Decreased adaptation at human disease genes as a possible result of low recombination between deleterious and advantageous variants

Chenlu Di
Enard lab
The University of Arizona

@DiChenlu

cdi@email.arizona.edu

23
Recombination rate variation:
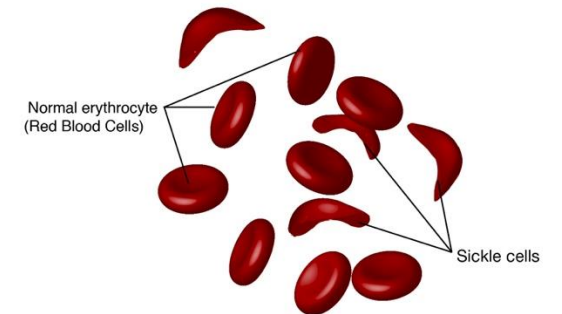implications to genome and organismal evolution

SMBEv 2021

4-8 July

What are disease genes?

What are disease genes?

- Human Disease genes have disease-associated variants in current human populations

What are disease genes?

- Human Disease genes have disease-associated variants in current human populations
- Mendelian and non-infectious disease (e.g. sickle cell anemia)

Normal erythrocyte
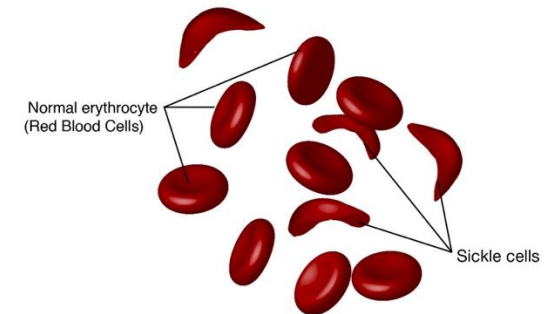(Red Blood Cells)

Sickle cells

What are disease genes?

- Human Disease genes have disease-associated variants in current human populations
- Mendelian and non-infectious disease (e.g. sickle cell anemia)

How adaptation processes connect to human disease?

Normal erythrocyte
(Red Blood Cells)

Sickle cells

What are disease genes?

- Human Disease genes have disease-associated variants in current human populations
- Mendelian and non-infectious disease (e.g. sickle cell anemia)

Have disease genes been associated with more or less adaptation?

What are disease genes?

- Human Disease genes have disease-associated variants in current human populations
- Mendelian and non-infectious disease (e.g. sickle cell anemia)

Have disease genes been associated with more or less adaptation?
What evolutionary processes cause differences in adaptation?

How does evolutionary process connect disease and adaptation?

Constraint:

Known disease genes are more constrained than non-disease genes (Blekhman et al., 2008).

Adaptation in disease genes VS. non-disease genes?

*Disease genes compared to non-disease genes

Adaptation in disease genes VS. non-disease genes?

*Disease genes compared to non-disease genes

Hypotheses	Recent adaptation
Constraint	Less

Adaptation in disease genes VS. non-disease genes?

*Disease genes compared to non-disease genes

Hypotheses	Recent adaptation
Null hypothesis (H_0): Constraint	**Equal** (control constraint)

Adaptation in disease genes VS. non-disease genes?

*Disease genes compared to non-disease genes

Hypotheses	Recent adaptation
H_0: Constraint	**Equal** (control constraint)
H_1: Adaptation increases the risk of diseases	**More**

Adaptation in disease genes VS. non-disease genes?

*Disease genes compared to non-disease genes

Hypotheses	Recent adaptation
H_0: Constraint	**Equal** (control constraint)
H_1: Adaptation increases the risk of diseases	**More**
H_2: Decreased adaptation	**Less**

Methods and data

- Restricted to strong recent adaptation
- Recent adaptation: Integrated Haplotype Score (iHS) and nSL

Voight et al., 2006;
Ferrer-Admetlla et al., 2014

Methods and data

- Restricted to strong recent adaptation
- Recent adaptation: Integrated Haplotype Score (iHS) and nSL
- 26 sub-populations in the 1000 Genome phase3

