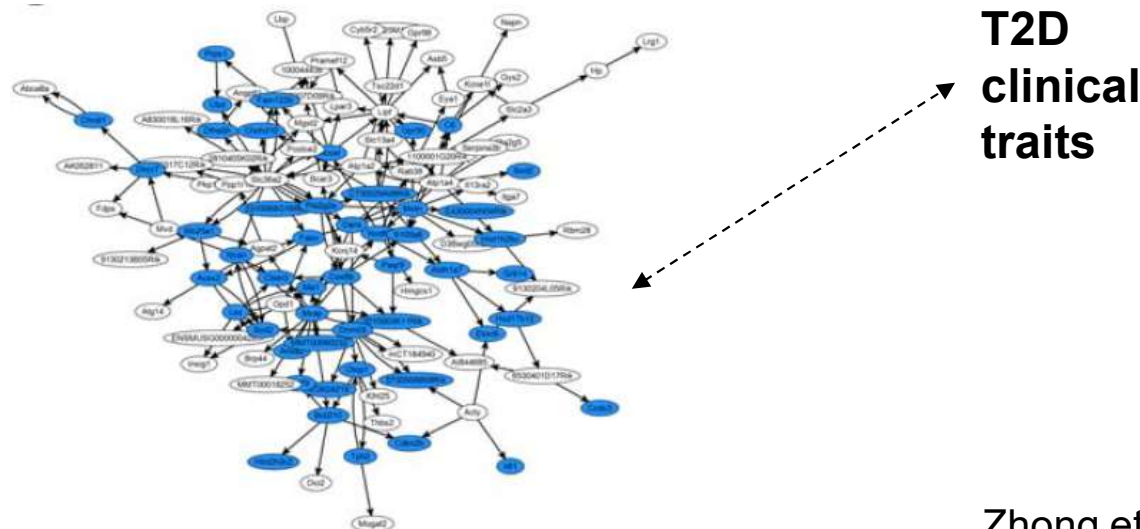
Adipose co-expression subnetwork in mice



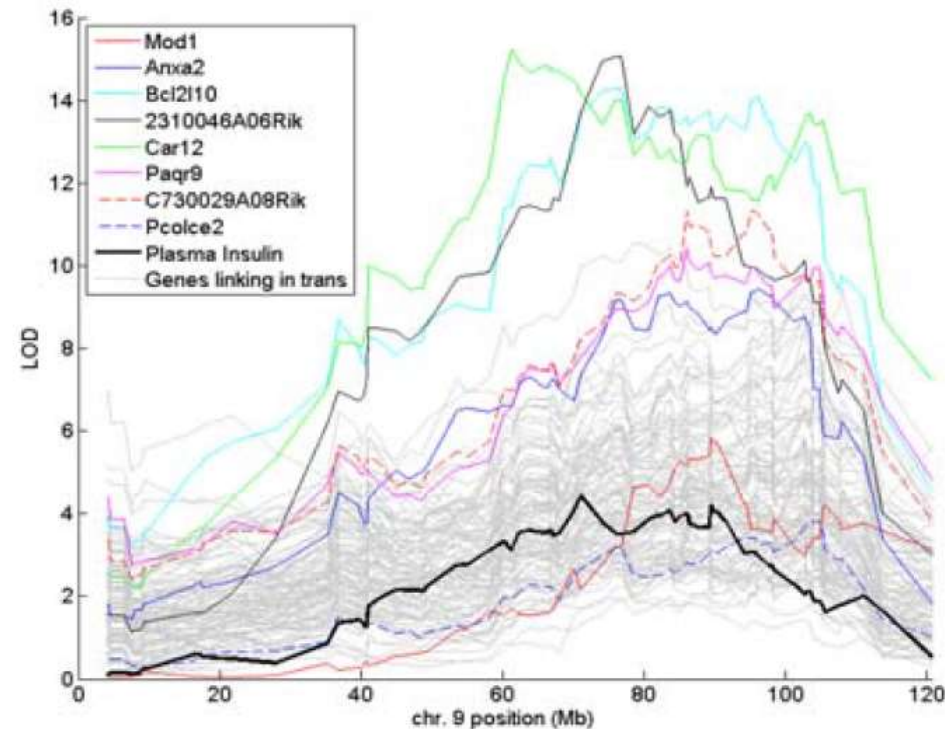
Highly enriched in lipid, fatty acid and steroid metabolism

Adipose co-expression subnetwork in mice: relation to T2D phenotype

- The first principal component of the gene expression traits making up this subnetwork:
 - explains 45.6% of the expression variation of the subnetwork
 - strongly positively correlated with T2D clinical traits in the same mice: number of islets ($R=0.52$, $p<10^{-70}$), plasma insulin levels ($R=0.70$, $p<10^{-70}$), and plasma glucose levels ($R=-0.57$, $p<10^{-4}$).

