

HPO Enrichment Analyze

Weize Xu

Huazhong Agriculture University

March 28, 2019

Outline

1 Data

2 Method

3 Implementation

4 Reference

HPO dataset

The Human Phenotype Ontology (HPO)¹ provides a standardized vocabulary of phenotypic abnormalities encountered in human disease. Each term in the HPO describes a phenotypic abnormality, such as Atrial septal defect.

¹Sebastian Köhler et al. "Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources". In: *Nucleic acids research* 47.D1 (2018), pp. D1018–D1027.

Enrichment



40 m&m's in the bag

Red	=	13%
Yellow	=	14%
Orange	=	21%
Green	=	20%
Brown	=	12%
Blue	=	21%



2

²<https://www.youtube.com/watch?v=udyAvvaMjfM>

Fisher's exact test

	succeed	failed	Total
studying	i	k-i	k
non-studying	m-i		N-k
Total	m	N-m	N

The probability of hypergeometric distribution.

$$p(x = i) = \frac{\binom{m}{i} \binom{N-m}{k-i}}{\binom{N}{k}} \quad (1)$$

$$pvalue = \sum_{x=i}^k p(x) \quad (2)$$

Implementation

A Python package for HPO enrichment analyze.

<https://github.com/Nanguage/BioTMCourse/tree/master/HPO%20enrich>

An HPO GSEA example

SAGD³ gene samples, SAGD_00055(human hypothalamus tissue).

https://nanguage.github.io/examples/hpo_enrich/example_sagd_00055.html

³Meng-Wei Shi et al. "SAGD: a comprehensive sex-associated gene database from transcriptomes". In: *Nucleic acids research* 47.D1 (2018), pp. D835–D840.

Thank you!

Reference

- Köhler, Sebastian, Leigh Carmody, Nicole Vasilevsky, Julius O B Jacobsen, Daniel Danis, Jean-Philippe Gourdine, Michael Gargano, Nomi L Harris, Nicolas Matentzoglou, Julie A McMurphy, et al. "Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources". In: *Nucleic acids research* 47.D1 (2018), pp. D1018–D1027.
- Shi, Meng-Wei, Na-An Zhang, Chuan-Ping Shi, Chun-Jie Liu, Zhi-Hui Luo, Dan-Yang Wang, An-Yuan Guo, and Zhen-Xia Chen. "SAGD: a comprehensive sex-associated gene database from transcriptomes". In: *Nucleic acids research* 47.D1 (2018), pp. D835–D840.