Voight et al., 2006;
Ferrer-Admetlla et al., 2014

Methods and data

- Restricted to strong recent adaptation

- Recent adaptation: Integrated Haplotype Score (iHS) and nSL

- 26 sub-populations in the 1000 Genome phase3

- Mendelian and non-infectious human disease genes curated from DisGeNET (4215)

Voight et al., 2006;
Ferrer-Admetlla et al., 2014
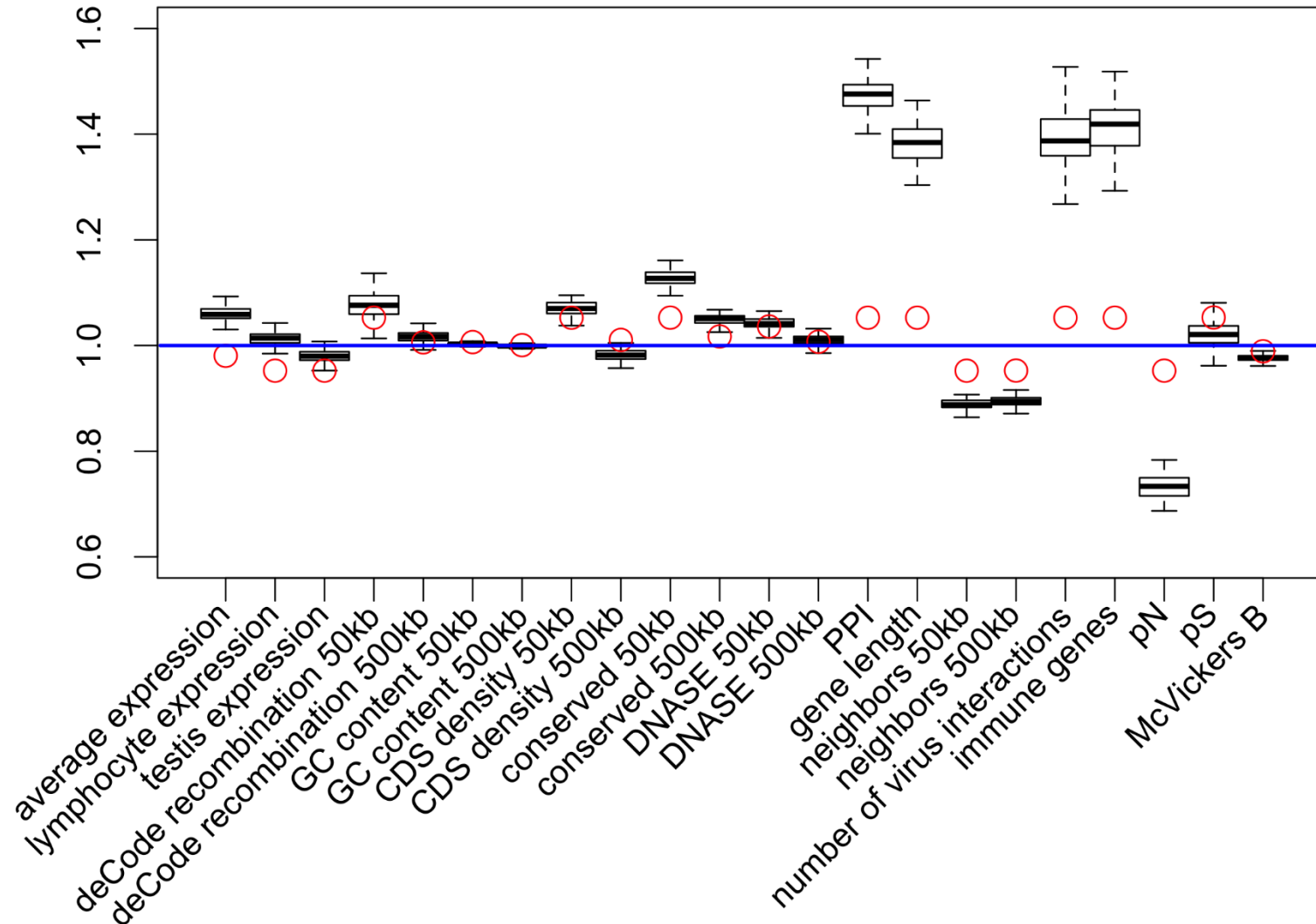Piñero et al., 2019

