



HAL
open science

Using eQTL networks to decipher the architecture of complex traits

Maud Fagny

► To cite this version:

Maud Fagny. Using eQTL networks to decipher the architecture of complex traits. Doctoral. NCMM PhD Course in Multi-Omics, NCMM, Oslo, Norway. 2023. <hal-04330996>

HAL Id: hal-04330996

<https://hal.inrae.fr/hal-04330996v1>

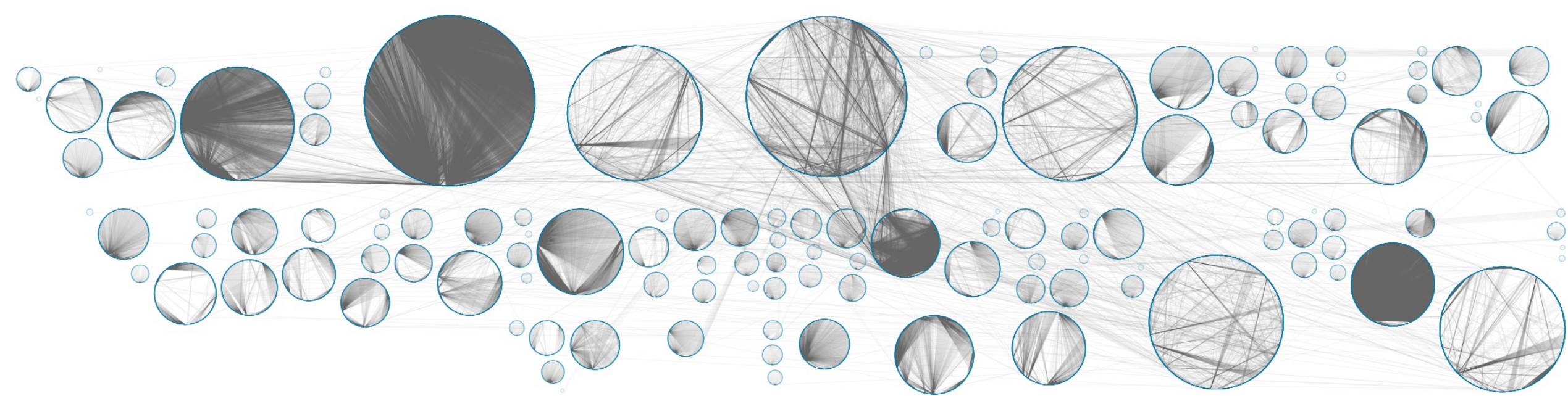
Submitted on 8 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License



Using eQTL networks to decipher the architecture of complex traits

Maud Fagny, PhD
INRAE, Gif-sur-Yvette

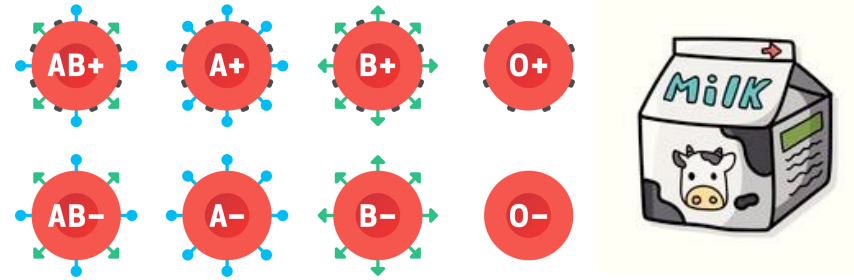
Outline

- Introduction:
 - The genetic architecture of complex traits
 - Why studying eQTL networks ?
- Biologically characterizing cancer risk SNP with eQTL networks
- Improving our understanding of complex trait heritability
- Detecting past selection events?
- Conclusion and future topics

Mendelian vs. Polygenic traits

Mendelian traits:

Gene \longrightarrow Phenotype

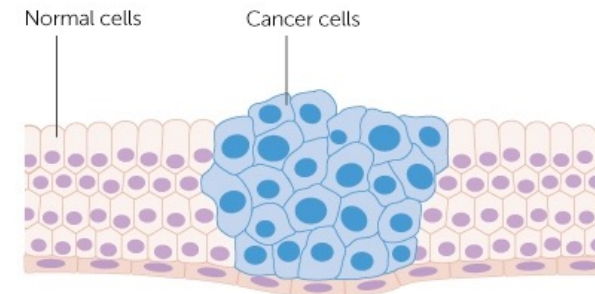


Complex traits:

Gene 1
Gene 2
Gene 2

\longrightarrow Phenotype

Environment \uparrow



Cancer Research UK

Genetic architecture of some often-studied traits

Monogenic

Mendelian

1 trait = 1 gene

Oligogenic

1 trait = 2-10 genes

Polygenic

1 trait = many genes

Genetic architecture of some often-studied traits

Monogenic

Mendelian

1 trait = 1 gene

Lactose tolerance

Sickle-cell disease

Oligogenic

1 trait = 2-10 genes

Skin/eye/hairs color

Risk to develop type II
diabetes

Risk to develop a cancer

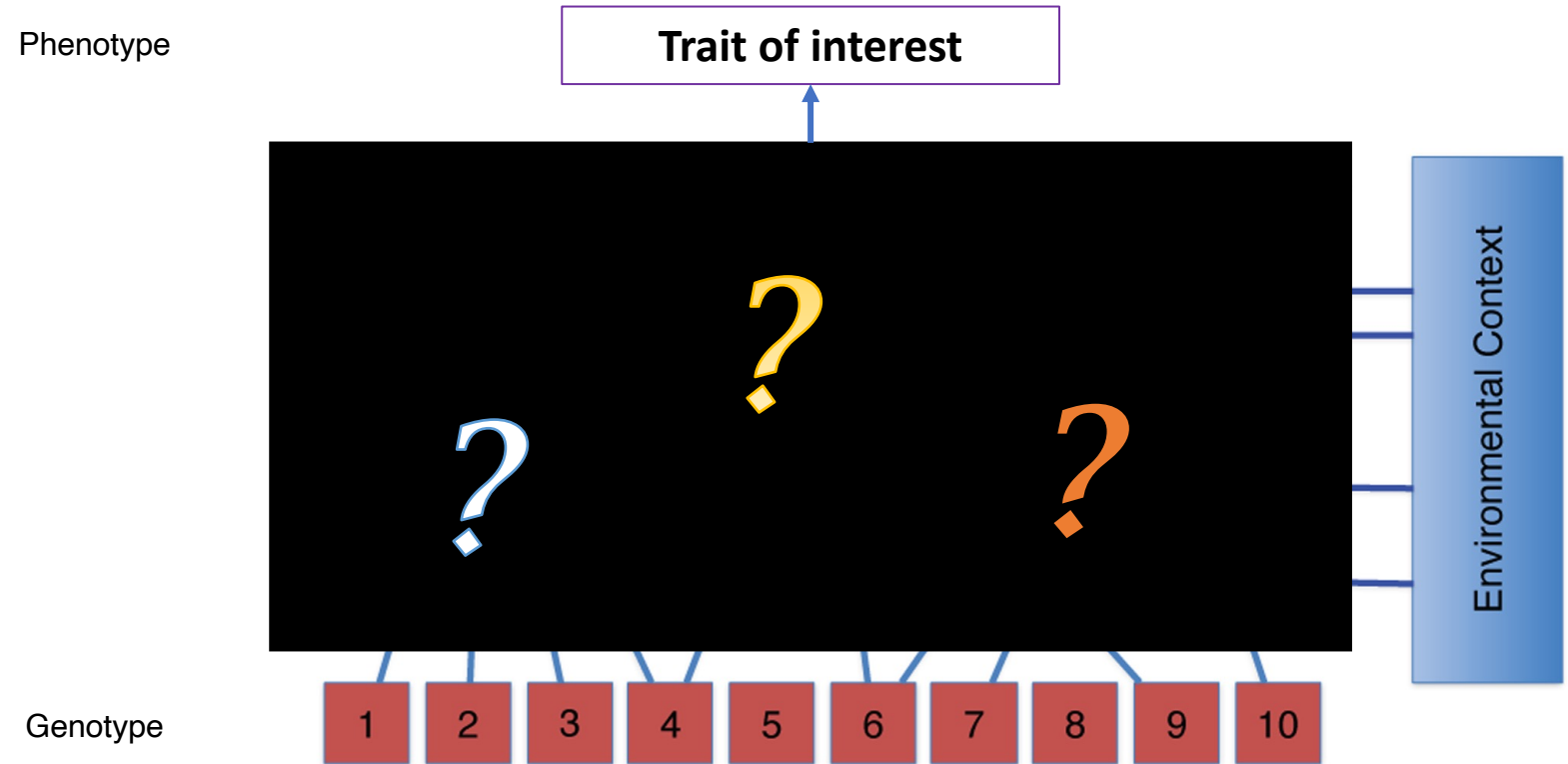
Polygenic

1 trait = many genes

Risk to develop
schizophrenia

Adult size

Diving deeper in the architecture of complex traits

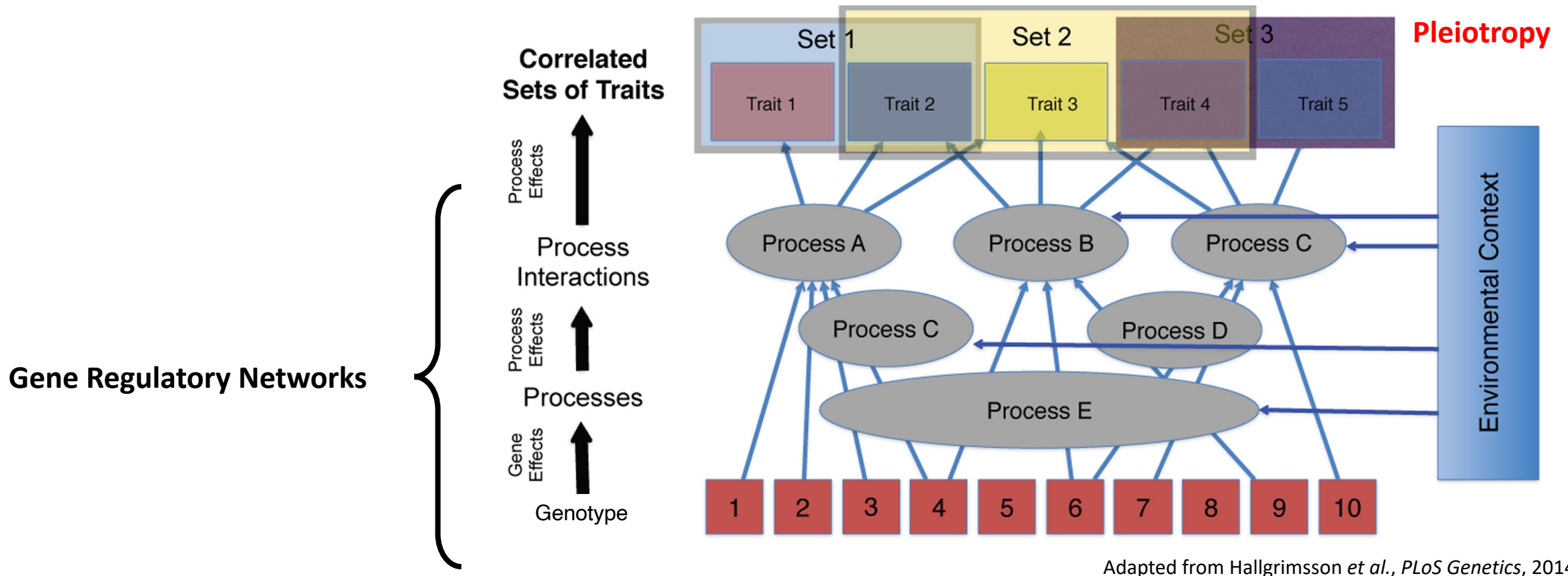


Adapted from Hallgrímsson *et al.*, *PLoS Genetics*, 2014

➤ The genotype-phenotype gap:

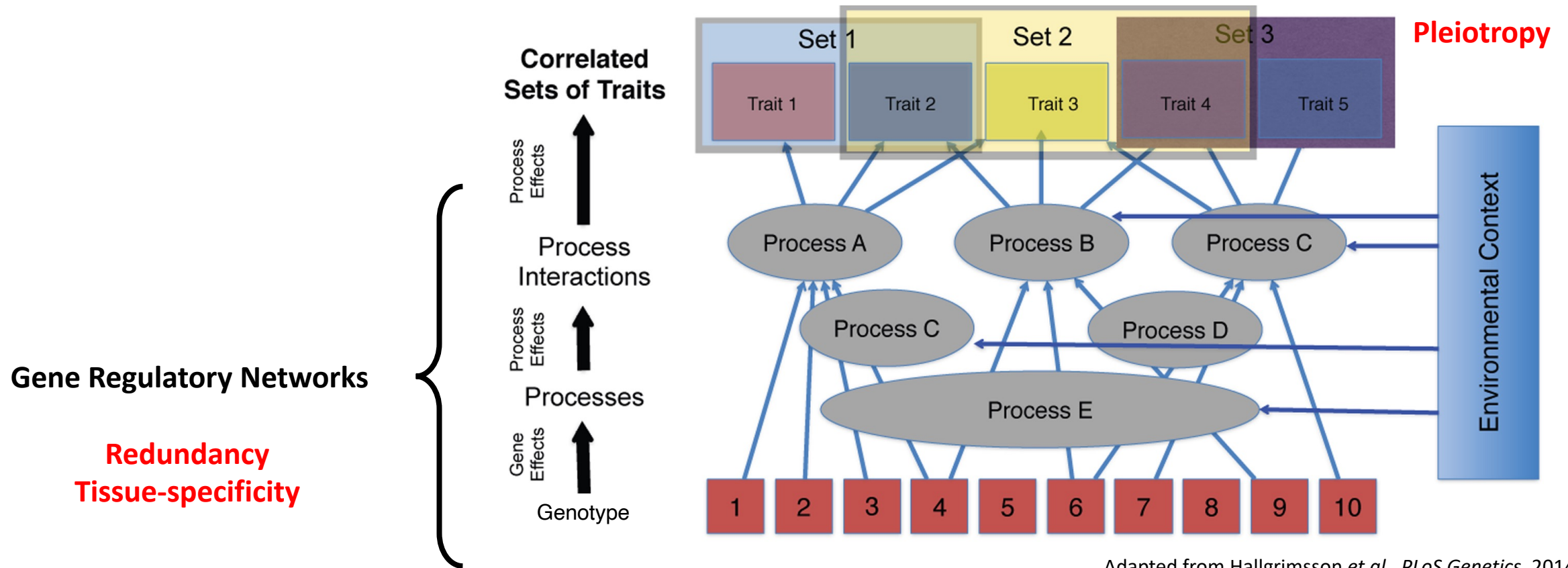
Mechanisms by which most genetic variation identified in GWAS affect the final phenotype **unknown**

Diving deeper in the architecture of complex traits



➤ The importance of gene regulatory networks

Diving deeper in the architecture of complex traits

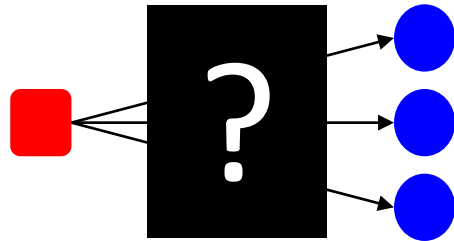


Adapted from Hallgrímsson *et al.*, *PLoS Genetics*, 2014

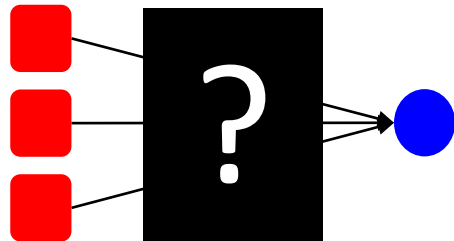
- Complex interactions at the basis of polygenic traits, including pleiotropy and redundancy

