Q&A

(3 points each; total 15 points; indicate all correct answers; only a correct set of answers given to a question adds up your score with 3 points – otherwise no marks are gained for that question; questions are quite similar to exam multiple choice questions)

1. Select the correct statement(s)
   a. Pleiotropic genes exhibit single phenotype
   b. Pleiotropy is caused by a gene that has multiple phenotypic effect
   c. Lethal genes causes the appearance of ancestral characters

2. Studies on eQTLs aim to
   a. discover genetic variants that explain variation in gene expression levels;
   b. discover environmental factors that implact a quantitative trait;
   c. discover exposures to a quantitiative trait.

3. Bayesian colocalizationa analysis, aiming to estimate the posterior probability that the same variant is causal in both a GWAS and eQTL study while accounting for the uncertainty of LD, may derive posterior support for the hypothesis
   a. no causal variants for either trait;
   b. two distinct causal variants, one for each trait;
   c. a single causal variant common to both traits (co-localization).

4. Which statement(s) about interactions in statistical models is(are) false:
   a. The scale on which the outcome variable is modelled matters
   b. Spurious interactions can arise when explanatory variables are correlated and there are unmodeled nonlinear relationships
   c. Interactions can be nonlinear such that the interaction term x_1x_2 will not capture all interactions of biological interest

5. Alice performed an expression quantitative trait locus (eQTL) analysis using left ventricular myocardium samples of 5204 individuals. She identified many statistically significant associations, amongst which an eQTL for the SNP rs7612445 and the gene *GNB4* (**Figure 1**). Den Hoed et al. published the results of a genome-wide association study on heart rate in 2013, in which they identified an associated locus on chromosome 3 (**Figure 2**). The statistically most strongly associated SNP is rs7612445. This SNP is associated with increased heart rate (effect allele = C).

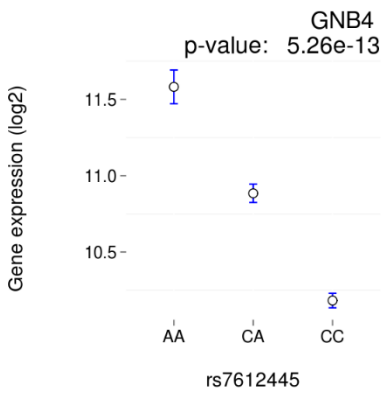


Figure 1: A mean and standard-error plot for an eQTL of the gene *GNB4*.

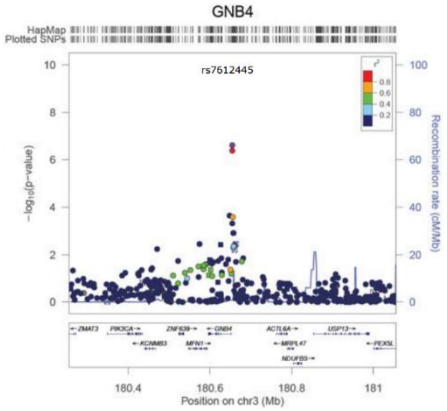


Figure 2: Overview of a heart rate associated locus.

Combining these data, which of the following statements are true:

(I) The more copies of the C allele a person possesses, the higher their heart rate and the higher their expression of *GNB4* in the left ventricular myocardium.

(II) The SNP rs7612445 is the causal SNP underlying increased heart rate.

a. Only statement I is true.

b. Both statements are true.

c. Both statements are false.

Literature style

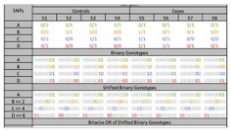
(15 points; Papers are organized around the following themes. Each group selects one paper. Further instructions: same as homework assignment 1)

Interactions

Research
Open Access
05 Aug 2021
Scientific Reports
Volume: 11, P: 1-12

Fast and accurate exhaustive higher-order epistasis search with BitEpi

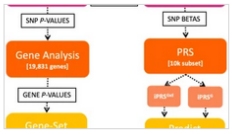
Arash Bayat, Brendan Hosking ... Denis C. Bauer



Research
Open Access
22 Mar 2021
Translational Psychiatry
Volume: 11, P: 1-13

Genome-wide gene-environment interactions in neuroticism: an exploratory study across 25 environments

Josefin Werme, Sophie van der Sluis ... Christiaan A. de Leeuw

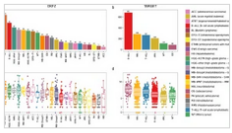


Research
Open Access
06 Oct 2021
Communications Biology
Volume: 4, P: 1-12

A systematic analysis of genetic interactions and their underlying biology in childhood cancer

Josephine Daub, Saman Amini, et al. developed a genetic interaction map of childhood cancers by analyzing over 2,500 tumors from 23 types of childhood cancer. Their results provide a valuable resource for investigating the biological underpinnings of pediatric cancer.