Confounding factors influence adaptation?

$$\frac{Disease}{non-disease}$$

Confounding factors
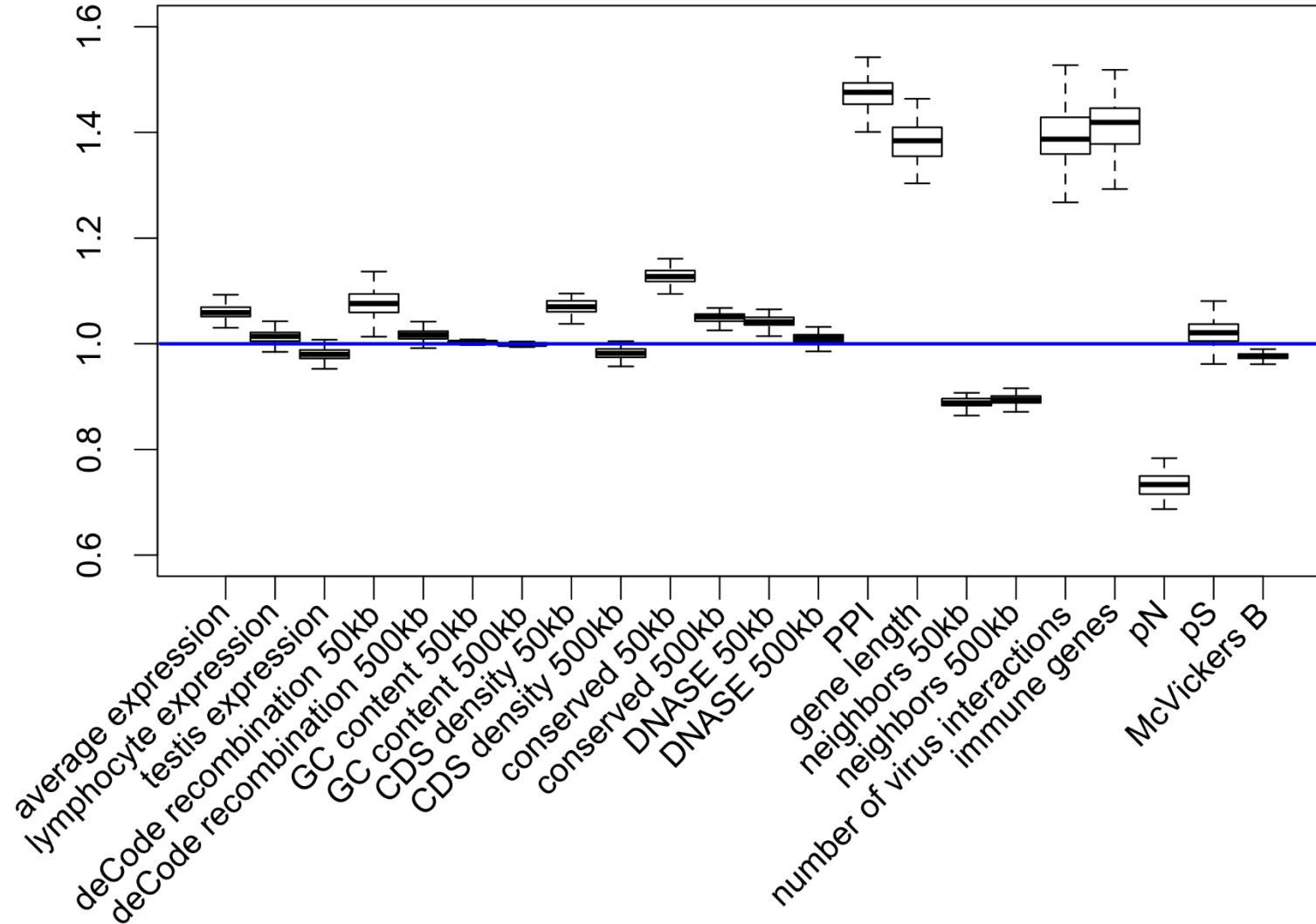
Bootstrap: Controlling for confounding factors

