COLLABORATIVE RESEARCH CENTER 1310

Predictability in Evolution

PROBABILISTIC STRUCTURES
IN EVOLUTION

DFG SPP 1590

Evolutionary pathways to antibiotic resistance

Joachim Krug
Institute for Biological Physics
University of Cologne

"Statistical Biological Physics", ICTS Bangalore, December 9, 2020

Evolution of antibiotic resistance

- Resistance evolution is a universal response of microbial pathogens to biomedical interventions

- Evolutionary considerations are key for developing strategies that prevent or delay resistance evolution in clinical or environmental settings

- At the same time microbial resistance evolution serves as a model system for addressing broader questions of evolutionary theory

Evolution of antibiotic resistance

- Resistance evolution is a universal response of microbial pathogens to biomedical interventions
- Evolutionary considerations are key for developing strategies that prevent or delay resistance evolution in clinical or environmental settings
- At the same time microbial resistance evolution serves as a model system for addressing broader questions of evolutionary theory

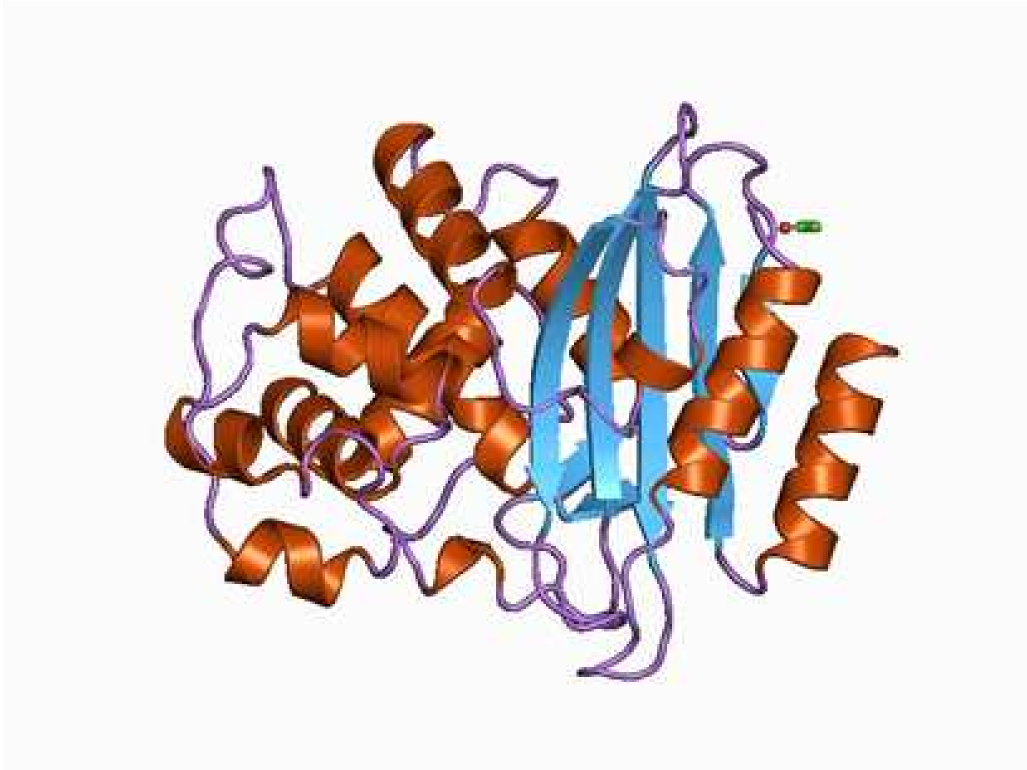
Two case studies

- The resistance landscape of TEM-1 β-lactamase
  joint work with de Visser lab (Wageningen)
- Concentration-dependent evolution of ciprofloxacin resistance
  joint work with Bartek Waclaw and Rosalind Allen (Edinburgh)

Quantifying antibiotic effects

- The effect of a drug is quantified by the dose-reponse curve, the growth rate of a (large) bacterial population as a function of drug concentration
- At the minimal inhibitory concentration (MIC) the growth rate drops to zero
- Resistance mutations increase the MIC
- The combination of multiple mutations along an evolutionary pathway leads to highly resistant strains

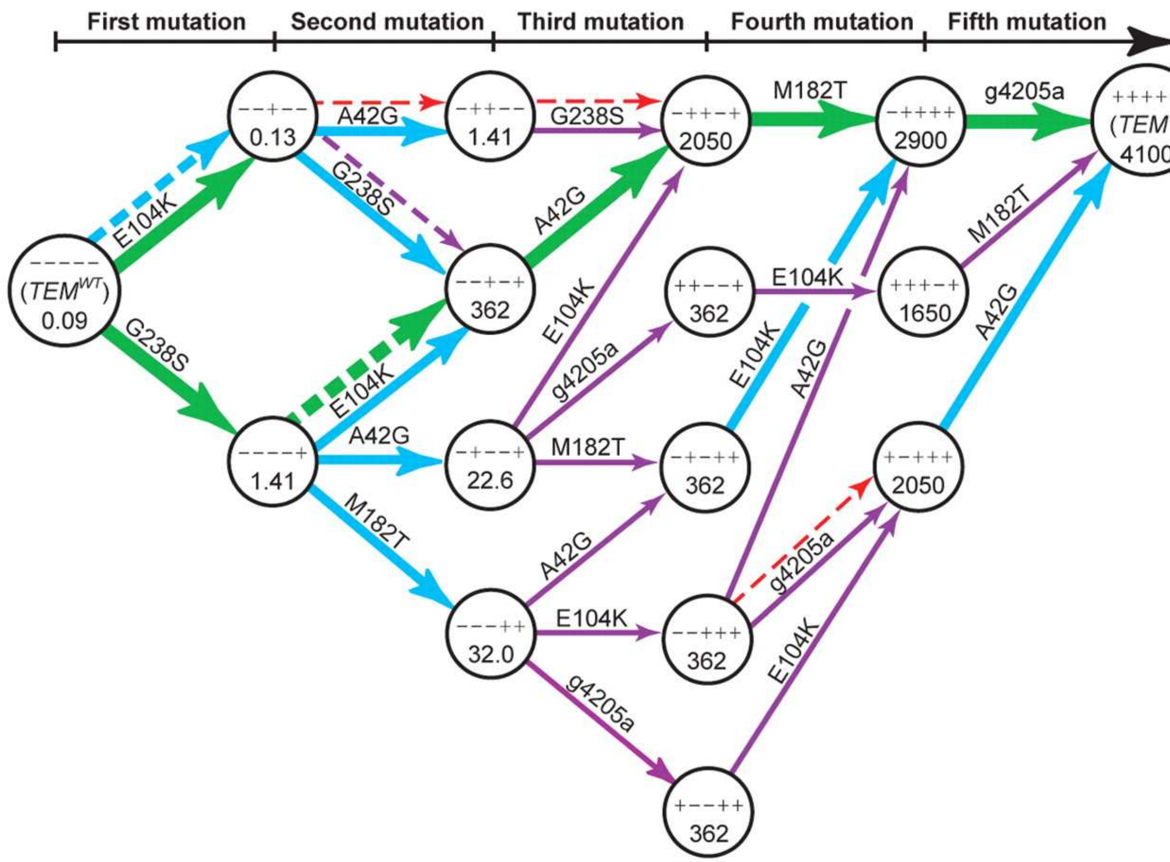
TEM-1 β-lactamase



- β-lactam antibiotics such as penicillin target cell wall synthesis
- TEM-1 β-lactamase confers resistance against ampicillin to *E. coli*
- Experiments study adaptation to novel antibiotic cefotaxime

