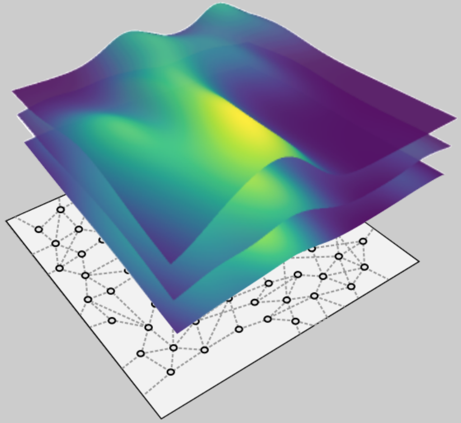
# Statistical mapping of structure-function landscapes in microbiomes

Djordje Bajić
Assistant Professor
Department of Biotechnology
**TU Delft**



School of AI applied to Microbiomes

Paris, October 1, 2025

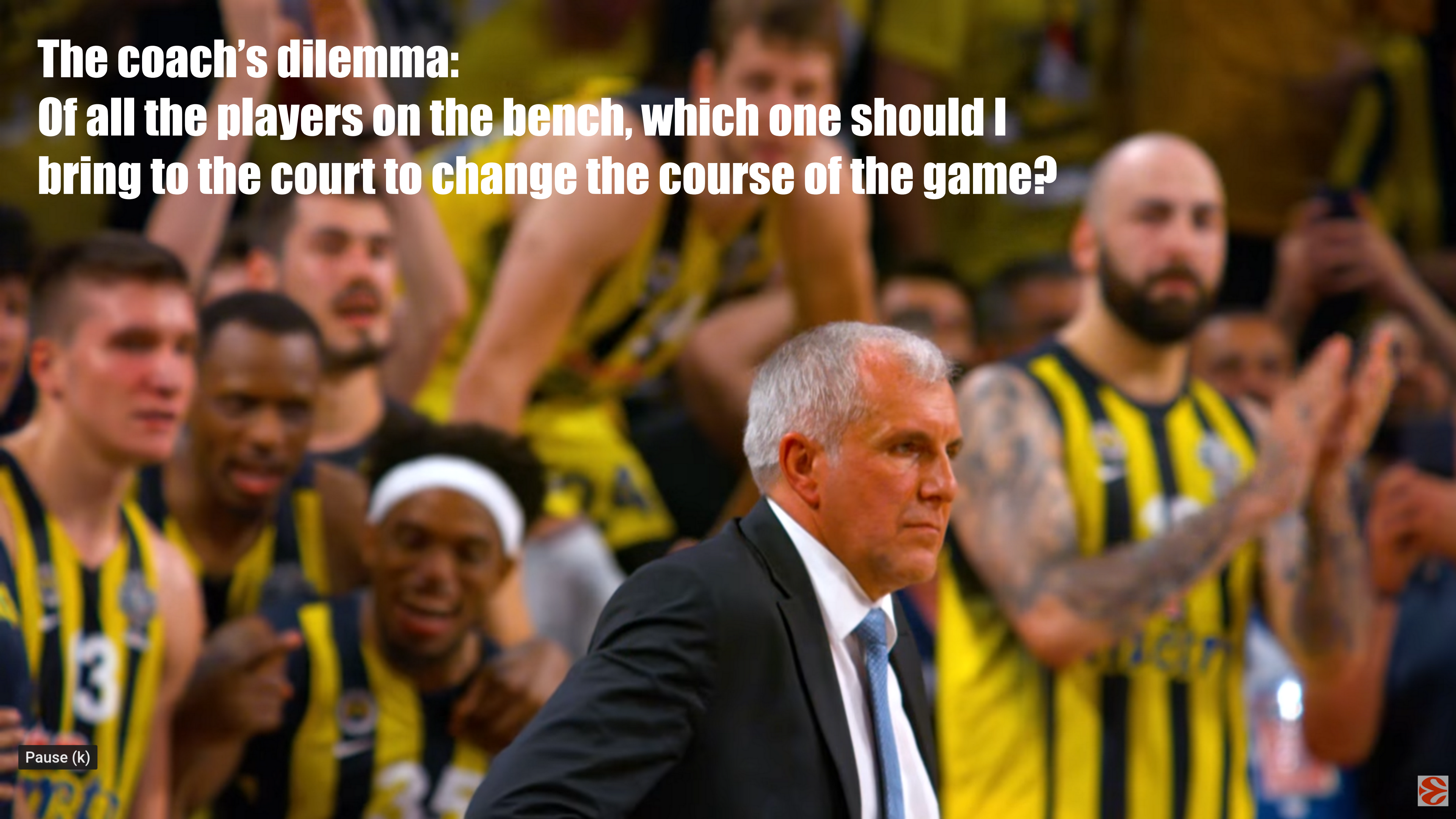
After today session, you will :

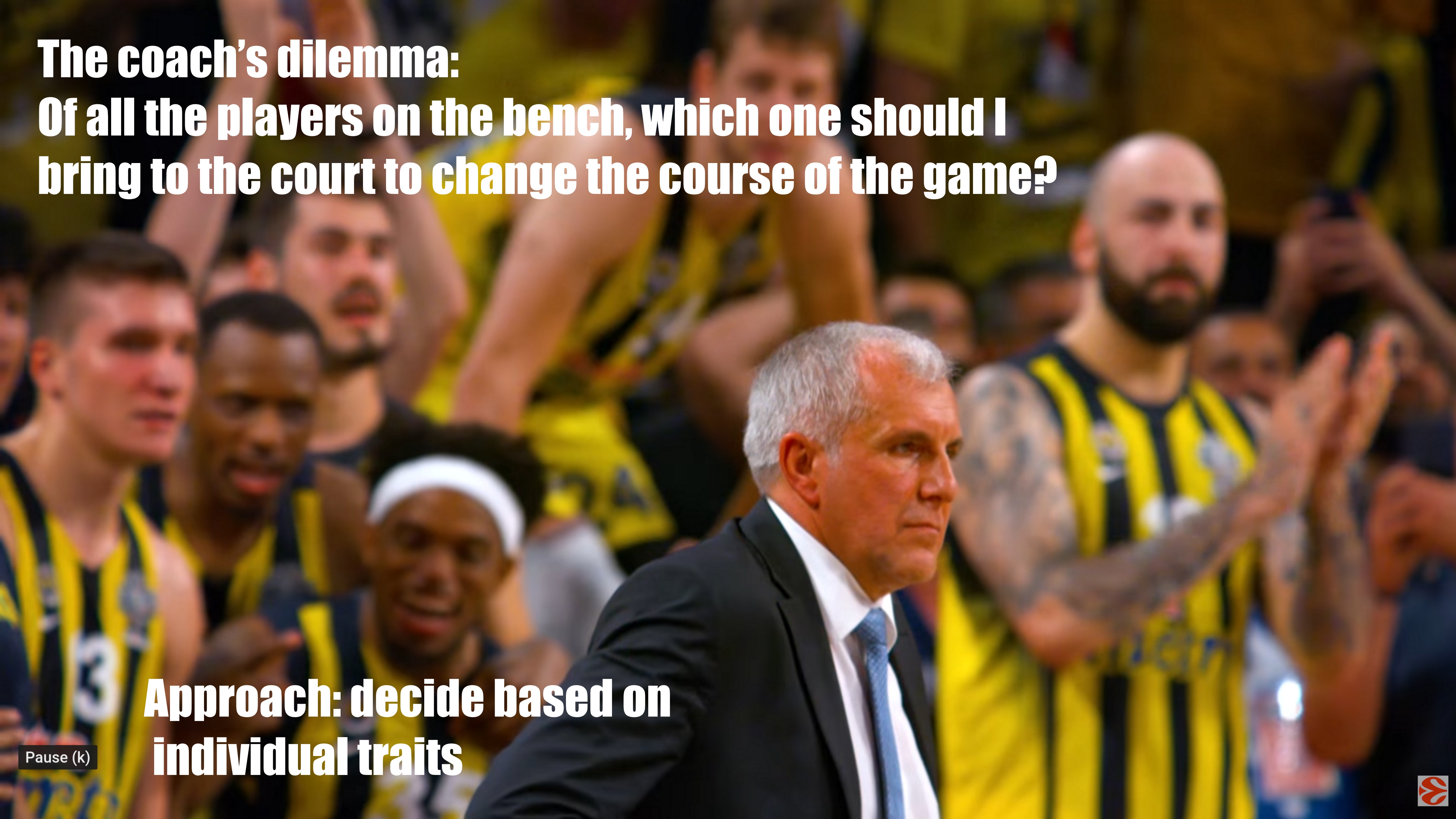
1. Know what a structure-function landscape is

2. Know what global epistasis is and why is useful

3. Touch some machine learning code
   make some predictions
   lose the fear ;)

# The coach's dilemma:
# Of all the players on the bench, which one should I bring to the court to change the course of the game?





# The coach's dilemma:
# Of all the players on the bench, which one should I bring to the court to change the course of the game?

**Approach: decide based on individual traits**

# BUT: The contribution of a player depends on its interaction with the rest of the team!

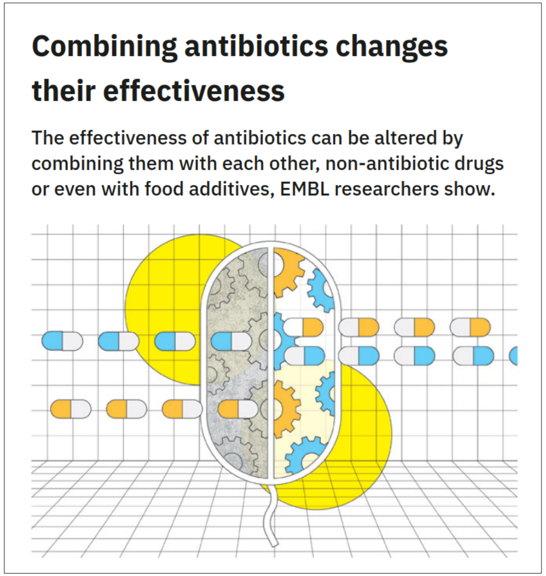
# In many systems, function arises from interacting components

Applied Soil Ecology
Volume 176, August 2022, 104503

Review

## Consortia-based microbial inoculants for sustaining agricultural activities

Shams Tabrez Khan



## Combining antibiotics changes their effectiveness

The effectiveness of antibiotics can be altered by combining them with each other, non-antibiotic drugs or even with food additives, EMBL researchers show.

Current Opinion in Chemical Engineering
Volume 22, December 2018, Pages 42-47

## Media formulation optimization: current and future opportunities

Shaun C. Galbraith [1], Hemlata Bhatia [2], Huolong Liu [1], Seongkyu Yoon [1 2]

# In many systems, function arises from interacting components

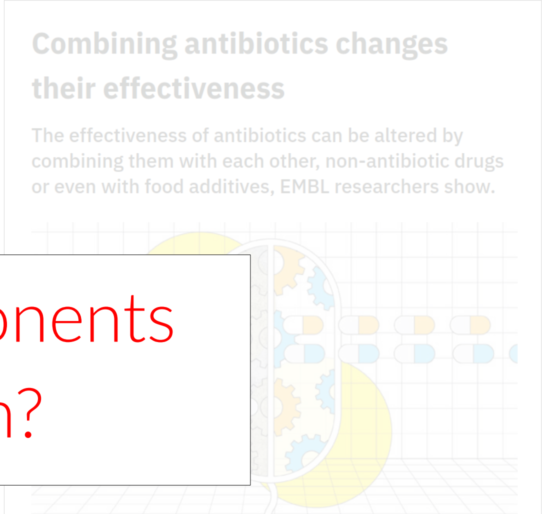
ELSEVIER

Applied Soil Ecology
Volume 176, August 2022, 104503

Review

Consortia-based microbial inoculants for sustaining agricultural activities

What combination of components maximizes a desired function?



**Combining antibiotics changes their effectiveness**

The effectiveness of antibiotics can be altered by combining them with each other, non-antibiotic drugs or even with food additives, EMBL researchers show.

ELSEVIER

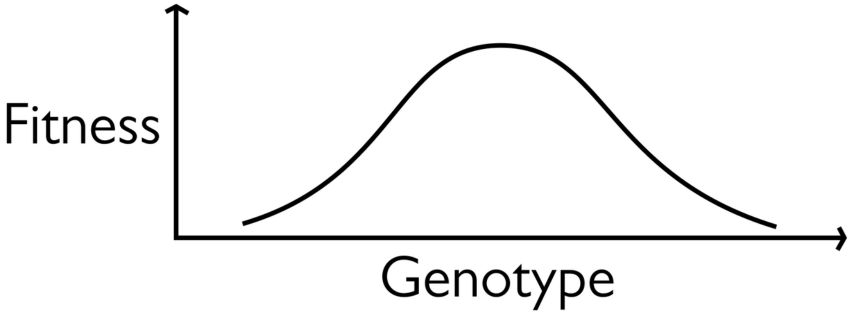
Current Opinion in Chemical Engineering
Volume 22, December 2018, Pages 42-47

Media formulation optimization: current and future opportunities

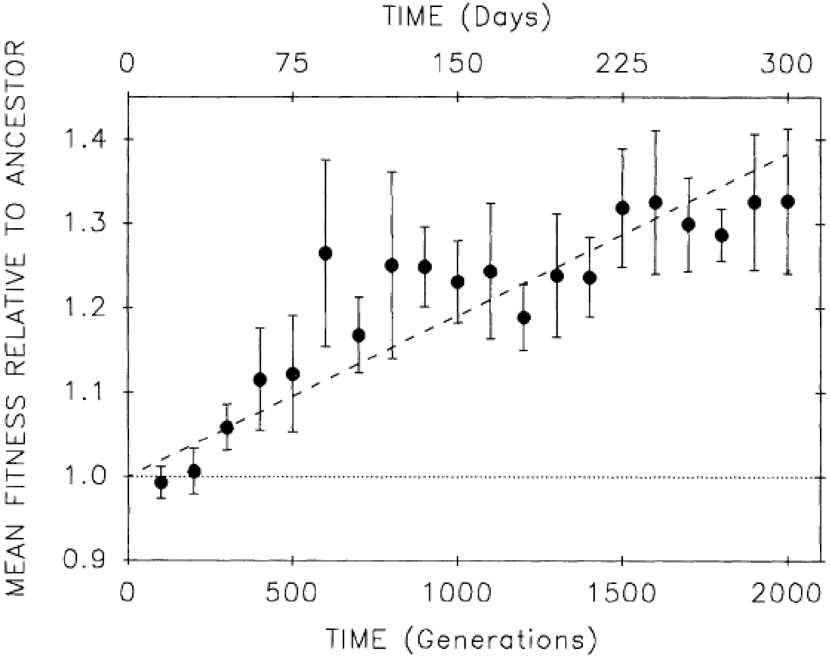
Shaun C. Galbraith [1], Hemlata Bhatia [2], Huolong Liu [1], Seongkyu Yoon [1 2]

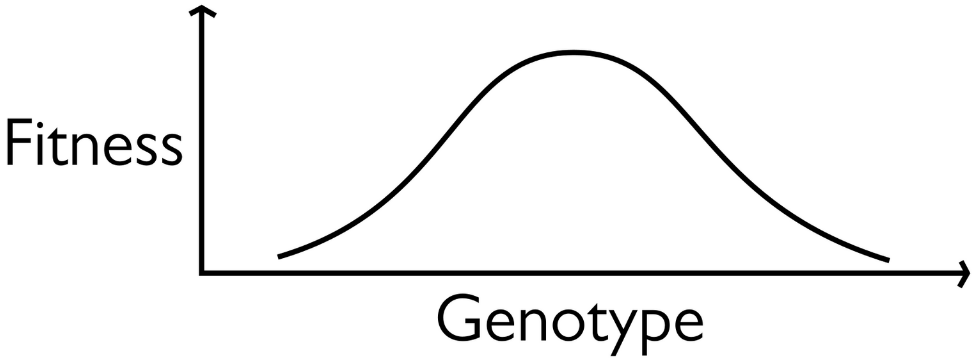
# Presenting our main tool:

## A **fitness landscape** is a map between genotypes and fitness

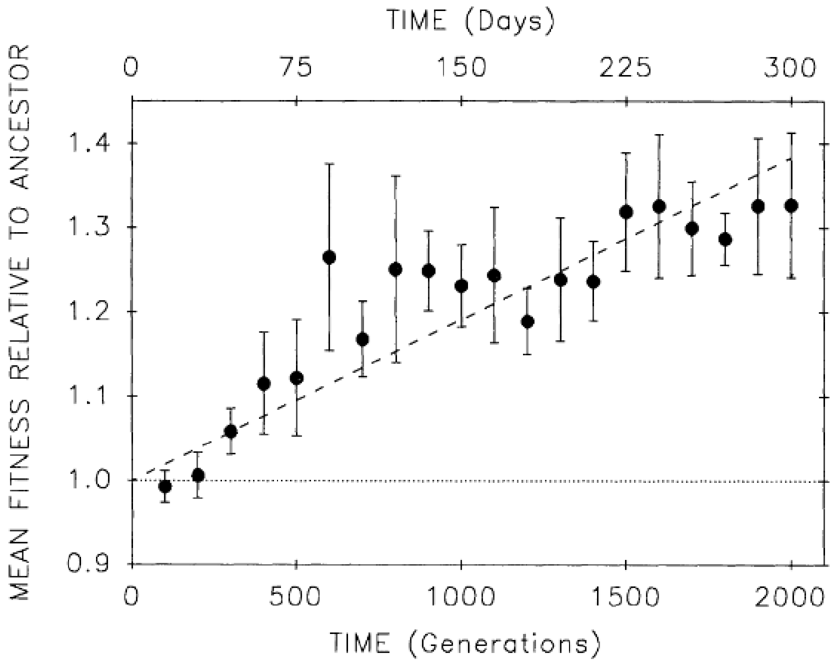


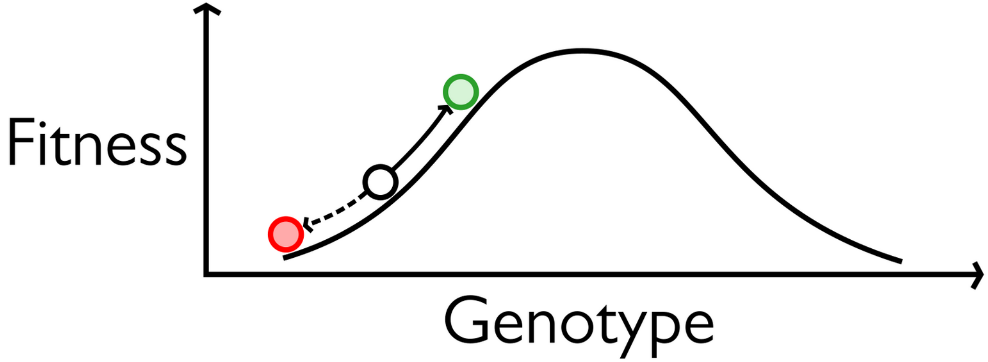
# Fitness landscapes have been historically a cornerstone concept to understand and predict evolution