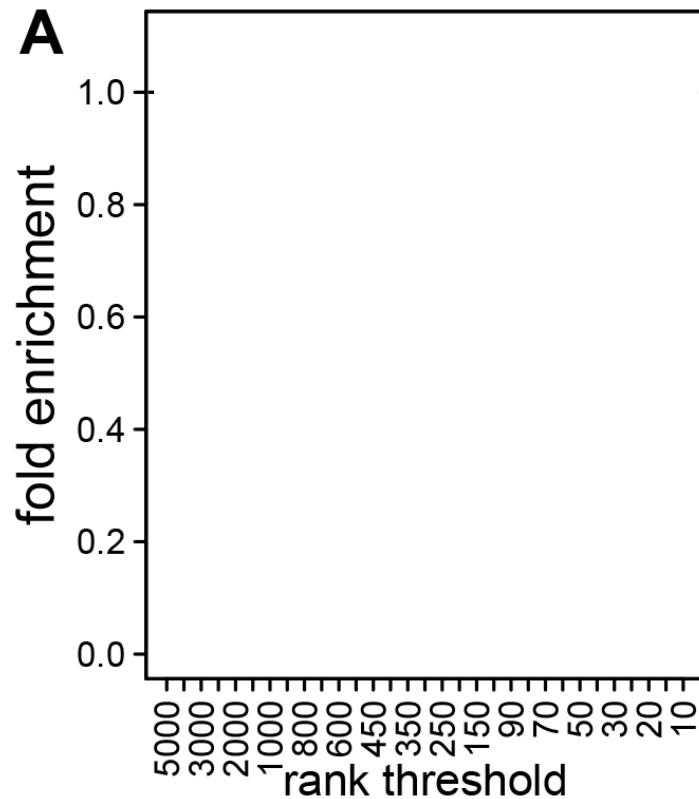
$$\frac{Disease}{non-disease}$$

Confounding factors

Di et al., biorxiv, 2021

Decreased adaptation at human disease genes as a possible consequence of interference between advantageous and deleterious variants

18

Results: Depletion of recent adaptation in human-disease genes

Africa iHS+nSL

$$\frac{disease}{non-disease}$$

☐ 95% confidence interval

A

fold enrichment

1.0
0.8
0.6
0.4
0.2
0.0

5000 3000 2000 1000 800 600 450 350 250 150 90 70 50 30 20 10

rank threshold

Africa iHS+nSL

$$\frac{disease}{non-disease}$$

A

fold enrichment

1.0

0.8

0.6

0.4

0.2

0.0

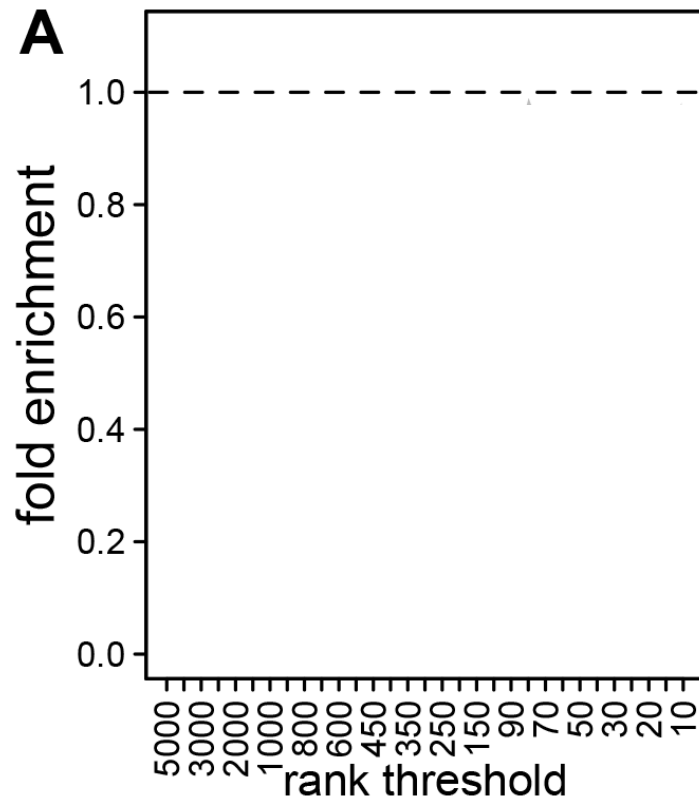5000 3000 2000 1000 800 600 450 350 250 150 90 70 50 30 20 10