Pathways to TEM-1 resistance

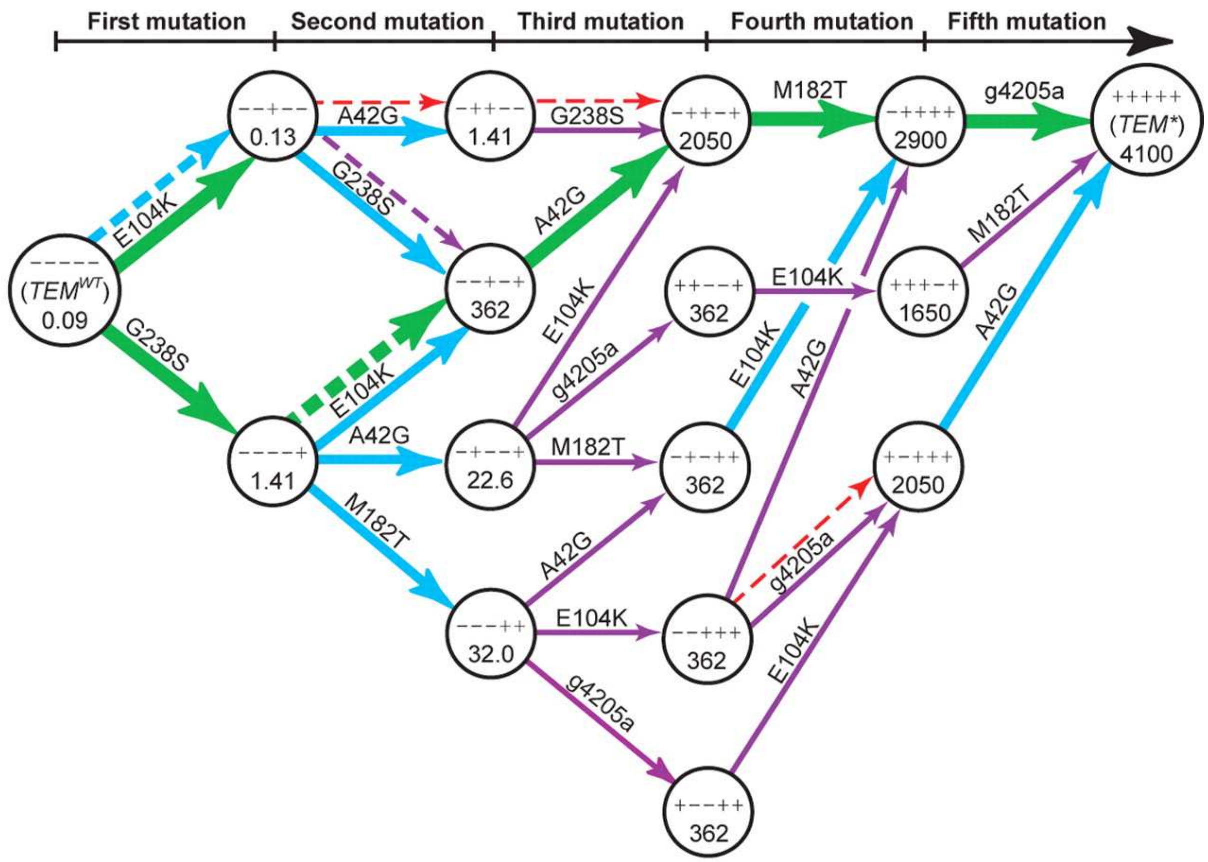
D.M. Weinreich et al., Science **312**, 111 (2006)



- 5 mutations increase the MIC by 4×10^4
- Construct all $2^5 = 32$ combinatorial mutants

Pathways to TEM-1 resistance

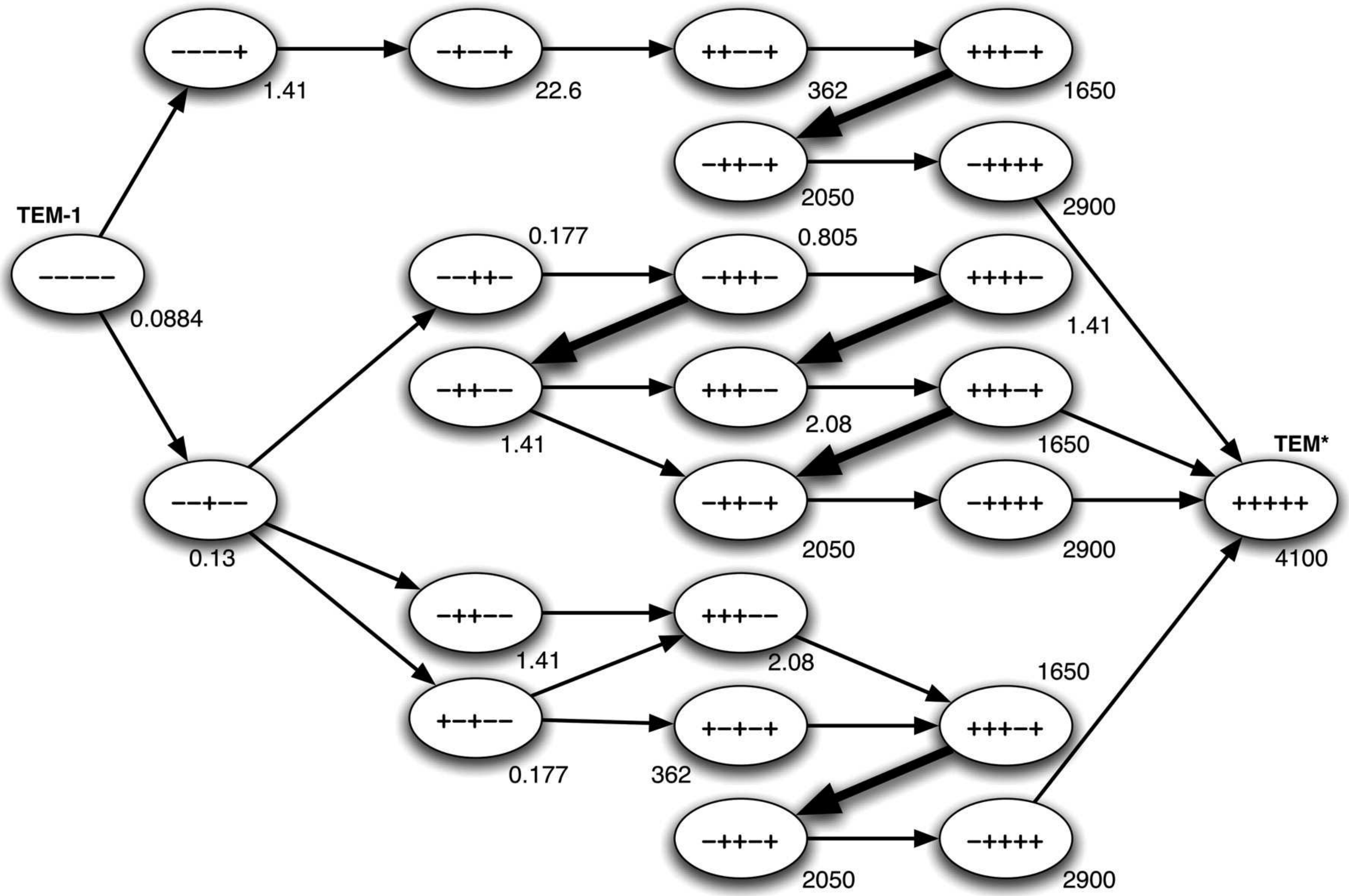
D.M. Weinreich et al., Science **312**, 111 (2006)