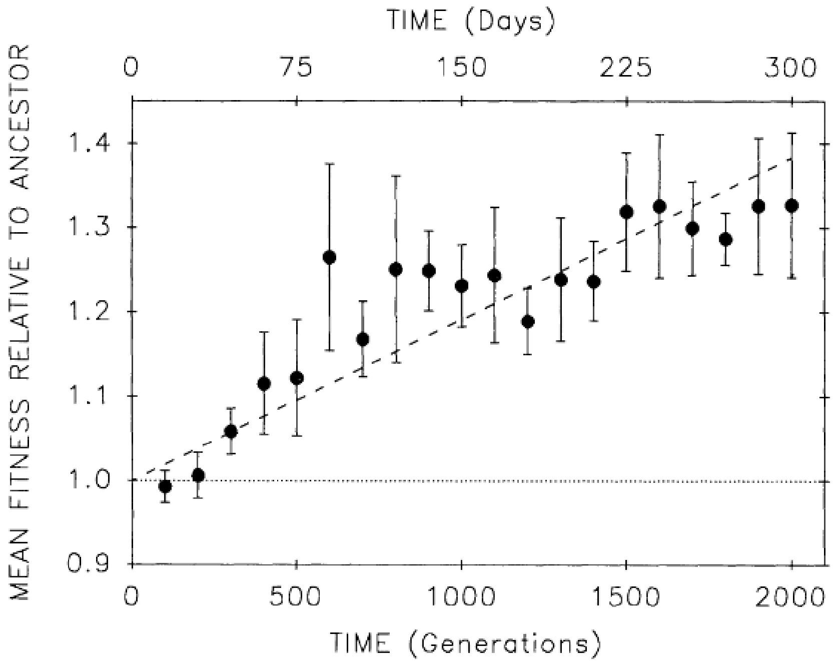


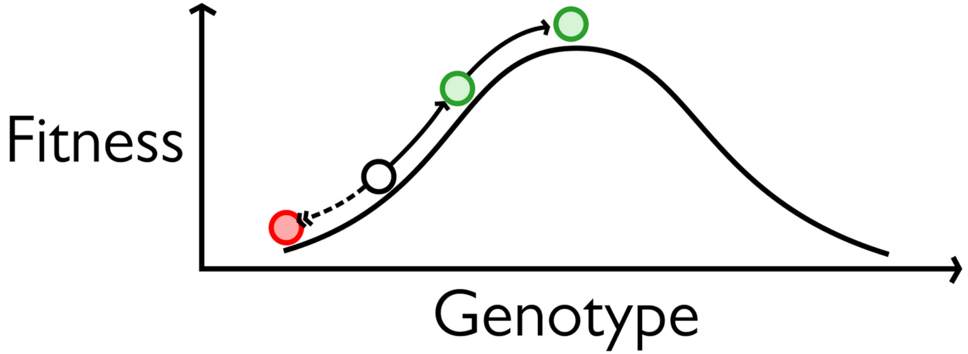
# When a spontaneous mutation has higher fitness, it has a chance of fixing. Deleterious mutations get outcompeted



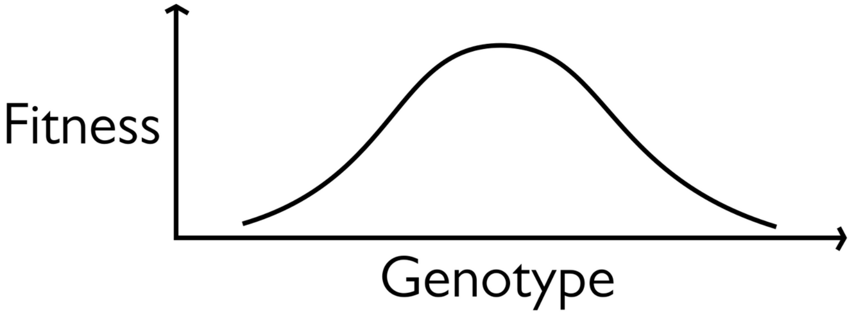


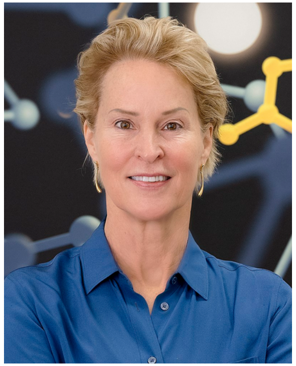
# In this way, populations “climb” the peak of the fitness landscape towards genotypes with high fitness



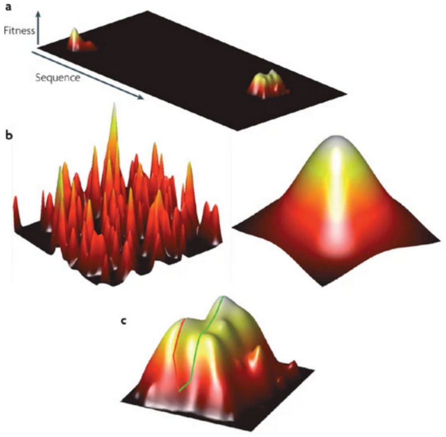


# Fitness landscapes have already been useful for engineering, e.g. proteins

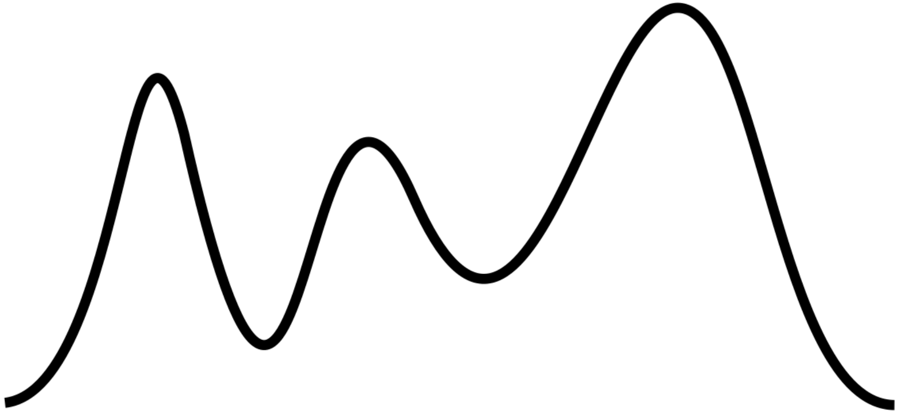




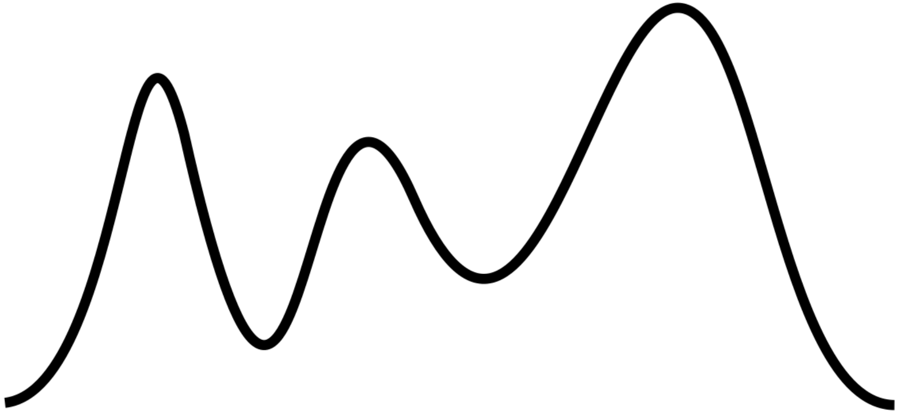
Frances Arnold, Nobel Prize for the directed evolution of enzymes

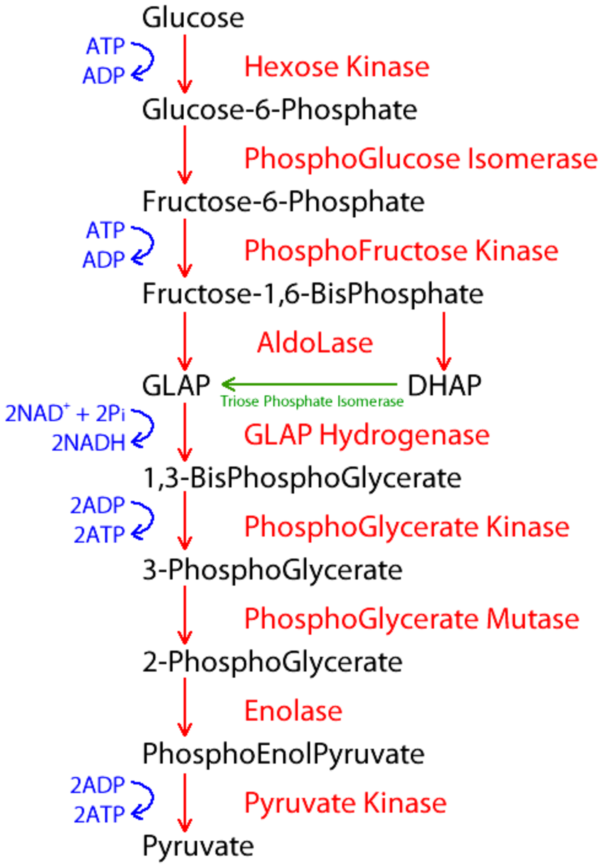


# What determines the shape of a fitness landscape?



# What determines the shape of a fitness landscape?





Glucose

ATP

ADP

Hexose Kinase

Glucose-6-Phosphate

PhosphoGlucose Isomerase

Fructose-6-Phosphate

ATP

ADP

PhosphoFructose Kinase

Fructose-1,6-BisPhosphate

AldoLase

GLAP

DHAP

Triose Phosphate Isomerase

$2NAD^+$ + 2Pi

2NADH

GLAP Hydrogenase

1,3-BisPhosphoGlycerate

2ADP

2ATP

PhosphoGlycerate Kinase

3-PhosphoGlycerate

PhosphoGlycerate Mutase

2-PhosphoGlycerate

Enolase

PhosphoEnolPyruvate

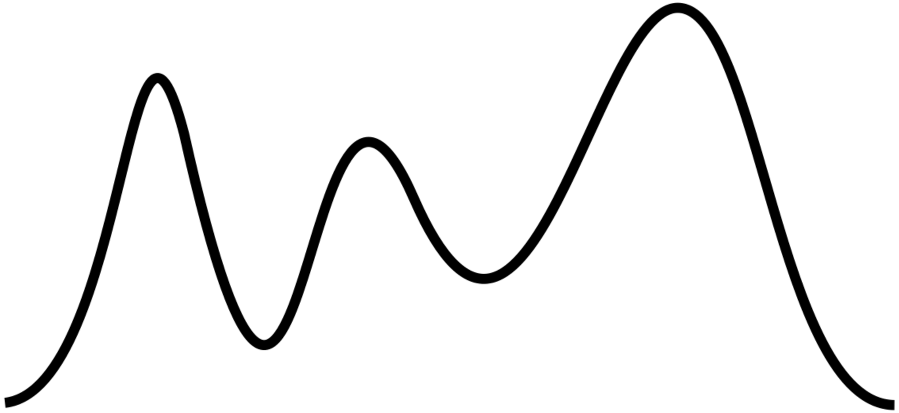
2ADP

2ATP

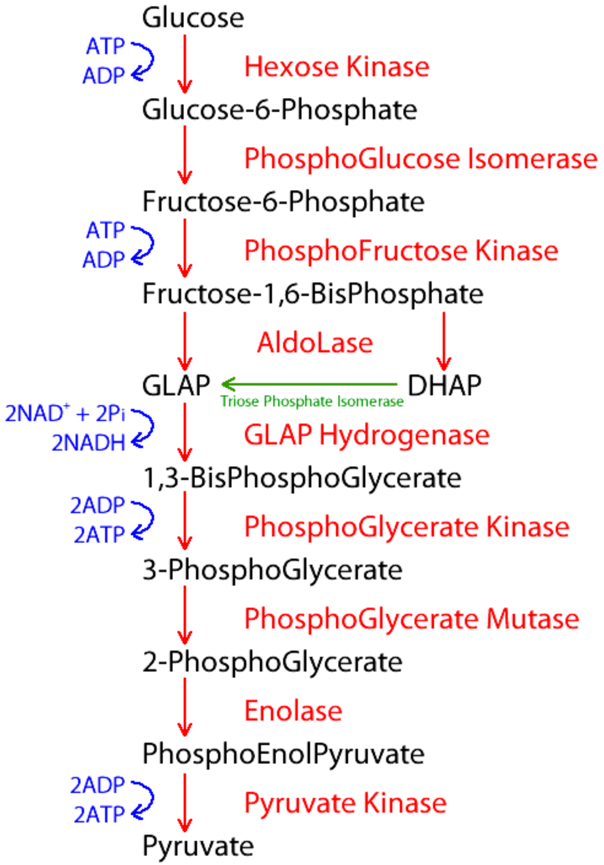
Pyruvate Kinase

Pyruvate

# What determines the shape of a fitness landscape?



- The environment



Glucose

ATP

ADP

Hexose Kinase

Glucose-6-Phosphate

PhosphoGlucose Isomerase

Fructose-6-Phosphate

ATP

ADP

PhosphoFructose Kinase

Fructose-1,6-BisPhosphate

AldoLase

GLAP

DHAP

Triose Phosphate Isomerase

$2NAD^+ + 2Pi$

2NADH

GLAP Hydrogenase

1,3-BisPhosphoGlycerate

2ADP

2ATP

PhosphoGlycerate Kinase

3-PhosphoGlycerate

PhosphoGlycerate Mutase

2-PhosphoGlycerate

Enolase

PhosphoEnolPyruvate

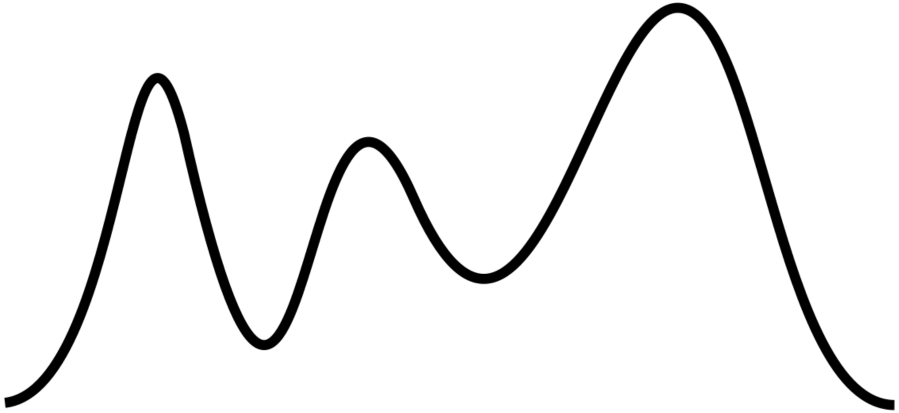
2ADP

2ATP

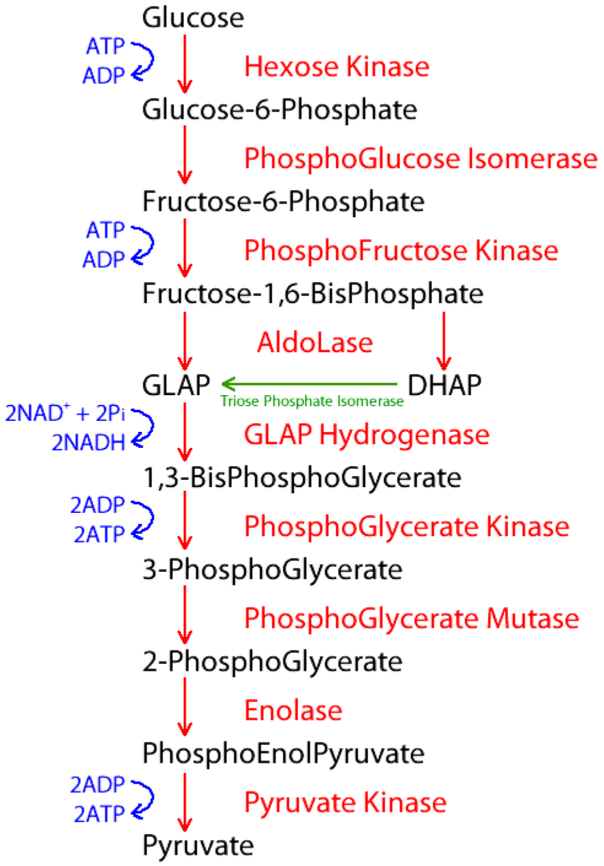
Pyruvate Kinase

Pyruvate

# What determines the shape of a fitness landscape?



- The environment
- Interactions between mutations - epistasis



Glucose

ATP → ADP — Hexose Kinase

Glucose-6-Phosphate

PhosphoGlucose Isomerase

Fructose-6-Phosphate

ATP → ADP — PhosphoFructose Kinase

Fructose-1,6-BisPhosphate

AldoLase

GLAP ← DHAP (Triose Phosphate Isomerase)

2NAD$^+$ + 2Pi → 2NADH — GLAP Hydrogenase

1,3-BisPhosphoGlycerate

2ADP → 2ATP — PhosphoGlycerate Kinase

3-PhosphoGlycerate

PhosphoGlycerate Mutase

2-PhosphoGlycerate

Enolase

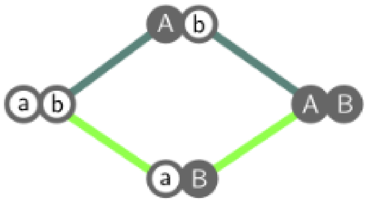
PhosphoEnolPyruvate

2ADP → 2ATP — Pyruvate Kinase

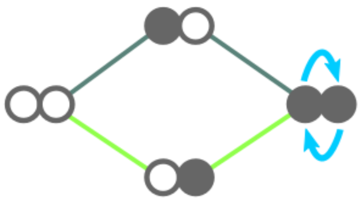
Pyruvate

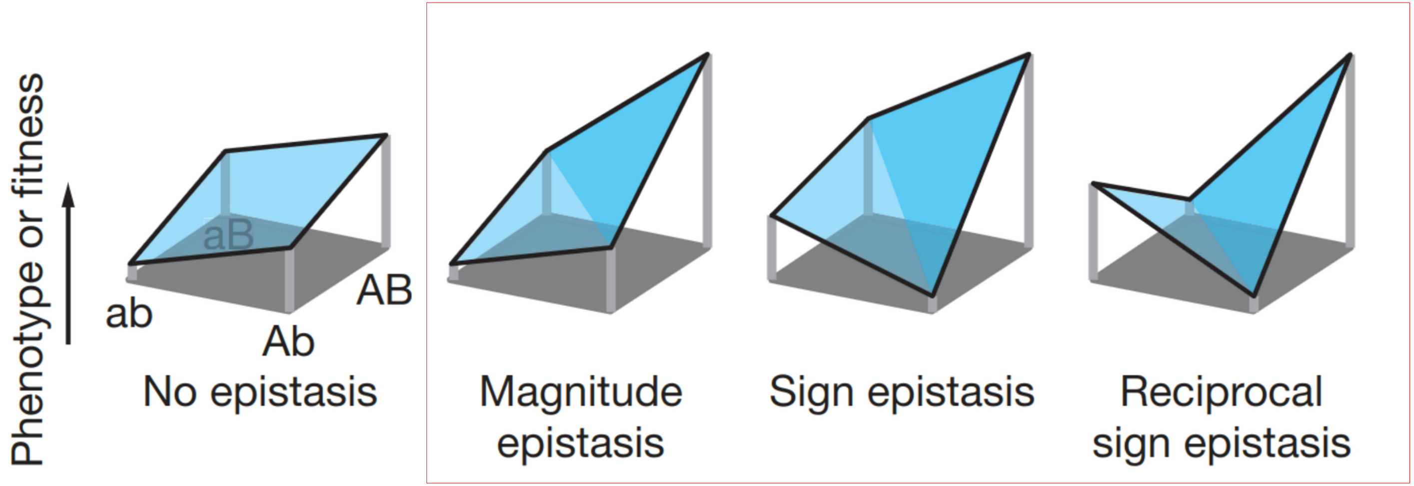
# When one mutation affects the fitness of another one, we say there is an **epistatic interaction** among them