rank threshold

☐ 95% confidence interval

Results: Depletion of recent adaptation in human-disease genes

Africa iHS+nSL
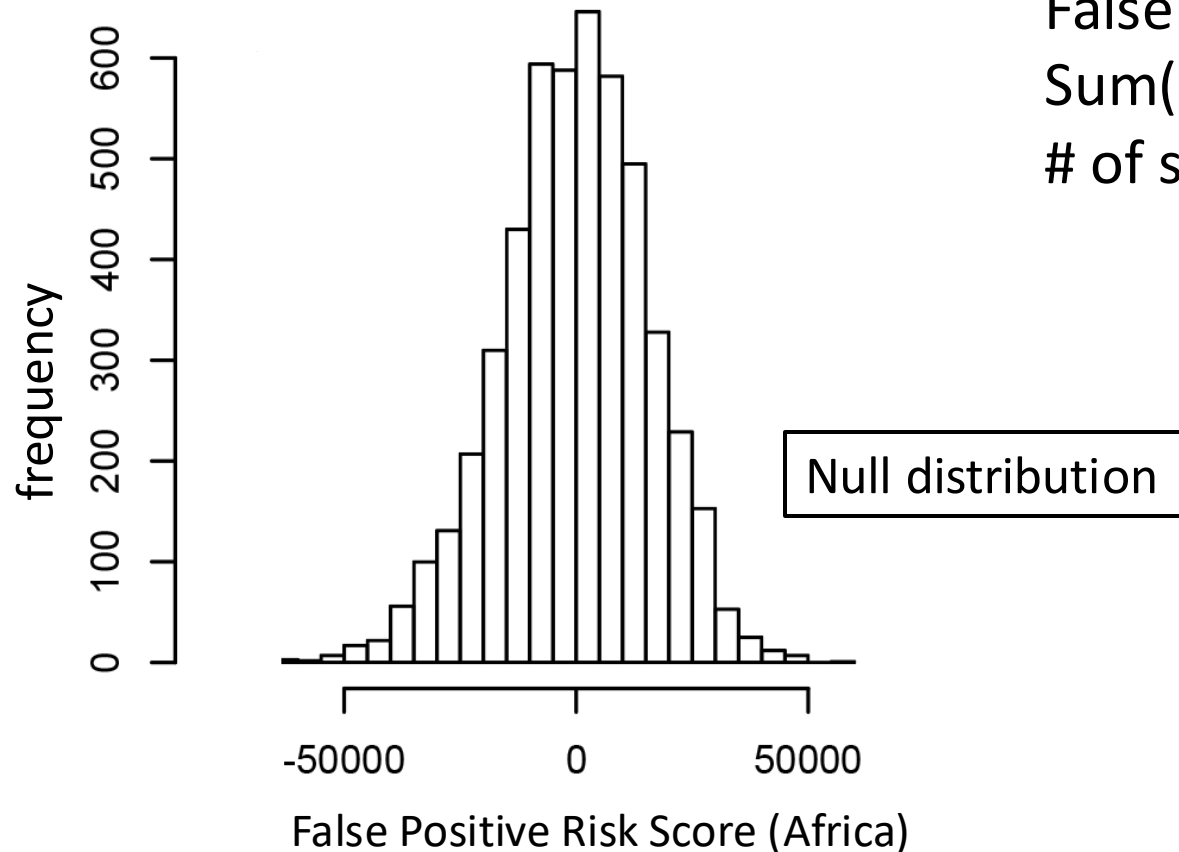

$$\frac{disease}{non-disease}$$
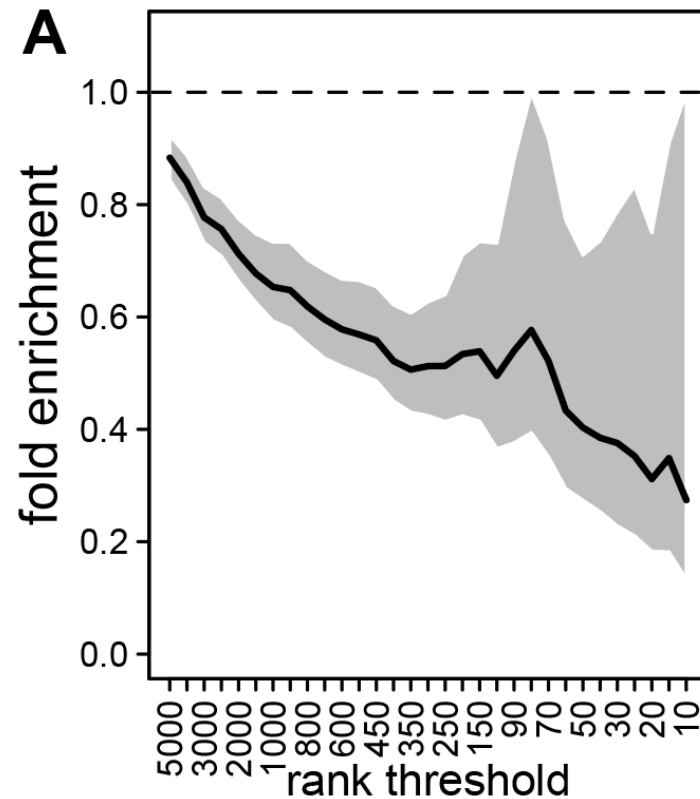