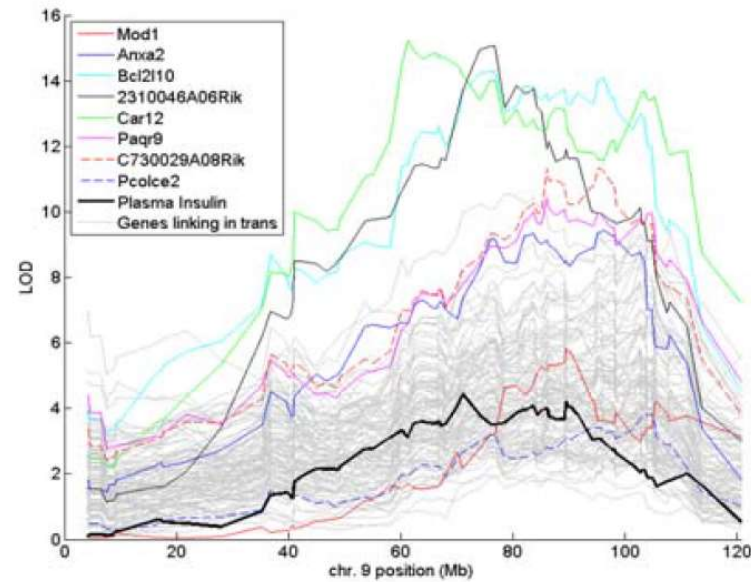


eQTL analysis



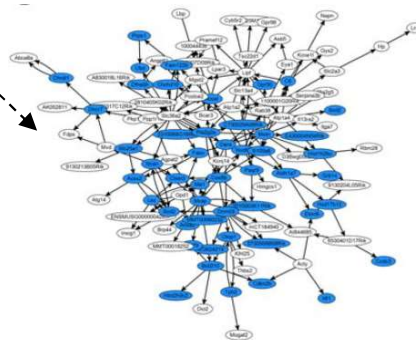
- Of 159 genes in the adipose subnetwork, 117 have an eQTL at the same region on mouse chromosome 9 (8 in cis)

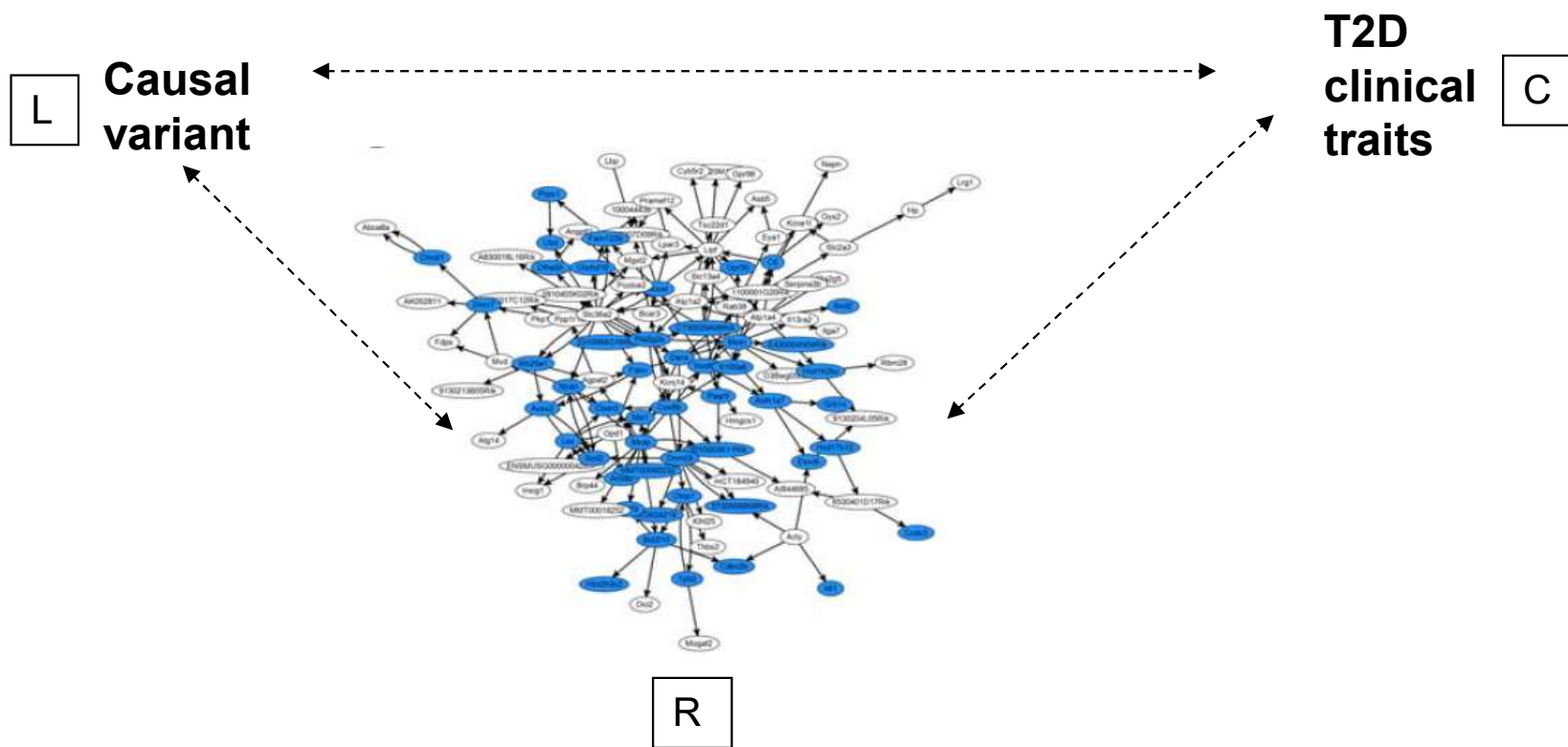
Adipose co-expression subnetwork in mice: relation to T2D phenotype