The different challenges

Complexity



Pleiotropy



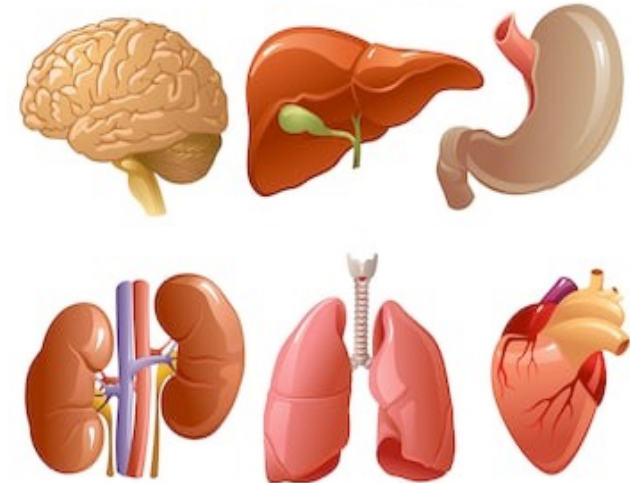
Redundancy

Mutations

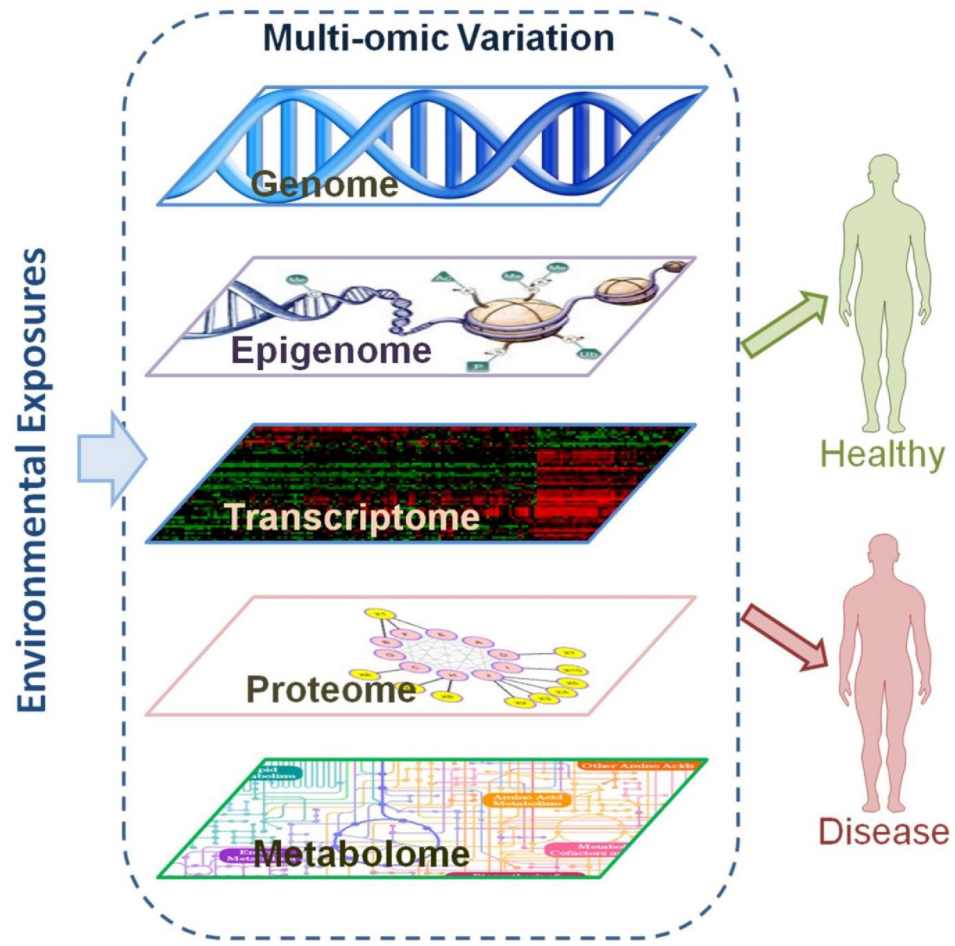
Traits

Genotype-Phenotype Gap

Tissue-Specificity



Exploring the genotype-phenotype gap



GWAS SNPs:

Intermediate frequencies

&

Small effect size on phenotype

&

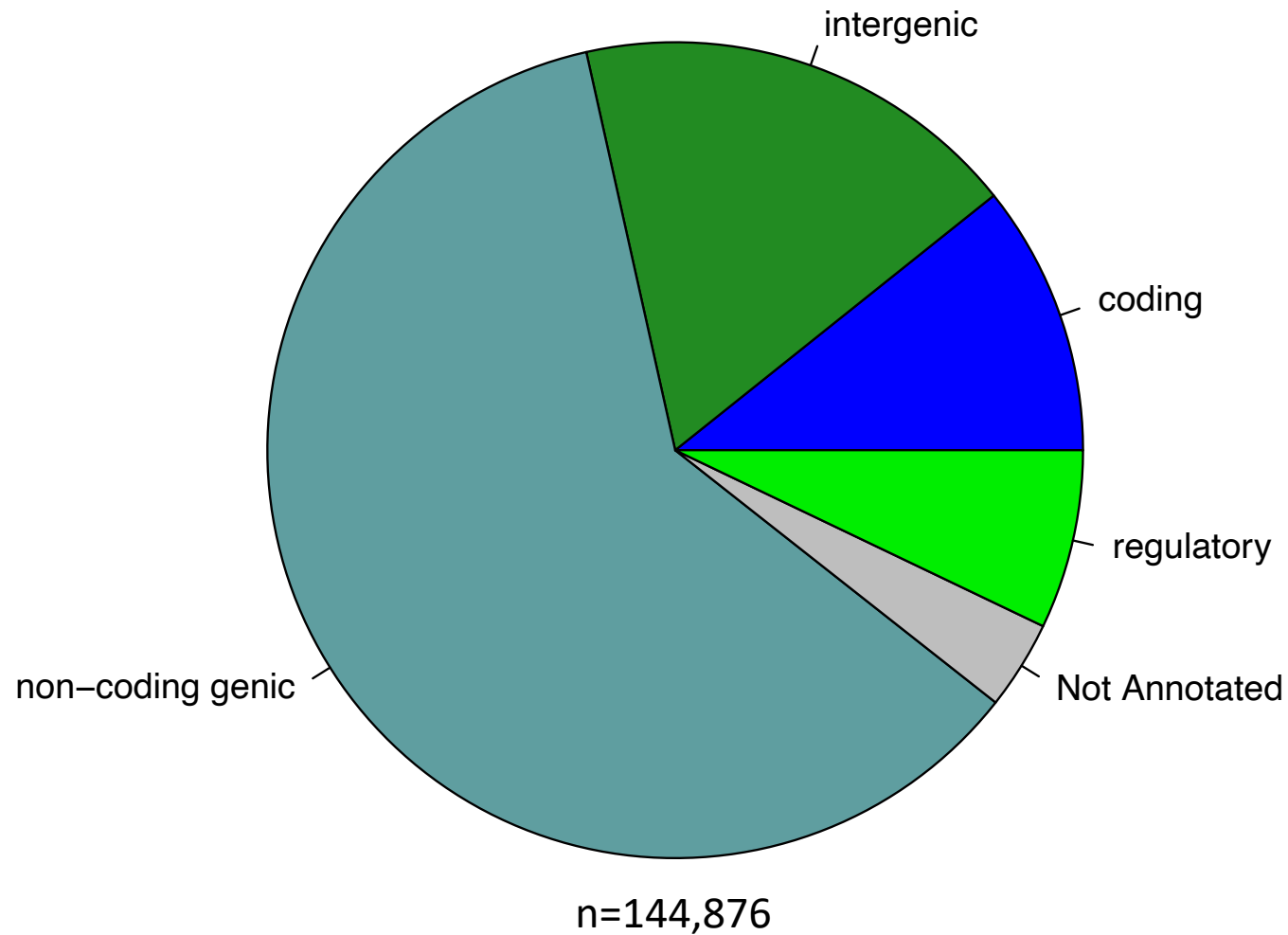
Tissue-specific effect

&

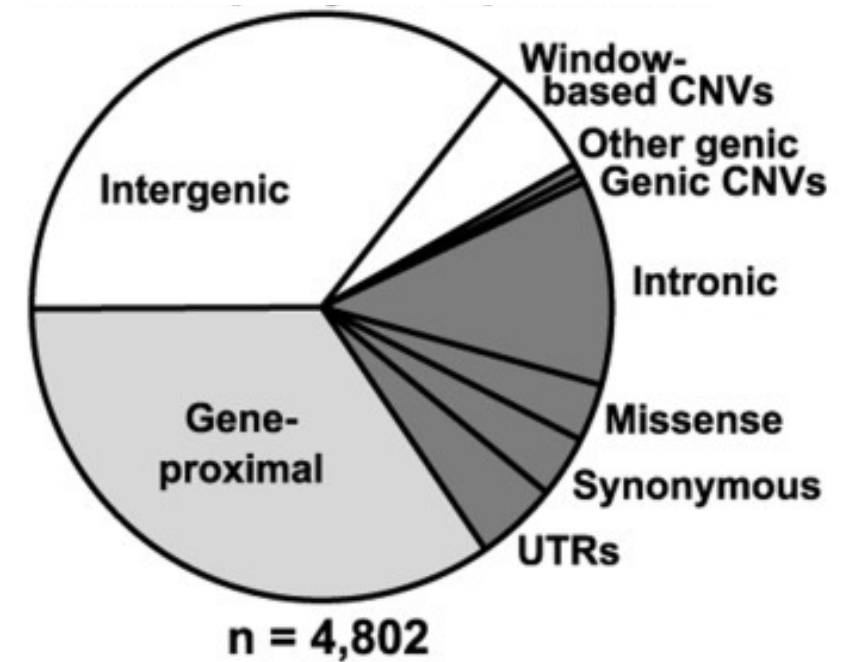
Located in regulatory regions

The importance of gene expression regulation in polygenic phenotypes

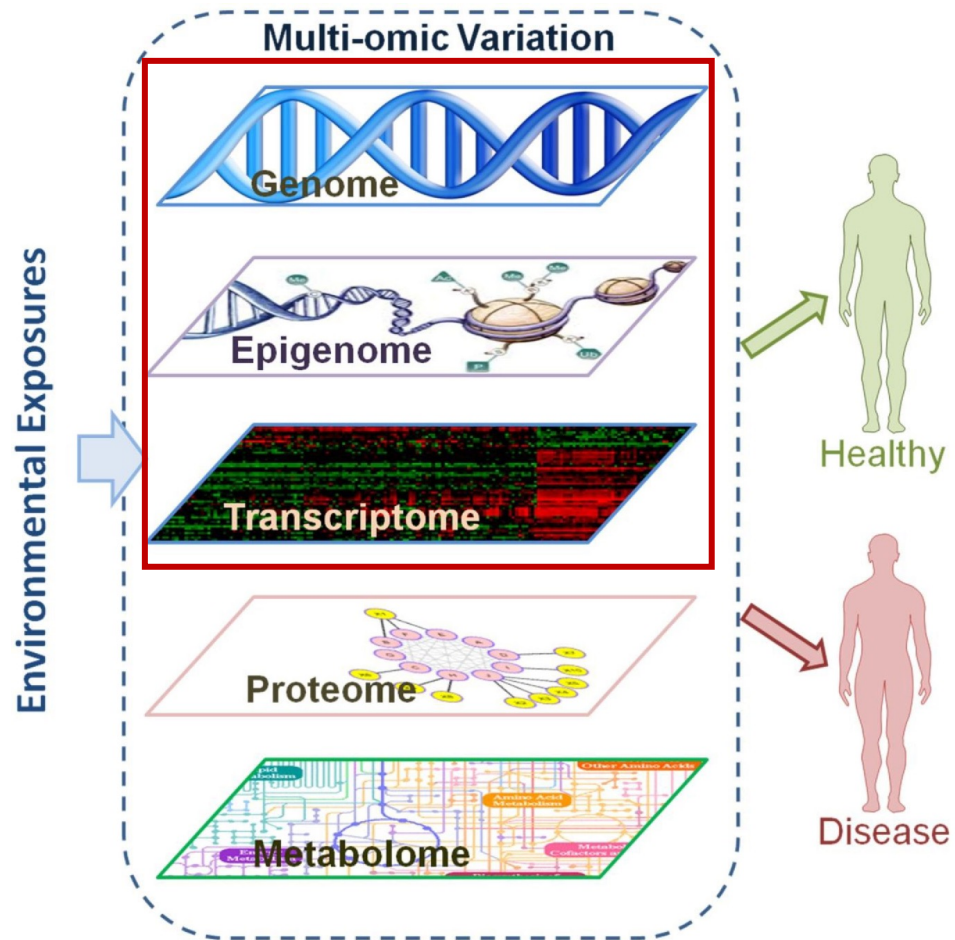
GWAS in humans



GWAS for developmental traits in maize



Exploring the genotype-phenotype gap



GWAS SNPs:

Intermediate frequencies

&

Small effect size on phenotype

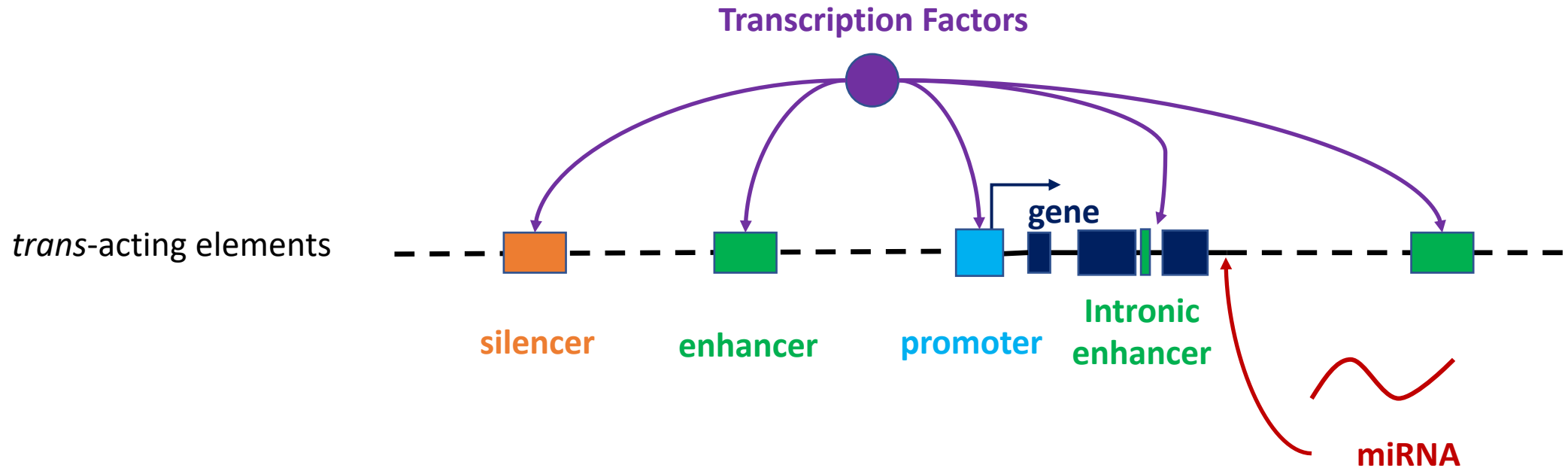
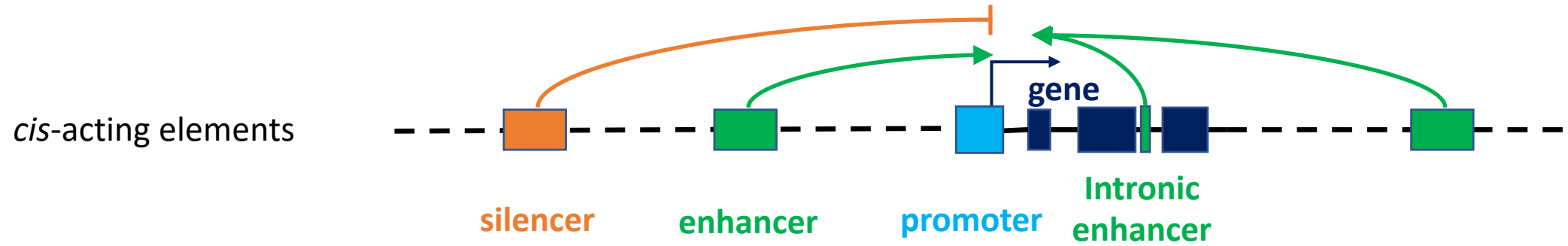
&

Tissue-specific effect

&

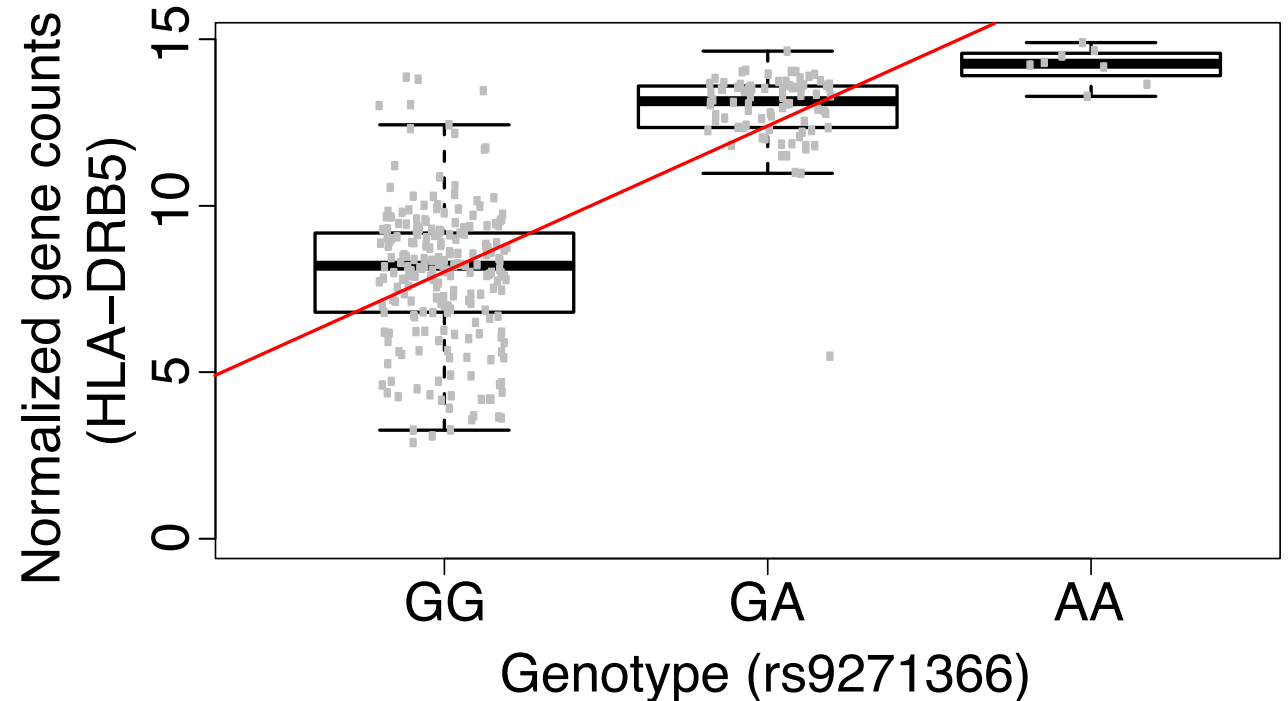
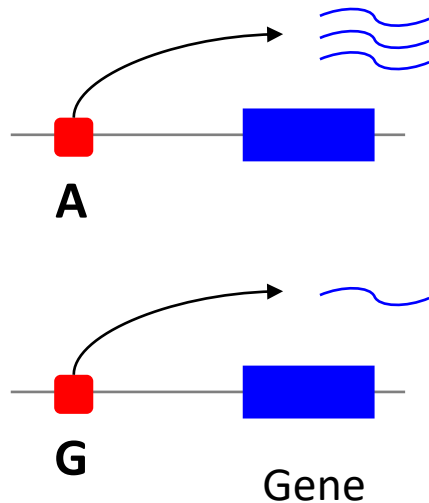
Located in regulatory regions