Josephine T. Daub, Saman Amini ... Patrick Kemmeren



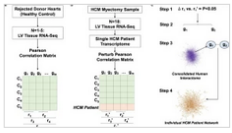
Networks

Research
Open Access
08 Feb 2021
Nature Communications
Volume: 12, P: 1-11

Individualized interactomes for network-based precision medicine in hypertrophic cardiomyopathy with implications for other clinical pathophenotypes

Understanding patient-specific pathobiological pathways is a critical step for advancing precision medicine. Here the authors show that individualized protein-protein interaction networks provide key insight on patient-level pathobiology and clinically relevant pathophenotypic characteristics in a complex disease.

Bradley A. Maron, Rui-Sheng Wang ... Joseph Loscalzo



Research
Open Access
14 Aug 2019
Scientific Data
Volume: 6, P: 1-18

HENA, heterogeneous network-based data set for Alzheimer's disease

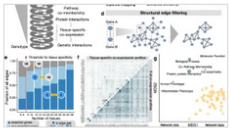
Elena Sügis, Jerome Dauvillier ... Hedi Peterson

Research
Open Access
09 Nov 2021
Nature Communications
Volume: 12, P: 1-15

Network analysis reveals rare disease signatures across multiple levels of biological organization

Despite the clear causal relationship between genotype and phenotype in rare diseases, identifying the pathobiological mechanisms at various levels of biological organization remains a practical and conceptual challenge. Here, the authors introduce a network approach for evaluating the impact of rare gene defects across biological scales.

Pisanu Buphamalai, Tomislav Kokotovic ... Jörg Menche

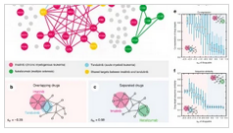


Research
Open Access
13 Mar 2019
Nature Communications
Volume: 10, P: 1-11

Network-based prediction of drug combinations

Combination therapy holds great promise, but discovery remains challenging. Here, the authors propose a method to identify efficacious drug combinations for specific diseases, and find that successful combinations tend to target separate neighbourhoods of the disease module in the human interactome.

Feixiong Cheng, István A. Kovács & Albert-László Barabási



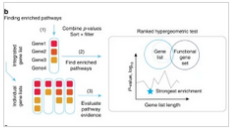
Multi-omics

Research
Open Access
05 Feb 2020
Nature Communications
Volume: 11, P: 1-16

Integrative pathway enrichment analysis of multivariate omics data

Multi-omics datasets pose major challenges to data interpretation and hypothesis generation owing to their high-dimensional molecular profiles. Here, the authors develop ActivePathways method, which uses data fusion techniques for integrative pathway analysis of multi-omics data and candidate gene discovery.

Marta Paczkowska, Jonathan Barenboim ... Jüri Reimand



Research
Open Access
16 Apr 2021
Nature Communications
Volume: 12, P: 1-11

multiSLIDE is a web server for exploring connected elements of biological pathways in multi-omics data

The integration and interpretation of different omics data types is an ongoing challenge for biologists. Here, the authors present a web-based, interactive tool called multiSLIDE for the visualization of protein, phosphoprotein, and RNA data presented as interlinked heatmaps.

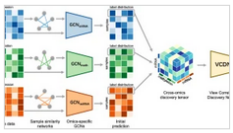
Soumita Ghosh, Abhik Datta & Hyungwon Choi

Research
Open Access
08 Jun 2021
Nature Communications
Volume: 12, P: 1-13

MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification

Our understanding of human disease can be improved by integrating the abundance of high throughput biomedical data. Here, the authors use deep learning methods successfully used on images to integrate various types of omics data to improve patient classification and identify disease biomarkers.

Tongxin Wang, Wei Shao ... Kun Huang



Research
Open Access
05 Jan 2021
Nature Communications
Volume: 12, P: 1-12

Benchmarking joint multi-omics dimensionality reduction approaches for the study of cancer

Advances in omics technology have resulted in the generation of multi-view data for cancer samples. Here, the authors compare dimensionality reduction techniques using simulated and TCGA data and identify the features of the methods with superior performance.

Laura Cantini, Pooya Zakeri ... Anaïs Baudot