- Only 18 out of $5! = 120$ directed mutational pathways are monotonically increasing in resistance, and only a few of them have appreciable weight

Pathways to TEM-1 resistance

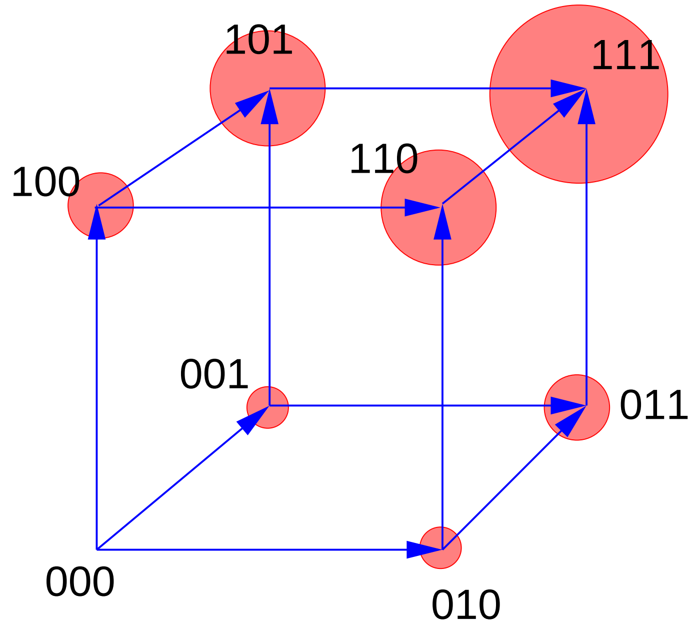
De Pristo et al. 2007



- 27 out of 18651552840 undirected pathways are monotonically increasing

Mathematical framework

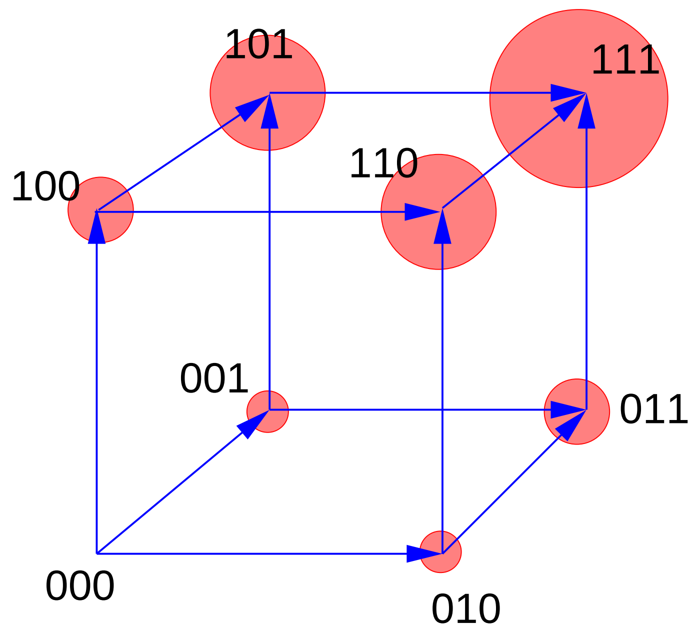
Pathways in fitness landscapes



- Genotypes are encoded by binary sequences $(\sigma_1, \ldots, \sigma_L)$ where $\sigma_i = 1$ ($\sigma_i = 0$) denotes the presence (absence) of a mutation at position i
- A fitness or resistance landscape is a function on the L-dimensional hypercube $\{0,1\}^L$ of genotypes
- The fitness graph is obtained by orienting the links in the direction of increasing fitness

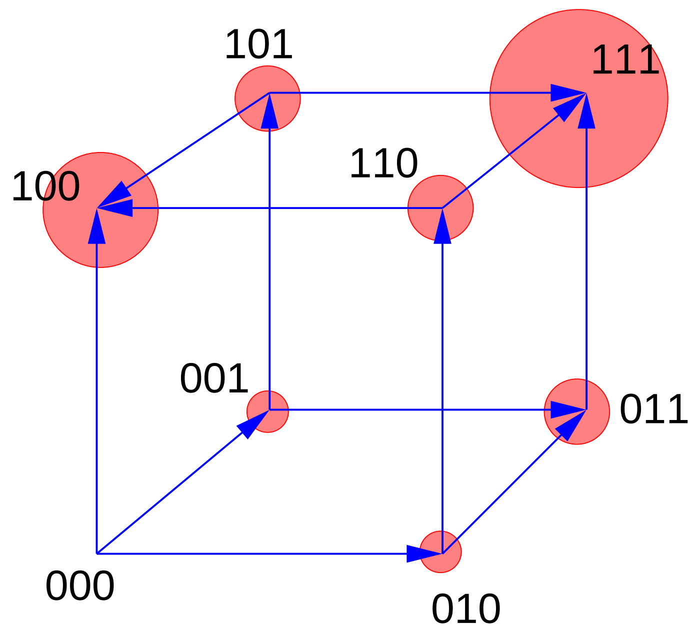
Crona et al. 2013

Pathways in fitness landscapes



- $L = 3$ mutational steps from the wild type 000 to the adapted mutant 111
- Mutations can occur in $3 \times 2 \times 1 = 3! = 6$ different orders corresponding to 6 possible directed pathways
- If all mutations are unconditionally beneficial all pathways are accessible (= increasing in fitness)

Pathways in fitness landscapes



- Sign epistasis occurs if mutations can be beneficial or deleterious depending on the genetic context Weinreich et al. 2005
- This implies that parallel arrows in the fitness graph point in opposite directions
- Sign epistasis reduces the number of direct accessible paths but may increase the number of evolutionary endpoints

Accessibility and predictability

- Pathways are accessible if fitness/resistance increases monotonically

- Existence of a small but nonzero fraction of accessible pathways implies high (retrospective) predictability

Accessibility and predictability

- Pathways are accessible if fitness/resistance increases monotonically
- Existence of a small but nonzero fraction of accessible pathways implies high (retrospective) predictability

Questions for theory

- How does accessibility depend on the structure of the fitness landscape and on the boundary conditions of the paths?
- How typical is it that a small but nonzero fraction of pathways are accessible?

Accessibility percolation

JK, arXiv:1903.11913

- Null model: assign fitness at random to genotypes — Kauffman & Levin 1987

- Probability of existence of accessible paths generically displays a sharp percolation transition from 0 to 1 at a critical value β^* of the fitness quantile $\beta \in [0,1]$ between initial and final genotype

- For directed paths on the hypercube $\beta^* = 1 - \frac{\ln L}{L} \to 1$ for $L \to \infty$
  Hegarty & Martinsson 2014

- Mutational reversions increase accessibility such that $\beta^* < 1$ for $L \to \infty$
  Berestycki et al. 2017

- For sequences with a alleles per site — B. Schmiegelt, JK, 2019

$$\beta^* \approx \frac{\ln a}{a} + \frac{1 + \ln a}{a^2} \to 0 \quad \text{for} \quad a \gg 1$$