No interaction



Epistasis





No epistasis

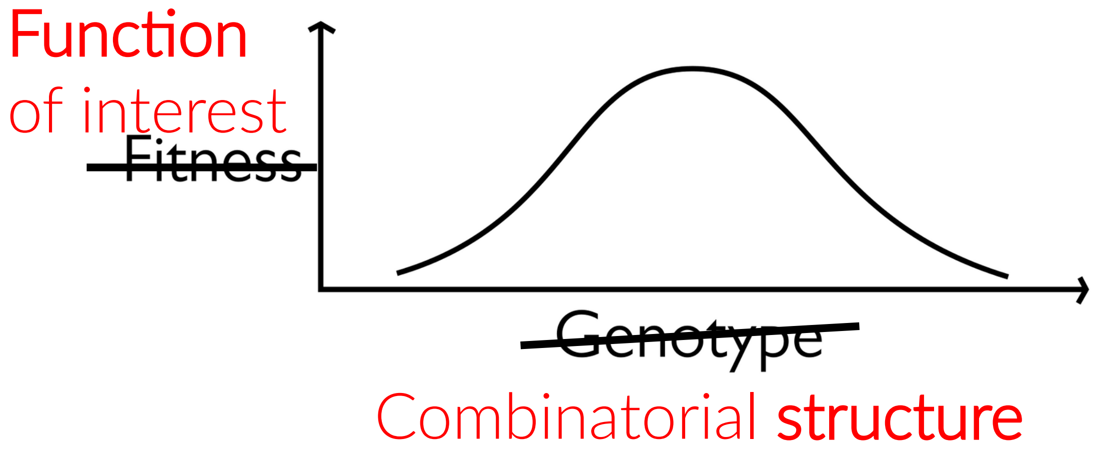
Magnitude epistasis

Sign epistasis

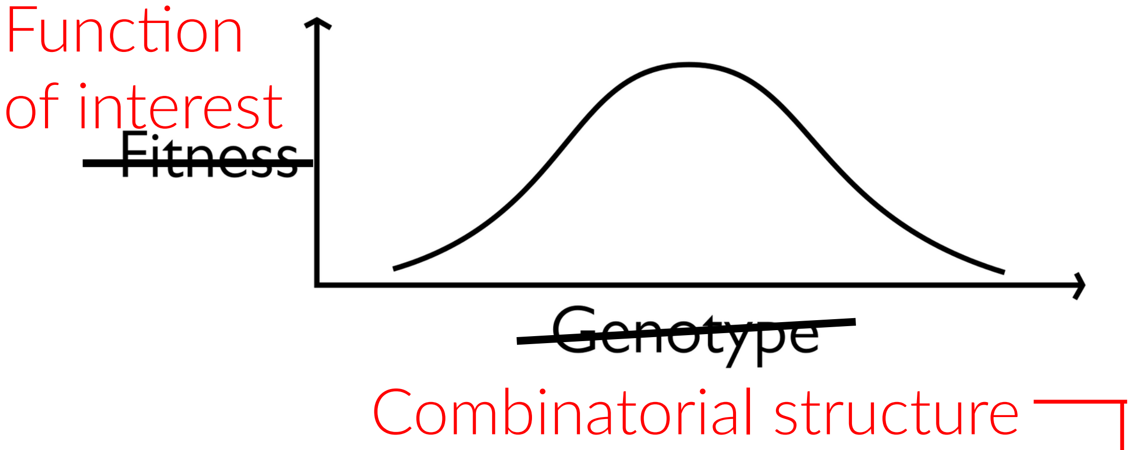
Reciprocal sign epistasis

Poelwijk et al, 2007 “**Empirical fitness landscapes reveal accessible evolutionary paths**”, Nature

# Can we extend the theory of fitness landscapes to engineer other systems with interacting components?



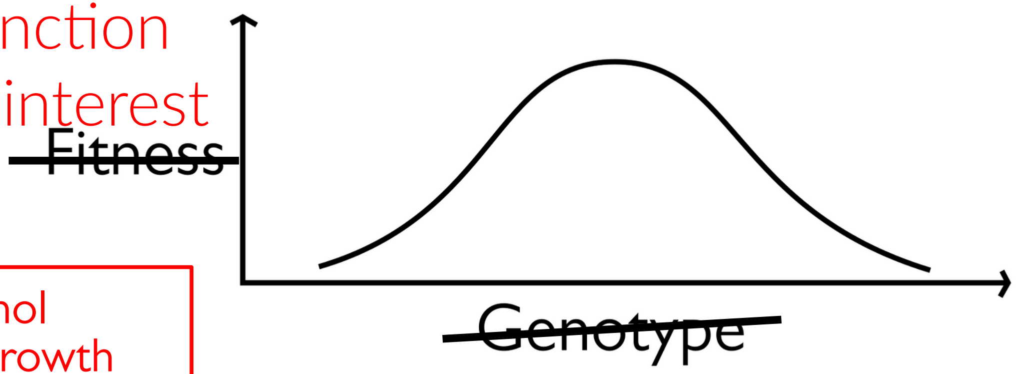
# Can we extend the theory of fitness landscapes to engineer other systems with interacting components?



The composition of a
synthetic microbial community

# Can we extend the theory of fitness landscapes to engineer other systems with interacting components?

Function of interest

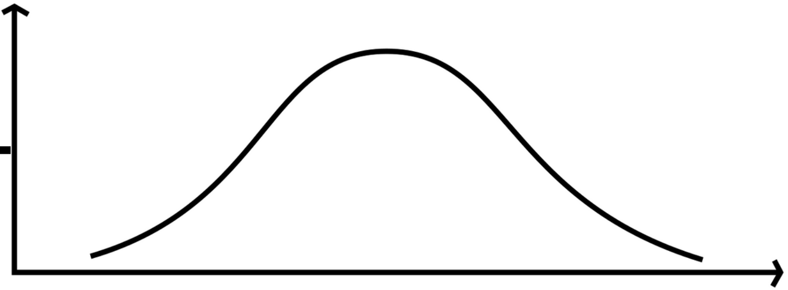


- production of ethanol
- effect on a plant's growth
- whatever you can quantify

Combinatorial structure

The composition of a synthetic microbial community

# A structure-function landscape



Function of interest

~~Fitness~~

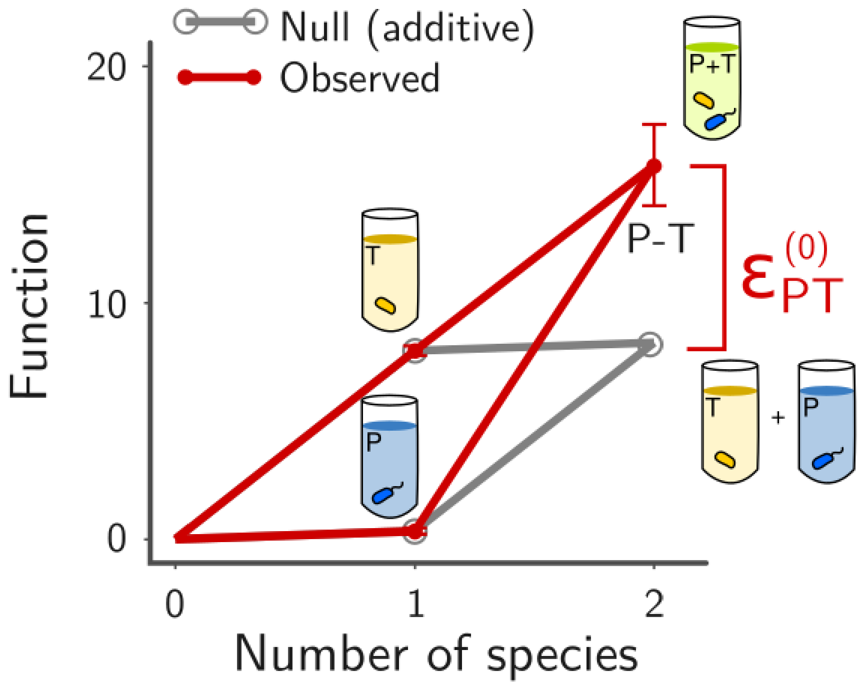
- production of ethanol
- effect on a plant's growth
- whatever you can quantify

~~Genotype~~

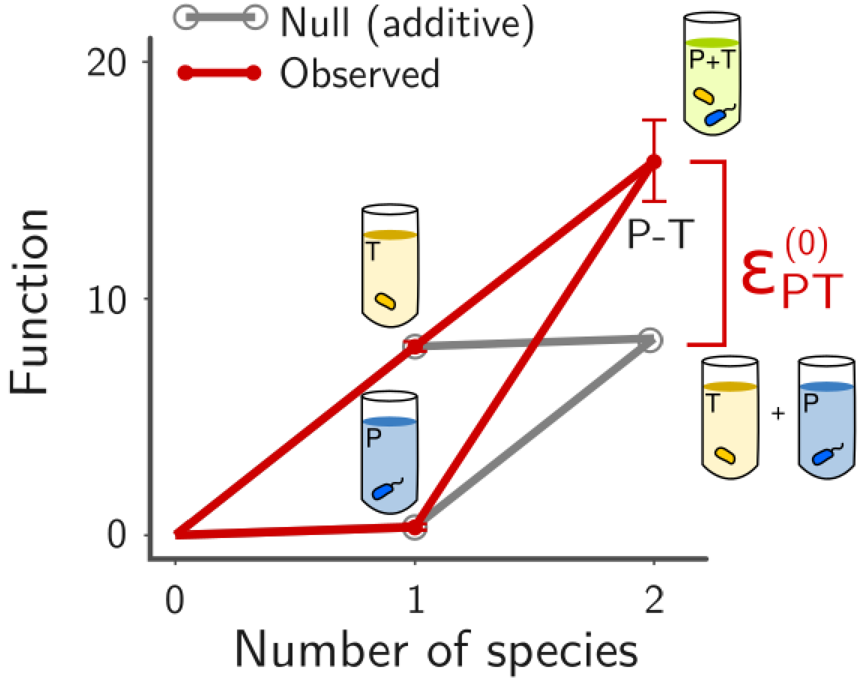
Combinatorial structure

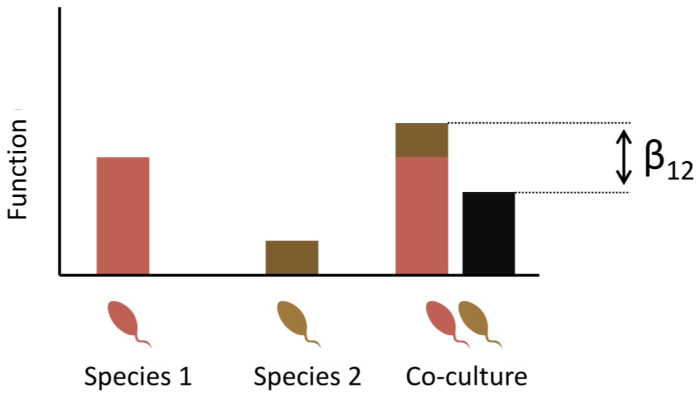
Example: the composition of a synthetic microbial community

# Same as fitness landscapes, structure-function landscapes are also shaped by **interactions**



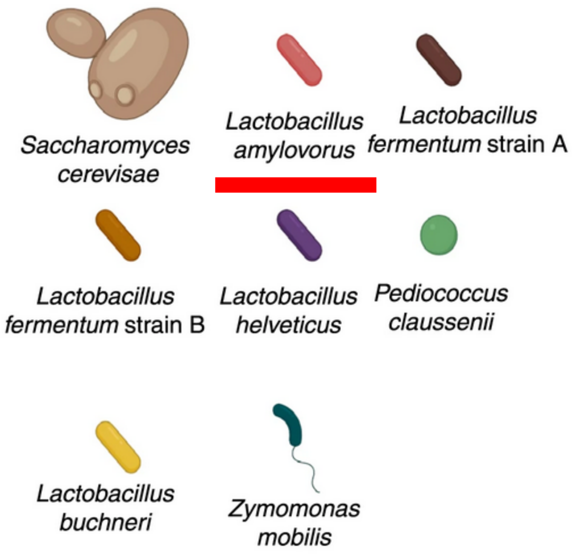
# Same as fitness landscapes, structure-function landscapes are also shaped by **interactions**

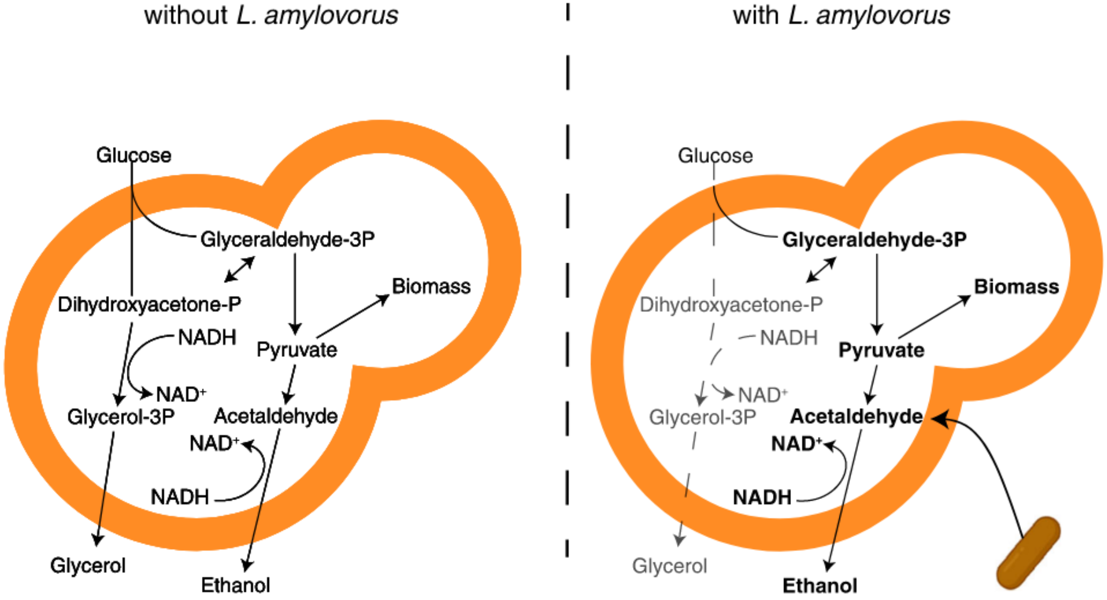




$$F(x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2$$

# Structure-function landscape modeling helped identify a potential industrial probiotic





nature COMMUNICATIONS

Complex yeast-bacteria interactions affect the yield of industrial ethanol fermentation

Felipe Senne de Oliveira Lino[1], Djordje Bajic[2,3], Jean Celestin Charles Vila[2,3], Alvaro Sánchez[2,3] & Morten Otto Alexander Sommer[1✉]

Felipe S. Lino

TU Delft

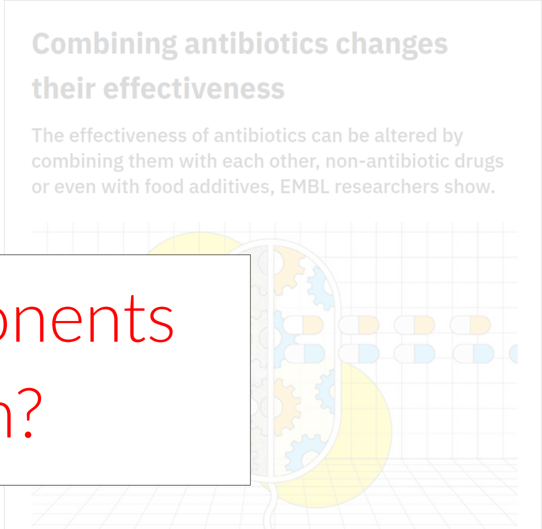
# In systems where function arises from interacting components…

ELSEVIER

Applied Soil Ecology
Volume 176, August 2022, 104503

Review

Consortia-based microbial inoculants for sustaining agricultural activities



**Combining antibiotics changes their effectiveness**

The effectiveness of antibiotics can be altered by combining them with each other, non-antibiotic drugs or even with food additives, EMBL researchers show.

ELSEVIER

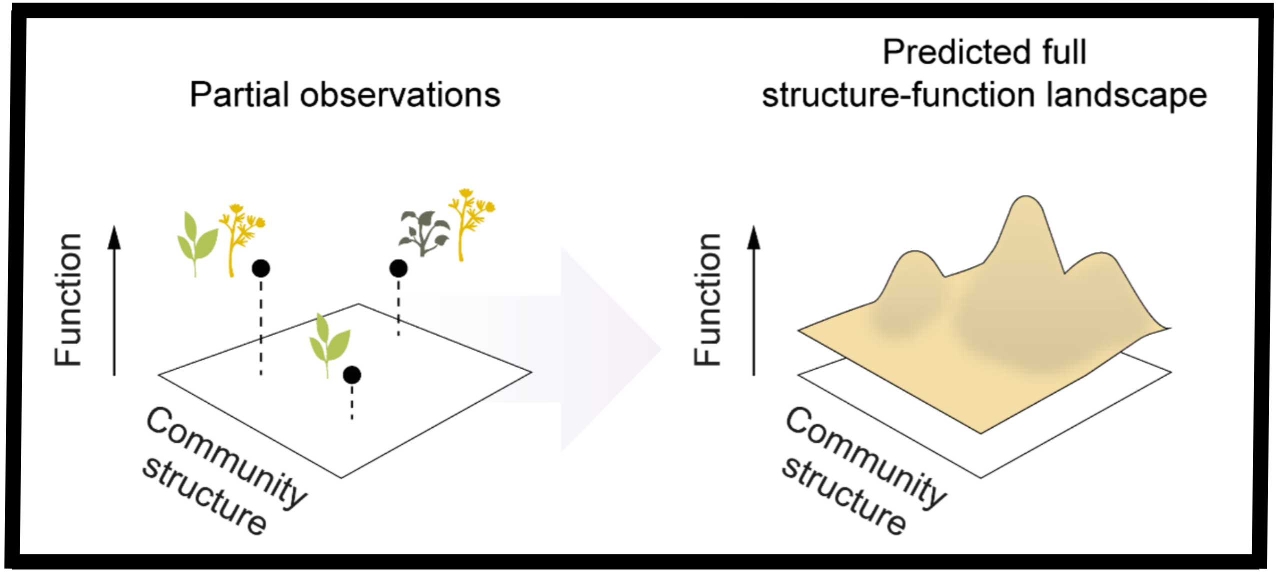
Current Opinion in Chemical Engineering
Volume 22, December 2018, Pages 42-47

Media formulation optimization: current and future opportunities

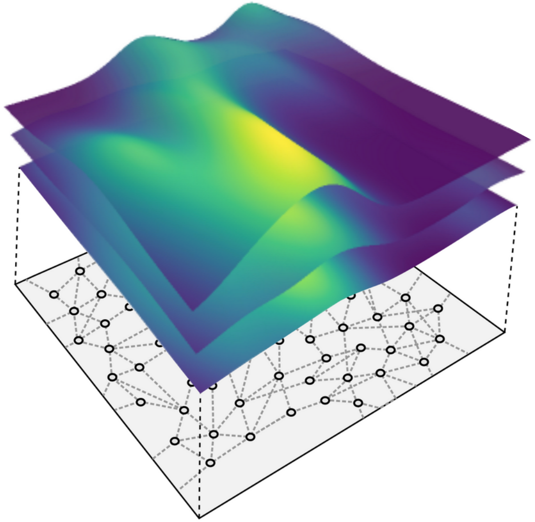
Shaun C. Galbraith[1], Hemlata Bhatia[2], Huolong Liu[1], Seongkyu Yoon[1,2]

What combination of components maximizes a desired function?