☐ 95% confidence interval

Results: Significant depletion of adaptation in human-disease genes

- Permutation test (Di et al., 2021)

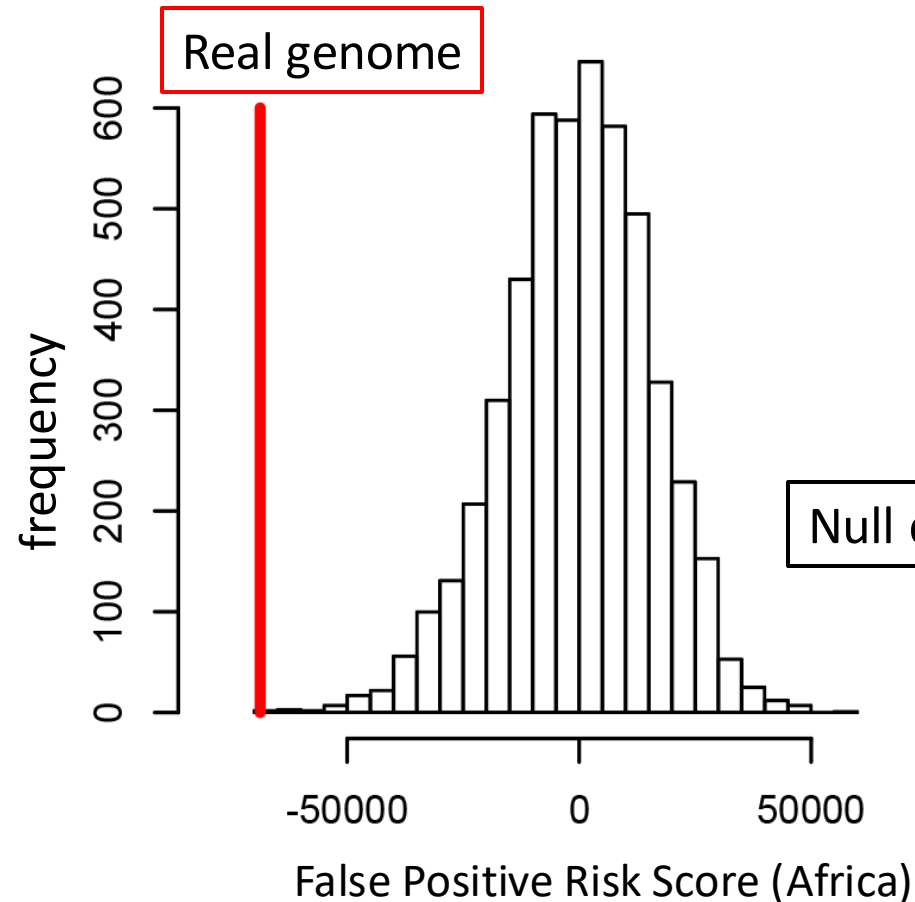# Results: Significant depletion of adaptation in human-disease genes

