Fitness constraints on horizontal gene transfer

Dan I Andersson

University of Uppsala,
Department of Medical Biochemistry and Microbiology,
Uppsala, Sweden

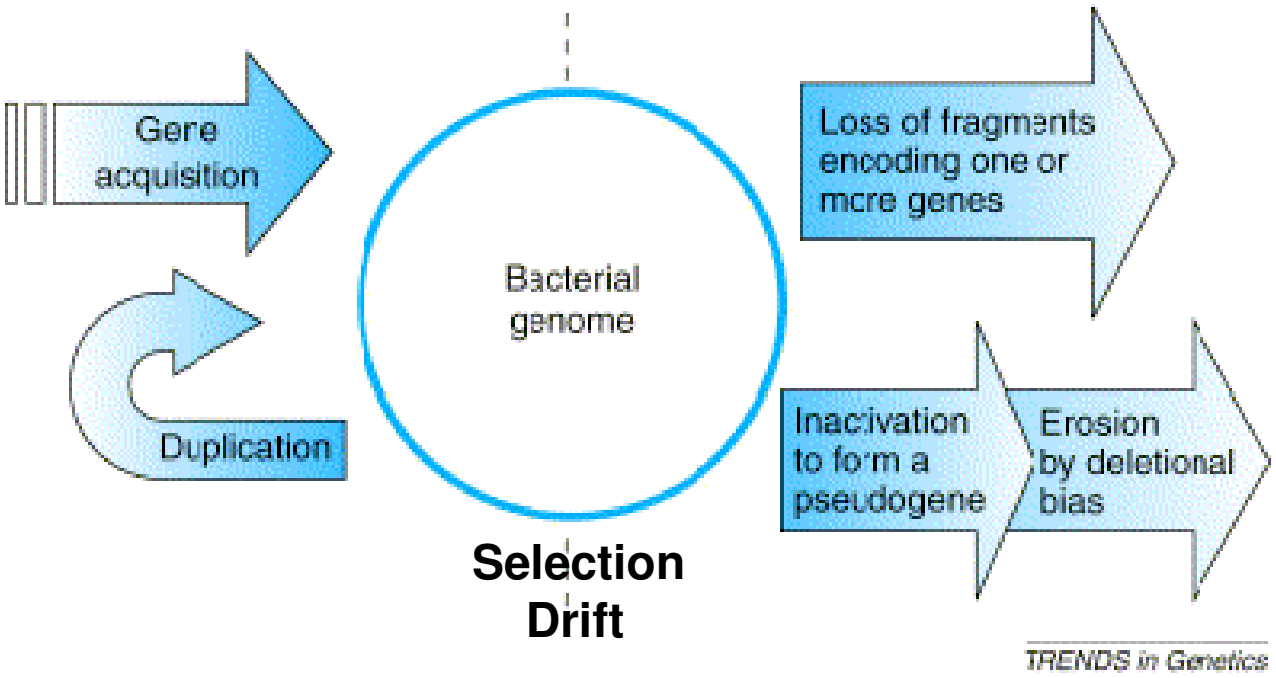
GMM 3, 30 Aug--2 Sep, Oslo, Norway

Acknowledgements:

Uppsala University
Peter Lind

Collaborators
Otto Berg
Chuck Kurland

Bacterial genome evolution



Mira et al. 2001

What genes are transferred and why?

What?

- Common: antibiotic resistance, pathogenicity, metabolic functions, transport: i.e. functions associated with niche-specific adaptations

- Rare: translation, transcription, replication: i.e. niche-independent essential functions

Why?

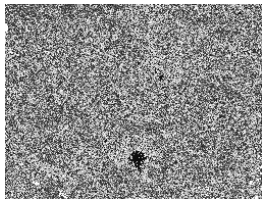
- Can give immediate benefit

- Complexity hypothesis: proteins with many interactions (structural, regulatory, functional) less likely to be transferred,

  i.e. a newly introduced orthologue is unlikely to outperform a gene that has co-evolved for a long time with its cognate partners