# Our goal: predict the structure-function landscape of a microbial community **with limited data, because the combinatorial space is infeasible!** ($2^n$)



# Simplest statistical approach: regression



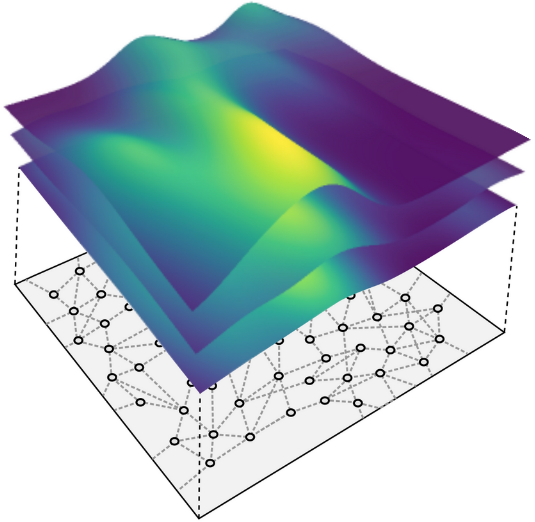
$$y = \beta_0 + \sum_i \beta_i x_i$$

Community function

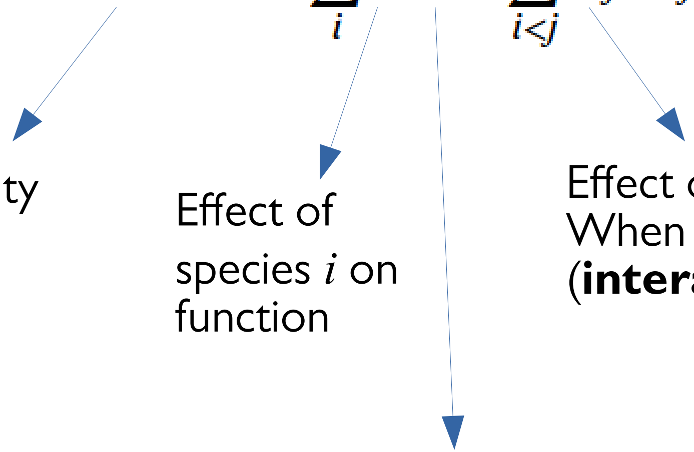
Effect of species $i$ on function

Presence/absence of species $i$ in the community

# Simplest statistical approach: regression



$$y = \beta_0 + \sum_i \beta_i x_i + \sum_{i<j} \beta_{ij} x_i x_j + \ldots ,$$



Community function

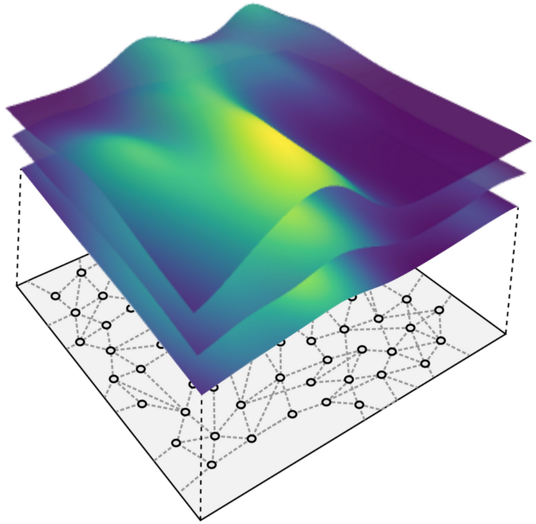
Effect of species $i$ on function

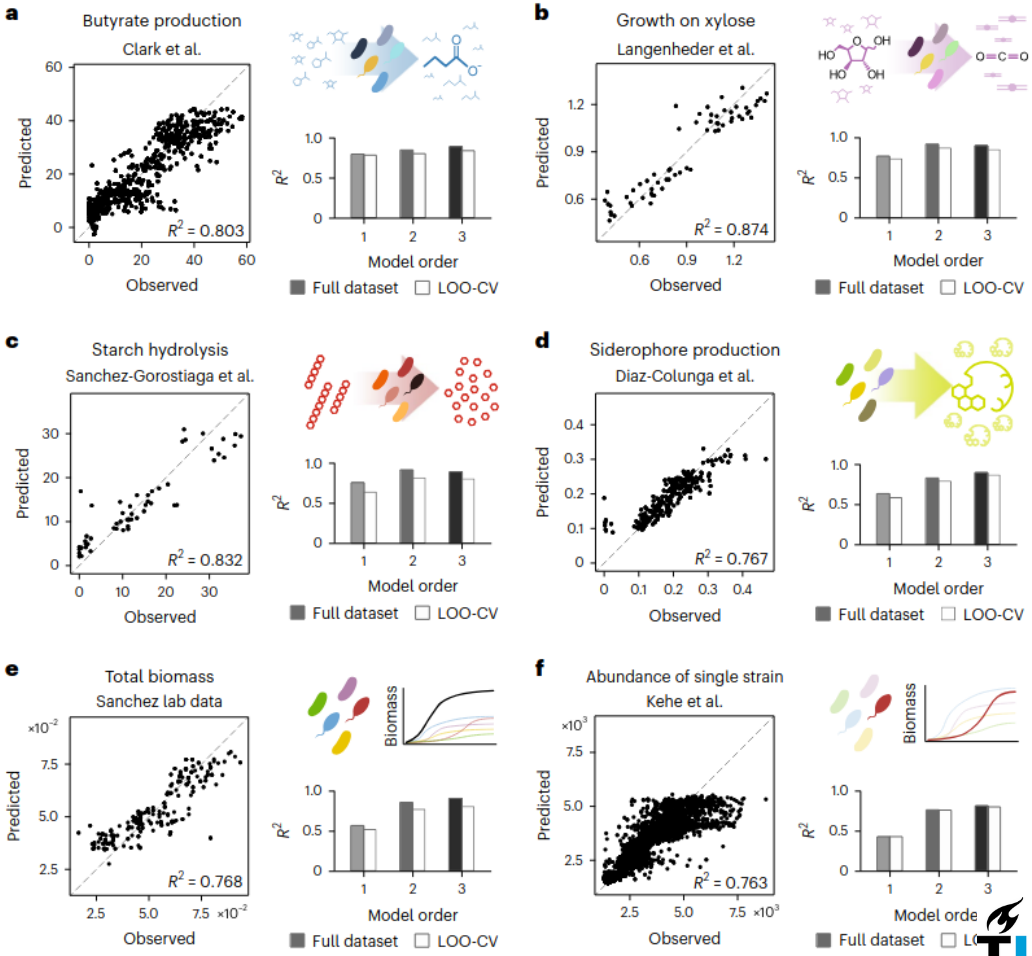
Effect of species $i$ and $j$ When both present (**interaction**)

Presence/absence of species $i$ in the community

# Linear regression works quite OK!

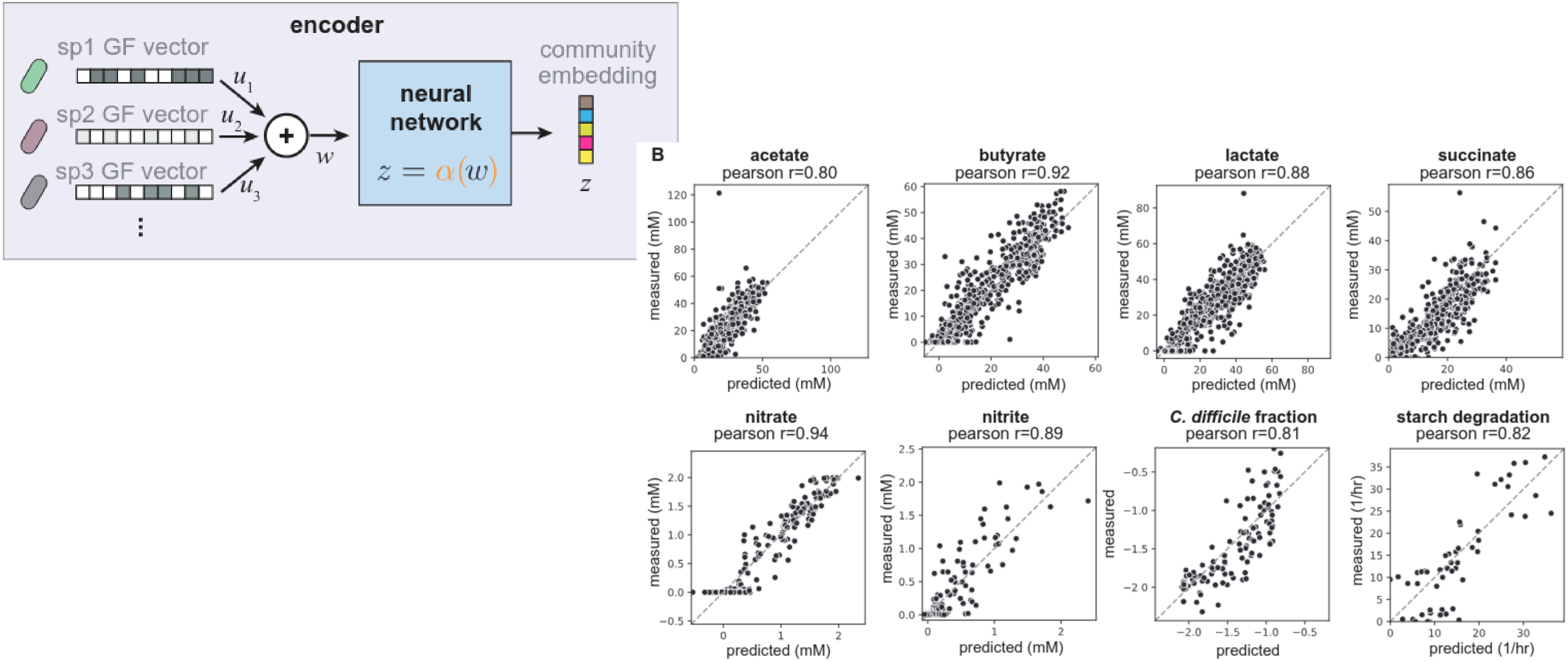
$$y = \beta_0 + \sum_i \beta_i x_i + \sum_{i<j} \beta_{ij} x_i x_j + \dots ,$$





Skwara et al, Nat. Eco. Evo. 2023

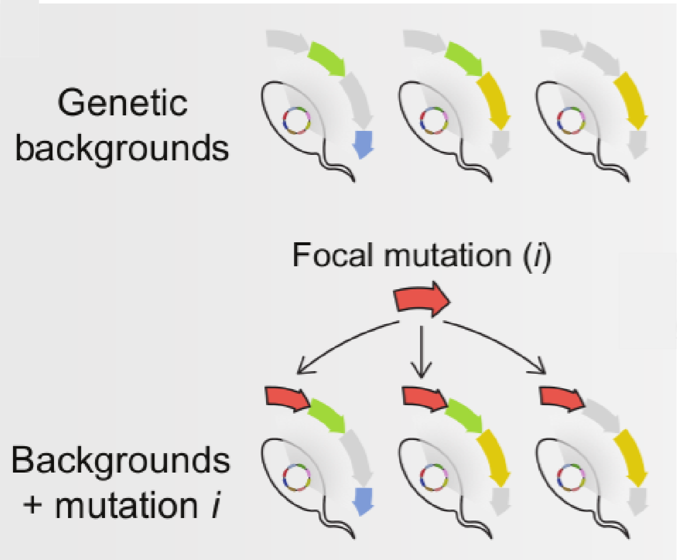
# More complex approaches: machine learning



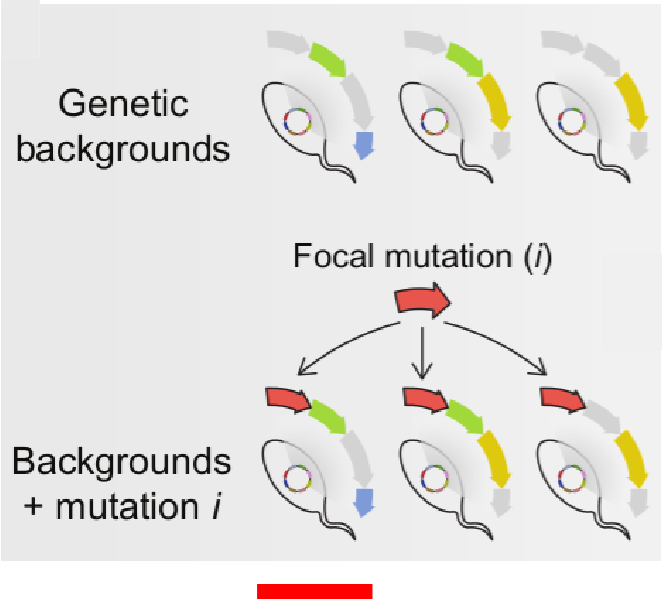
Qian et al, biorxiv 2025

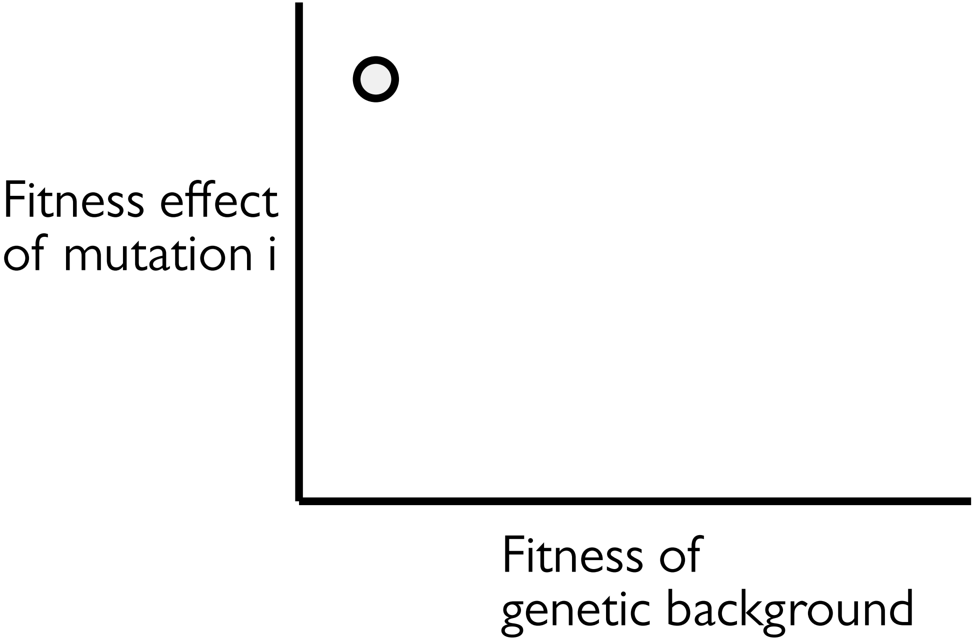
# *Global* epistasis:
# A new way to look at interactions

# Back to genomes: let's test the effect of a mutation, one background at a time

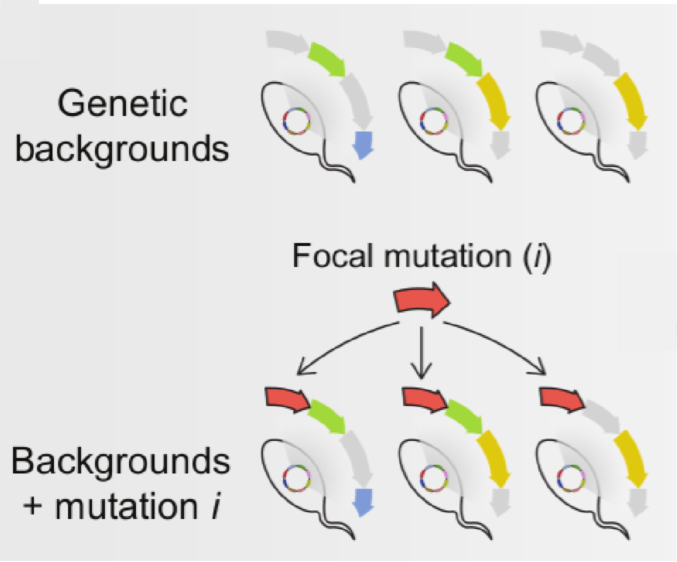


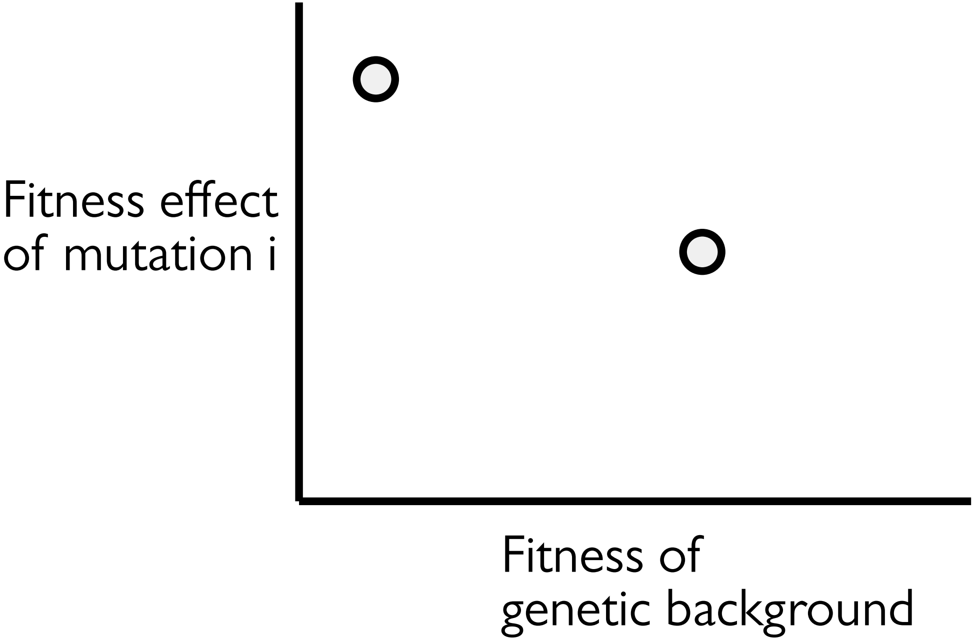
# Back to genomes: let's test the effect of a mutation, one background at a time