- Permutation test (Di et al., 2021)

False positive Risk Score:
Sum(# of selected disease genes minus
# of selected control non-disease genes)

Null distribution

Results: Significant depletion of adaptation in human-disease genes

- Permutation test (Di et al., 2021)

False positive Risk Score:
Sum(# of selected disease genes minus
# of selected control non-disease genes)

False Positive Risk =0.0003 (**Africa**)

Conclusion: Decreased adaptation

- Disease genes compared to non-disease genes

Hypotheses	Recent adaptation
H_0: Constraint	**Equal** (control constraint)
H_1: Adaptation increases the risk of diseases	**More**
H_2: Decreased adaptation	**Less**

The cause of sweep deficit at disease genes?

Maybe **Not**: pleiotropy or biological function (GO ontology).

The cause of sweep deficit at disease genes?

Maybe **Not**: pleiotropy or biological function (GO ontology).

Possible reason: Interference due to deleterious variants can decrease the fixation probability of adaptive mutations in human. (Uricchio et al. 2019)

The cause of sweep deficit at disease genes?

Interference due to deleterious variants can decrease the fixation probability of adaptive mutations in human. (Uricchio et al. 2019)

	Many disease variants (≥5/500kb)	Few disease variants (<5/500kb)
Low recombination rate (≥1.14 cM/Mb)		
High recombination rate (<1.14 cM/Mb)		

The cause of sweep deficit at disease genes?

Interference due to deleterious variants can decrease the fixation probability of adaptive mutations in human. (Uricchio et al. 2019)

■ Strong deficit
□ Weak deficit
□ No deficit

	Many disease variants (≥5/500kb)	Few disease variants (<5/500kb)
Low recombination rate (<1.14 cM/Mb)		
High recombination rate (≥1.14 cM/Mb)		

The cause of sweep deficit at disease genes?

Interference due to deleterious variants can decrease the fixation probability of adaptive mutations in human. (Uricchio et al. 2019)

■ Strong deficit
□ Weak deficit
□ No deficit

	Many disease variants (≥5/500kb)	Few disease variants (<5/500kb)
Low recombination rate (<1.14 cM/Mb)		deficit=12.91 FPR=0.021 (*)
High recombination rate (≥1.14 cM/Mb)	deficit=15.36 FPR=0.026 (*)	

The cause of sweep deficit at disease genes?

Interference due to deleterious variants can decrease the fixation probability of adaptive mutations in human. (Uricchio et al. 2019)

■ Strong deficit

□ Weak deficit

□ No deficit

	Many disease variants (≥5/500kb)	Few disease variants (<5/500kb)
Low recombination rate (≥1.14 cM/Mb)	deficit=66.8 FPR=0.0002 (***)	deficit=12.91 FPR=0.021 (*)
High recombination rate (<1.14 cM/Mb)	deficit=15.36 FPR=0.026 (*)	deficit=2.16 FPR=0.74

Summary

- Depletion of recent adaptation in human non-infectious diseases genes.

- Adaptation might have been slowed down by the presence of interfering deleterious variants at disease genes.

- More work!!!

Acknowledgements

David Enard

Elise Lauterbur

Diego Salazar Tortosa

Thank you!

https://github.com/DavidPierreEnard/Gene_Set_Enrichment_Pipeline

@DiChenlu

cdi@email.arizona.edu

The causes of strongly depleted recent **adaptation** in human **disease** genes

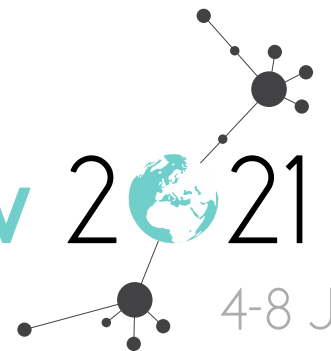Chenlu Di
The University of Arizona

23
Recombination rate variation: implications to genome and organismal evolution

@DiChenlu

cdi@email.arizona.edu

SMBEv 2021

4-8 July