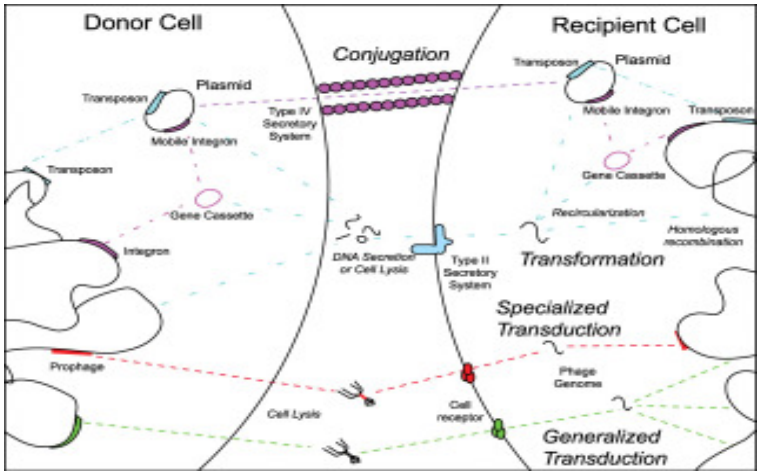
Potential constraints on HGT

1. Ecological opportunity



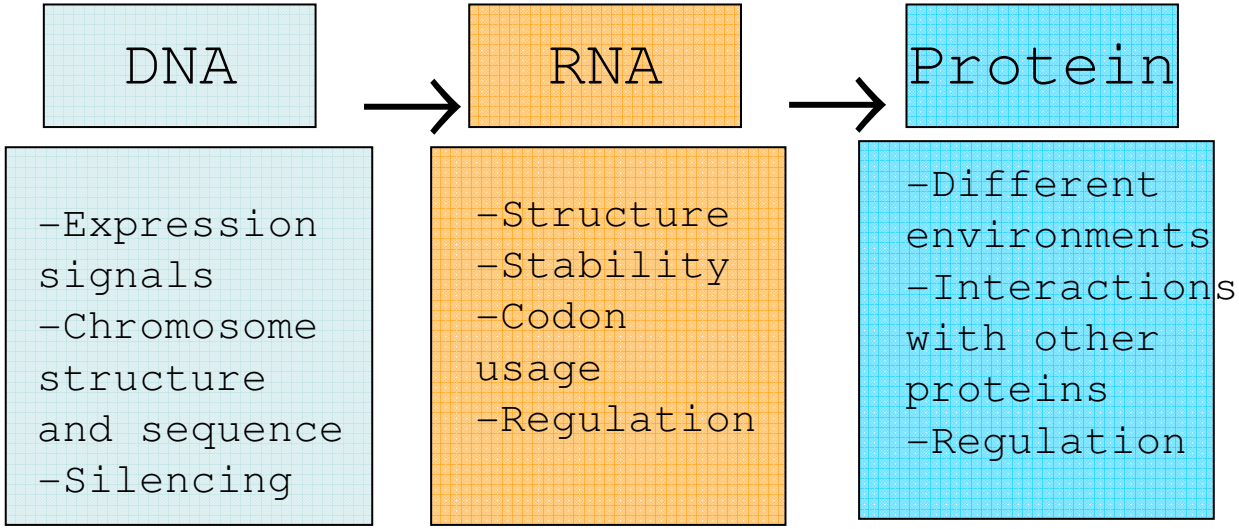
***Buchnera aphidicola* within aphid bacteriocyte**

2. Transfer systems



Zaneveld et al. Microbiology 2008

3. Function and expression of transferred gene, i.e. fitness

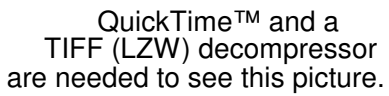


Questions addressed

I. Examine barriers to HGT in terms of the distribution and magnitude of fitness constraints at RNA and protein level

II. Determine the potential and mechanism of compensatory evolution to ameliorate the fitness costs of HGT

Model system: 70S ribosome



50S large subunit:
23S, 5S, 33 r-proteins

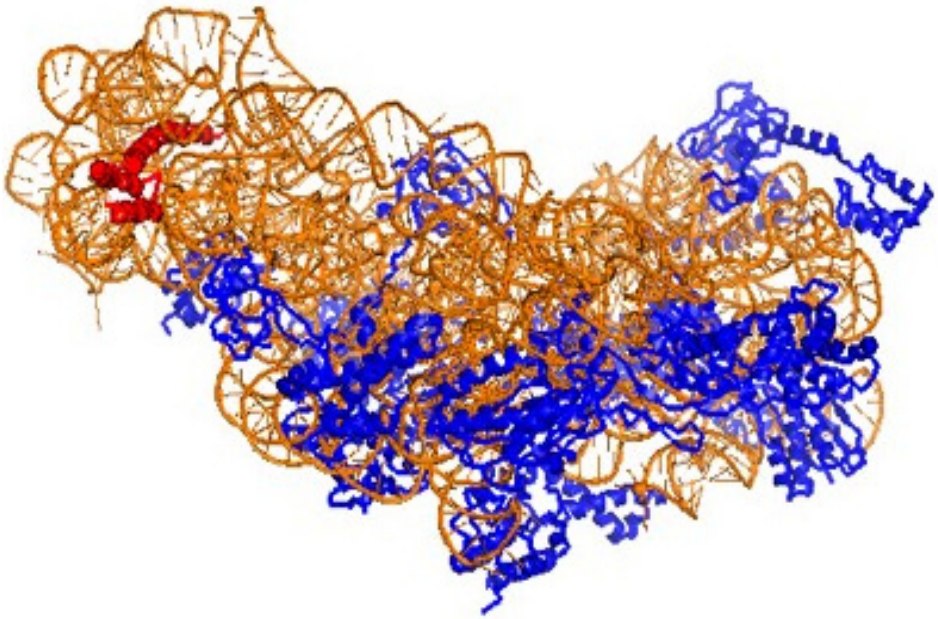
30S small subunit:
16S rRNA, 21 r-proteins

+ several translation factors

Complex structure with high functional and structural integration
Ribosome function directly related to fitness (growth rate)