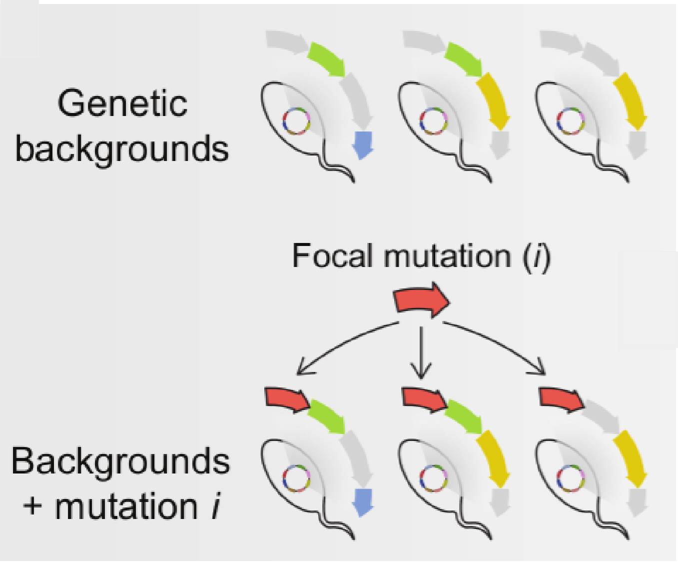


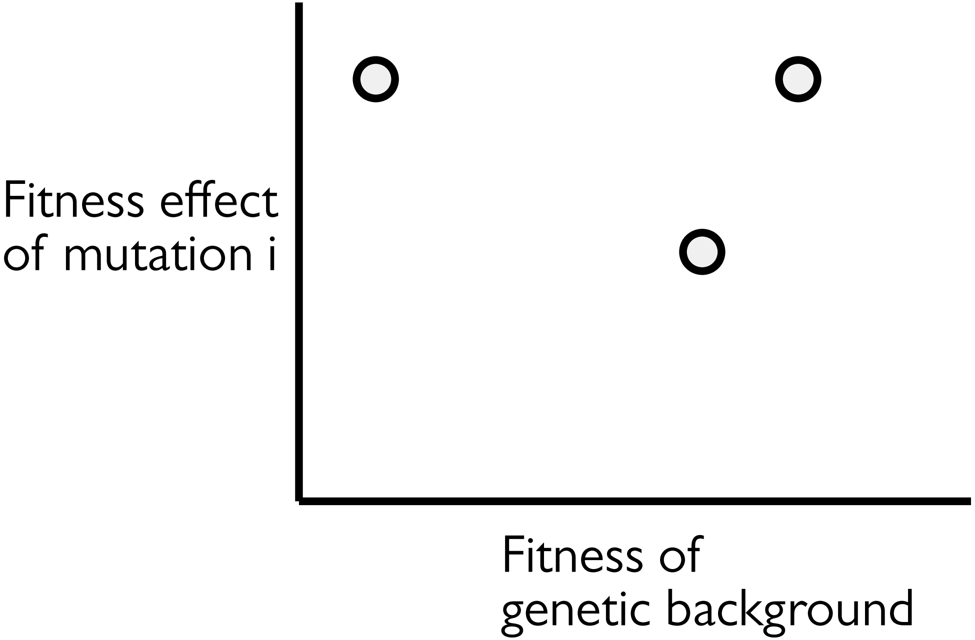
# Testing a mutation, one background at a time



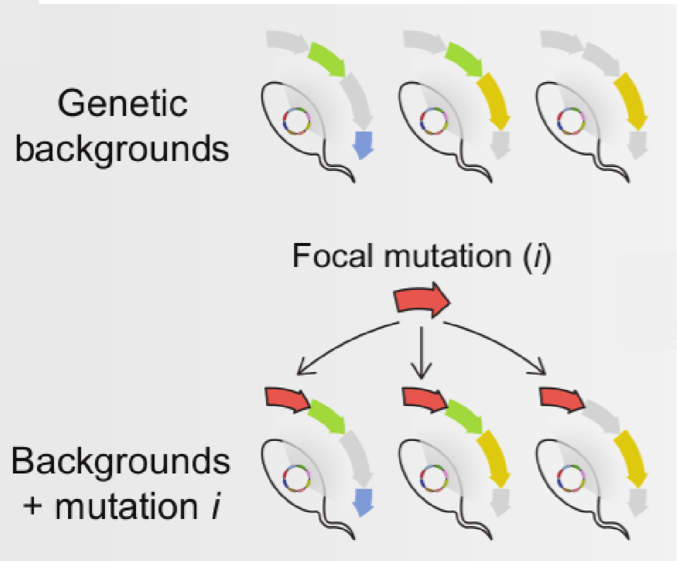


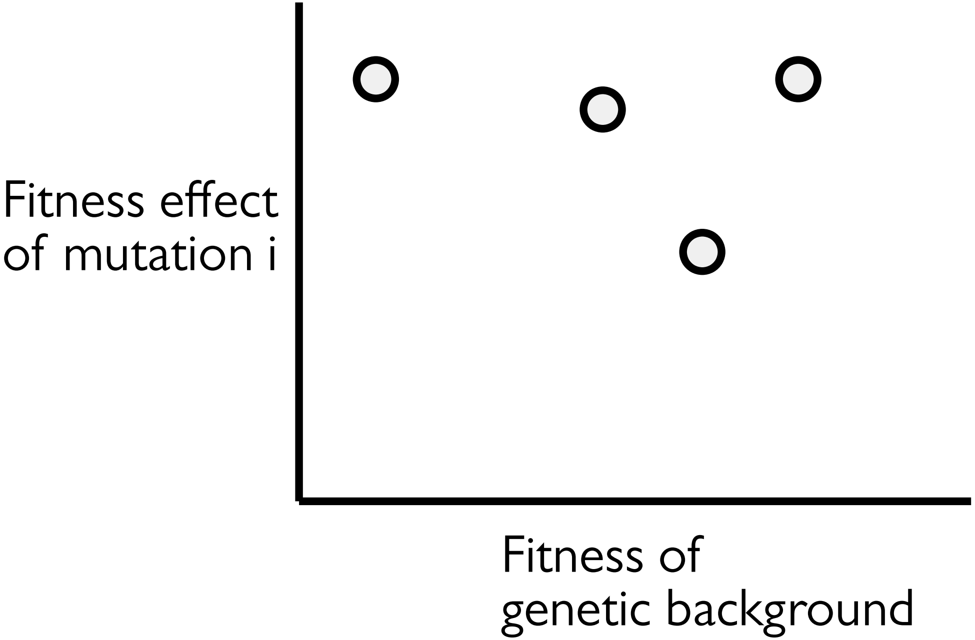
# Testing a mutation, one background at a time



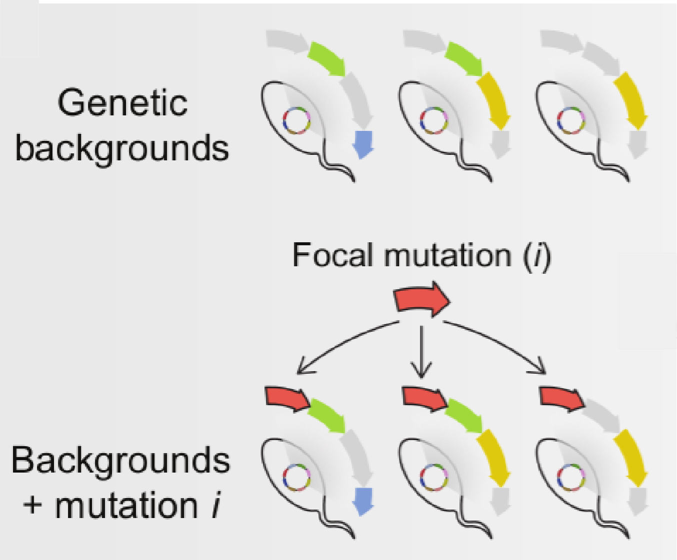


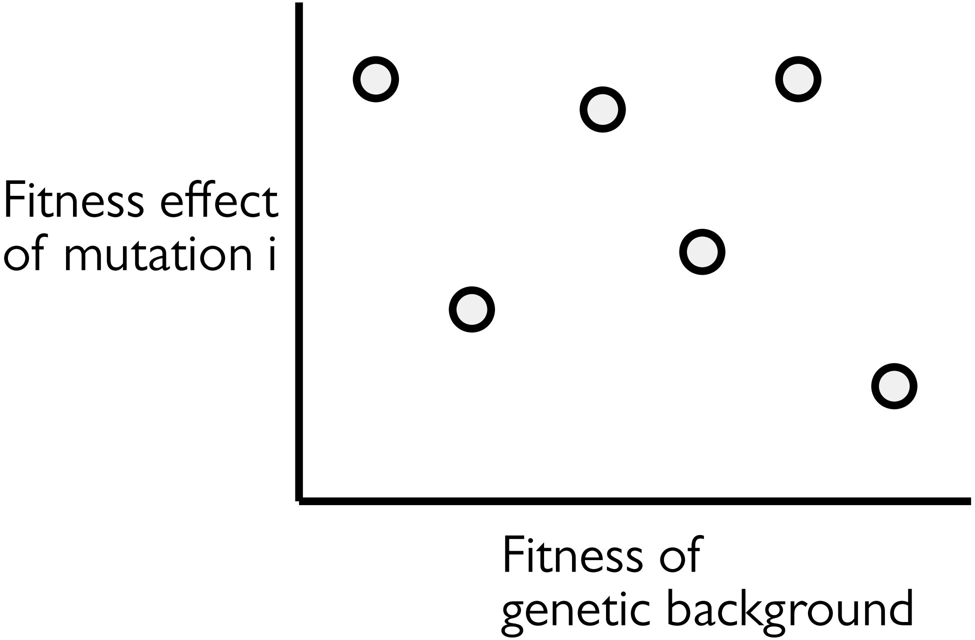
# Testing a mutation, one background at a time



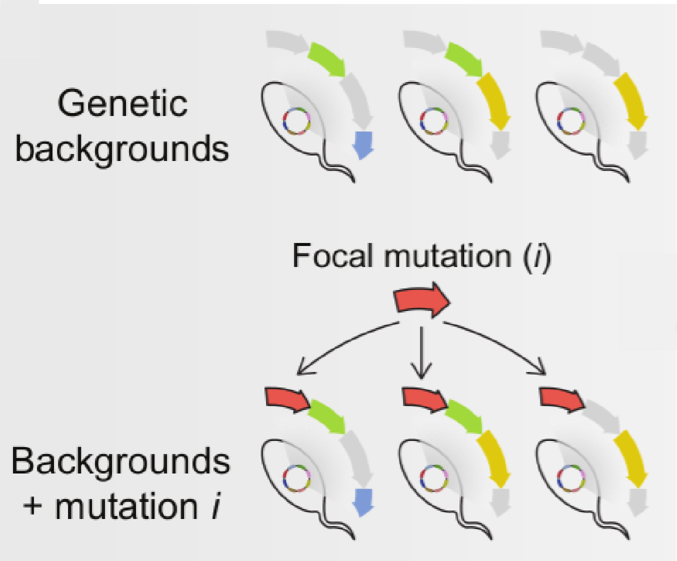


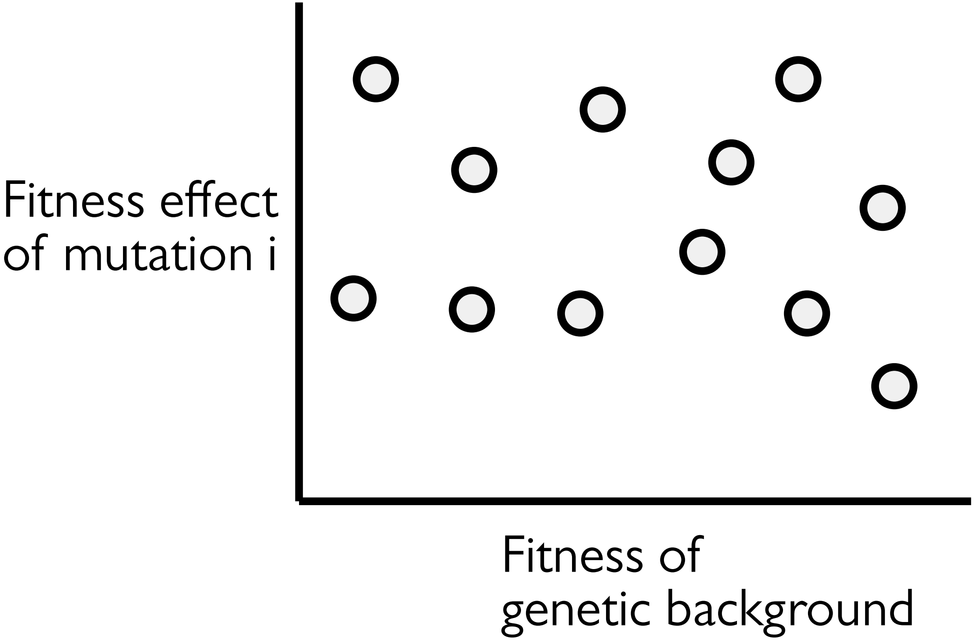
# Testing a mutation, one background at a time



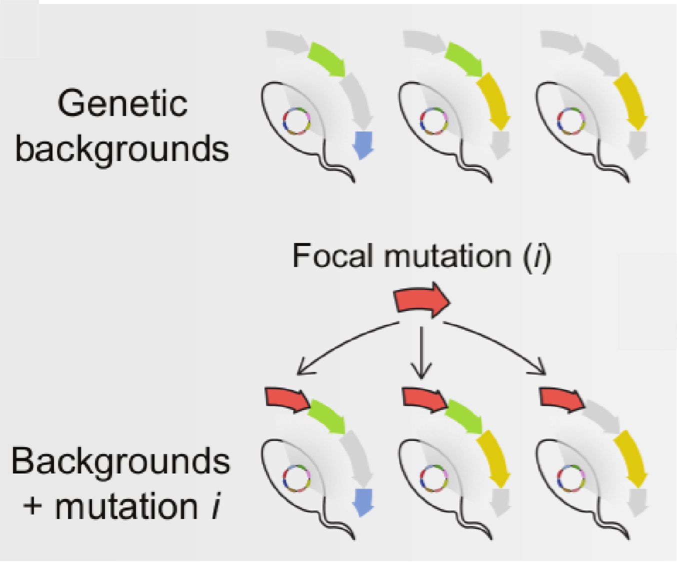


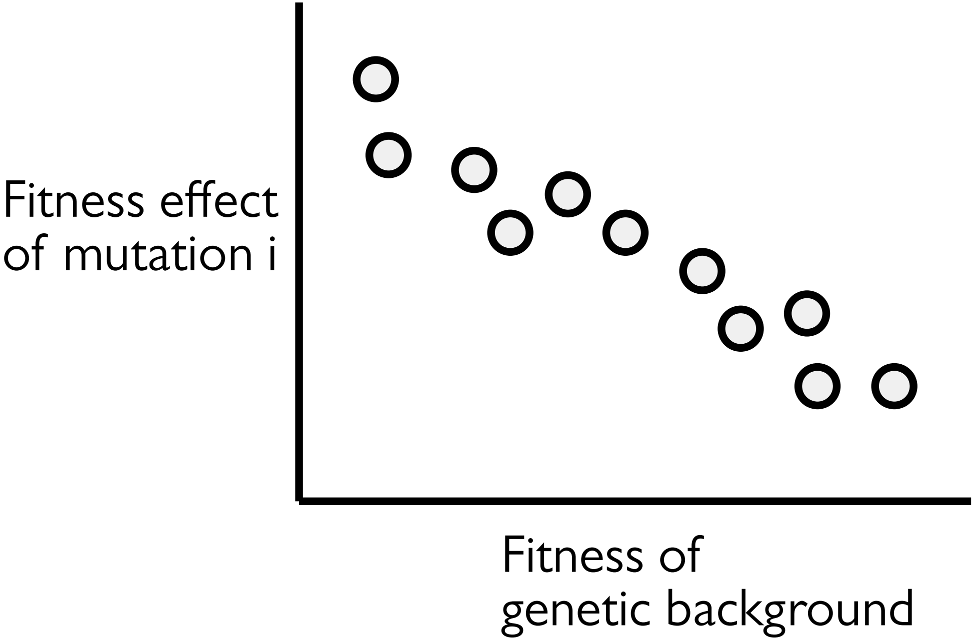
# Intuitively, interactions will lead to rugged landscapes, where making predictions is hard



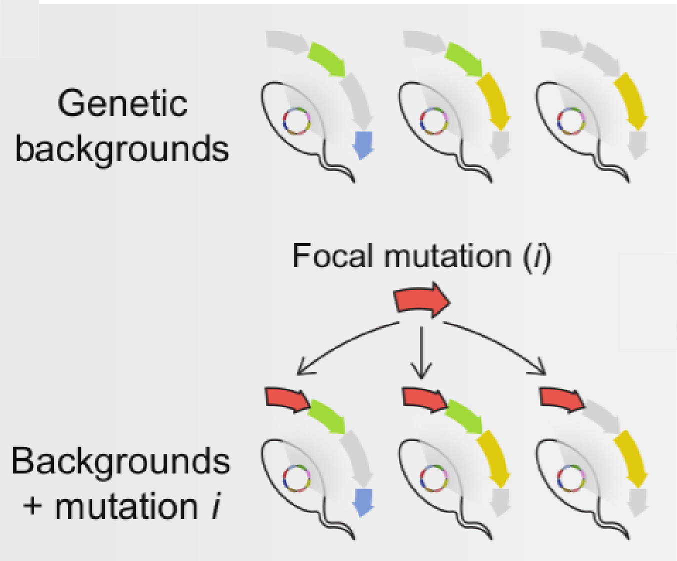


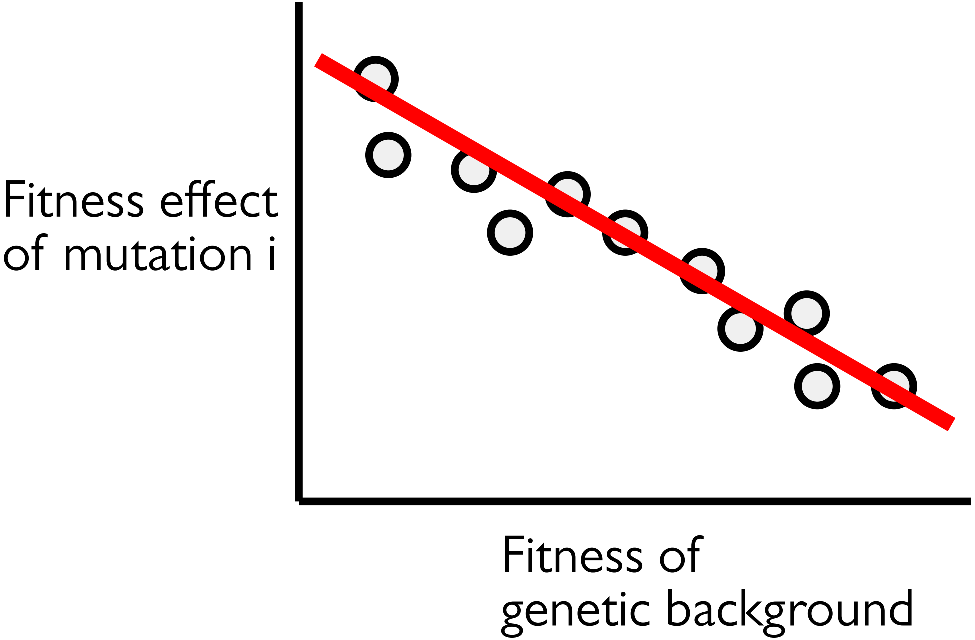
# **Global epistasis** is an emergent linear scaling pattern found in many empirical fitness landscapes