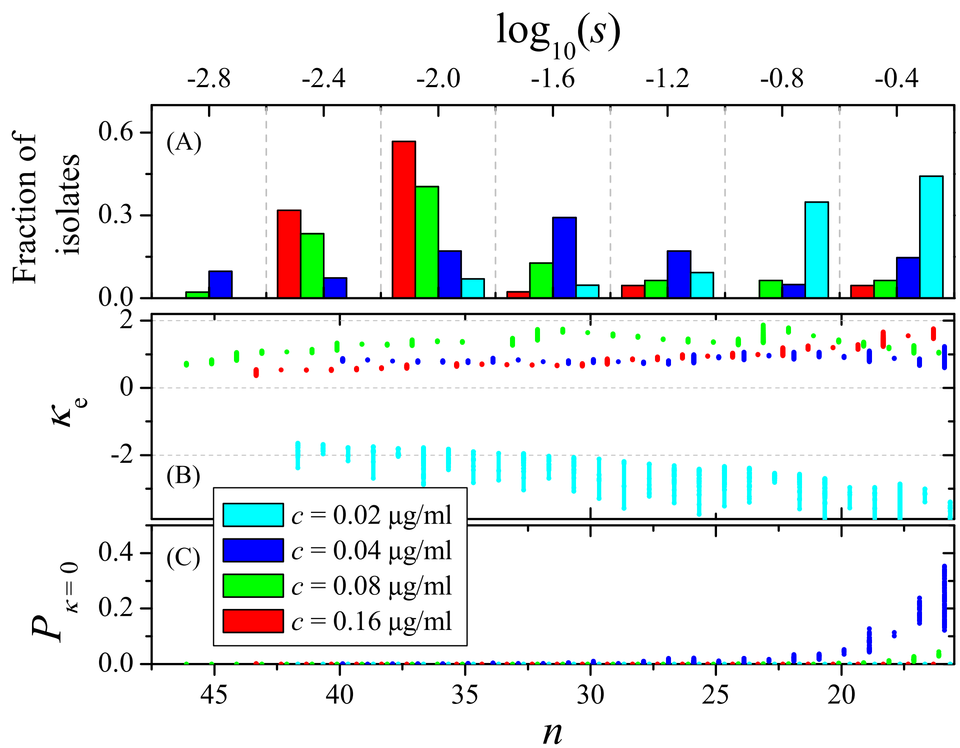
- Near $\beta = \beta^*$ the number of accessible paths is small and hence predictability is high

Exploring the TEM-1 resistance landscape

A panel of resistance mutations

M.F. Schenk et al., PLoS Genet. 2012

- At least 48 out of 2583 point mutations increase resistance against cefotaxime



- Analysis using extreme value theory yields power law distribution of mutational effects with an exponent ~ 1

Construction of combinatorial resistance landscapes

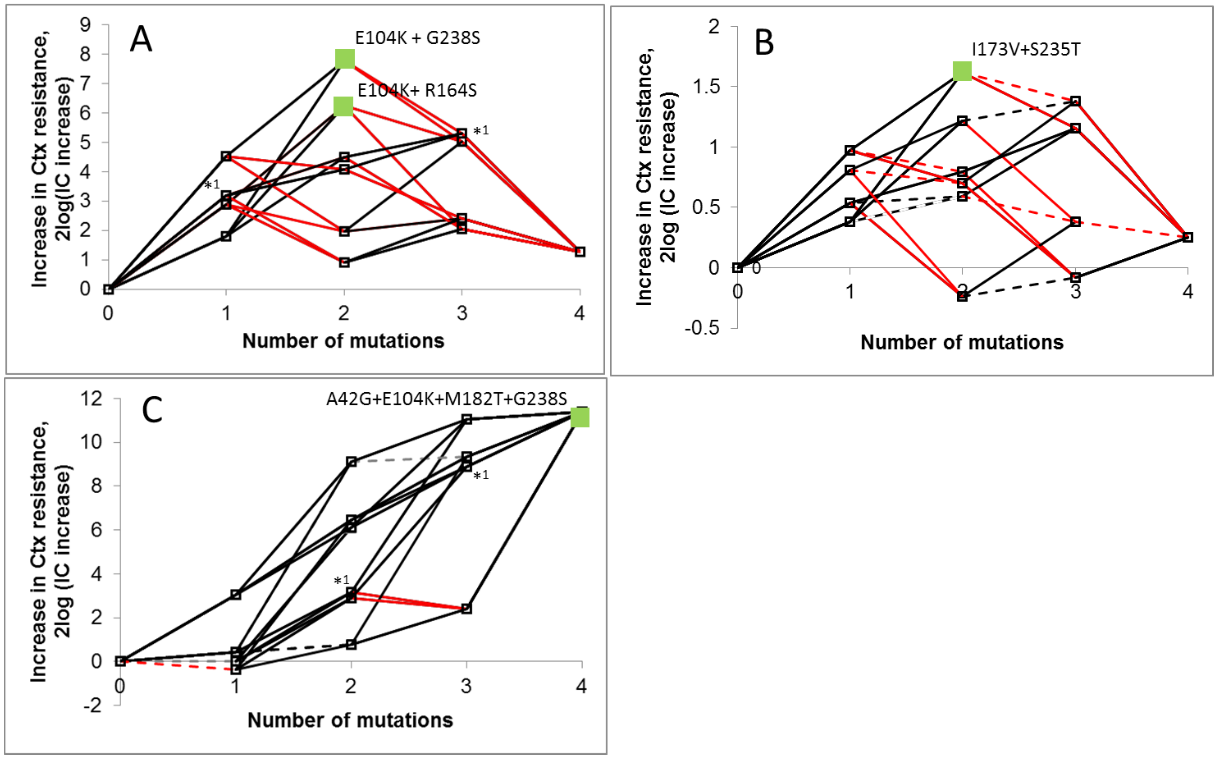
- Constructing all possible $2^{48} \approx 2.8 \times 10^{14}$ combinatorial mutants is obviously unfeasible

- The choice of a subset of mutations is expected to bias the structure of the fitness landscape: de Visser & Krug, Nat. Rev. Genet. 2014
  - singly beneficial vs. singly deleterious mutations
  - mutations chosen for individual or collective effects
  - mutations occurring along an adaptive trajectory

- Here we consider 4-dimensional landscapes constructed from three subsets of individually beneficial mutations: Schenk et al. 2013
  - non-synonymous mutations of strong effect
  - non-synonymous mutations of weak/typical effect
  - synonymous mutations of strong effect

Mutations chosen for individual vs. collective effect

M.F. Schenk et al., Mol. Biol. Evol. (2013)