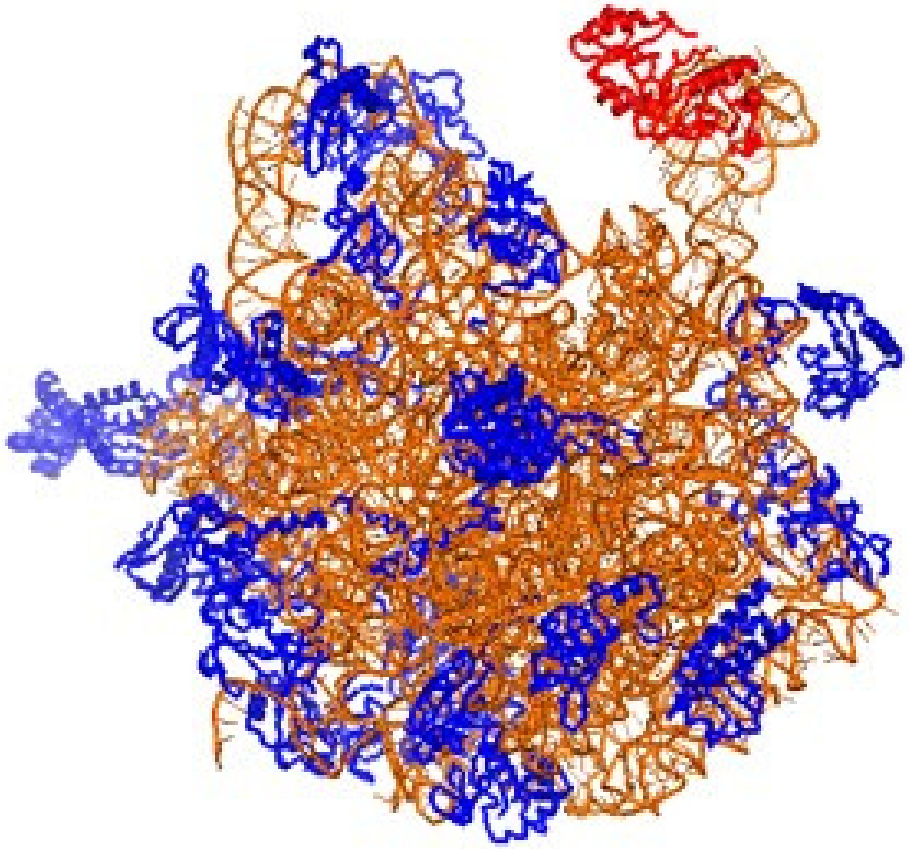
Ribosomal protein S20 (*rpsT*)

- 86 amino acids
- Binds to 16S rRNA
- Deletion mutants defective in assembly of subunits
- Translational repressor
- Non-essential



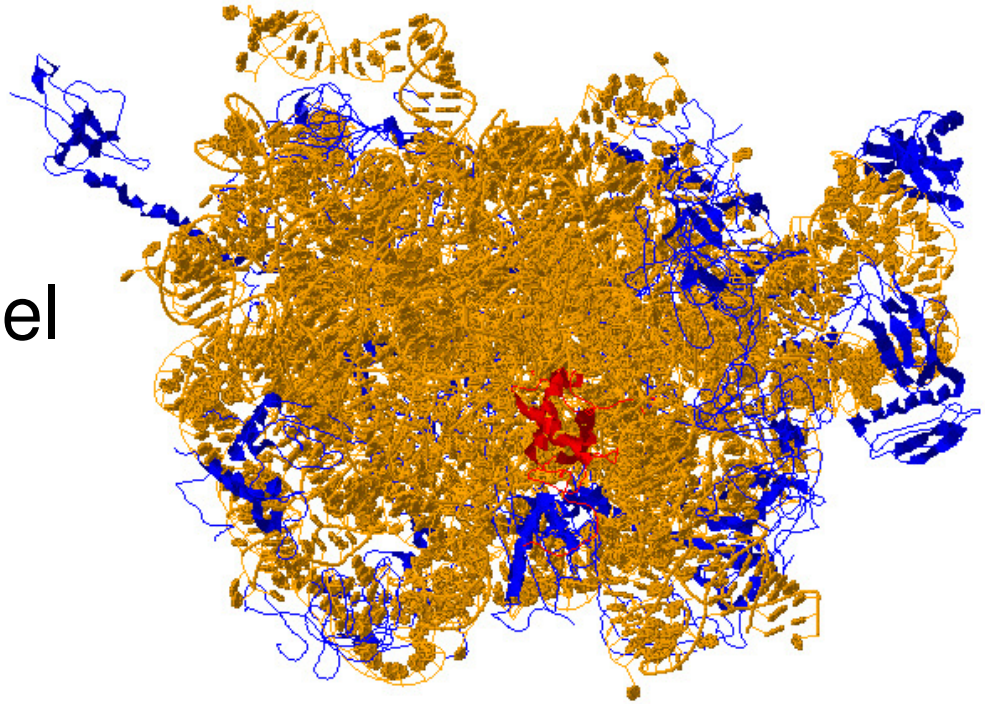
Ribosomal protein L1 (*rplA*)

- 234 amino acids
- Binds to 23S rRNA
- Ejection of deacetylated tRNA
- Deletion mutants defective in translocation
- Translational repressor - L1/L11
- Non-essential



Ribosomal protein L17 (*rplQ*)