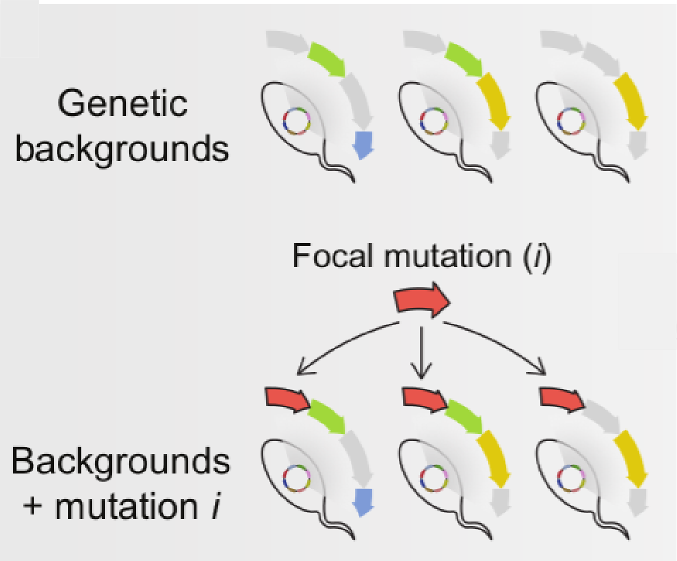


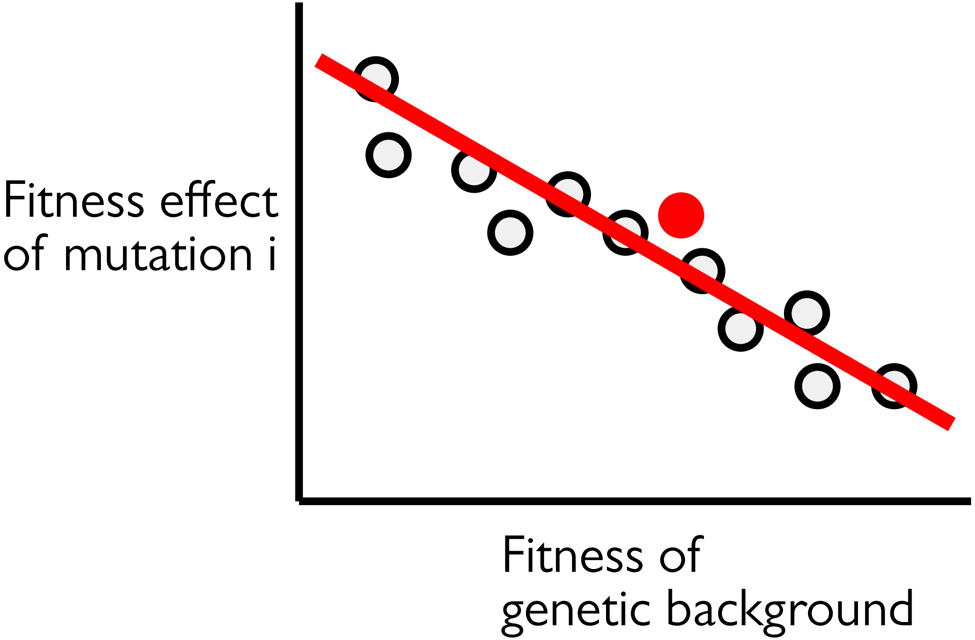
# **Global epistasis** is an emergent linear scaling pattern found in many empirical fitness landscapes



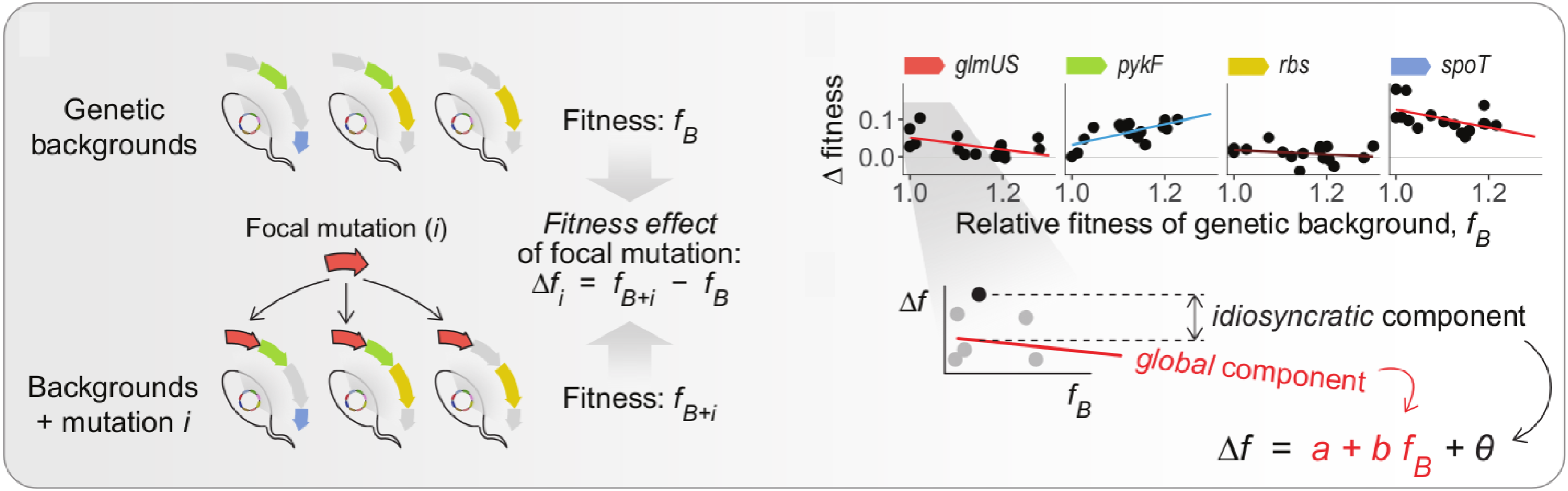


# Global epistasis can help us make **predictions**





# Global epistasis is very common in empirical fitness landscapes



Does global epistasis exist also in structure function landscapes of microbial communities?

# Does global epistasis exist also in structure function landscapes of microbial communities?


Juan Diaz
Colunga
(IPLA, CSIC)


Abby
Skwara
(Yale)


Álvaro
Sánchez
(CBGP, CSIC)

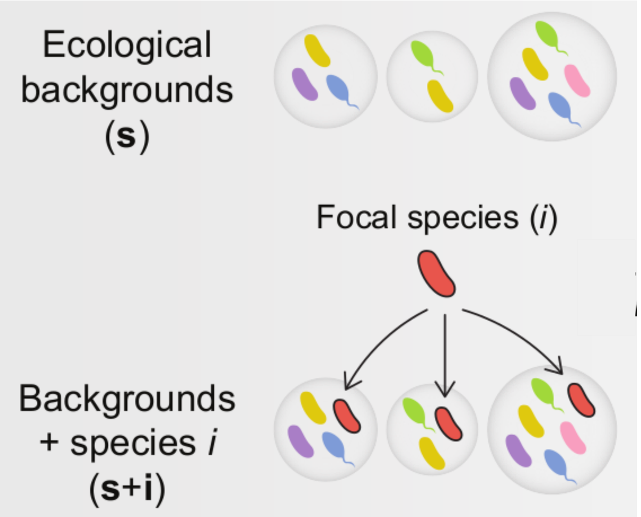

Theory

Cell

**Global epistasis and the emergence of function in microbial consortia**

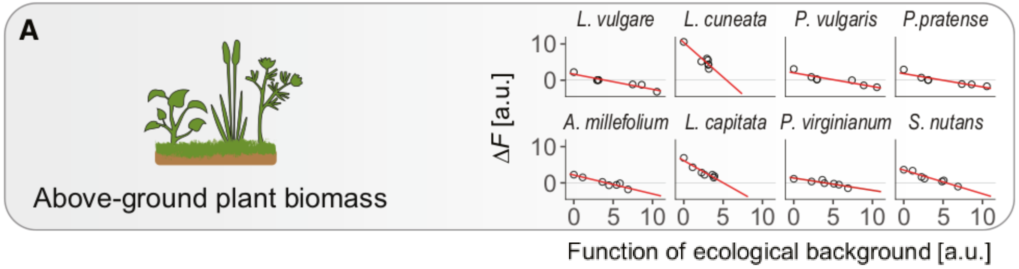
Juan Diaz-Colunga,[1,2,3,4,7,*] Abigail Skwara,[1,2,7] Jean C.C. Vila,[1,2,5] Djordje Bajic,[1,2,6,*] and Alvaro Sanchez[1,2,3,4,8,*]

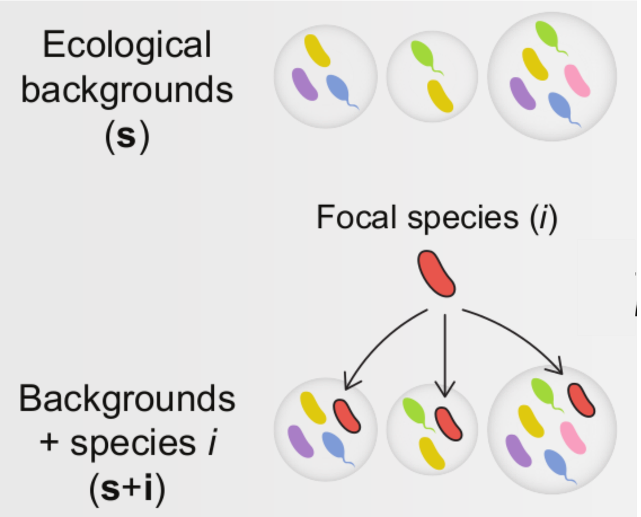
# Is there global epistasis also in ecological structure-function landscapes?



Diaz-Colunga*, Skwara* et al, Cell, 2024

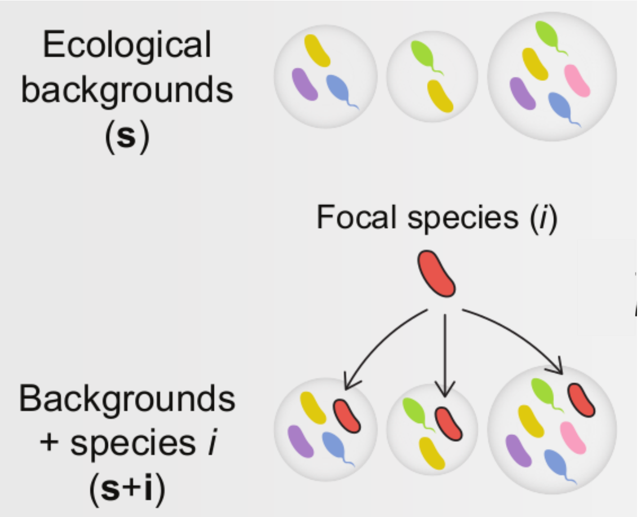
# Is there global epistasis also in ecological structure-function landscapes?

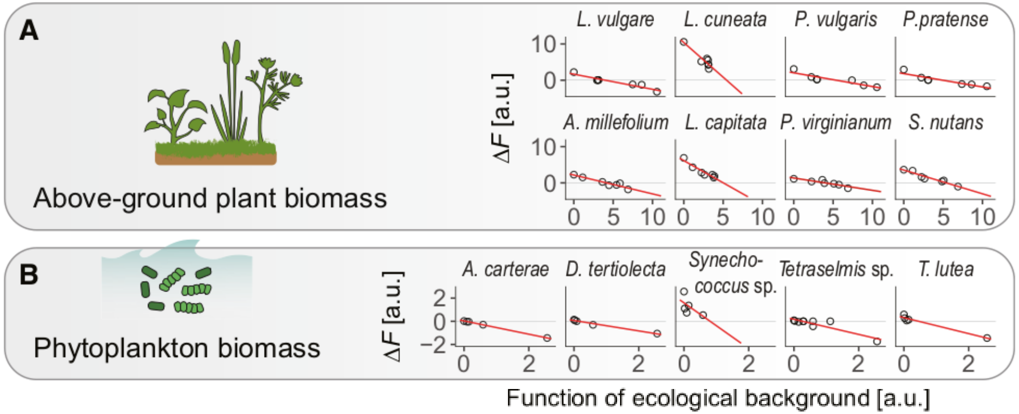




Diaz-Colunga*, Skwara* et al, Cell, 2024

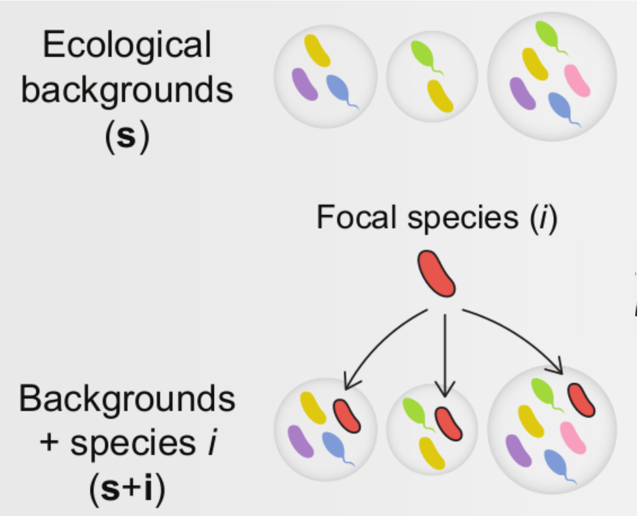
# Is there global epistasis also in ecological structure-function landscapes?

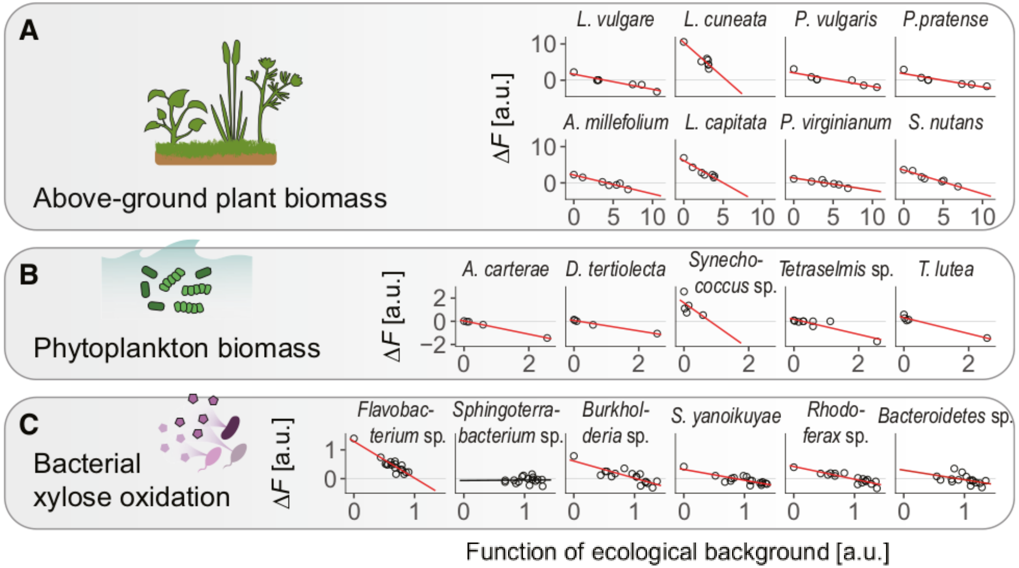




Diaz-Colunga*, Skwara* et al, Cell, 2024

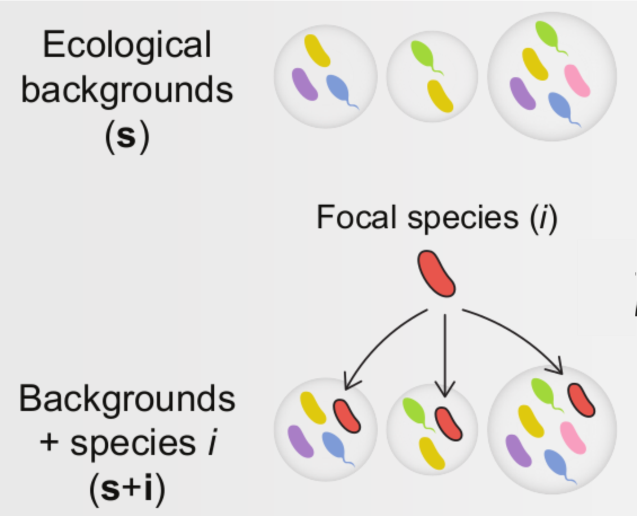
# Is there global epistasis also in ecological structure-function landscapes?

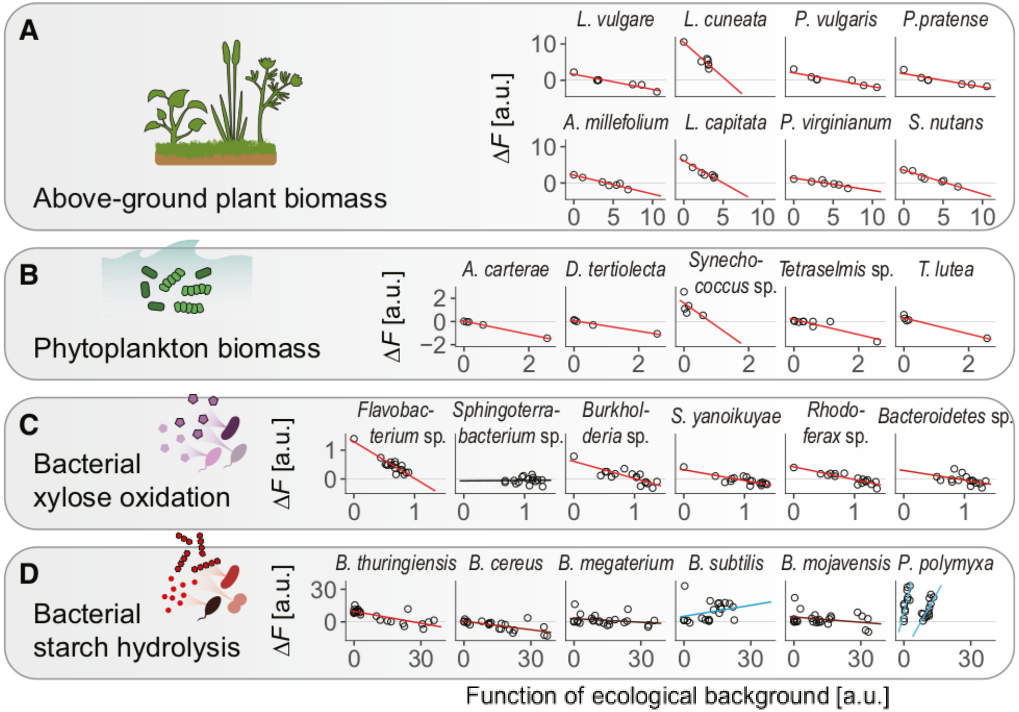




Diaz-Colunga*, Skwara* et al, Cell, 2024

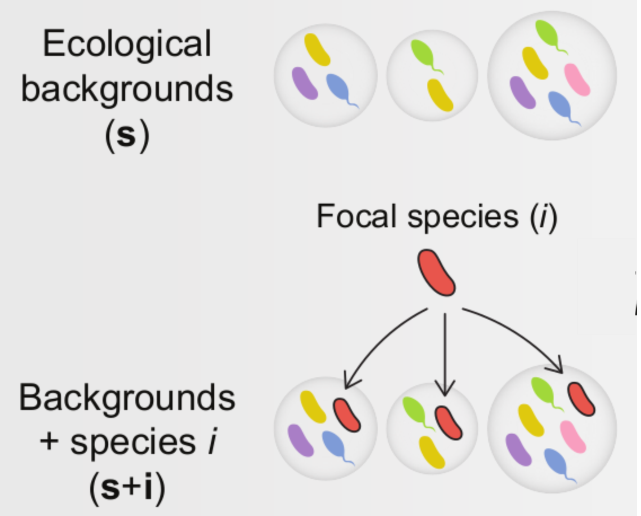
# Is there global epistasis also in ecological structure-function landscapes?