Gene Expression Regulation in cis and in trans



Building a bridge between genotype and phenotype: Step 1 – the expression Quantitative Trait Loci (eQTLs)

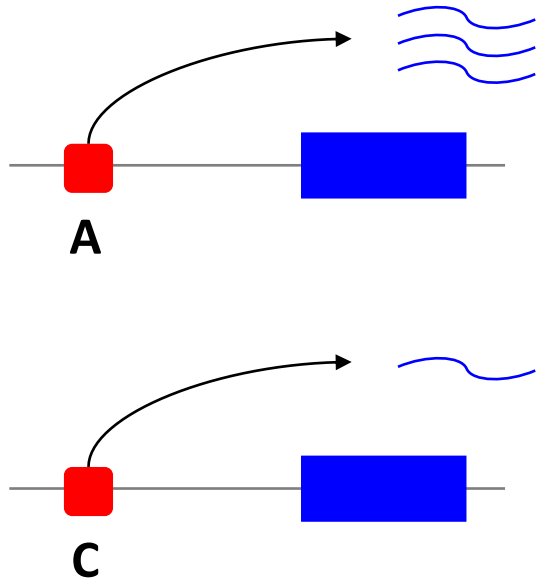
Expression quantitative trait locus (eQTL)



- Expression quantitative traits loci are potential regulatory SNPs
- SNPs associated to traits or diseases are enriched for eQTLs

Organising the complexity: a graph representation

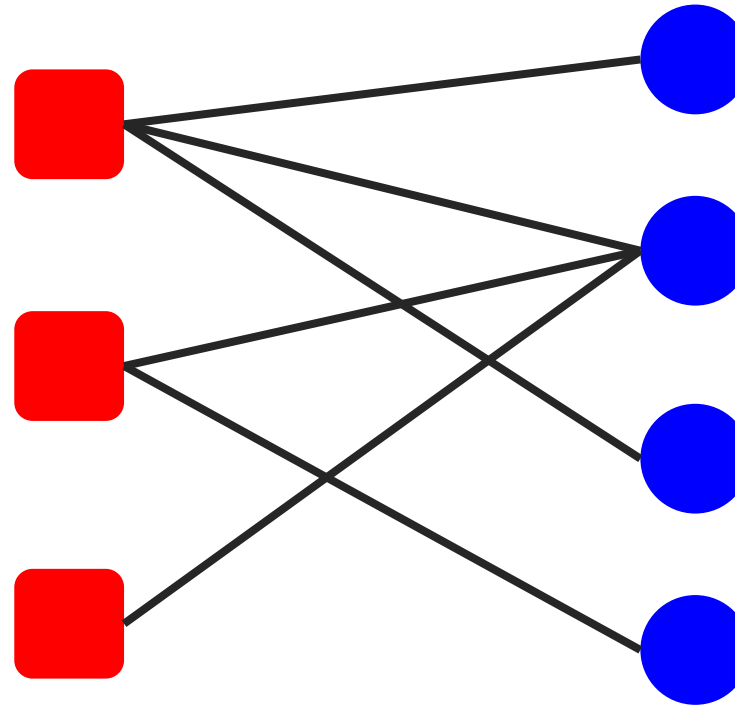
Expression quantitative trait locus



Cis- and trans-eQTL bipartite networks

SNPs

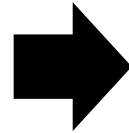
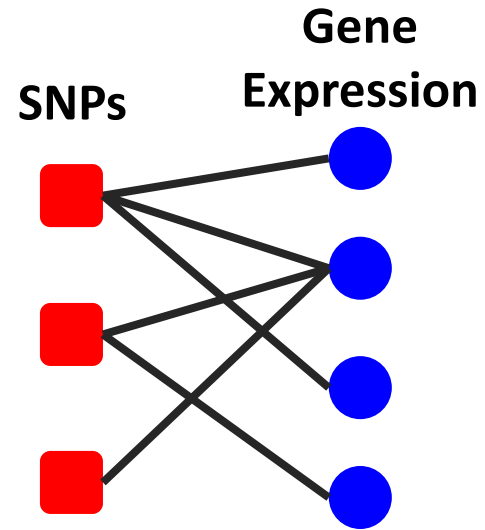
Gene Expression



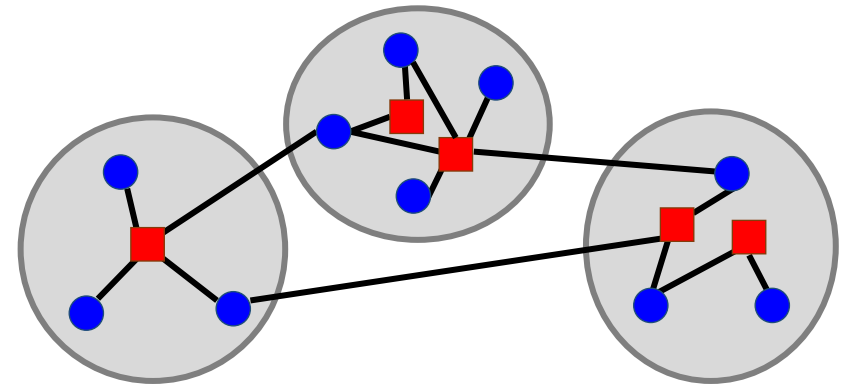
- Using a systems biology network approach to groups SNPs influencing the expression of the same genes

Grouping SNPs using a network property : modularity

Cis- and trans-eQTL bipartite network

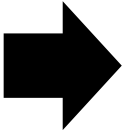
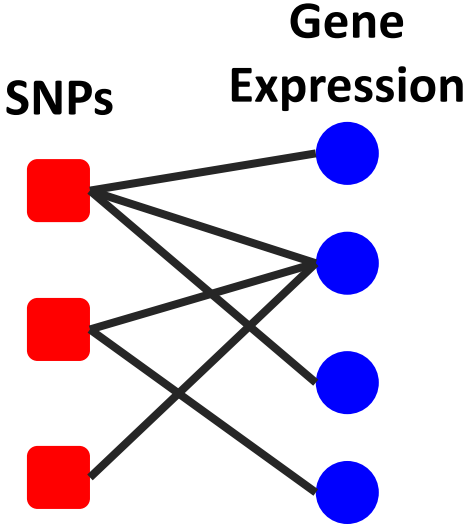


Community detection

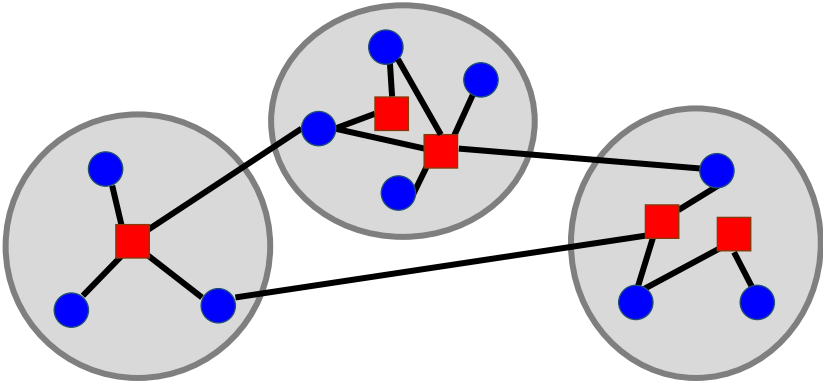


Grouping SNPs using a network property : modularity

Cis- and trans-eQTL bipartite network



Community detection



Bipartite Modularity Maximization

$$Q = \sum_i (e_{ii} - a_i^2)$$

Fraction of network links in community i

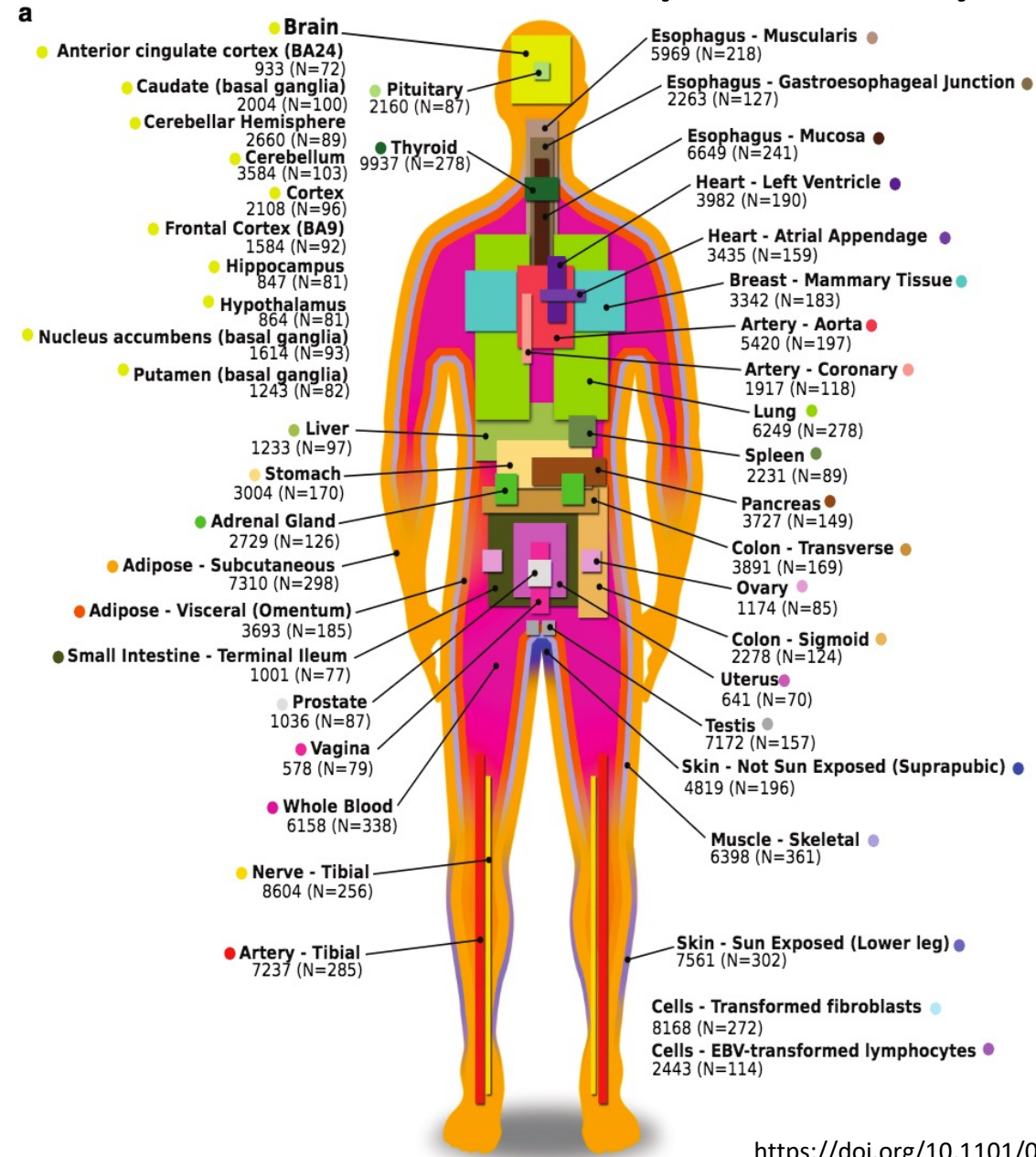
Fraction of links expected by chance

Taking into account the tissue-specificity

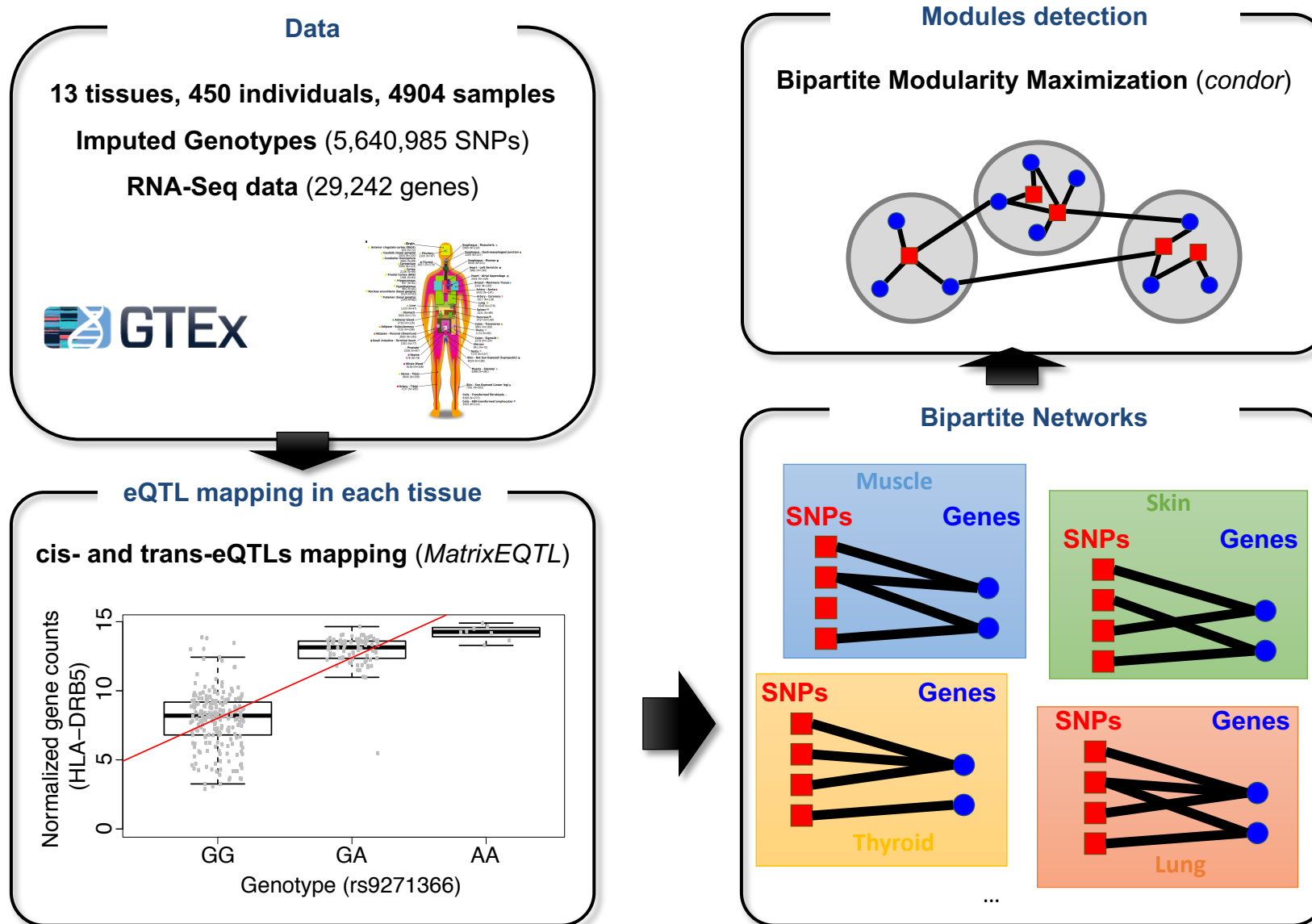
Get tissue-specific expression data !



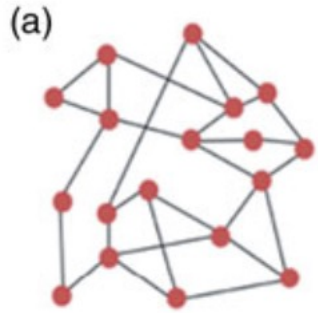
- 449 individuals.
- Genotyping data:
 - 84.3% European Am.
 - 13.7% African Am.
 - 1% Asian Am.
- RNA sequencing data:
 - 13 tissues.