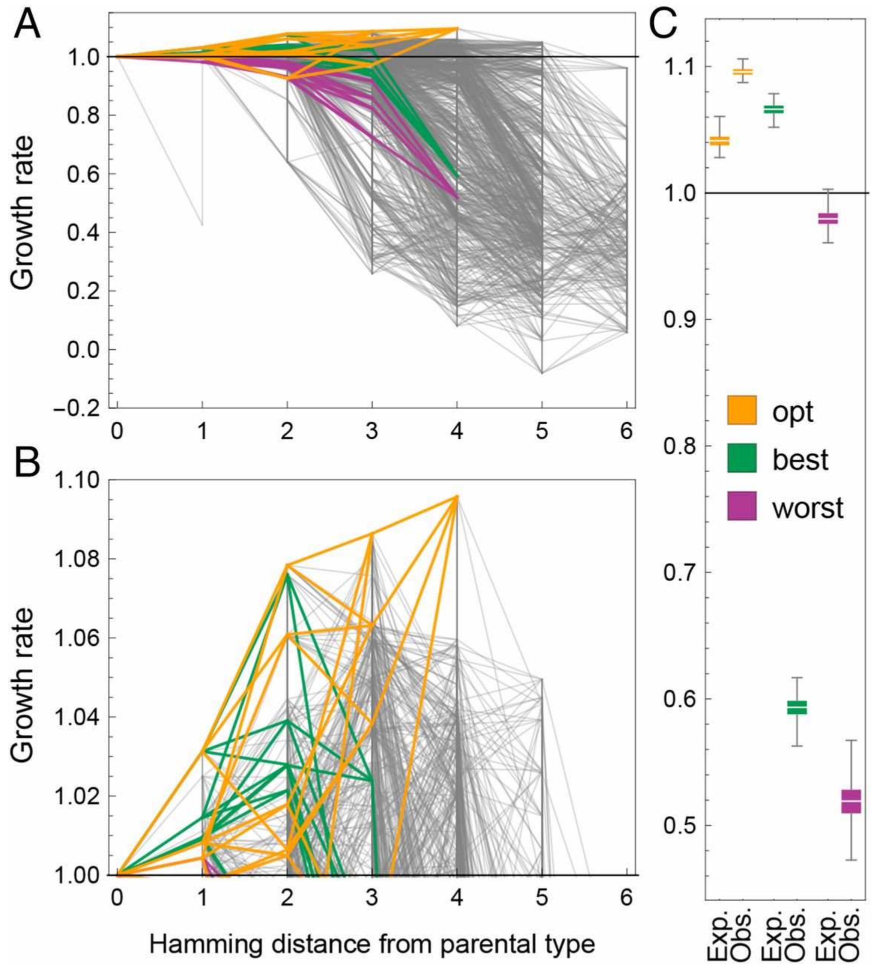


A: Large effect B: Small effect C: Weinreich 2006

- Mutations chosen for individual effect interact more strongly and negatively than mutations chosen "with hindsight" because of their collective effect

Similar patterns observed in yeast

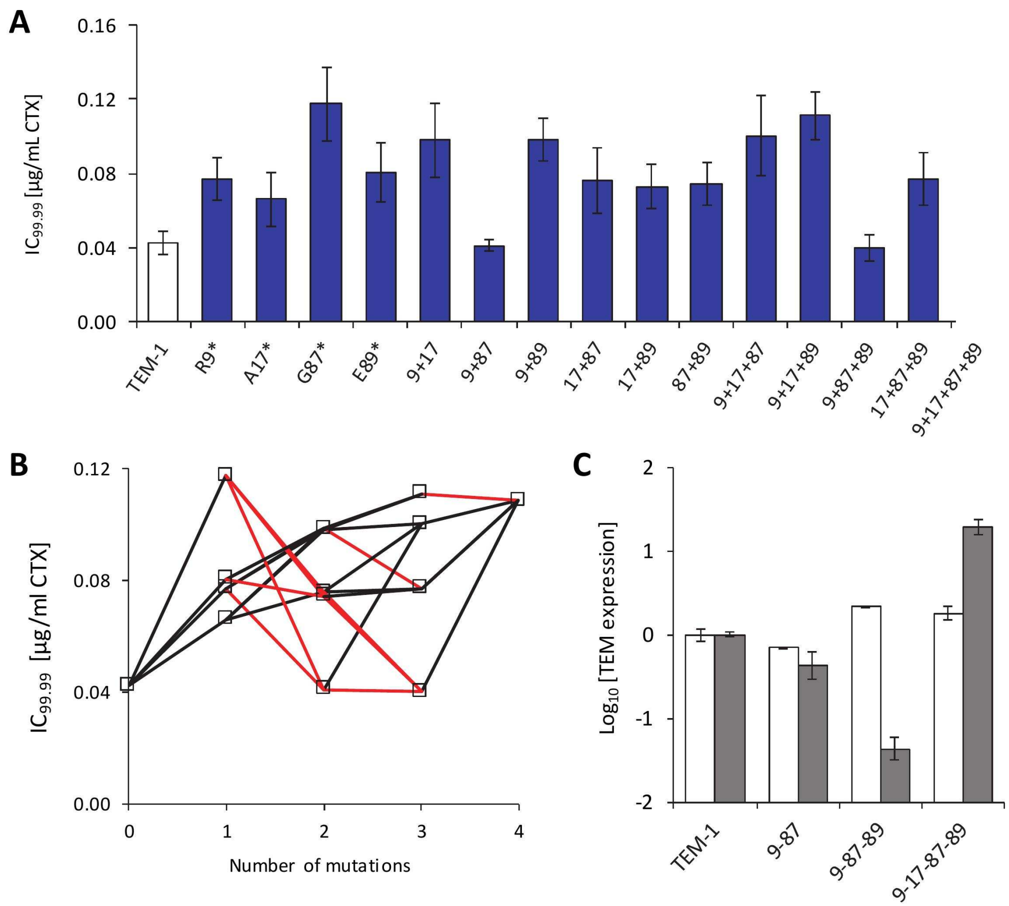
Bank et al., PNAS 2016



- Combinatorial study of 13 mutations in the Hsp90 heat shock protein

Synonymous resistance landscape

M.P. Zwart et al., Heredity (2018)



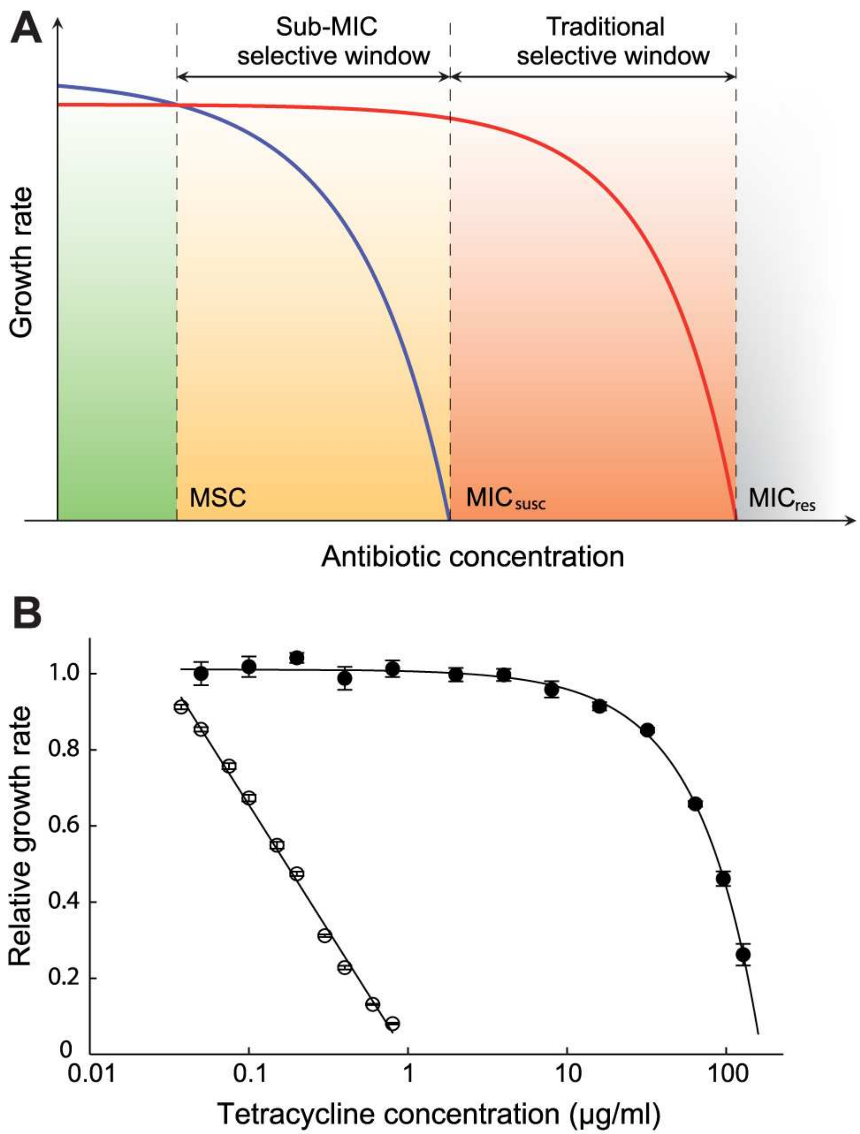
- Landscape displays a “layered” structure that may be related to translational bottlenecks