Diaz-Colunga*, Skwara* et al, Cell, 2024

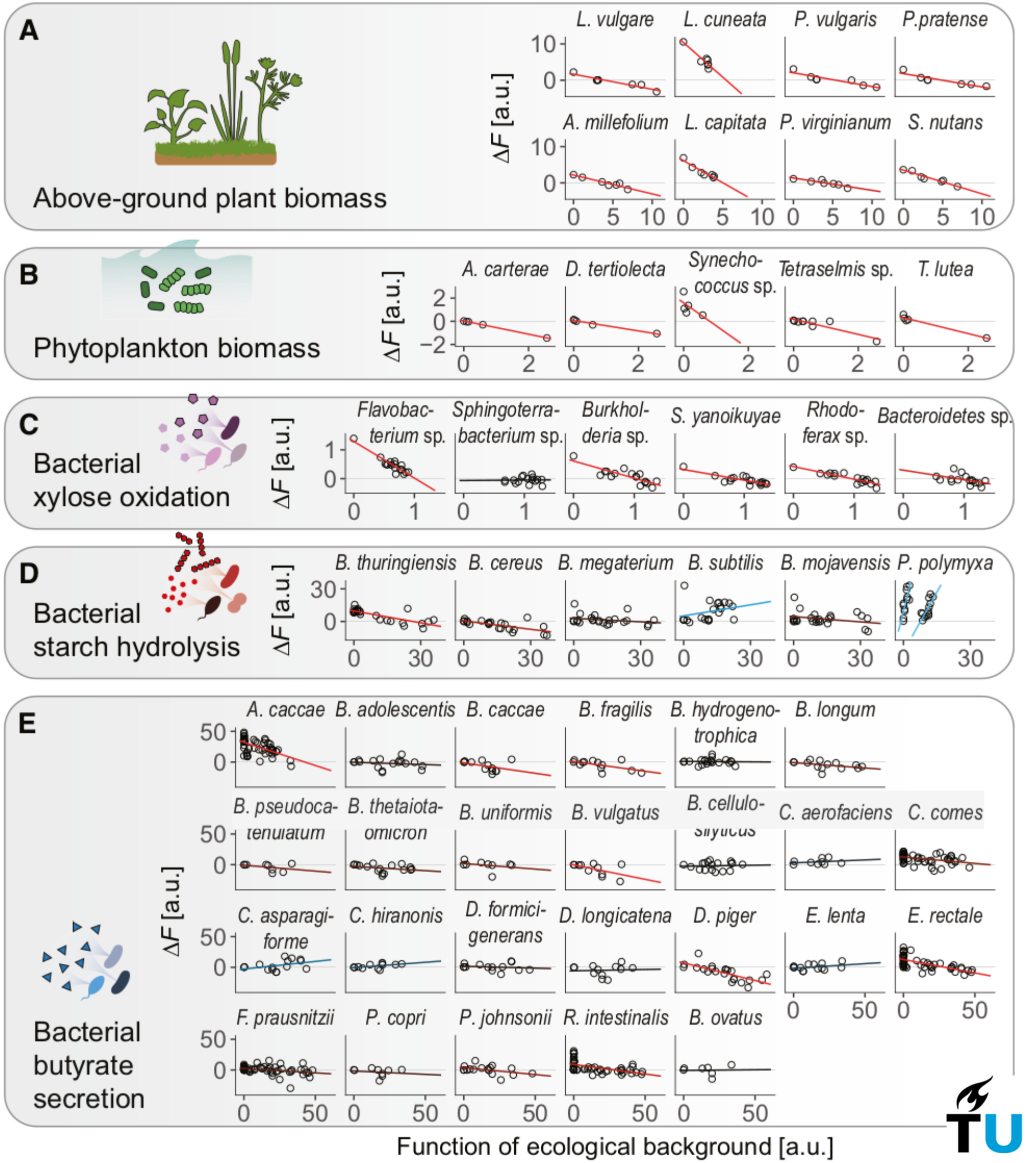
# Is there global epistasis also in ecological structure-function landscapes?



**Ecological backgrounds (s)**

Focal species (*i*)

**Backgrounds + species *i* (s+i)**



Diaz-Colunga*, Skwara* et al, Cell, 2024

# How can we use global epistasis to make predictions?

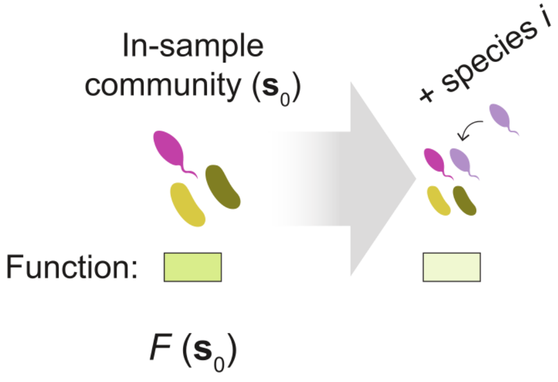
In-sample community ($\mathbf{s}_0$)

Function: 🟩

$F(\mathbf{s}_0)$

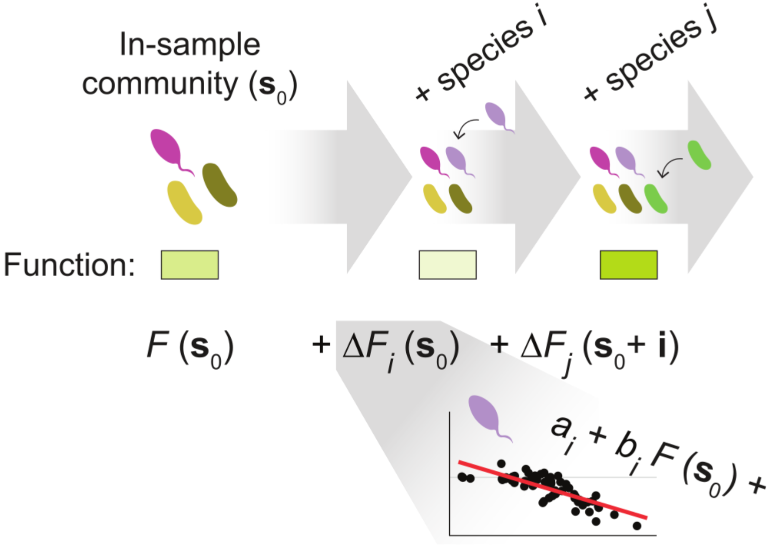
Diaz-Colunga*, Skwara* et al, Cell, 2024

# How can we use global epistasis to make predictions?



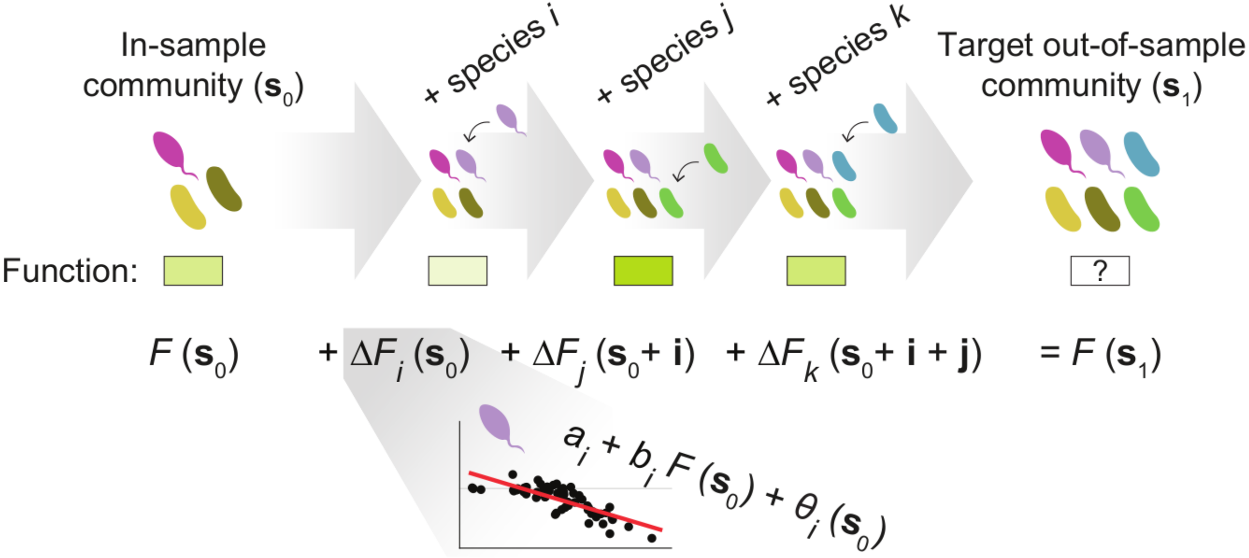
Diaz-Colunga*, Skwara* et al, Cell, 2024

# How can we use global epistasis to make predictions?



Diaz-Colunga*, Skwara* et al, Cell, 2024

# How can we use global epistasis to make predictions?

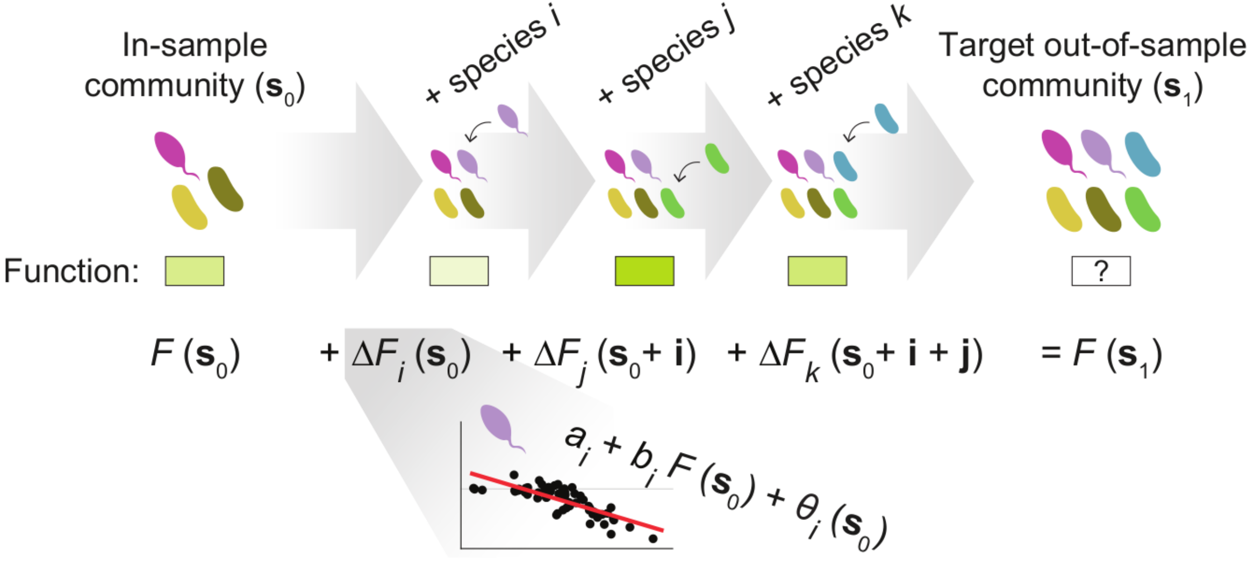


$$F(\mathbf{s}_0) \quad + \Delta F_i(\mathbf{s}_0) \quad + \Delta F_j(\mathbf{s}_0 + \mathbf{i}) \quad + \Delta F_k(\mathbf{s}_0 + \mathbf{i} + \mathbf{j}) \quad = F(\mathbf{s}_1)$$

$$a_i + b_i F(\mathbf{s}_0) + \theta_i(\mathbf{s}_0)$$

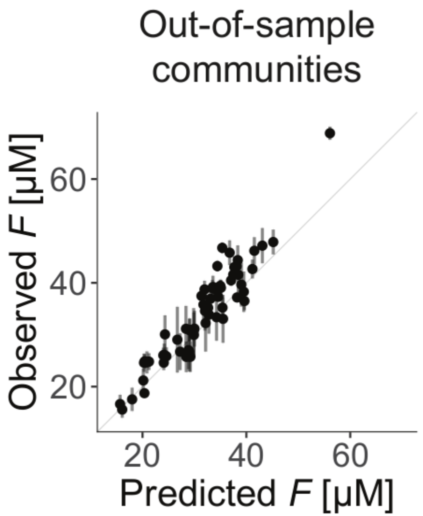
Diaz-Colunga*, Skwara* et al, Cell, 2024

# How can we use global epistasis to make predictions?



$$F(\mathbf{s}_0) \quad + \Delta F_i(\mathbf{s}_0) \quad + \Delta F_j(\mathbf{s}_0 + \mathbf{i}) \quad + \Delta F_k(\mathbf{s}_0 + \mathbf{i} + \mathbf{j}) \quad = F(\mathbf{s}_1)$$

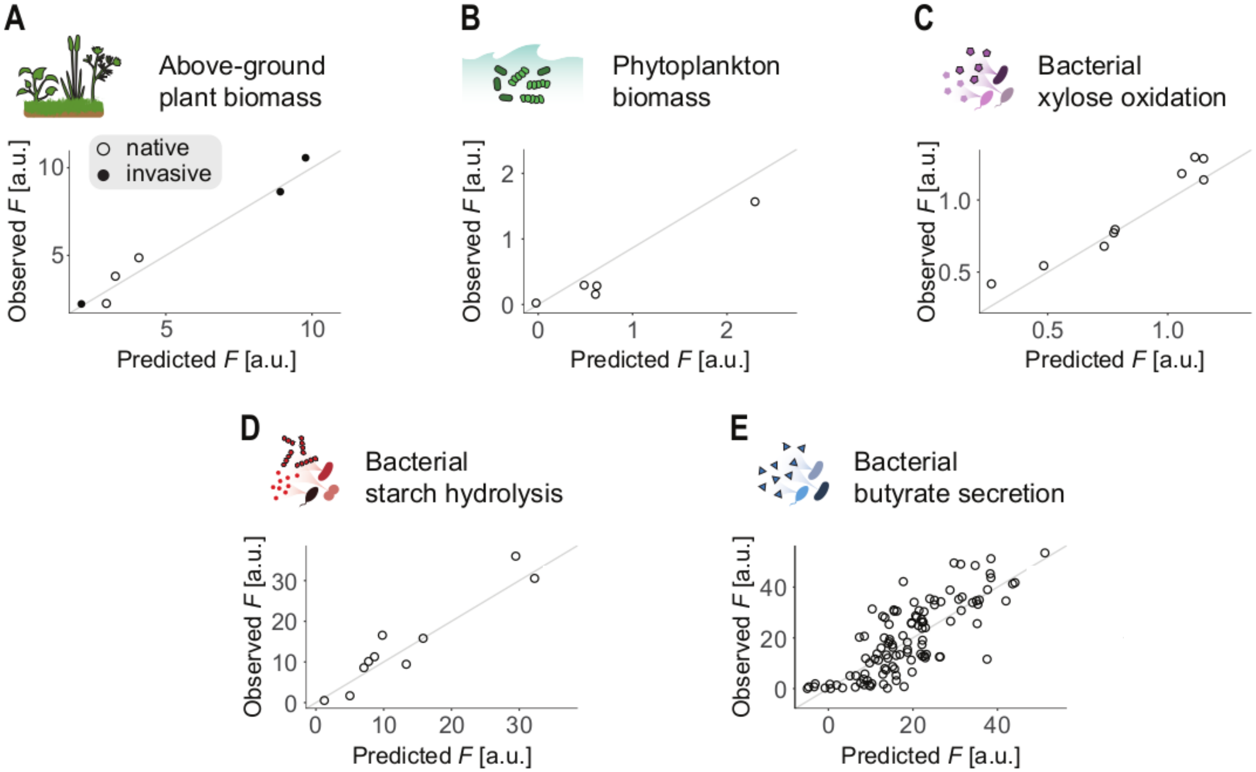
$$a_i + b_i F(\mathbf{s}_0) + \theta_i(\mathbf{s}_0)$$



$F$ = Pyoverdine production,
$R^2$=0.81

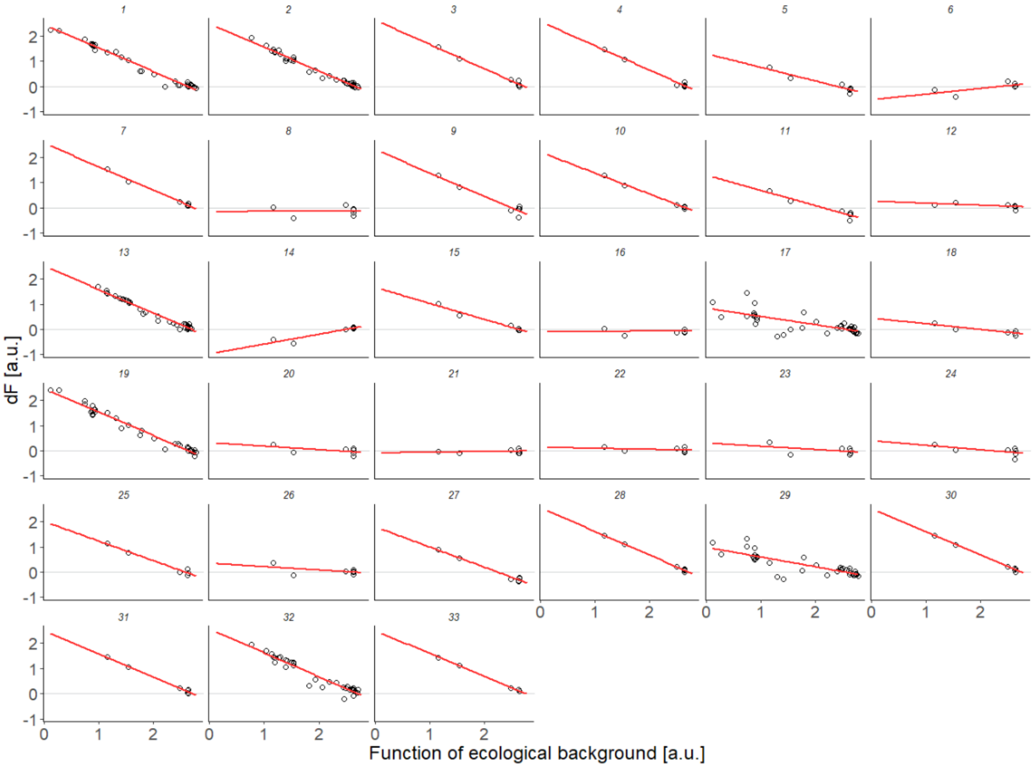
Diaz-Colunga*, Skwara* et al, Cell, 2024

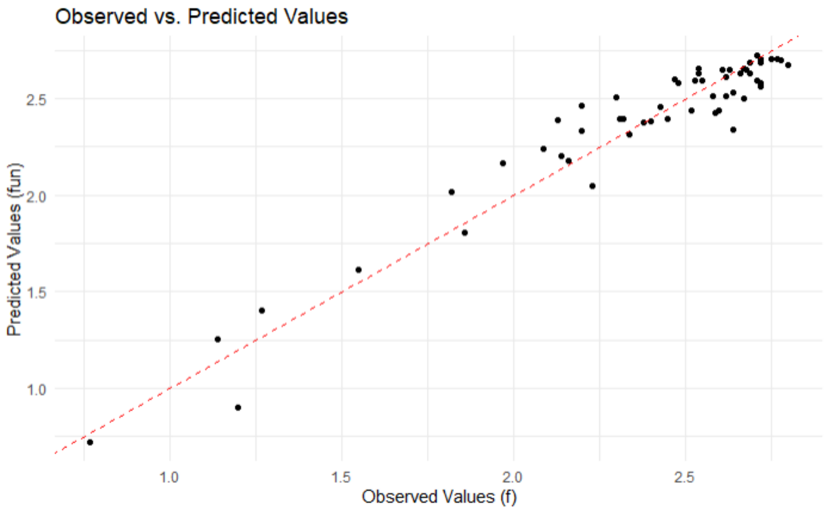
# Global epistasis shows very promising predictive power

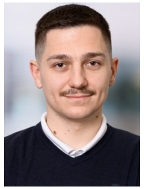


Diaz-Colunga*, Skwara* et al, Cell, 2024

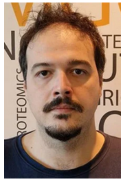
# Global epistasis is able to predict **acidification** in synthetic soymilk-fermenting communities






Uros Gojković


Nemanja
Stanisavljevic
(IMGGE)


Nemanja
Kljajević
(IMGGE)