Approach summary: building tissue-specific eQTL networks

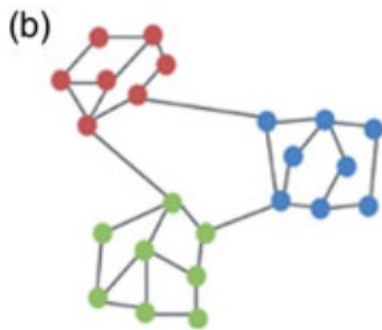


eQTL networks have high modularity



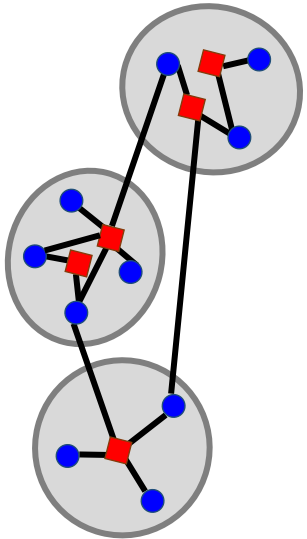
A random network

Modularity ~ 0

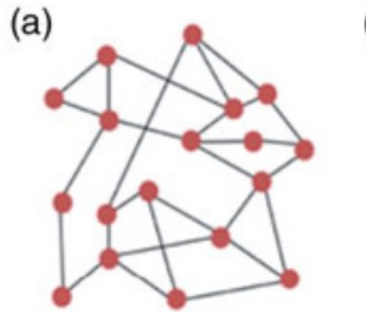


A modular network

Modularity $\gg 0$

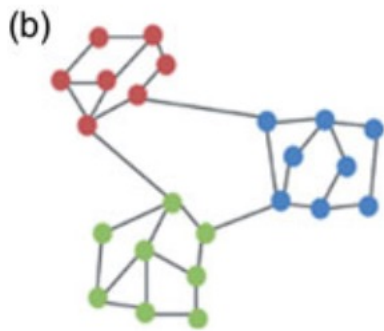


eQTL networks have high modularity



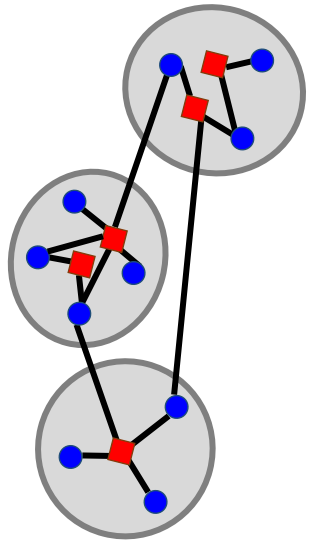
A random network

Modularity ~ 0

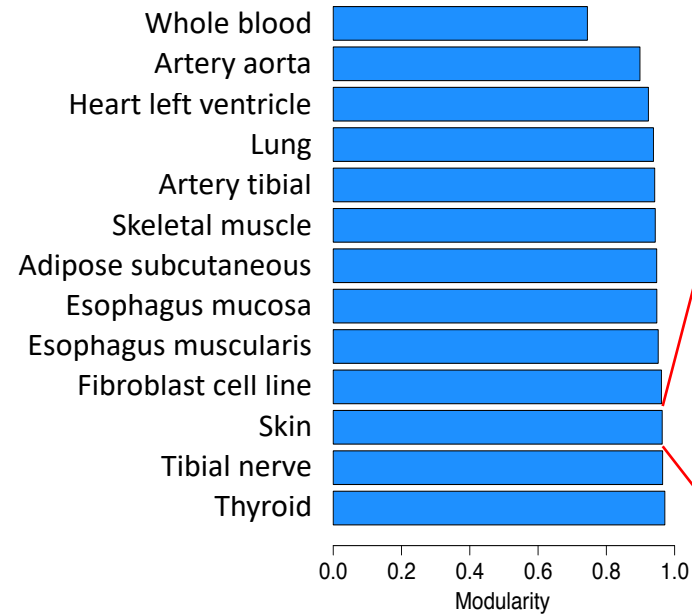


A modular network

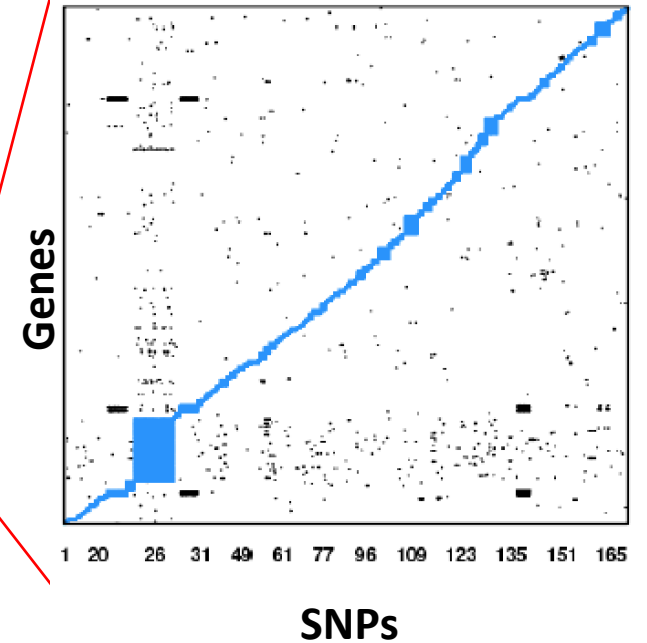
Modularity $\gg 0$



Network Modularity



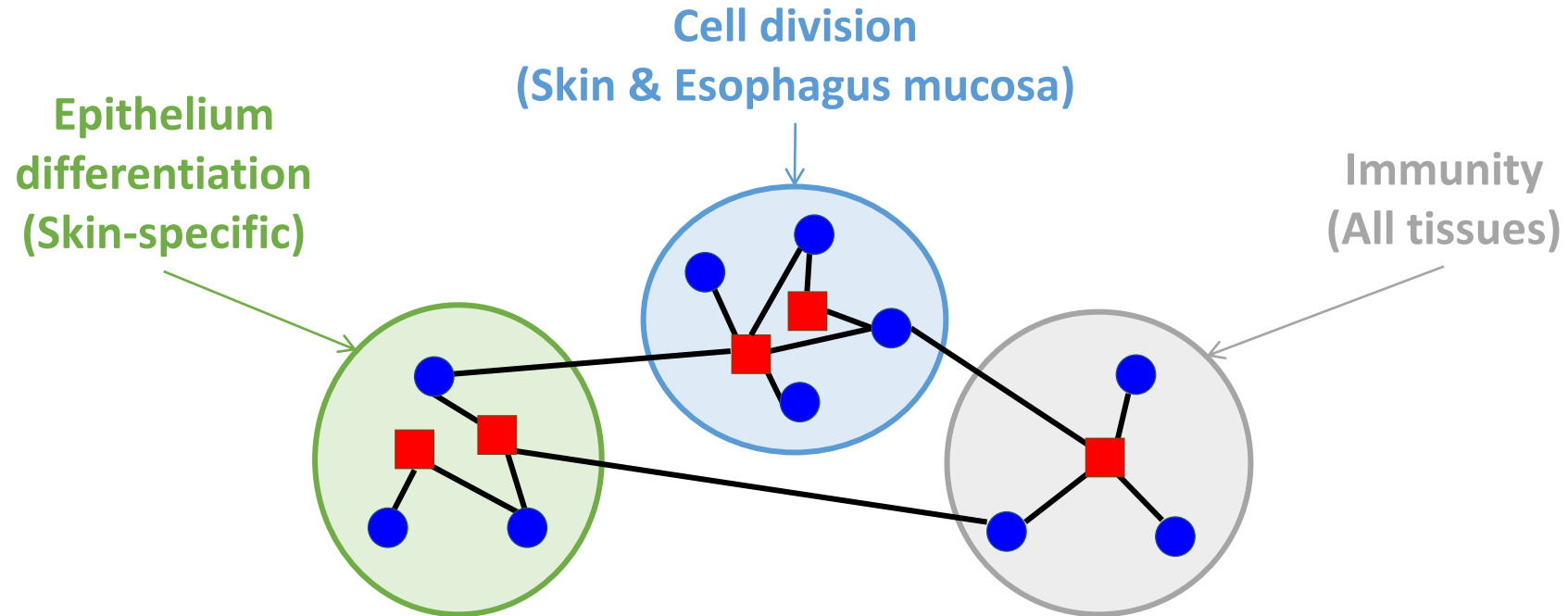
Skin network



Fagny et al. (2017) PNAS

Communities correspond to biological functions of tissues

Example of communities identified in skin eQTL network



- **Groups of SNPs regulate groups of genes involved in similar functions.**
- **Communities can be tissue-specific or shared across tissues**

Research question: Understanding the genetic architecture of complex traits

- **Biological characterization of SNPs identified by GWAS: How do genetic variation influence a trait ? Which biological pathways are involved?**
- **Heritability: Which mutations most affect the trait? Where are they located in the regulatory network**
- **Natural Selection: How do such a complex trait evolve?**

Research question: Understanding the genetic architecture of complex traits

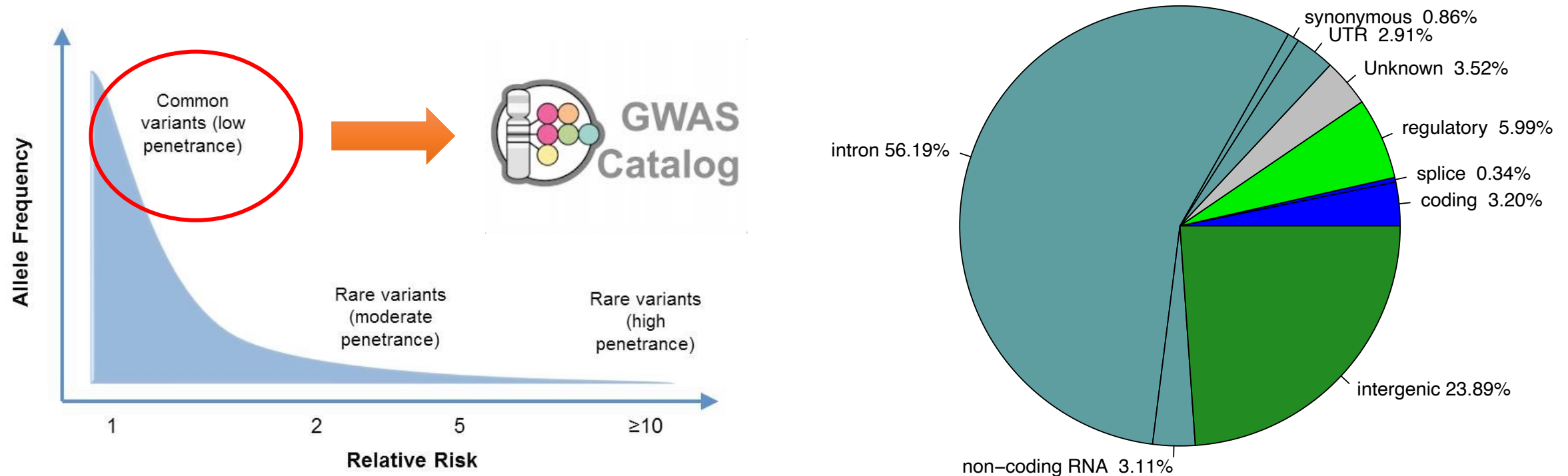
- **Biological characterization of SNPs identified by GWAS: How do genetic variation influence a trait ? Which biological pathways are involved?**
- **Heritability: Which mutations most affect the trait? Where are they located in the regulatory network**
- **Natural Selection: How do such a complex trait evolve?**



John Platig,
Assistant Professor,
Univ. Of Virginia

A particularly complex trait: the risk to develop cancer(s)

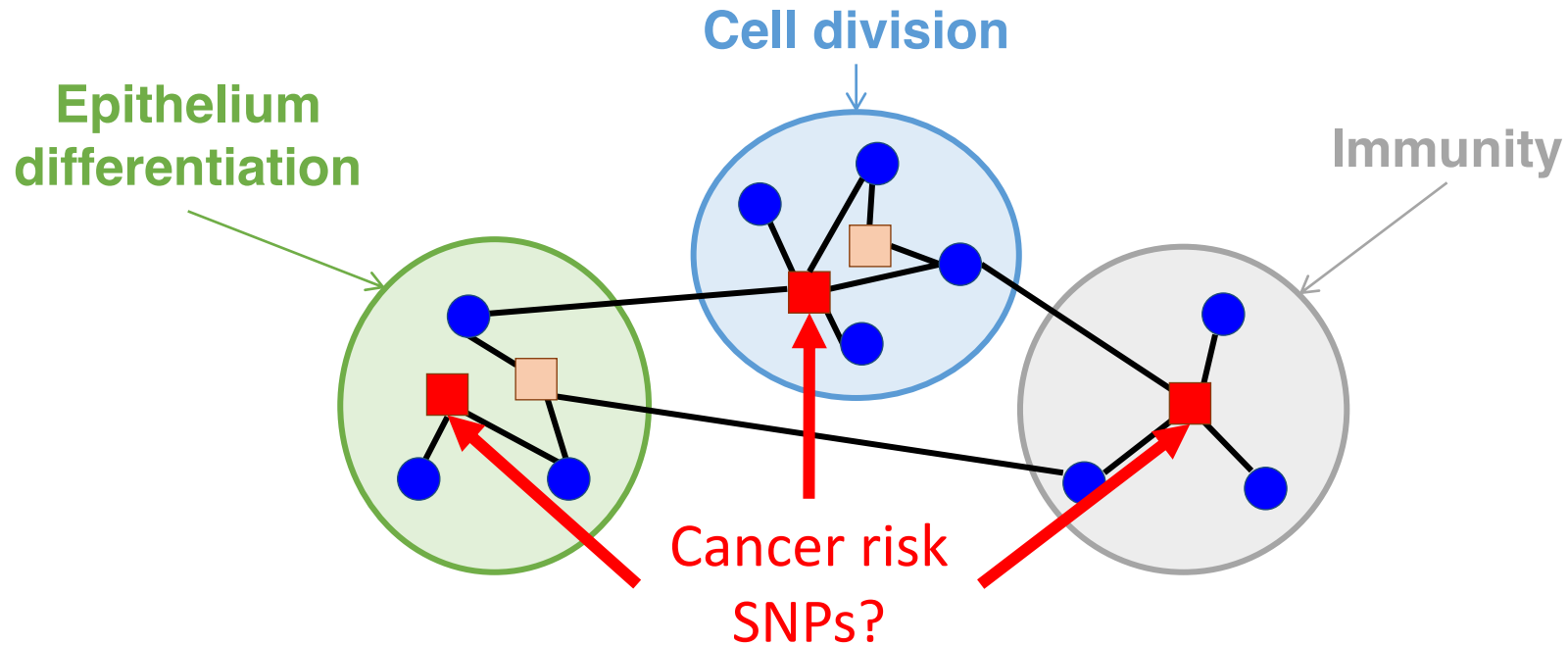
Genetic Architecture of Cancer Risk



Adapted from *Cancer Genetics Overview* (PQD®)