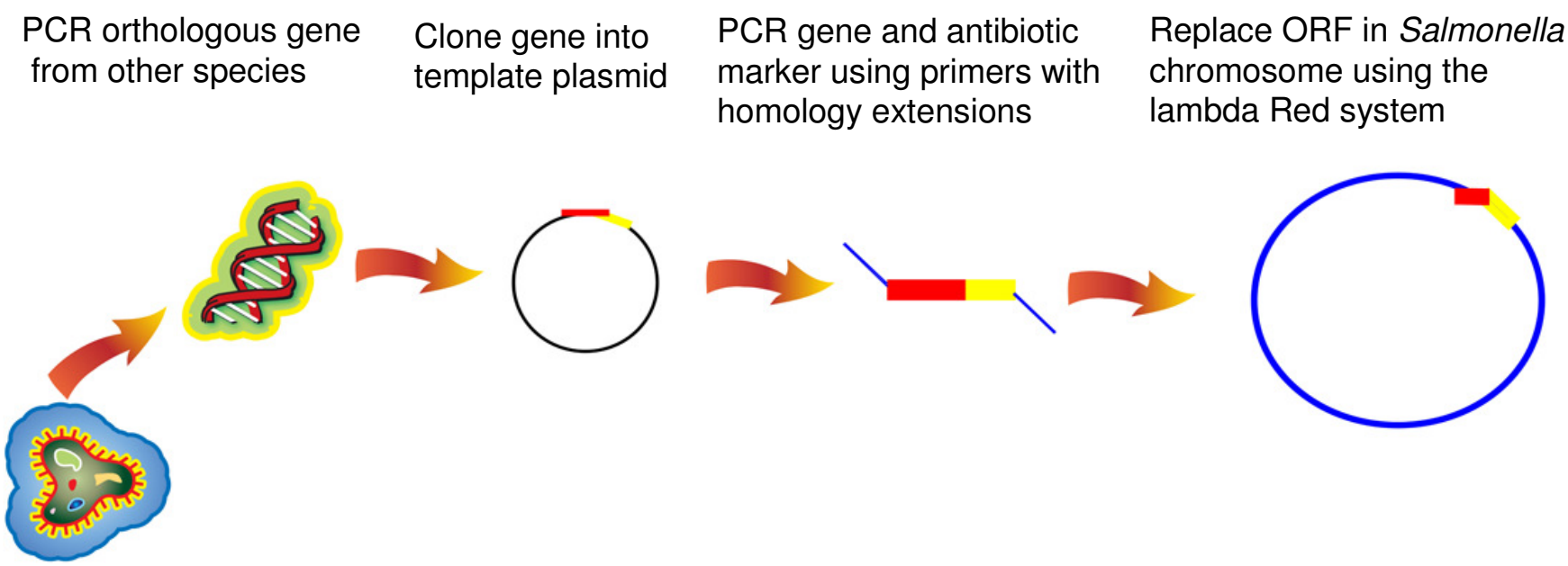
- 127 amino acids
- Binds to 23S rRNA
- Located at exit tunnel opening
- Essential



Experimental design

- Model organism *Salmonella typhimurium* LT2
- Two non-essential and one essential ribosomal protein used as model for orthologous transfer
- Highly sensitive measurements of fitness

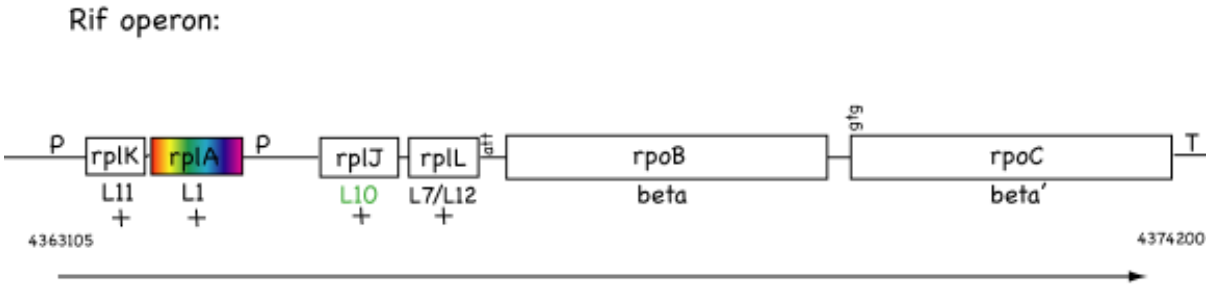
Strain construction



- Genes from other species cloned in plasmid used as PCR template
- Genes introduced together with antibiotic marker to replace the native Salmonella gene by Lambda Red recombineering --> **regulatory and chromosomal context unaltered**

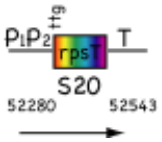
Ribosomal protein genes replaced with orthologues from other species