M. Josupeit, JK, arXiv:2009.10621

Concentration-dependent fitness landscapes

S. Das, S. Direito, B. Waclaw, R. Allen, JK, eLife 9:e55155 (2020)

Dose-response curves



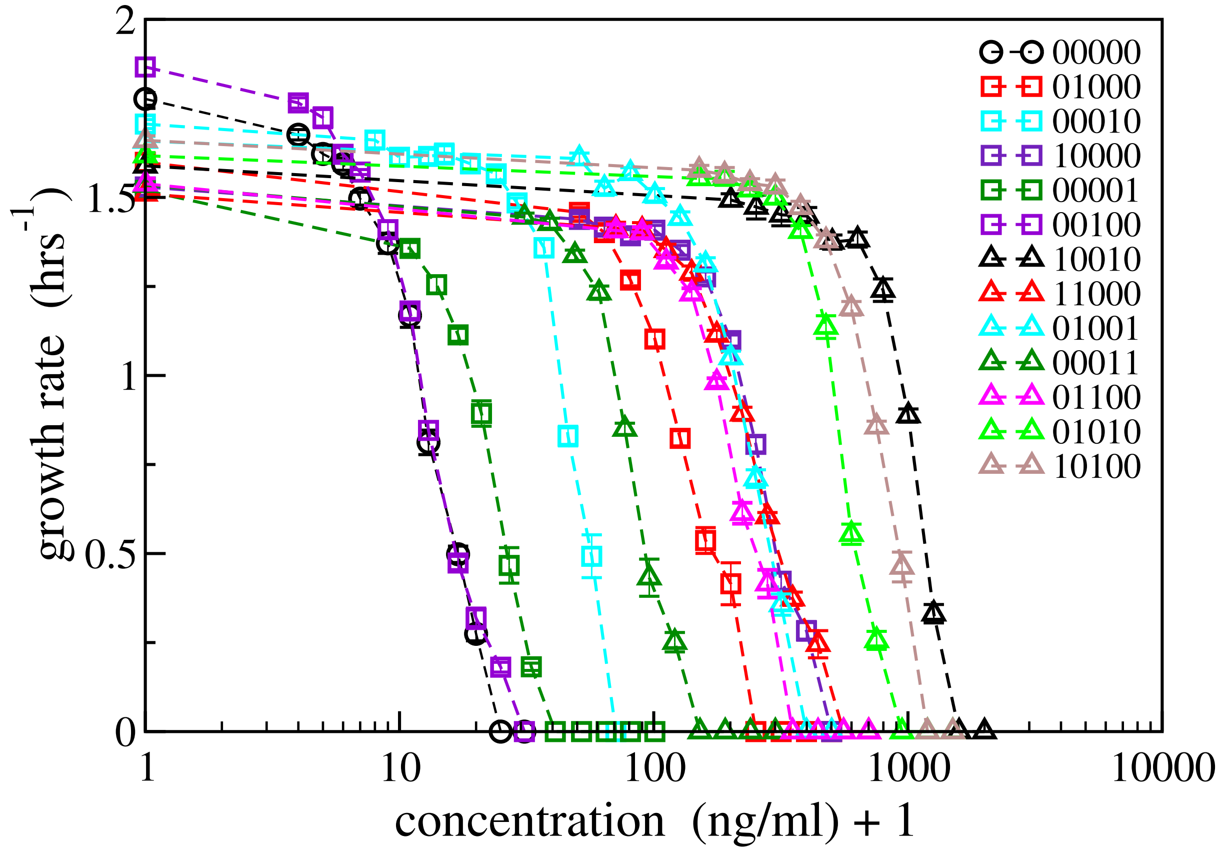
Gullberg et al., PLoS Pathogens 2011

- Mutations that increase resistance often decrease growth rate in the absence of antibiotic (null fitness)
- As a consequence the dose-response curves of susceptible and mutant strains cross at the minimal selective concentration (MSC)
- The mutant selection window is the concentration range

$$\text{MSC} < c < \text{MIC}_{\text{res}}$$

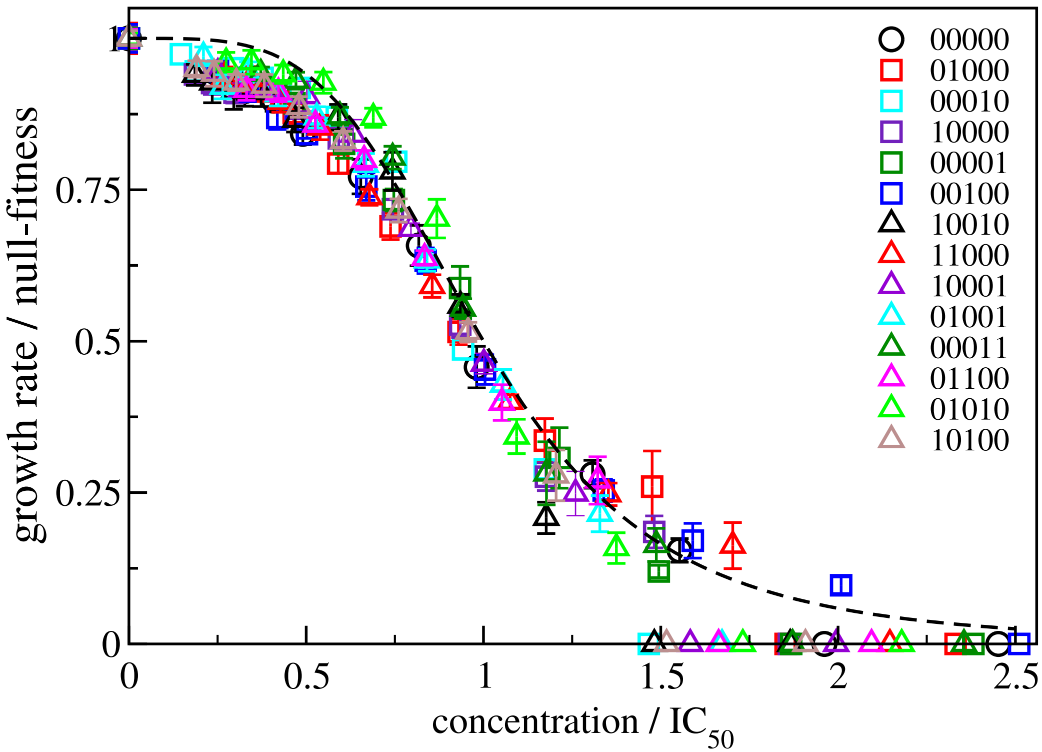
Observation 1: Scaling of dose-response curves

- Single and double mutations in *Escherichia coli* conferring resistance against ciprofloxacin S. Direito, B. Waclaw, R. Allen



Observation 1: Scaling of dose-response curves

- Single and double mutations in *Escherichia coli* conferring resistance against ciprofloxacin S. Direito, B. Waclaw, R. Allen