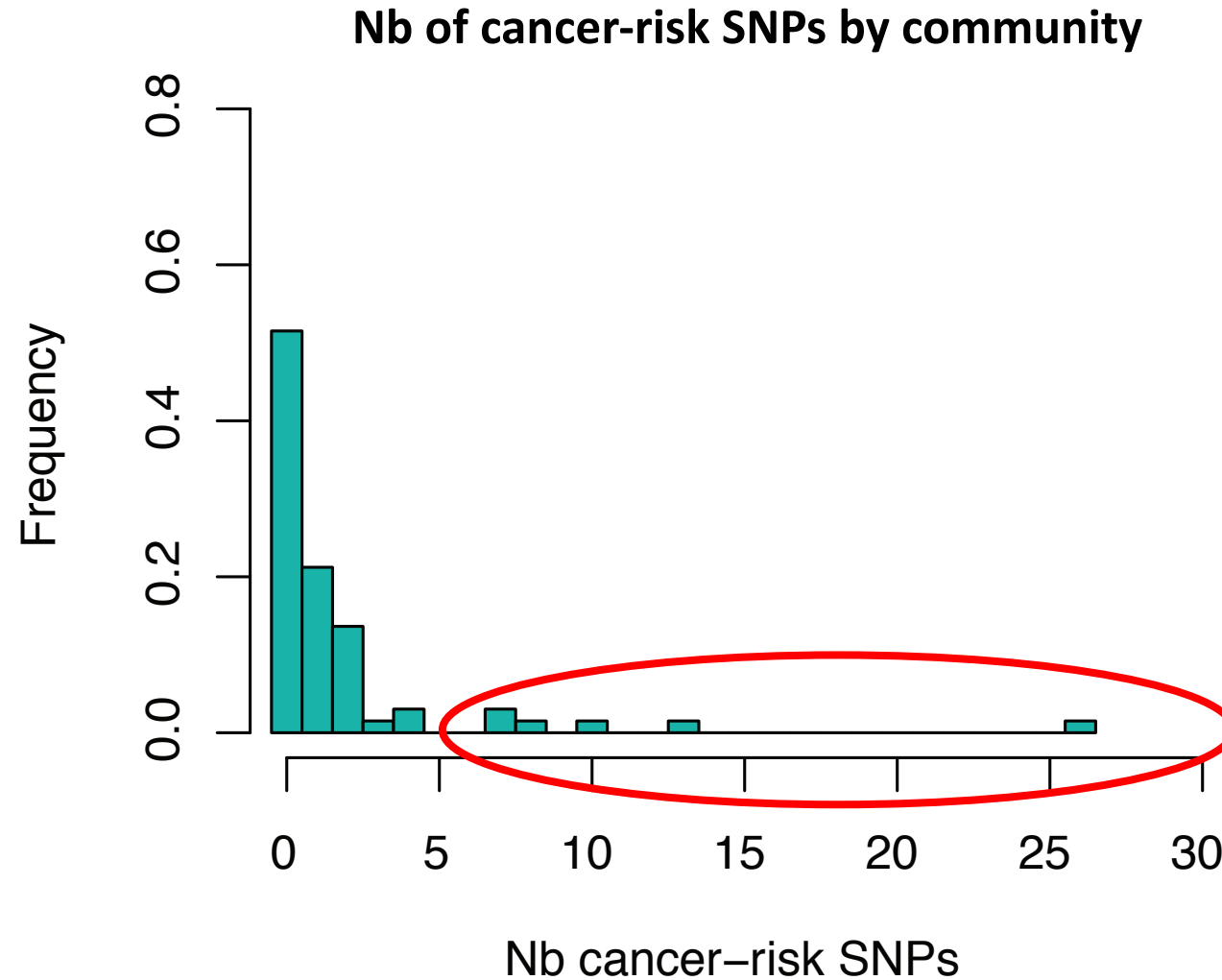
- 4,587 SNPs associated to 265 cancer-related traits
- 87% of cancer-risk SNPs with an odds ratio under 3

Mapping cancer-risk SNPs to eQTL networks



- **Where do the cancer-risk SNPs map in these networks?**
- **In which community are they located?**
- **What are their properties?**

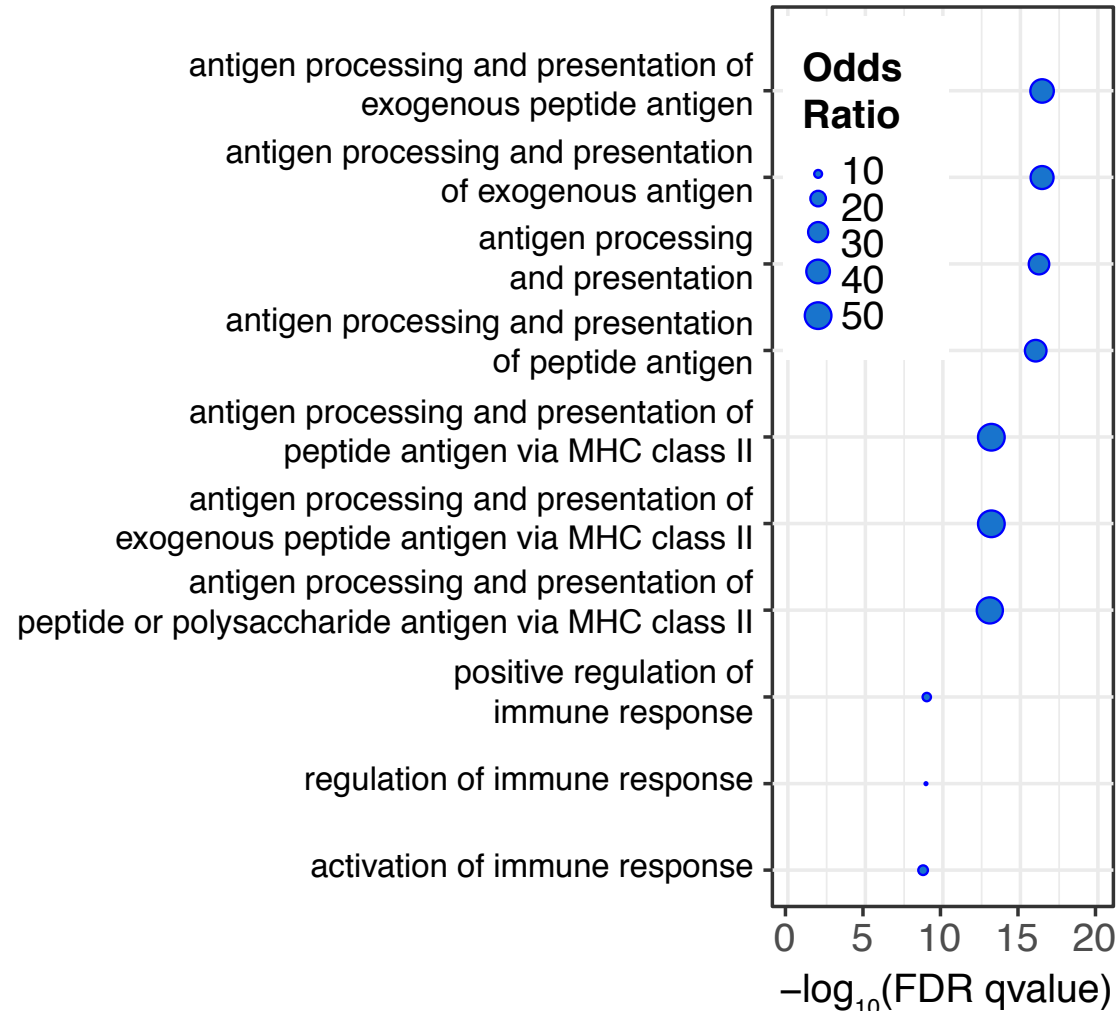
Mapping cancer-risk SNPs to eQTL networks



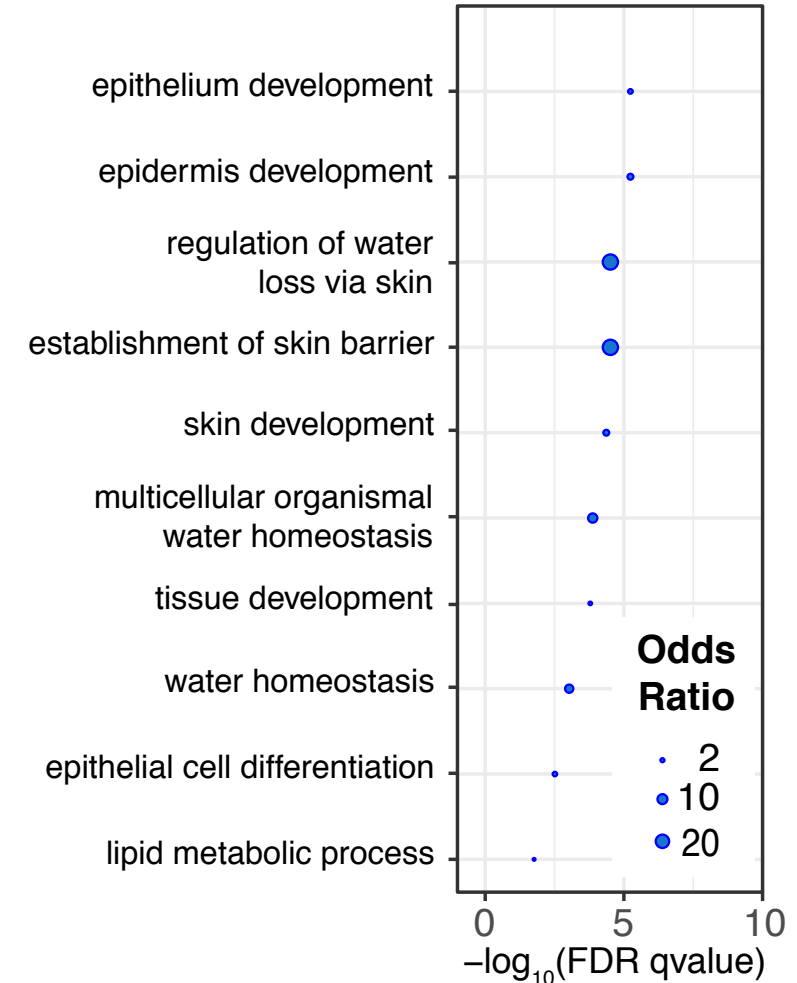
➤ **2-6 communities enriched in cancer risk SNPs by tissue**

Biological function of cancer-risk SNPs-enriched communities

Enriched in multiple cancer risk SNPs (all tissues)

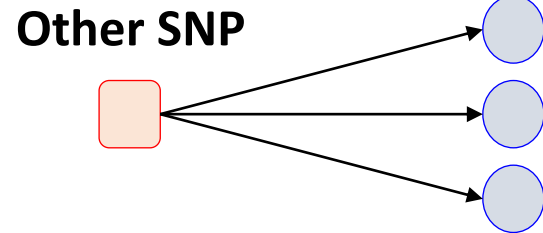
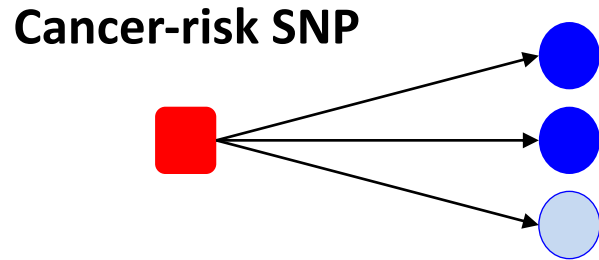


Enriched in breast cancer risk SNPs (skin)



Cancer-risk SNPs preferentially target oncogenes

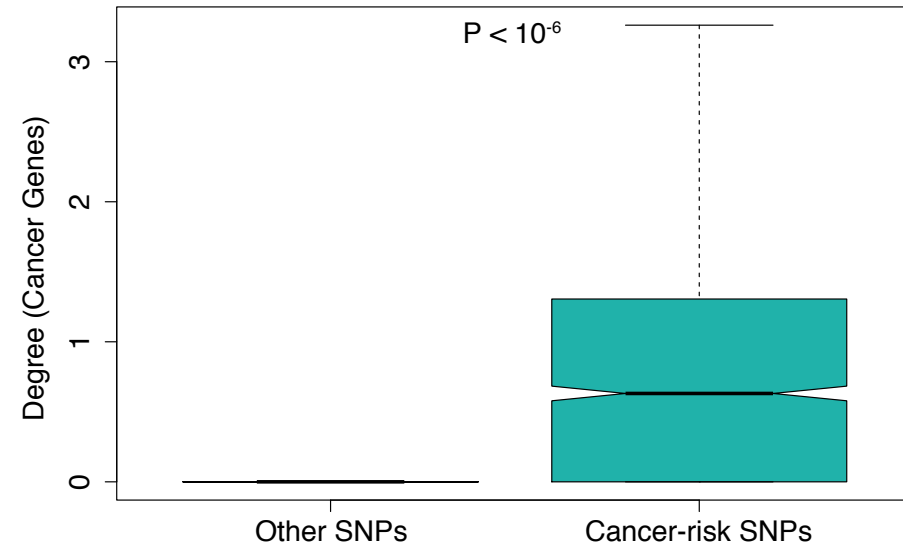
Genes targeted by cancer risk SNPs?



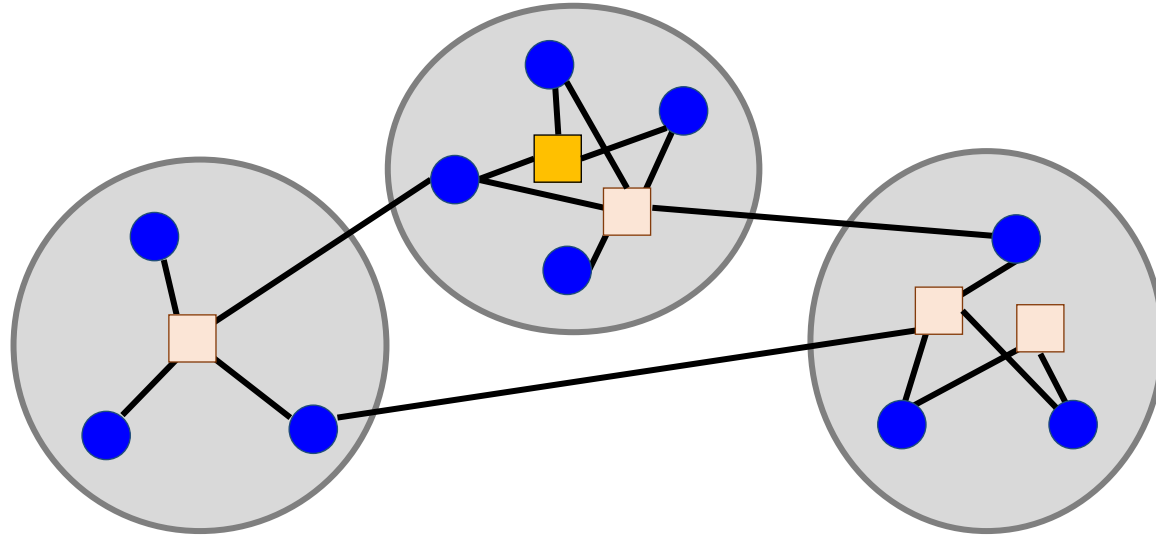
Others



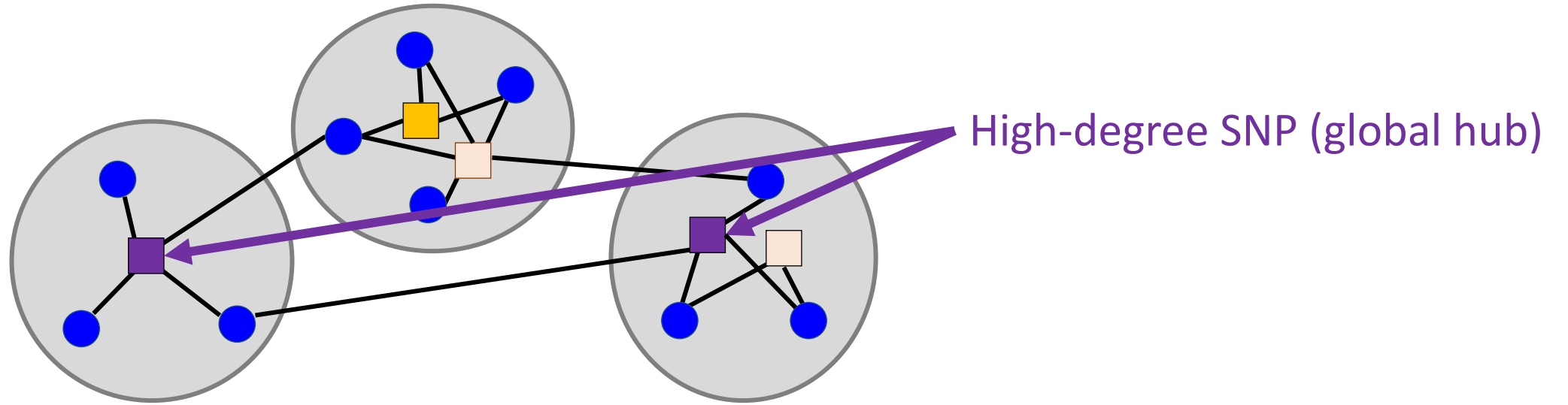
Oncogenes
Tumor suppressors



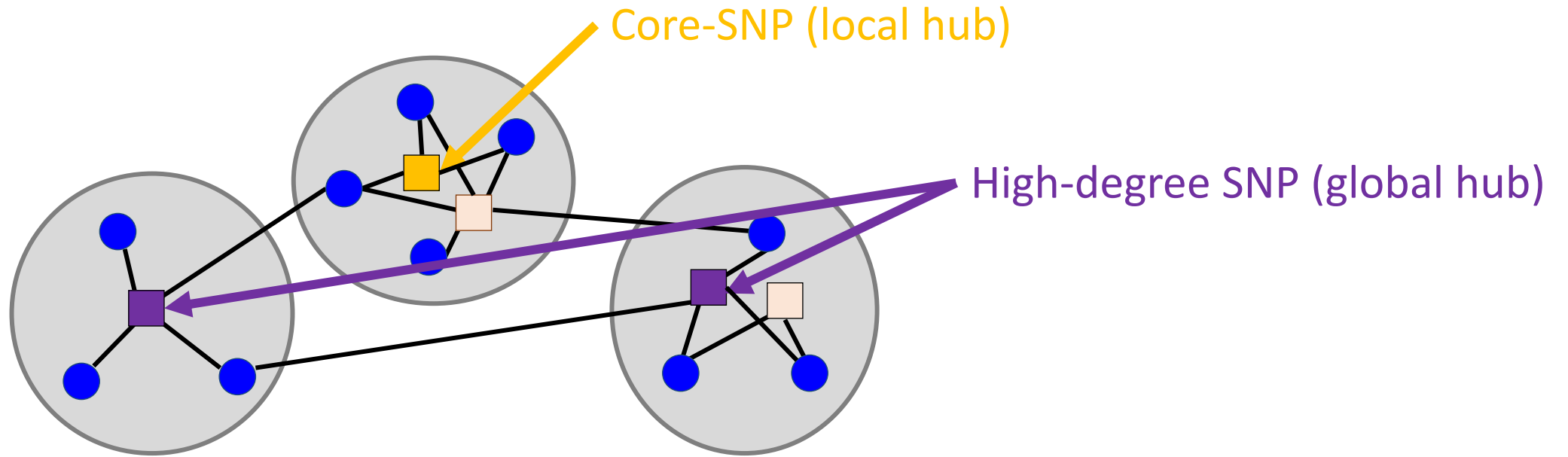
What are the properties of cancer-risk SNPs in the eQTL networks?