L1



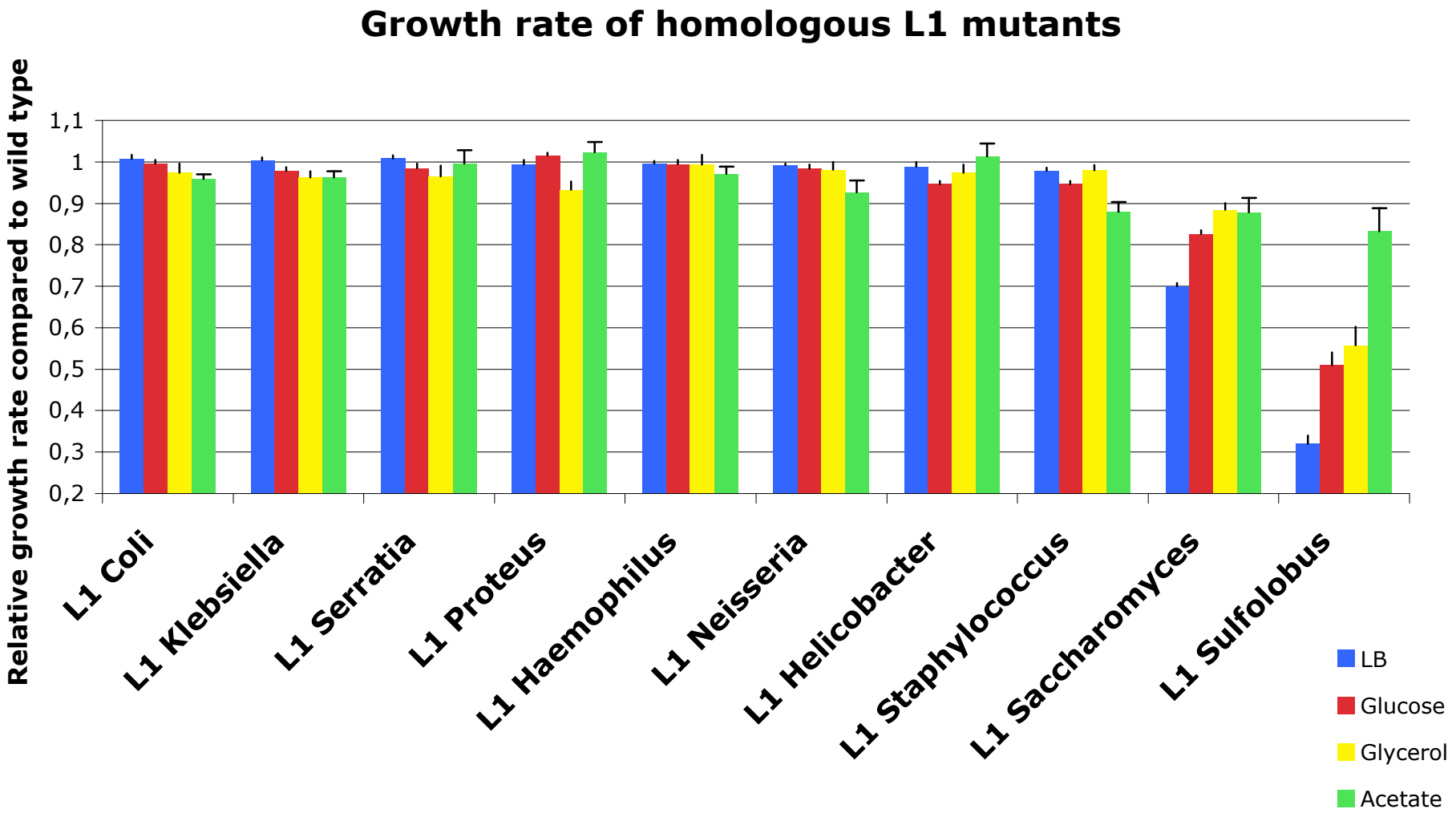
L17

S20

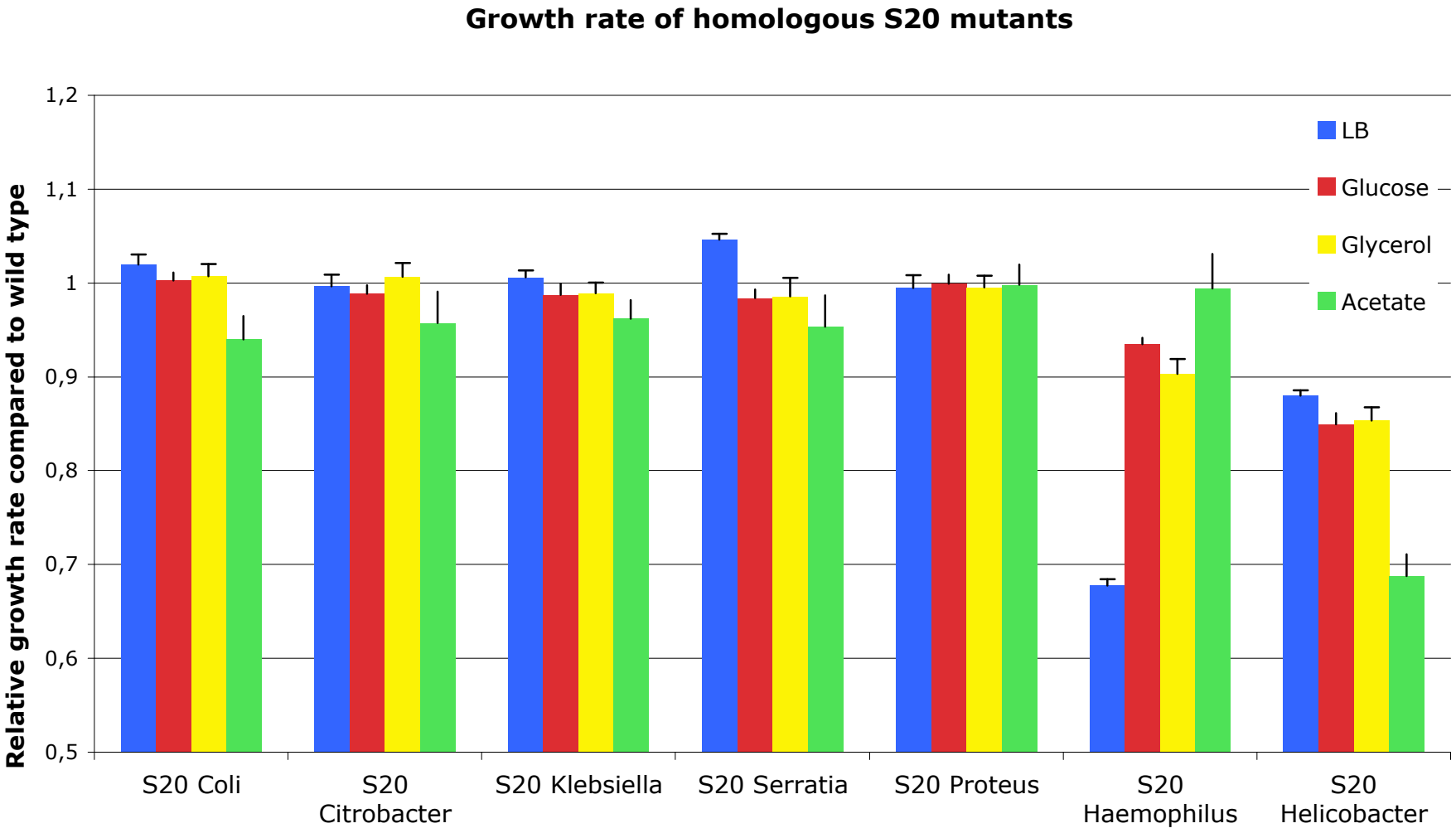


Species comparison of ribosomal proteins

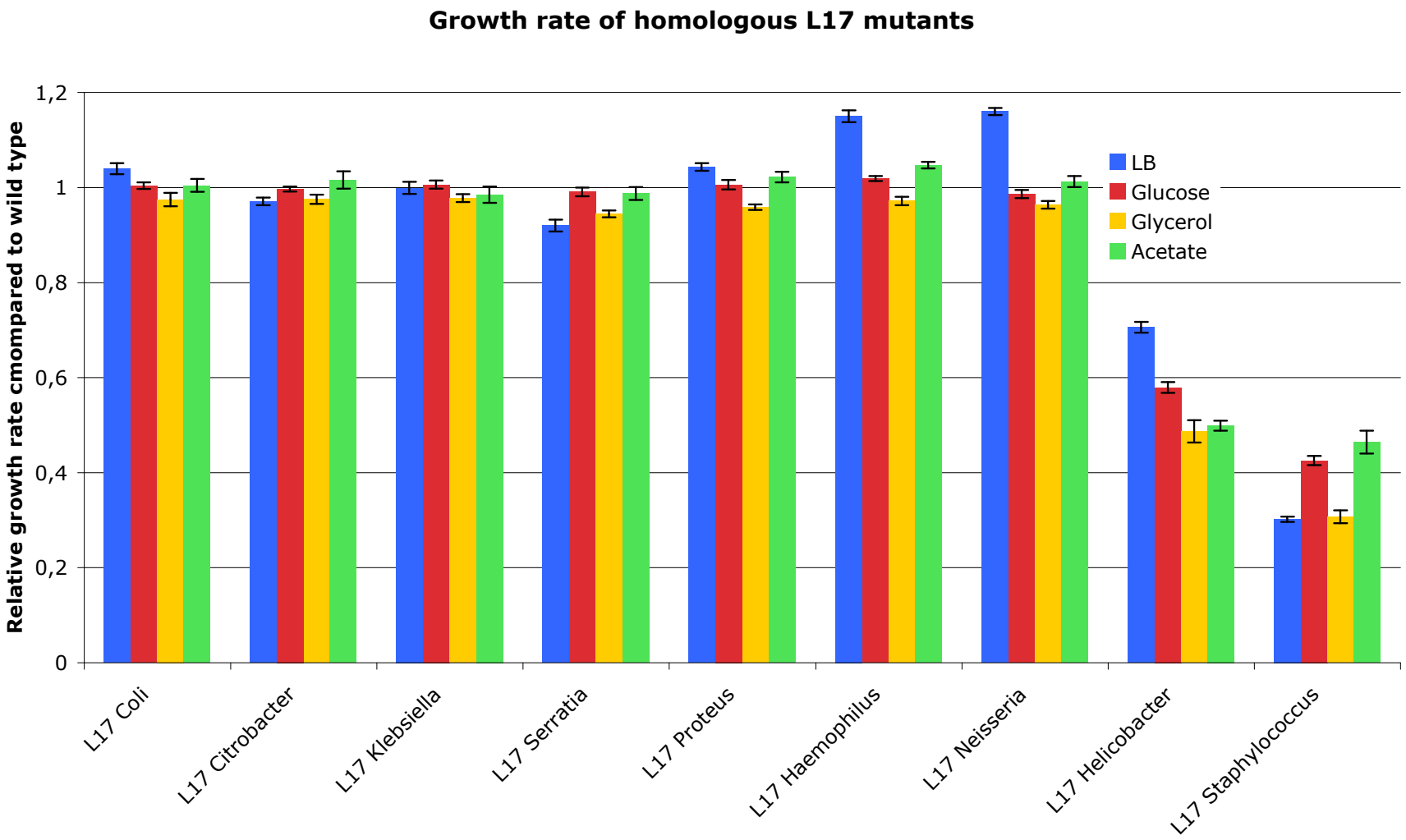
Gene/protein species	Nucleotide substitutions	Nucleotide identity (%)	Amino acid substitutions	**Amino acid identity (%)**	Codon adaptation index	GC content (%)
***rpsT*/S20 (Å264 bp/86 aa)**						
Salmonella typhimurium	-	-	-	**-**	0.611	46.9
Escherichia coli	3	98.9	2	**97.7**	0.607	47.1
Citrobacter freundii	6	97.7	1	**98.9**	0.608	46.9
Klebsiella pneumoniae	7	97.3	1	**98.9**	0.616	47.9
Serratia marcescens	34	87.1	14	**83.7**	0.615	50.6
Proteus mirabilis	47	82.2	14	**83.5**	0.592	39.9
Haemophilus influenzae	68	74.2	23	**73.6**	0.580	40.3