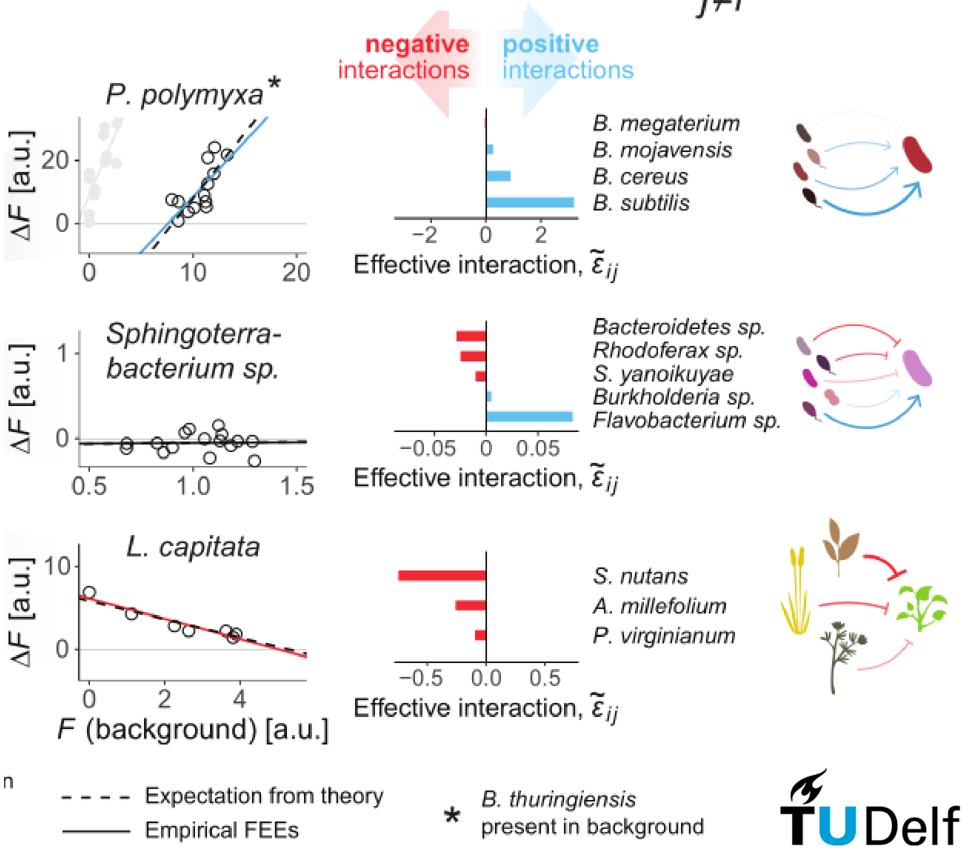
Gojković et al, in prep

TU Delft

# Global epistasis is a spotlight to identify relevant interactions

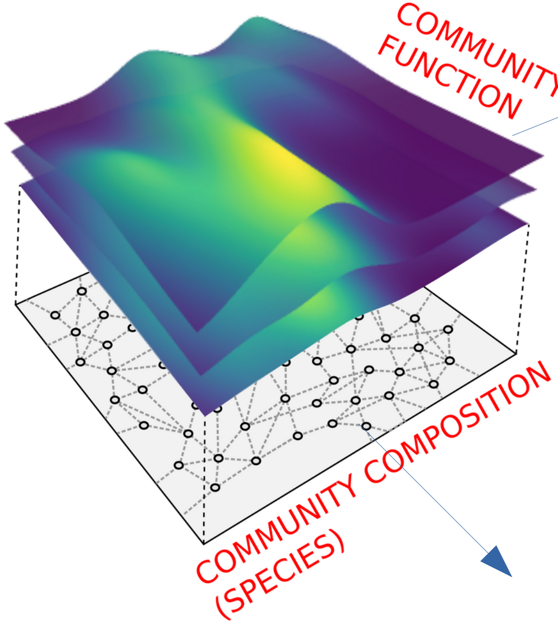
slope of species *i*
expected from theory

$$b_i \approx \sum_{j \neq i} \tilde{\varepsilon}_{ij}$$



Diaz-Colunga*, Skwara* et al, Cell, 2024

Can we generalize this even further, beyond microbial communities?

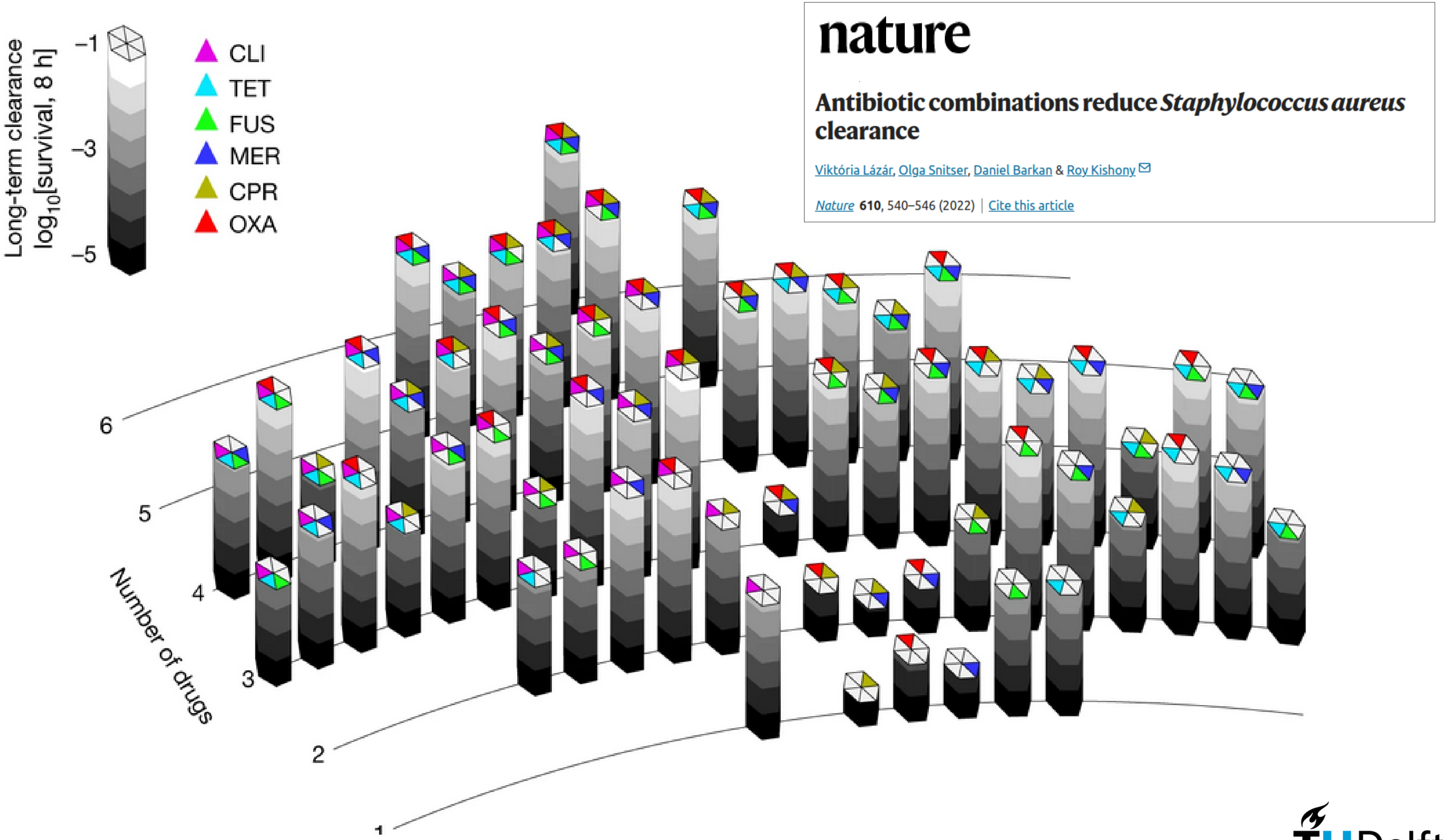


Function: How much it kills
*Staphylococcus aureus*?

Structure: composition of an
antibiotic cocktail

Floris
Klein Obbink

Minke
Gabriëls



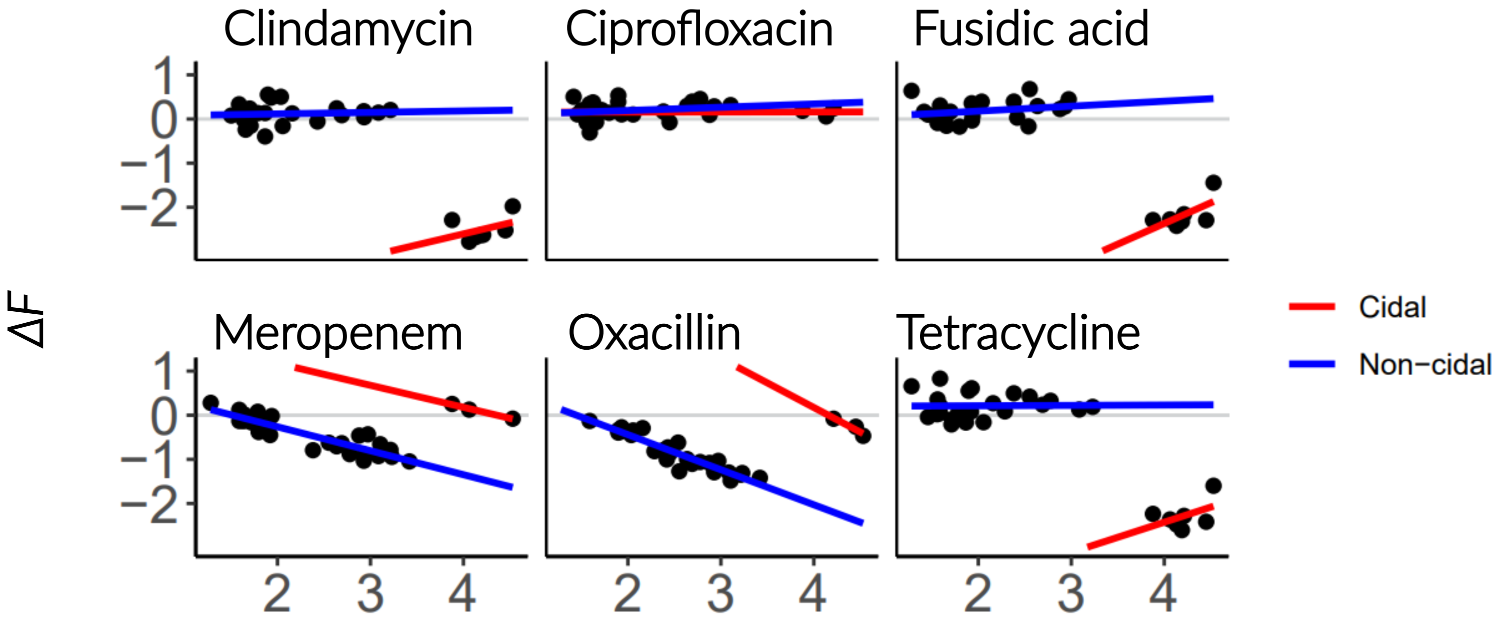
**nature**

**Antibiotic combinations reduce *Staphylococcus aureus* clearance**

Viktória Lázár, Olga Snitser, Daniel Barkan & Roy Kishony

*Nature* **610**, 540–546 (2022) | Cite this article

# **Antibiotic combinations** exhibit global epistasis in their effects on *Staphylococcus aureus* clearance

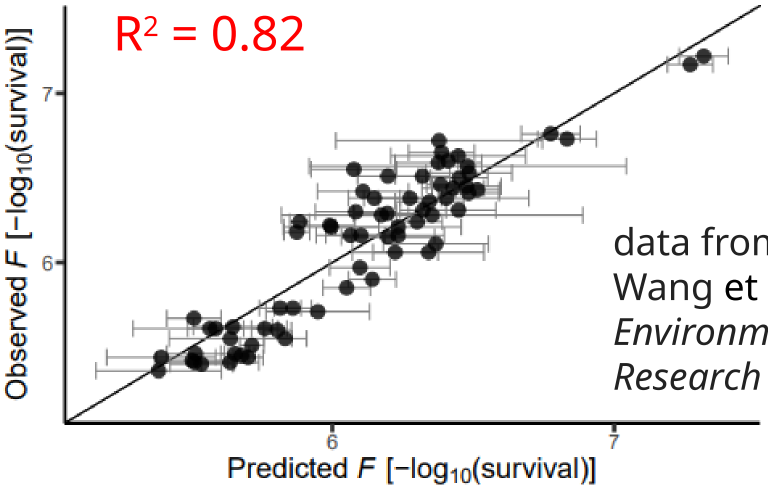


Background function

$F$ = -log(*S. aureus* survival)

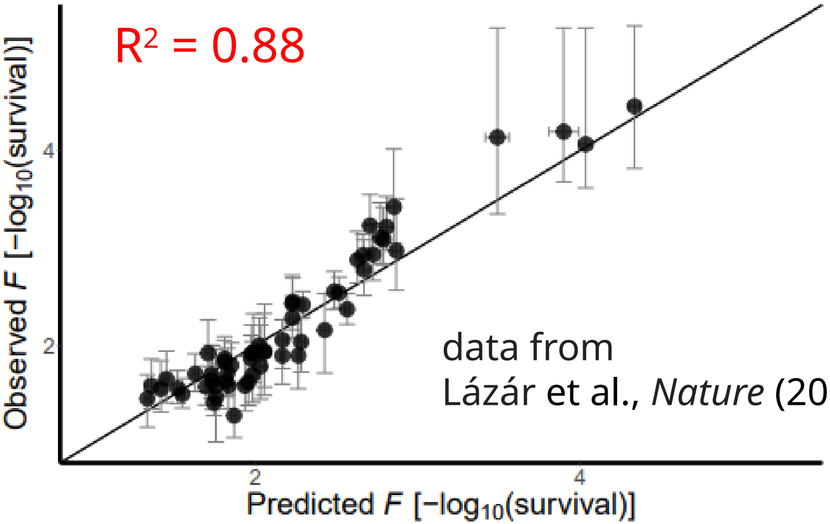
(atricle in prep)

# The effect of unknown antibiotic cocktails can be predicted using global epistasis!



$R^2 = 0.82$

data from
Wang et al.,
*Environmental Research* (2018)

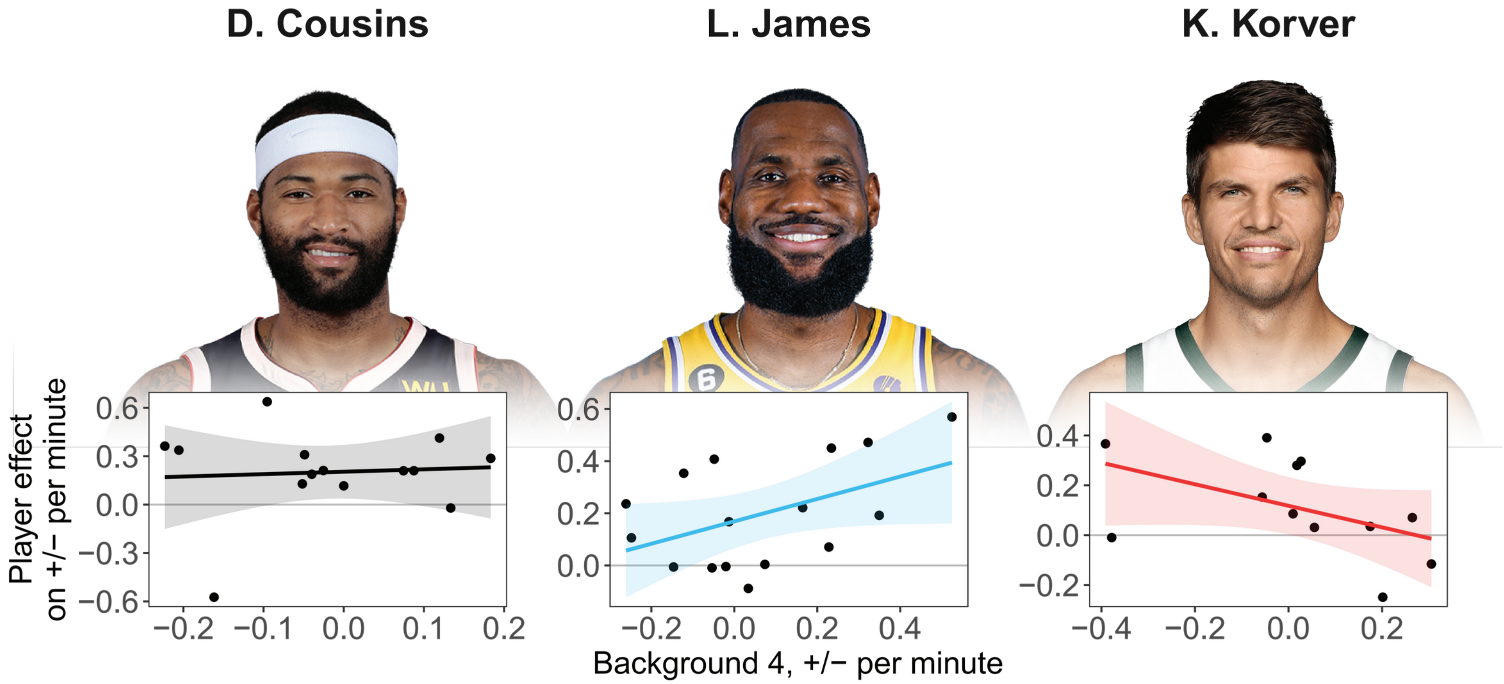


$R^2 = 0.88$

data from
Lázár et al., *Nature* (2022)

(article in prep)

# And we can take it even further!



(Real data for NBA season 2015-16)

# Collaborators

Alvaro Sanchez (IBFG, CBGP, CSIC, Spain)
Juan Diaz Colunga (IBFG, IPLA, CSIC, Spain)
Abby Skwara (Yale University, USA)
Nemanja Stanisavljević (IMGGE, Serbia)
Nemanja Kljajević (IMGGE, Serbia)
Viktoria Lazar (BRC, Szeged, Hungary)
Ivica Dimkić (University of Belgrade, Serbia)
Ilija Dukovski (Boston University, USA)
Daniel Segre (Boston University, USA)
Sotaro Takano (AIST, Japan)
Jean Vila (Stanford, USA)
Ludovic Jourdin (TU Delft, The Netherlands)
Jean-Marc Daran (TU Delft, Netherlands)
Robbert Kleerebezem (TU Delft, Netherlands)

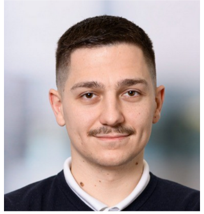
# Funding

Zero Emission Biotechnology Program @ TU Delft

Thank you!

## Group @ TU Delft


Minke
Gabriëls


Uroš
Gojković


Rienk
Van Vliet

# Collaborators

Alvaro Sanchez (IBFG, CBGP, CSIC, Spain)
Juan Diaz Colunga (IBFG, IPLA, CSIC, Spain)
Abby Skwara (Yale University, USA)
Nemanja Stanisavljević (IMGGE, Serbia)
Nemanja Kljajević (IMGGE, Serbia)
Viktoria Lazar (BRC, Szeged, Hungary)
Ivica Dimkić (University of Belgrade, Serbia)
Ilija Dukovski (Boston University, USA)
Daniel Segre (Boston University, USA)
Sotaro Takano (AIST, Japan)
Jean Vila (Stanford, USA)
Ludovic Jourdin (TU Delft, The Netherlands)
Jean-Marc Daran (TU Delft, Netherlands)
Robbert Kleerebezem (TU Delft, Netherlands)

# Funding

Zero Emission Biotechnology Program @ TUDelft

Thank you!

## Group @ TUDelft

Minke Gabriëls
Uroš Gojković
Rienk Van Vliet

Practical Exercise

# Statistical prediction and optimization of microbial community functions

Courtesy of: Juan Díaz-Colunga

Microbial Systems Biology Lab (MSBlab), Spanish National Research Council

10 minute break?