2 hypotheses: high-degree?



Or high core-score?

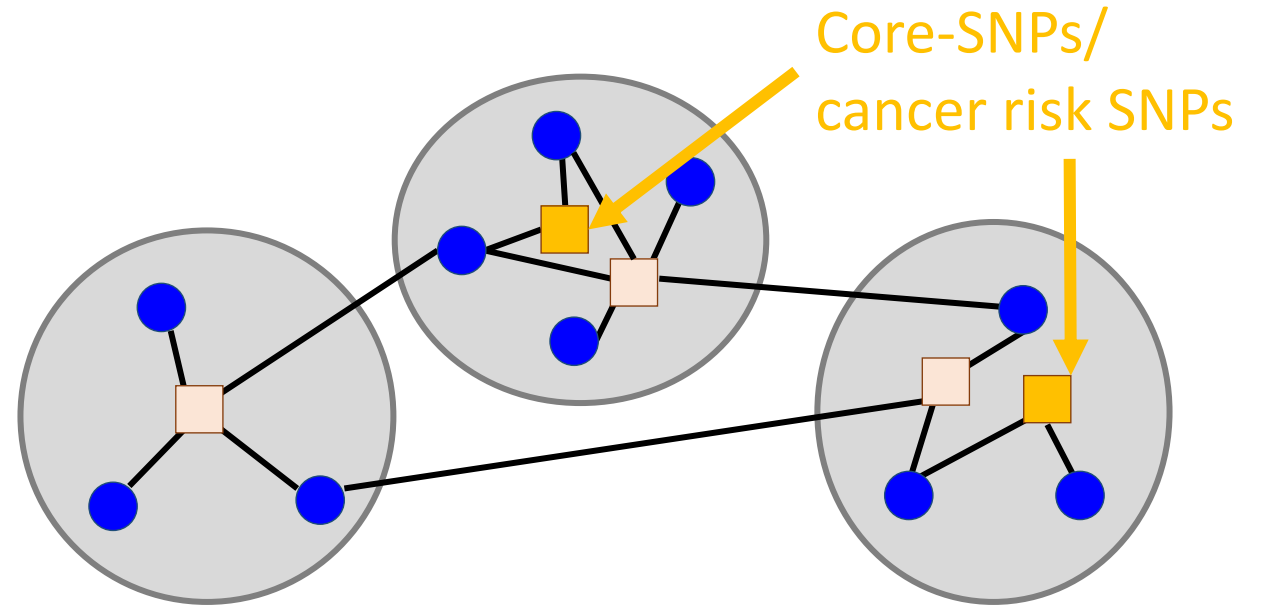
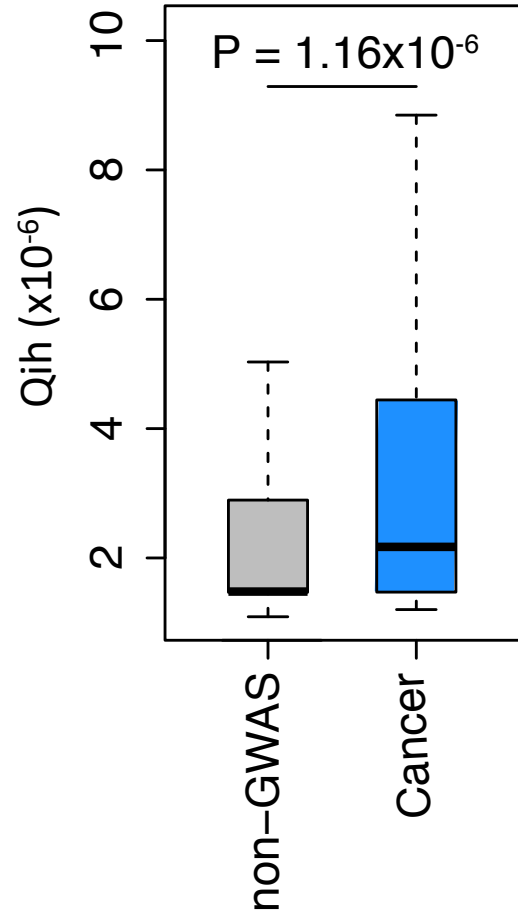


Core-score for SNP i in community h

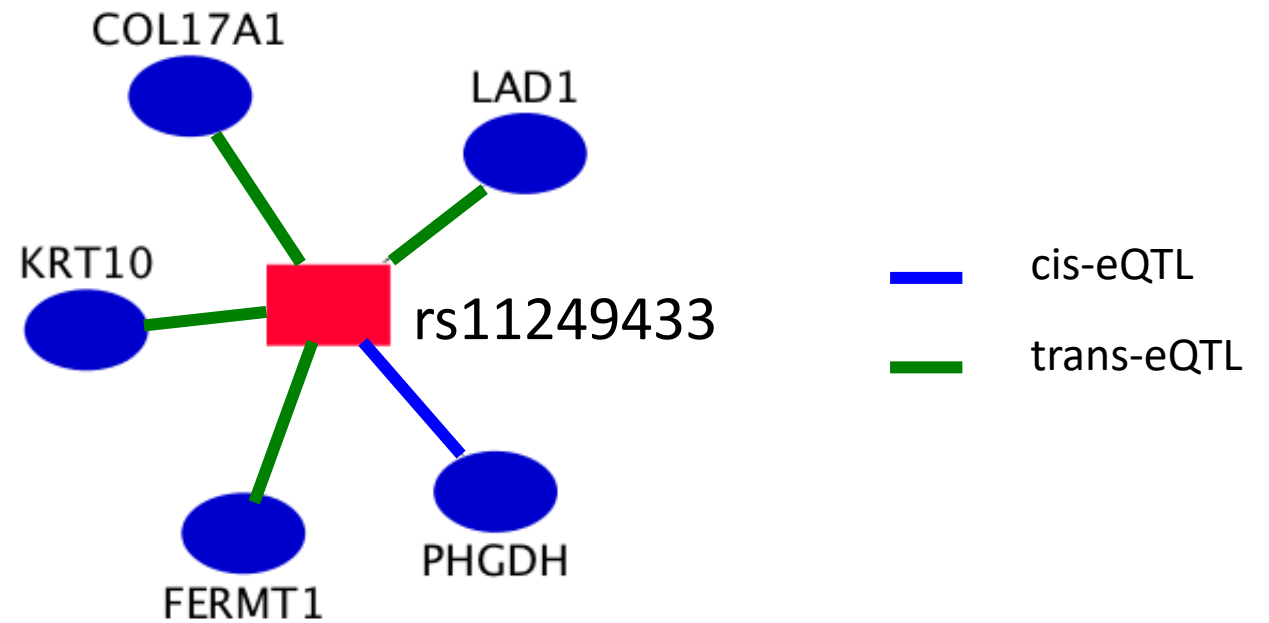
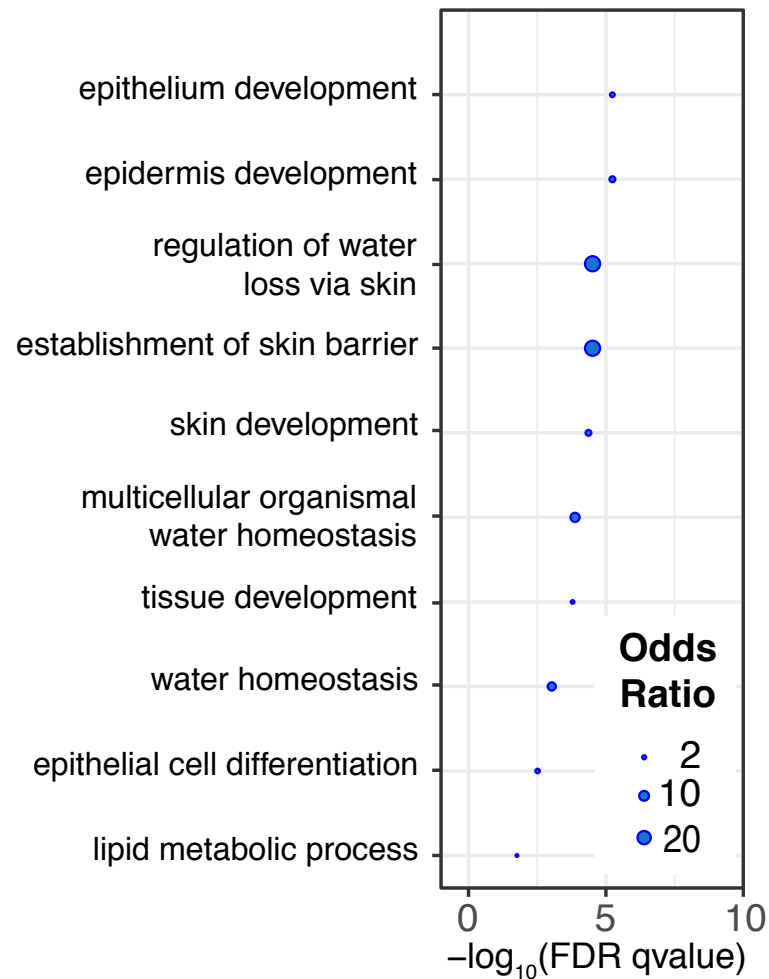
$$Q_{ih} = \frac{1}{m} \sum_j \left(\tilde{A}_{ij} - \frac{k_i d_j}{m} \right) \delta(C_i, h) \underbrace{\delta(C_j, h)}_{\text{Is gene } j \text{ in community } h?}$$

Observed – expected edge
between SNP i and gene j

Cancer-risk SNPs are local hubs



An example of breast cancer risk SNP



- rs11249433 is associated to breast cancer
- It targets genes that are deregulated in epithelium cancers (EMS).

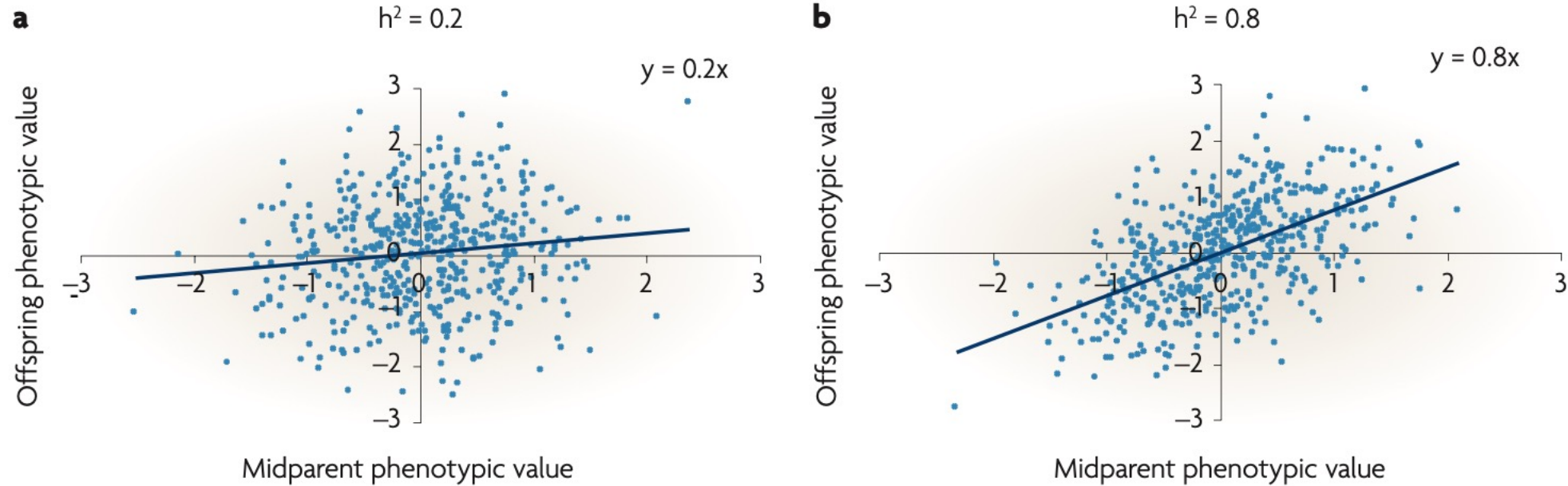
Research question: Understanding the genetic architecture of complex traits

- **Biological characterization of SNPs identified by GWAS: How do genetic variation influence a trait ? Which biological pathways are involved?**
- **Heritability: Which mutations most affect the trait? Where are they located in the regulatory network**
- **Natural Selection: How do such a complex trait evolve?**



Katherine Stone
(Bachelor student)

Genetic heritability



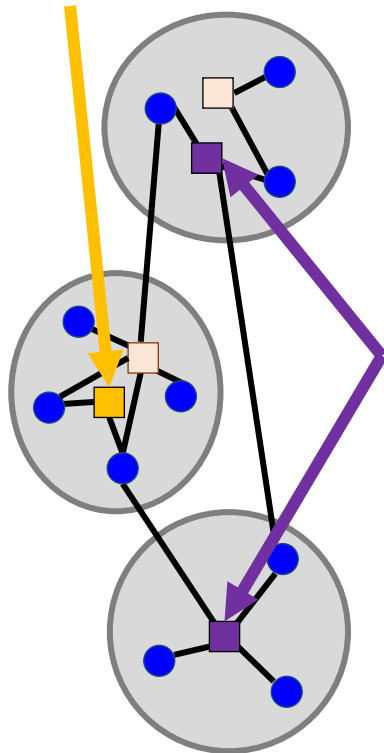
h^2 = heritability, proportion of variance explained by additive genetic value

Visscher *et al.* (2008) *Nature Review Genetics*

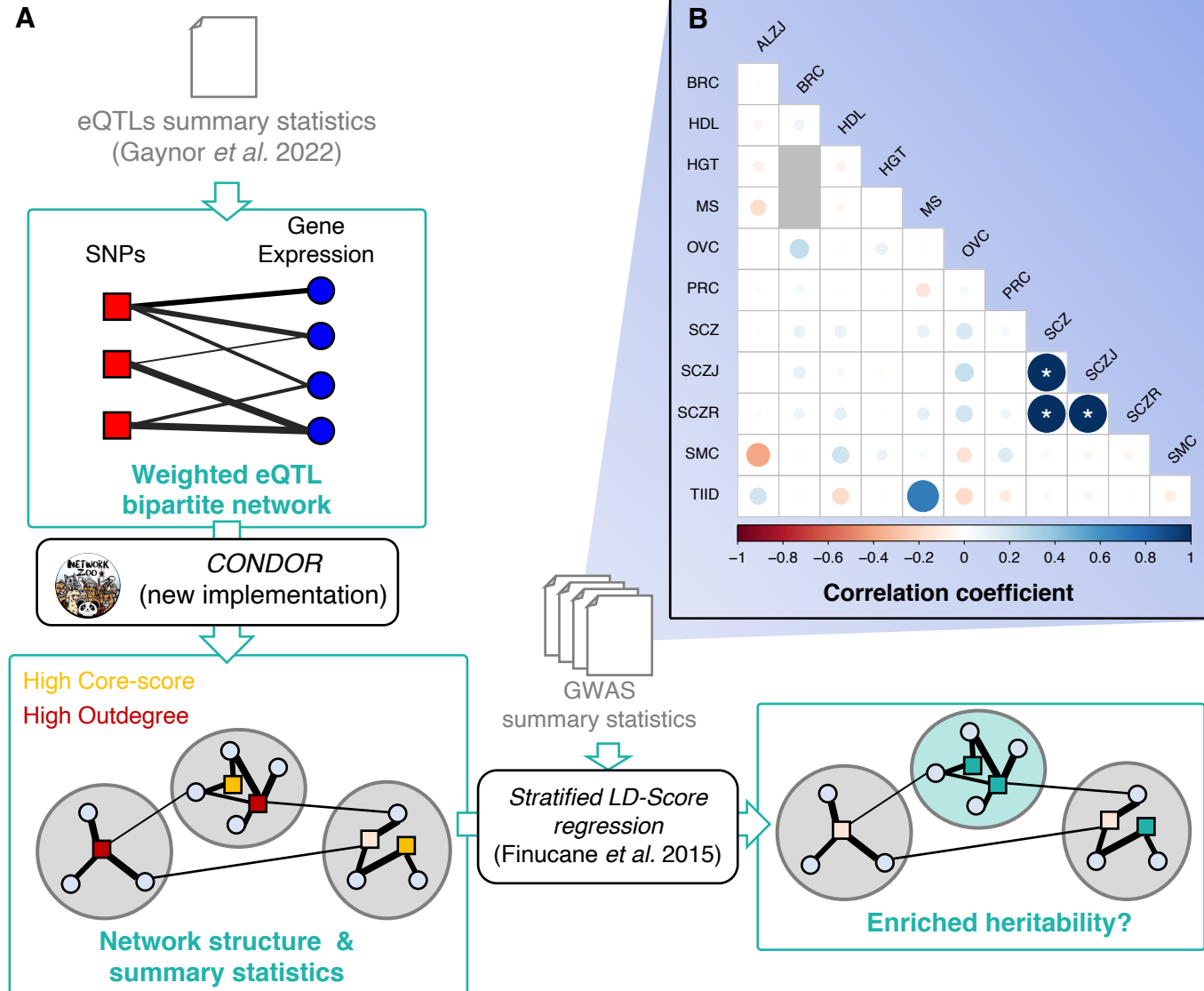
Genetic heritability in a network

Where is most of the heritability located?