Helicobacter pylori	179	32.0	46	**47.0**	0.561	39.8
***rplA*/L1 (Å705 bp/234 aa)**						
Salmonella typhimurium	-	-	-	**-**	0.719	53.3
Escherichia coli	45	93.4	4	**98.3**	0.688	51.9
Klebsiella pneumoniae	43	93.9	5	**97.9**	0.712	52.6
Serratia marcescens	71	89.9	14	**94.8**	0.678	53.0
Proteus mirabilis	98	86.0	25	**88.0**	0.578	43.1
Haemophilus influenzae	190	73.0	37	**84.0**	0.549	39.0
Neisseria menigitidis	255	63.9	77	**67.1**	0.578	43.9
Helicobacter pylori	313	55.6	117	**50.0**	0.609	43.6
Staphylococcus aureus	316	55.0	110	**53.0**	0.536	36.4
Saccharomyces cerevisiae	439	37.7	185	**19.7**	0.466	36.6
Sulfolobus acidocaldarius	418	40.7	179	**23.5**	0.456	35.8
***rplQ*/L17 (Å384 bp/128 aa)**						
Salmonella typhimurium					0.549	54.6
Escherichia coli	8	97.9	1	**99.2**	0.559	53.8
Citrobacter freundii	8	97.9	2	**98.4**	0.560	54.9
Klebsiella pneumoniae	11	97.2	3	**97.7**	0.597	55.7
Serratia marcescens	26	93.3	8	**93.8**	0.563	55.3
Proteus mirabilis	45	88.6	8	**93.7**	0.475	49.7
Haemophilus influenzae	95	75.4	15	**88.3**	0.338	41.2
Neisseria menigitidis	125	66.1	24	**80.3**	0.341	46.7
Helicobacter pylori	198	43.6	65	**44.0**	0.212	45.7
Staphylococcus aureus	206	44.2	63	**48.4**	0.340	37.5



Relative growth rate of L1 deletion mutant = 0.3



Relative growth rate of S20 deletion mutant = 0.3



Relative growth rate of L17 deletion mutant = 0

Competition assays

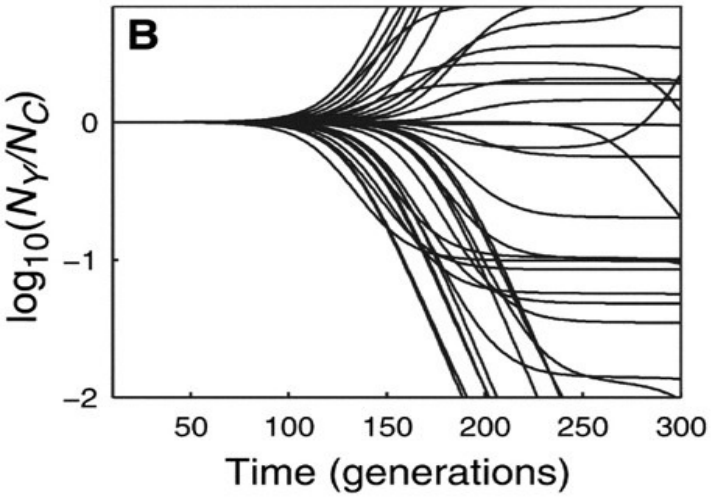
- Need more sensitive assay than single culture growth rate