- Shape of dose-response curve is a Hill function $f(x) = (1 + x^4)^{-1}$

Observation 2: Independent marginal phenotypes

- Null-fitness and MIC of multiple mutants combine multiplicatively or display negative interactions data from Marcusson et al., PLoS Pathogens 2009

Strain	String	log(null-fitness)	Non-epistatic	log(MIC)	Non-epistatic
MG1655	00000	0.00 (± .004)		0.00 (± .35)	
LM378	10000	0.01 (± .016)		3.17 (± .70)	
LM534	01000	-0.01 (± .018)		2.75 (± .70)	
LM202	00010	-0.19 (± .020)		0.69 (± .70)	
LM351	00001	-0.094 (± .014)		1.08 (± .70)	
LM625	11000	-0.030 (± .011)	0.0 (± .038)	3.17 (± .70)	5.92 (± 1.1)
LM421	10010	-0.15 (± .019)	-0.18 (±.040)	4.13 (± .70)	3.56 (± 1.1)
LM647	10001	-0.051 (± .013)	-0.084 (± .034)	3.44 (± .70)	4.65 (± 1.1)
LM538	01010	-0.19 (± .020)	-0.20 (± .042)	4.13 (± .70)	3.46 (± 1.1)
LM592	01001	-0.083 (± .015)	-0.10 (± .036)	3.16 (± .70)	3.83 (± 1.1)
LM367	00011	-0.20 (± .026)	-0.28 (± .038)	2.06 (± .70)	1.77 (± 1.1)
LM695	11010	-0.24 (± .017)	-0.19 (± .058)	3.85 (±. 70)	6.61 (± 1.1)
LM691	11001	-0.073 (± .013)	-0.094 (± .052)	3.85 (±. 70)	7.00 (± 1.4)
LM709	10011	-0.24 (± .027)	-0.274 (± .054)	4.54 (±. 70)	4.94 (± 1.4)
LM595	01011	-0.51 (± .051)	-0.294 (± .056)	4.54 (±. 70)	4.52 (± 1.4)
LM701	11011	-0.42 (± .037)	-0.284 (±.072)	4.83 (±. 70)	7.69 (± 1.8)

- This pattern was first observed for resistance mutations in *Salmonella enterica* and *E. coli* Knopp & Andersson, mBio 2018

Model with predictable concentration dependence

- L resistance mutations $i = 1, \ldots, L$ characterized by null-fitness $r_i < 1$ and resistance $m_i > 1$ relative to the wild type

- Dose-response curve of a mutant $\sigma = (\sigma_1, \ldots, \sigma_L)$ is given by

$$w_\sigma(x) = r_\sigma f(x/m_\sigma)$$

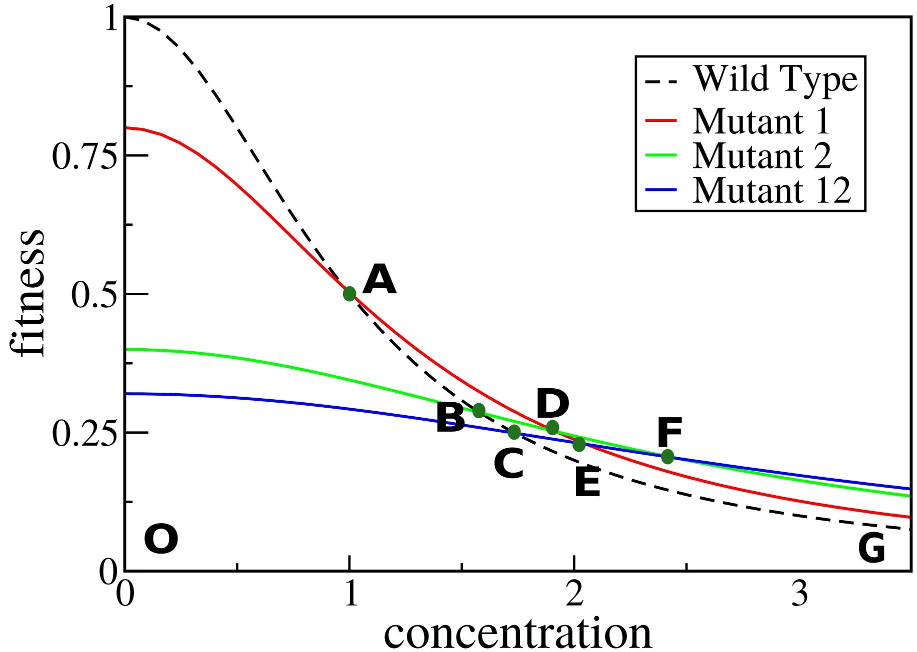
  where the function $f(x)$ is independent of σ and the marginal phenotypes combine multiplicatively as

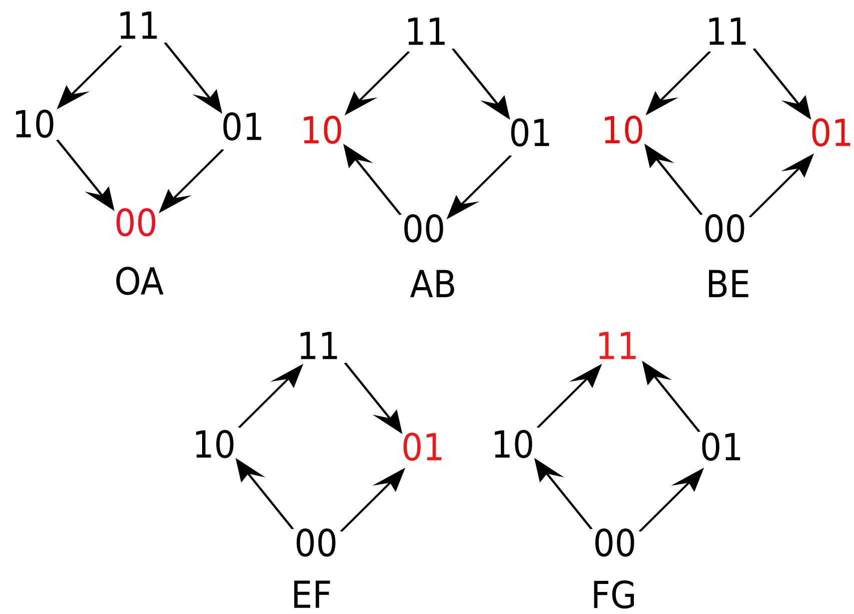
$$r_\sigma = \prod_{j=1}^{L} (r_j)^{\sigma_j} \quad \text{and} \quad m_\sigma = \prod_{j=1}^{L} (m_j)^{\sigma_j}$$

- Resistance is quantified by the concentration at which growth drops by 50% (IC_{50}), which implies that $f(1) = \frac{1}{2}$

- In this way 2^L concentration-dependent fitness values can be predicted from $2L$ single mutant phenotypes and one shape function