Core-SNP
(local hub)



High-degree SNP
(global hub)



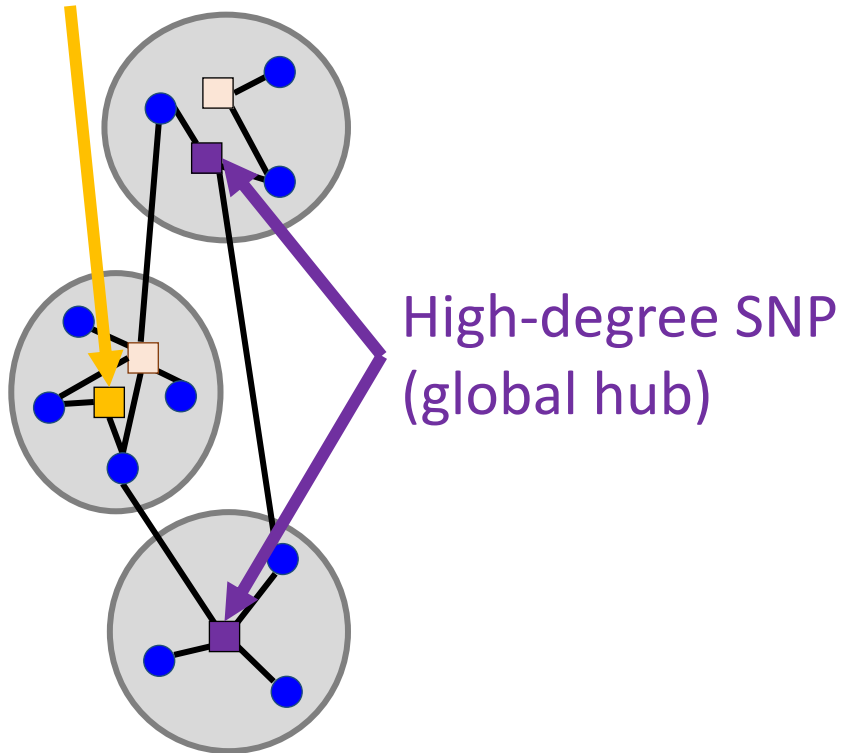
Traits

Trait or disease	Abbreviation	Genetic structure	Estimated genetic heritability
Alzheimer's disease	ALZJ	Oligogenic	58-90%
Breast Cancer	BRC	Polygenic	31%
HDL	HDL	Polygenic	40-60%
Height	HGT	Omnigenic	50%
Multiple Sclerosis	MS	Polygenic	64%
Ovarian Cancer	OVC	Polygenic	39%
Prostate Cancer	PRC	Polygenic	57%
Schizophrenia	SCZP	Polygenic	79%
Schizophrenia	SCZR	Polygenic	79%
Schizophrenia	SCZ	Polygenic	79%
Smoking Cessation	SMC	Polygenic	75%
Type 2 diabetes	TIID	Oligogenic	25-72%

Genetic heritability in a network

Is there a link between the amount of heritability explained and the network topology?

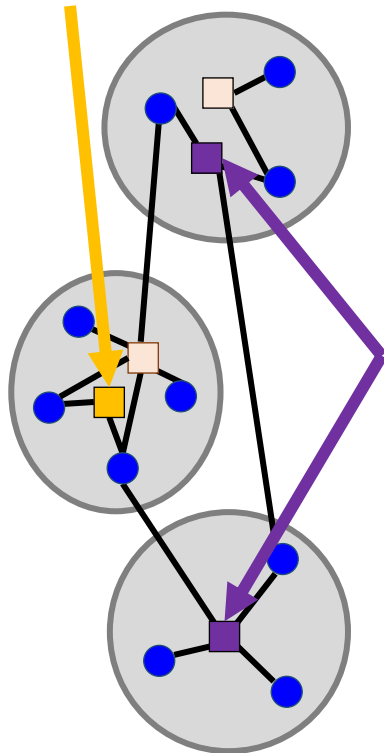
Core-SNP
(local hub)



Genetic heritability in a network

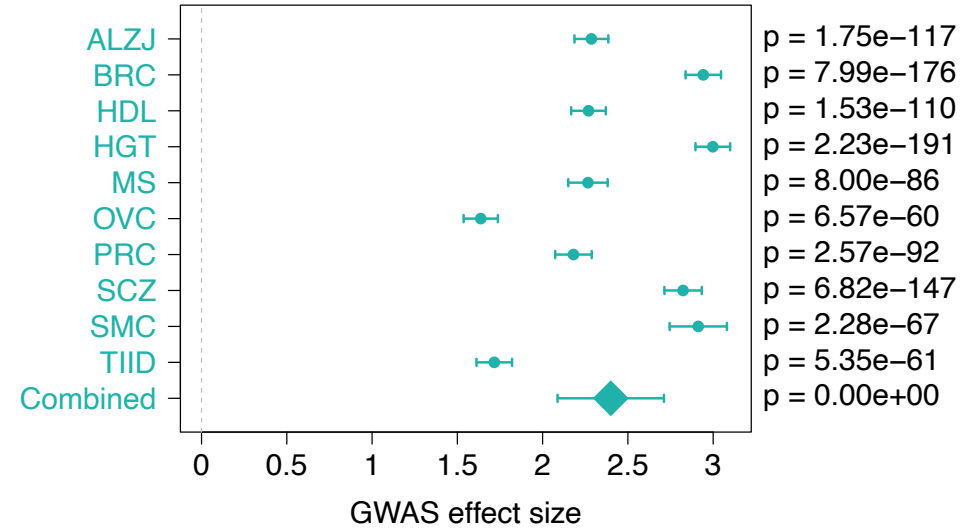
Is there a link between the amount of heritability explained and the network topology?

Core-SNP
(local hub)

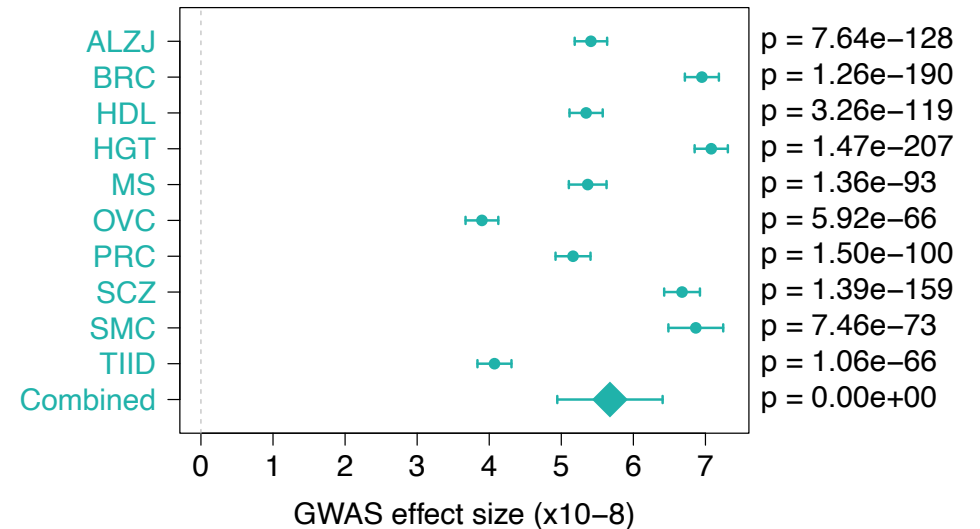


High-degree SNP
(global hub)

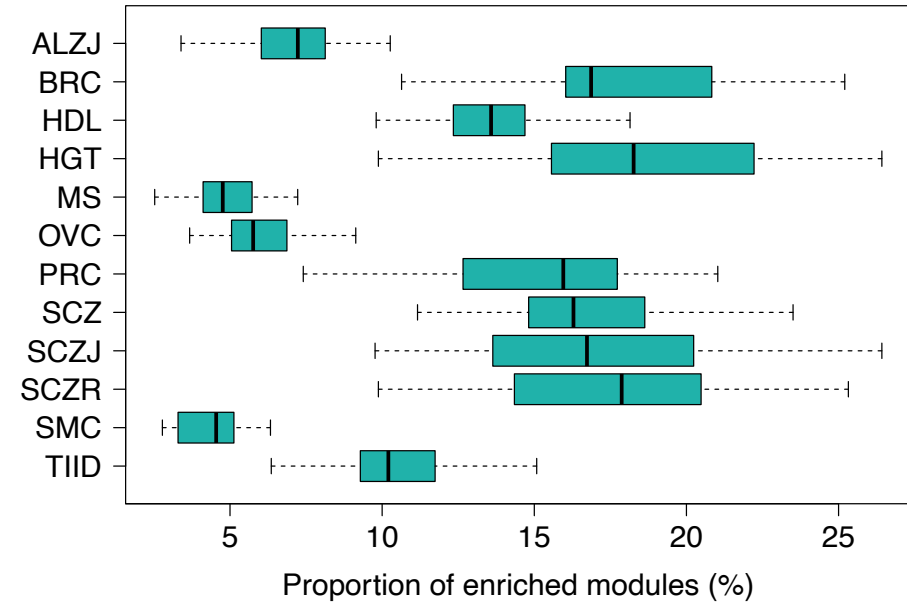
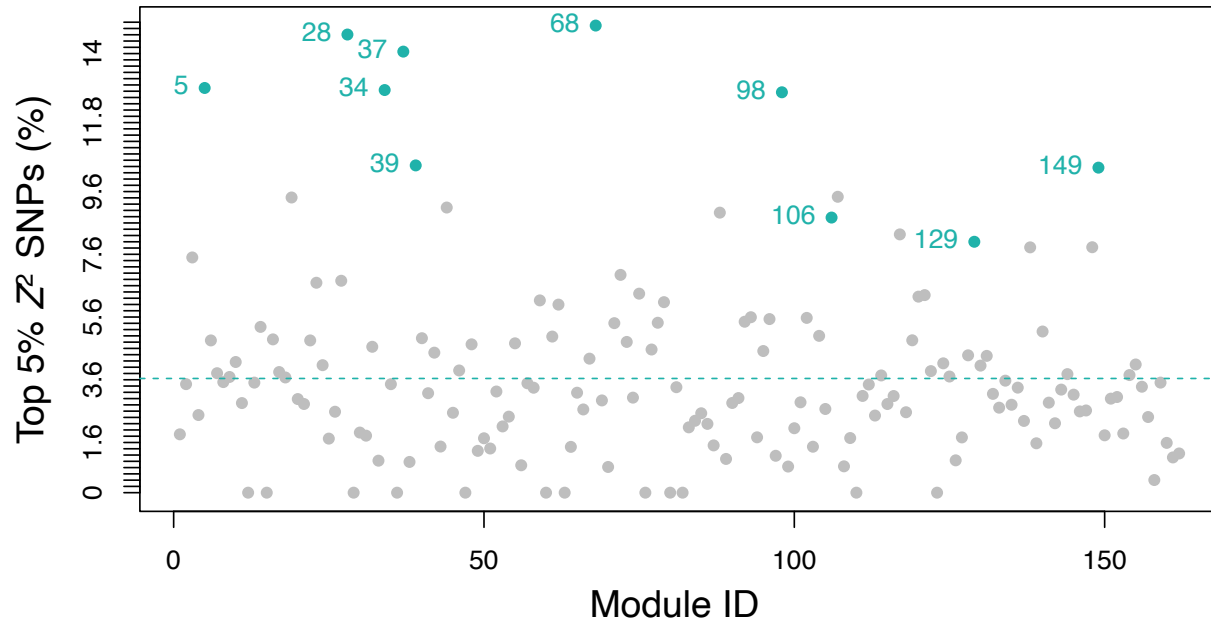
Global hubs



Local hubs



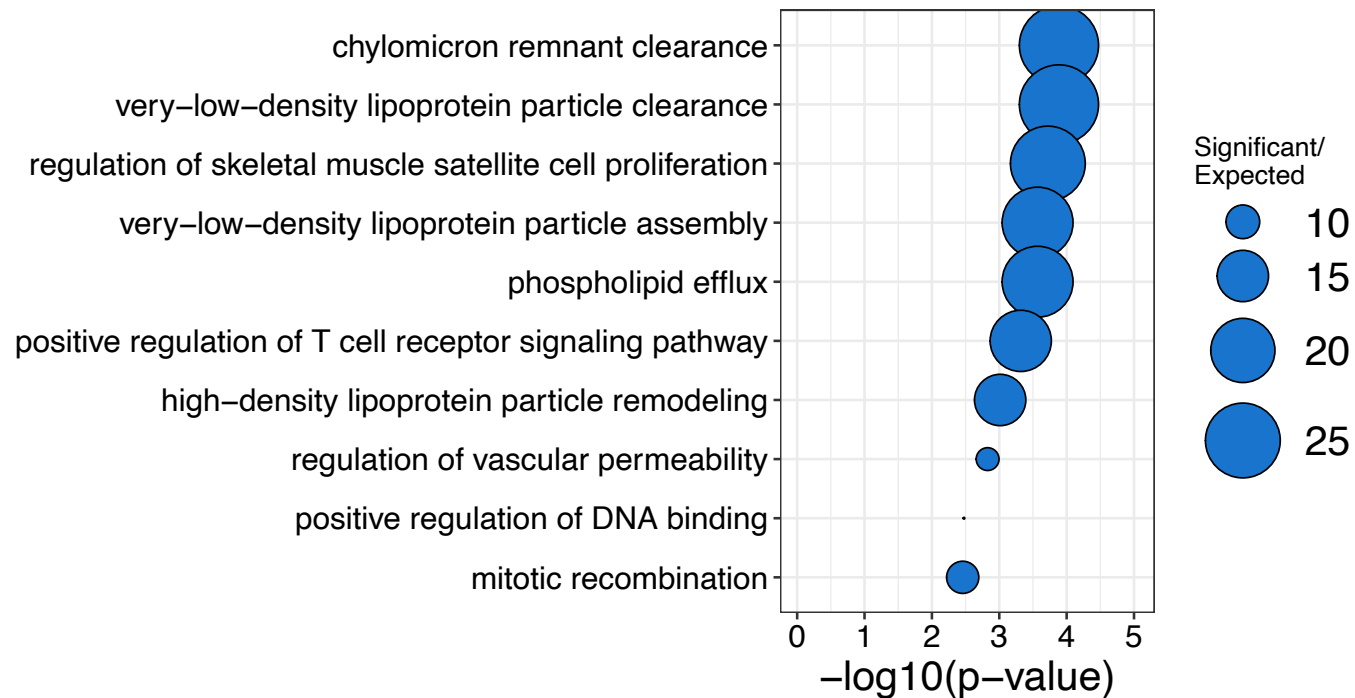
Genetic heritability among communities



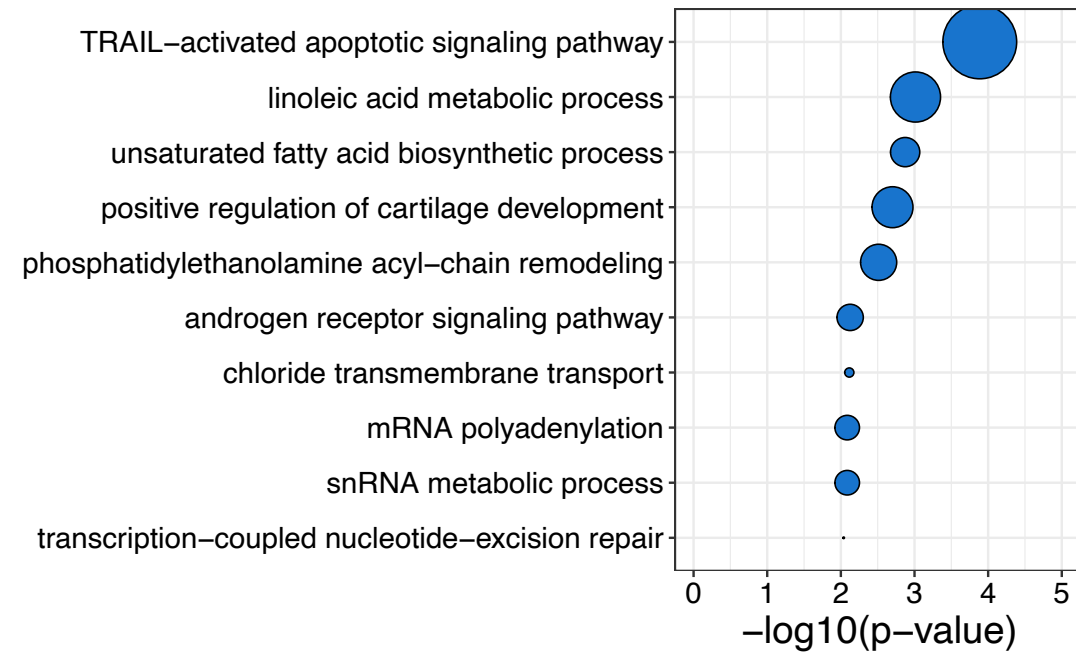
➤ Most of the heritability is concentrated in a few communities