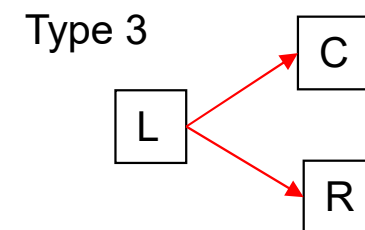
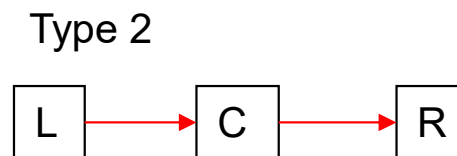
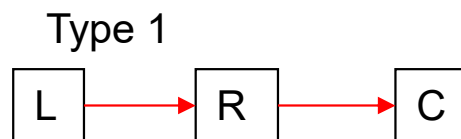
**Causal
variant**

**T2D
clinical
traits**

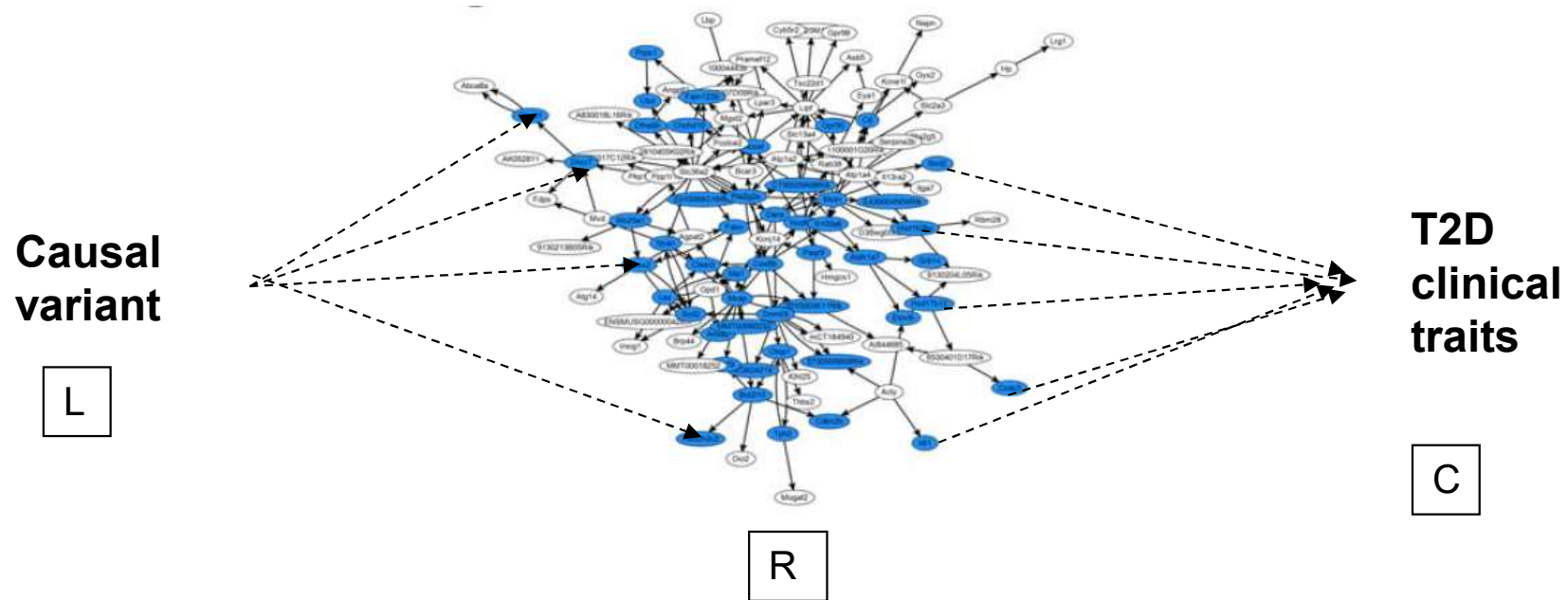




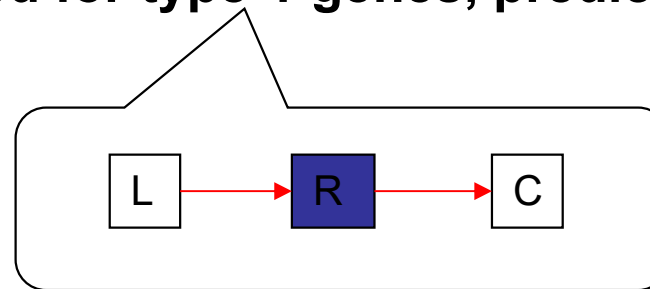
Possible (simple) relations:



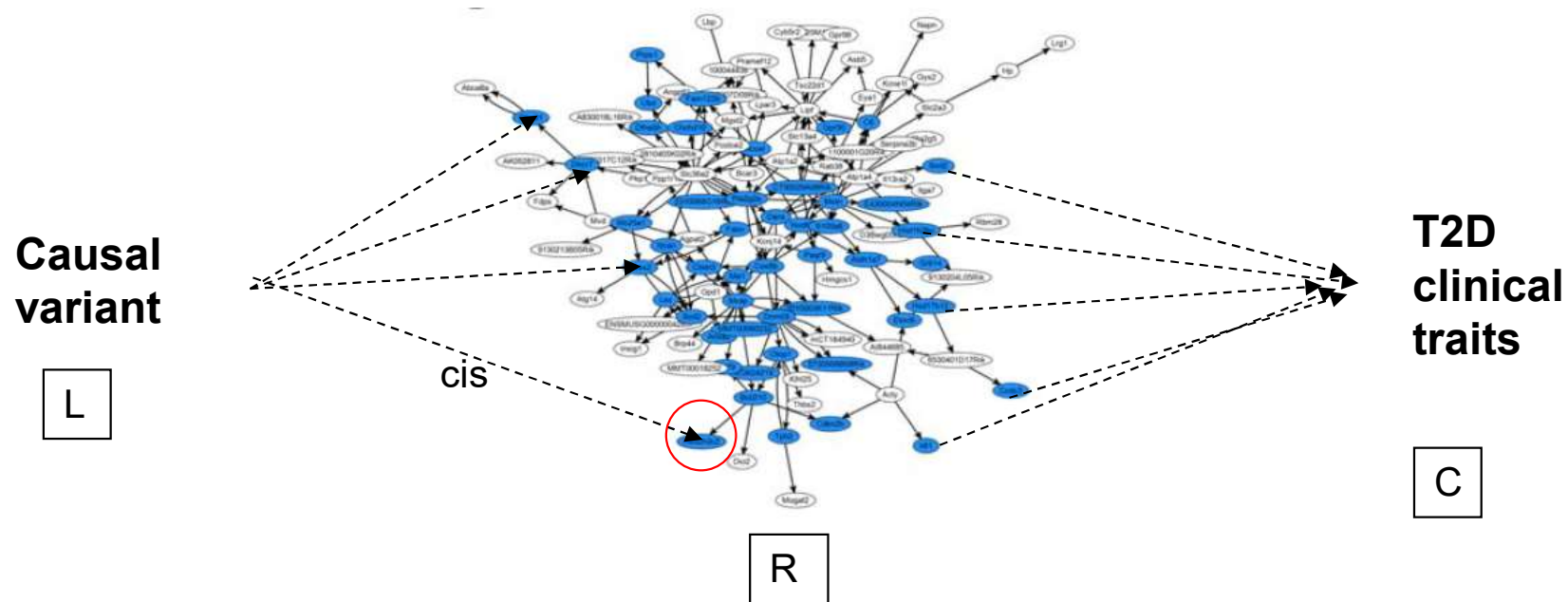
Inferring causal relations (mediation analysis)



The subnetwork is enriched for type-1 genes, predicted to cause T2D trait variation



eQTL filtering lead to the identification of *Me1*



***Me1* is predicted to cause T2D trait variation ('type 1')**

***Me1* was identified with strong cis-eQTL in both human and mouse adipose tissue.**

Validation: *Me1*^{-/-} mice fed a high fat diet demonstrate significantly lower insulin levels compared to the controls