Two resistance mutations



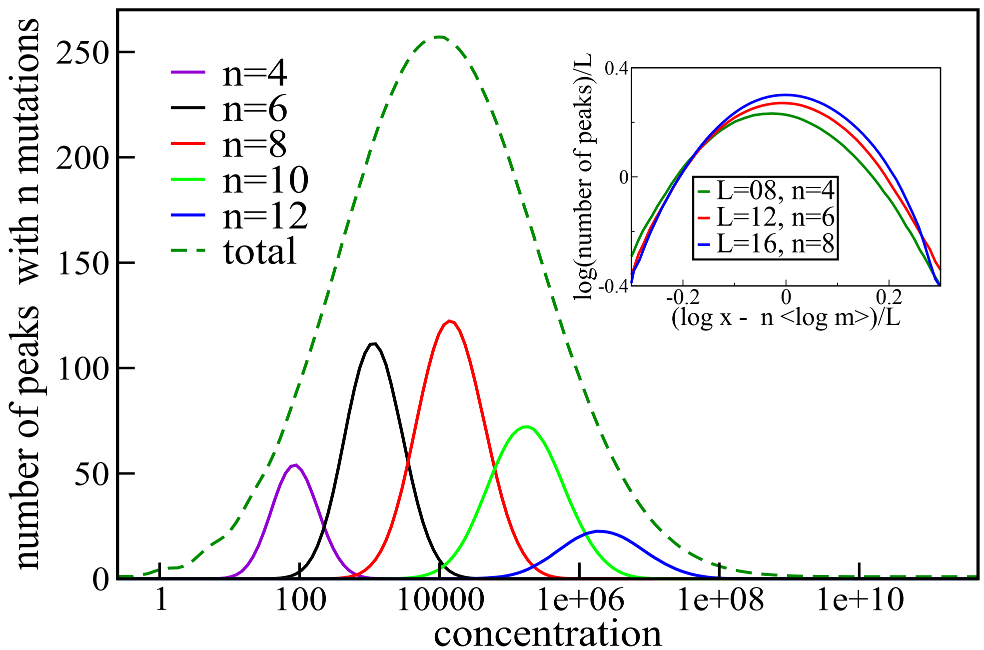


- Fitness landscape evolves from single-peaked to two-peaked and back
- Not all rank orders can appear in this process

Maximal ruggedness at intermediate concentrations

- $L = 16$ mutations with randomly distributed r_i, m_i

- Quantify ruggedness by the number of local fitness peaks



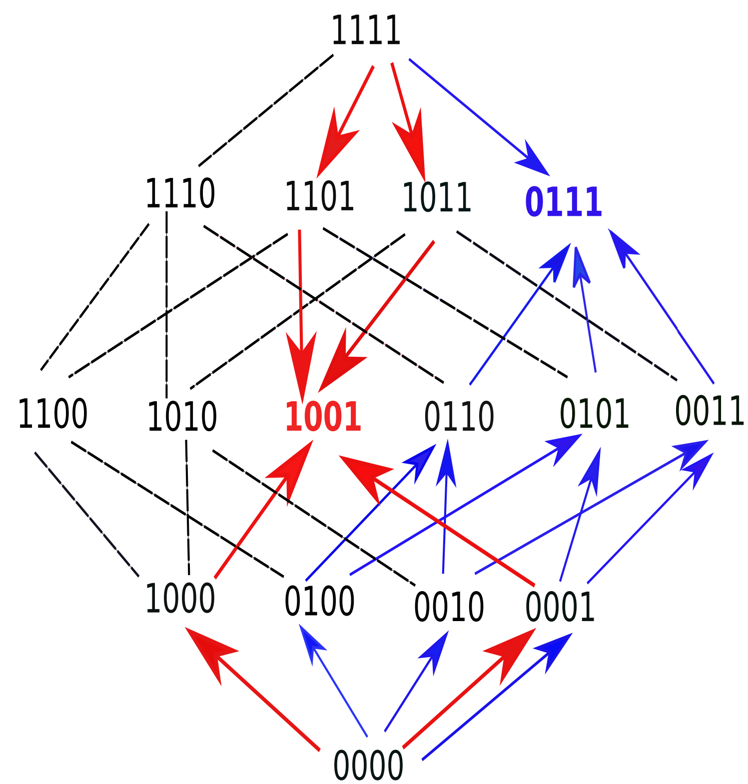
- Typical fitness peaks carry n mutations at $\ln x \sim n\langle \ln m_i \rangle$, and the maximal number of peaks grows exponentially with L

Landscapes are nevertheless highly accessible

- Scaling and absence of positive marginal epistasis imply that certain rank orders are forbidden at any concentration

- As a consequence, any peak genotype is accessible from all its sub- and supersets

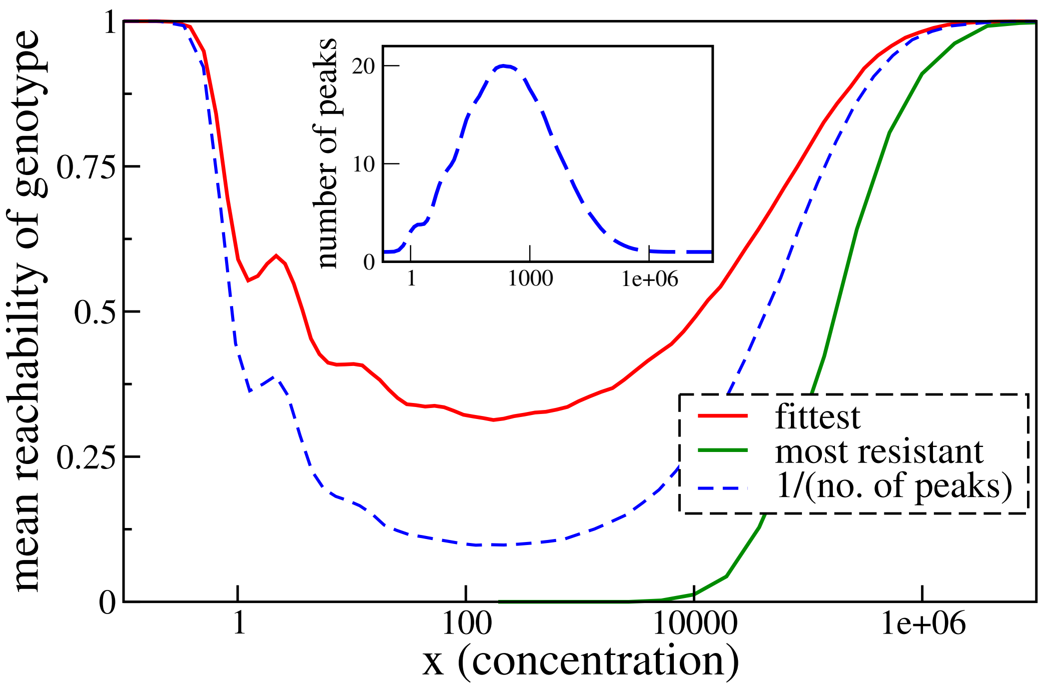
- In particular, the fittest type is always accessible from the wild type



Two-peaked landscape at an intermediate concentration

Reachability of the fittest and most resistant mutant

- $L = 10$ mutations with randomly distributed r_i, m_i



- Probability of reaching the fittest/most resistant mutant from the wild type using strong selection/weak mutation dynamics

Summary

- Evolution of antimicrobial resistance through multiple mutational steps is a model for evolutionary predictability

- Combinatorial construction of resistance landscapes reveals a systematic dependence on the choice of the combined mutations

- Tradeoff between resistance and growth rate induces rugged fitness landscapes at intermediate antibiotic concentration

- Despite their ruggedness these landscapes are remarkably accessible and the evolution of high levels of resistance remains facile

- Outlook: Consider time-dependent antibiotic concentrations as a model for evolution in changing enviroments

Thanks to

- Cologne:
  Jasper Franke, Stefan Nowak, Benjamin Schmiegelt, Ivan Szendro, Sungmin Hwang, Suman Das

- Wageningen:
  Martijn Schenk, Mark Zwart, Manja Saebelfeld, Arjan de Visser

- Edinburgh:
  Susana Direito, Bartek Waclaw, Rosalind Allen