Genetic heritability and biological functions

**Adipose Visceral Omentum community
HDL levels**



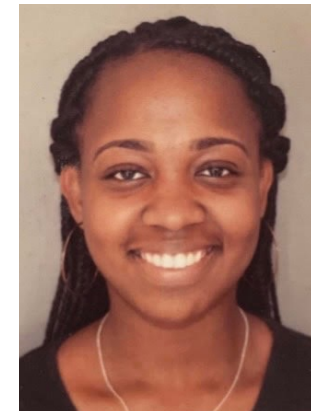
**Colon sigmoid community
prostate cancer**



➤ Genetic heritability is concentrated in tissue-specific, biologically relevant communities

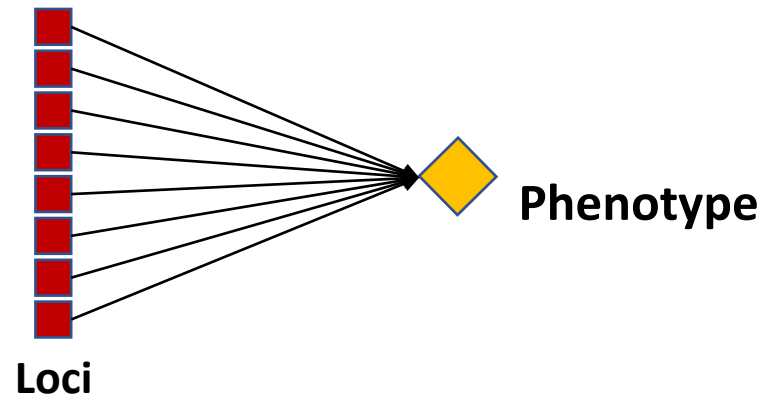
Research question: Understanding the genetic architecture of complex traits

- **Biological characterization of SNPs identified by GWAS: How do genetic variation influence a trait ? Which biological pathways are involved?**
- **Heritability: Which mutations most affect the trait? Where are they located in the regulatory network**
- **Natural Selection: How do such a complex trait evolve?**



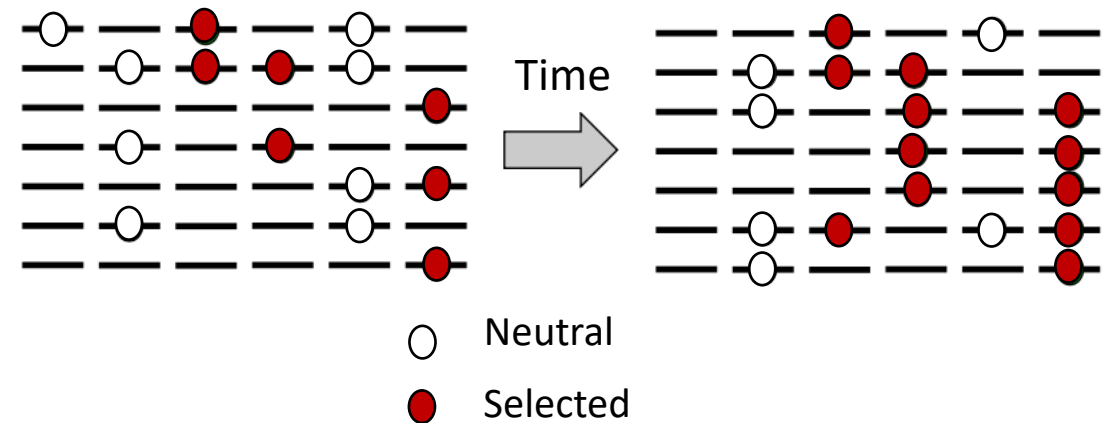
Rosanne Phebe (M1)

Polygenic adaptation: from phenotype to molecules



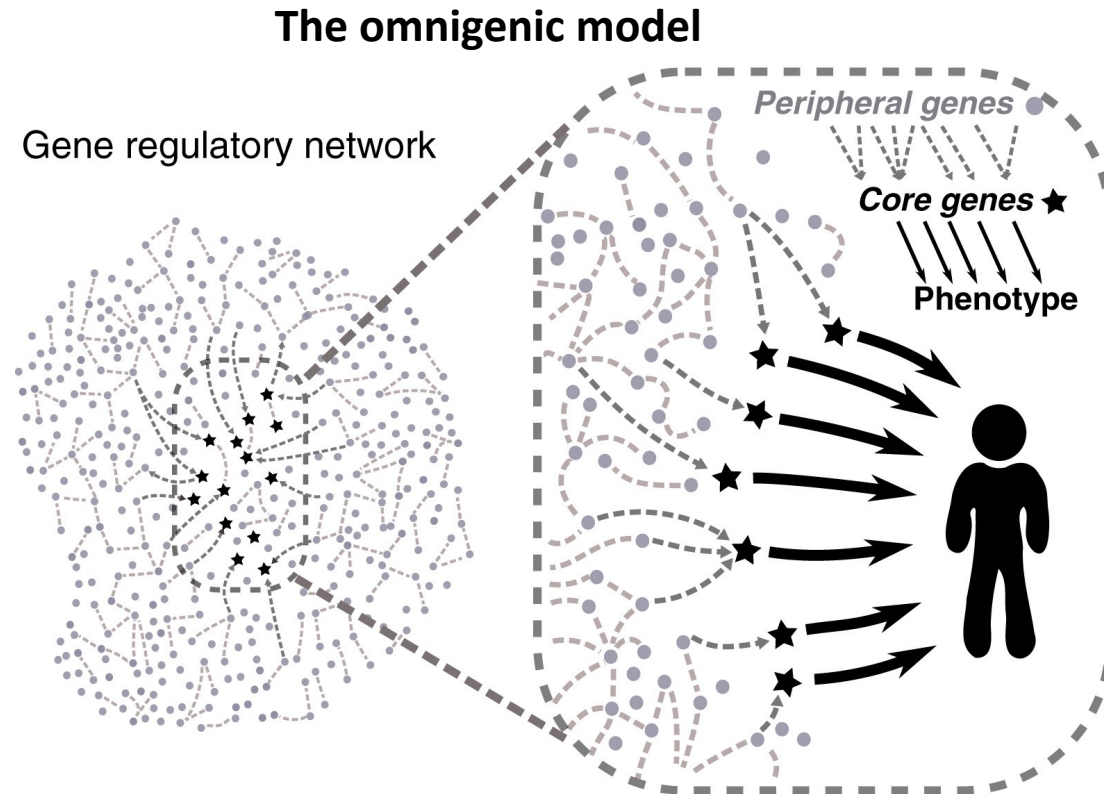
Adapted from Hallgrímsson *et al.*, *PLoS Genetics*, 2014

Polygenic adaptation at the molecular level



- The omnigenic model
- Pleiotropic loci

The genetic architecture of complex traits may limit adaptation



Boyle *et al.*, *Cell*, 2017
Liu *et al.*, *Cell*, 2019

➤ **High pleiotropy**

But... Many examples of polygenic adaptation

PLOS GENETICS

Arabidopsis & Light

RESEARCH ARTICLE

Polygenic adaptation of rosette growth in *Arabidopsis thaliana*

Benedict Wieters¹, Kim A. Steige¹, Fei He¹, Evan M. Koch^{2,3}, Sebastián E. Ramos-Onsins⁴, Hongya Gu⁵, Ya-Long Guo⁶, Shamil Sunyaev^{2,3}, Juliette de Meaux^{1*}

PLOS BIOLOGY

Drosophila & Temperature

RESEARCH ARTICLE

Genetic redundancy fuels polygenic adaptation in *Drosophila*

Neda Barghi¹, Raymond Tobler^{1,2ab}, Viola Nolte¹, Ana Marija Jakšić^{1,2ab}, François Mallard^{1bc}, Kathrin Anna Otte¹, Marlies Dolezal^{1,3}, Thomas Taus^{1,2}, Robert Kofler¹, Christian Schlötterer^{1*}

Maize & Altitude

Molecular Parallelism Underlies Convergent Highland Adaptation of Maize Landraces

Li Wang^{*,1,2,3}, Emily B. Josephs^{4,5}, Kristin M. Lee³, Lucas M. Roberts², Rubén Rellán-Álvarez^{6,7}, Jeffrey Ross-Ibarra^{*,3,8} and Matthew B. Hufford^{*,2}

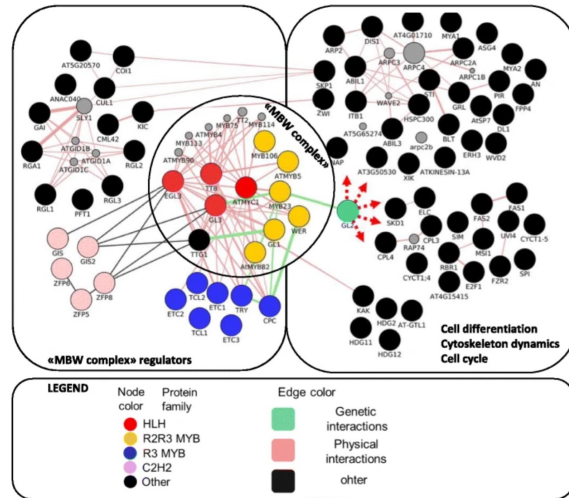
Humans & Immunity

Evidence for Polygenic Adaptation to Pathogens in the Human Genome

Josephine T. Daub^{*,1,2}, Tamara Hofer^{1,2}, Emilie Cutiver¹, Isabelle Dupanloup^{1,2}, Lluís Quintana-Murci^{3,4}, Marc Robinson-Rechavi^{1,2,5} and Laurent Excoffier^{*,1,2}

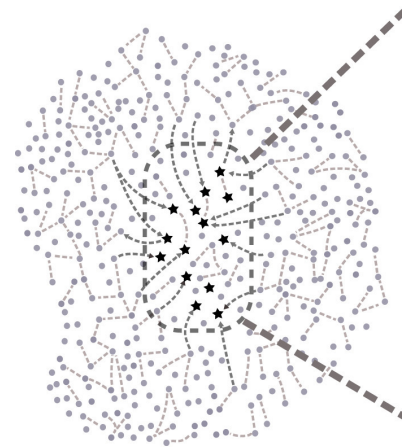
Structuration & Tissue-specificity & Redundancy

Structured GRN (Arabidopsis)

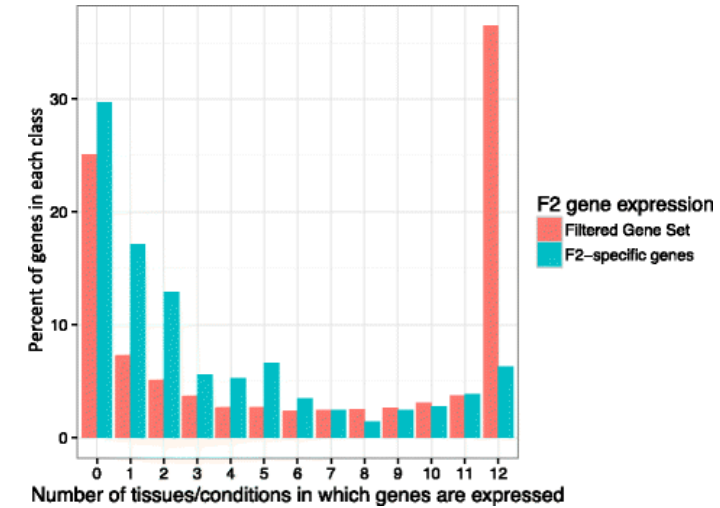


Doroshkov *et al.*, *BMC Plant Biology*, 2019

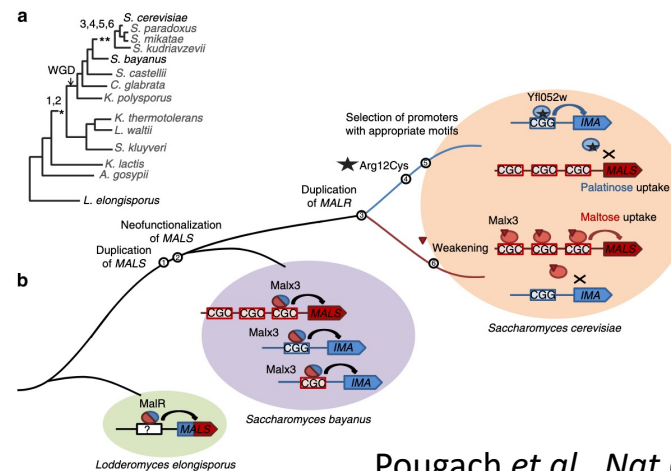
The omnigenic model



Tissue-specific gene expression (maize)



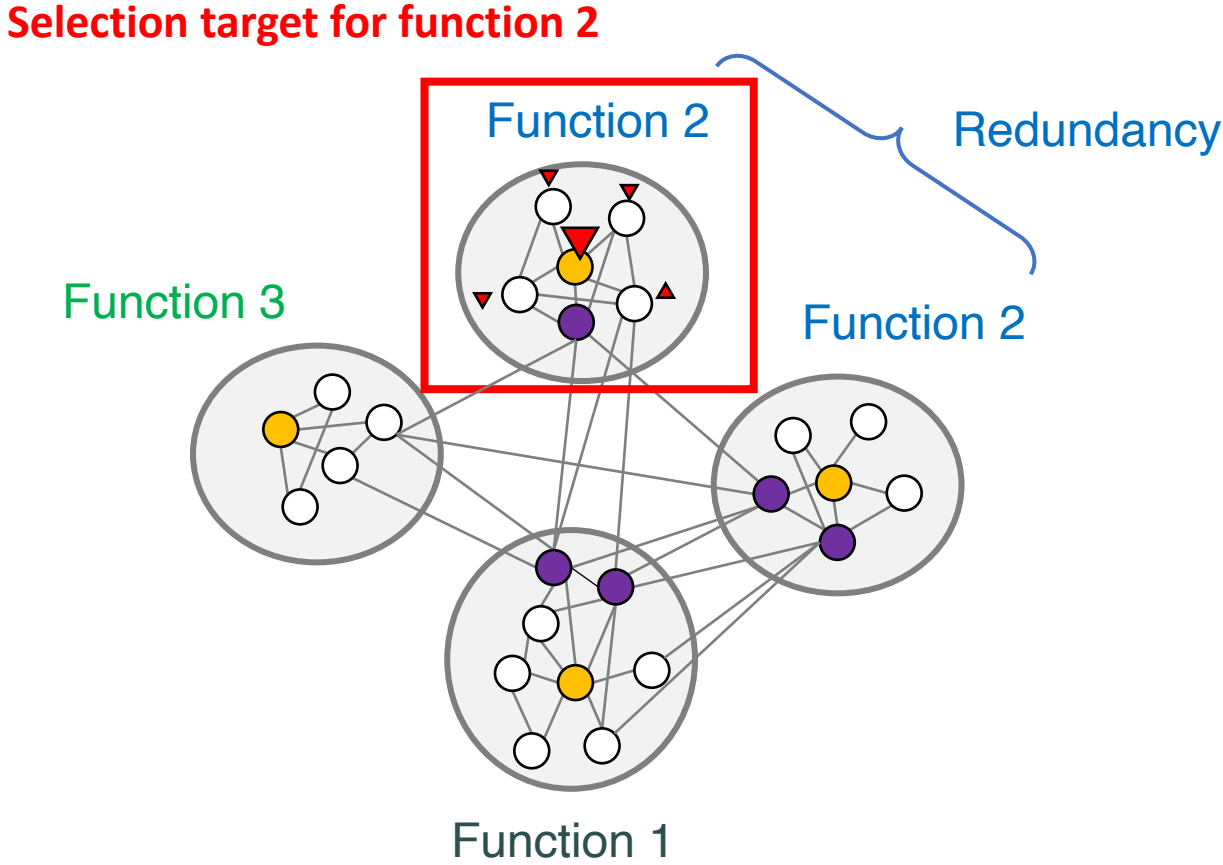
Darracq *et al.*, *BMC Genomics*, 2018



Pougach *et al.*, *Nat Comm*, 2014

Redundancy and neofunctionalization (yeast)

Proposed model: a major role for redundancy and node topology in evolvability



Take-home message

- Expression quantitative trait loci (eQTL) bipartite networks can help functionally annotating SNPs associated with complex traits
- Cancer-risk SNPs are :
 - located preferentially in local hubs and communities related to immune (several cancers), or tissue-specific (cancer-specific) functions.
 - impacting the expression of oncogenes and tumor suppressor genes
- Most of complex trait heritability is :
 - Located in local and global hubs,
 - Concentrated in a few, tissue-specific and biologically relevant communities.
- To go further:
 - eQTL network structure may help us understand how complex trait evolve despite a high level of pleiotropy