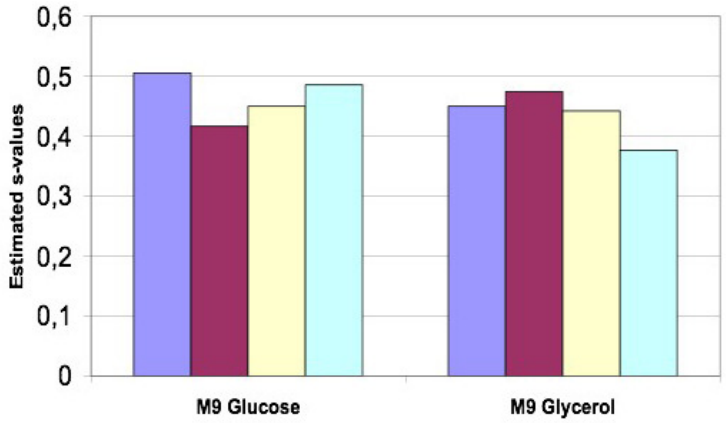
- Problems with traditional competitions
  - Many colonies needed for statistical significance
  - Adaptive mutations during experiment (periodic selection)

  ---> ***s < 0.01 not detectable-->***

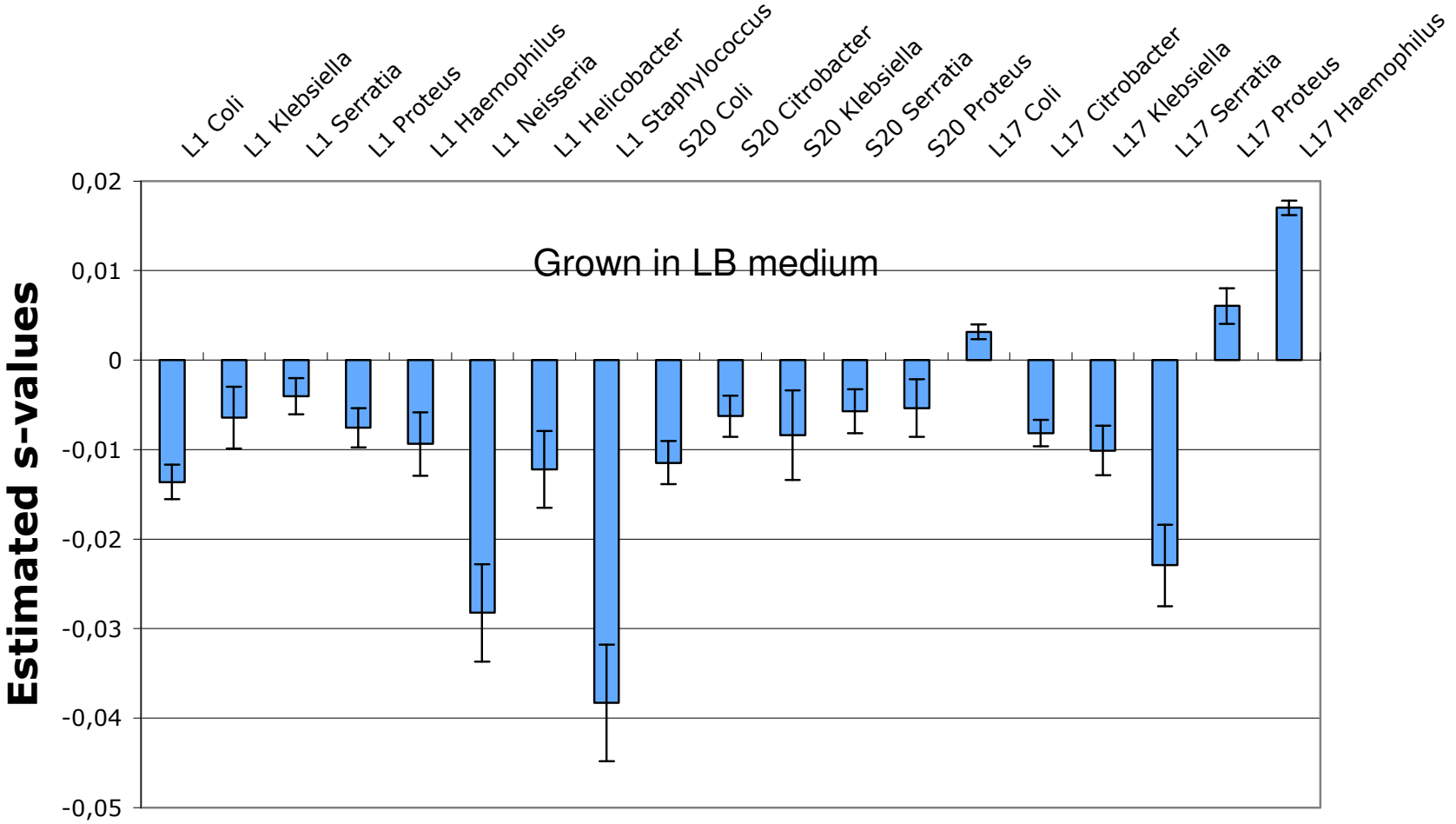
  Experimental detection vs. population genetics dilemma

- Solutions
  - Label cells with YFP/CFP and count with flow cytometry
  - Pre-adapt strains for 1000 generations





Competitive fitness of homologous mutants



Conclusions replacement experiments

- Replacement with orthologues from other eubacerial species generally confer small fitness costs (at least from a geneticists view!)

- Differences in amino acid sequence, codon usage and GC content have small effects on function

- Fitness effects still large enough to prevent HGT (at least from a population biologists view!)

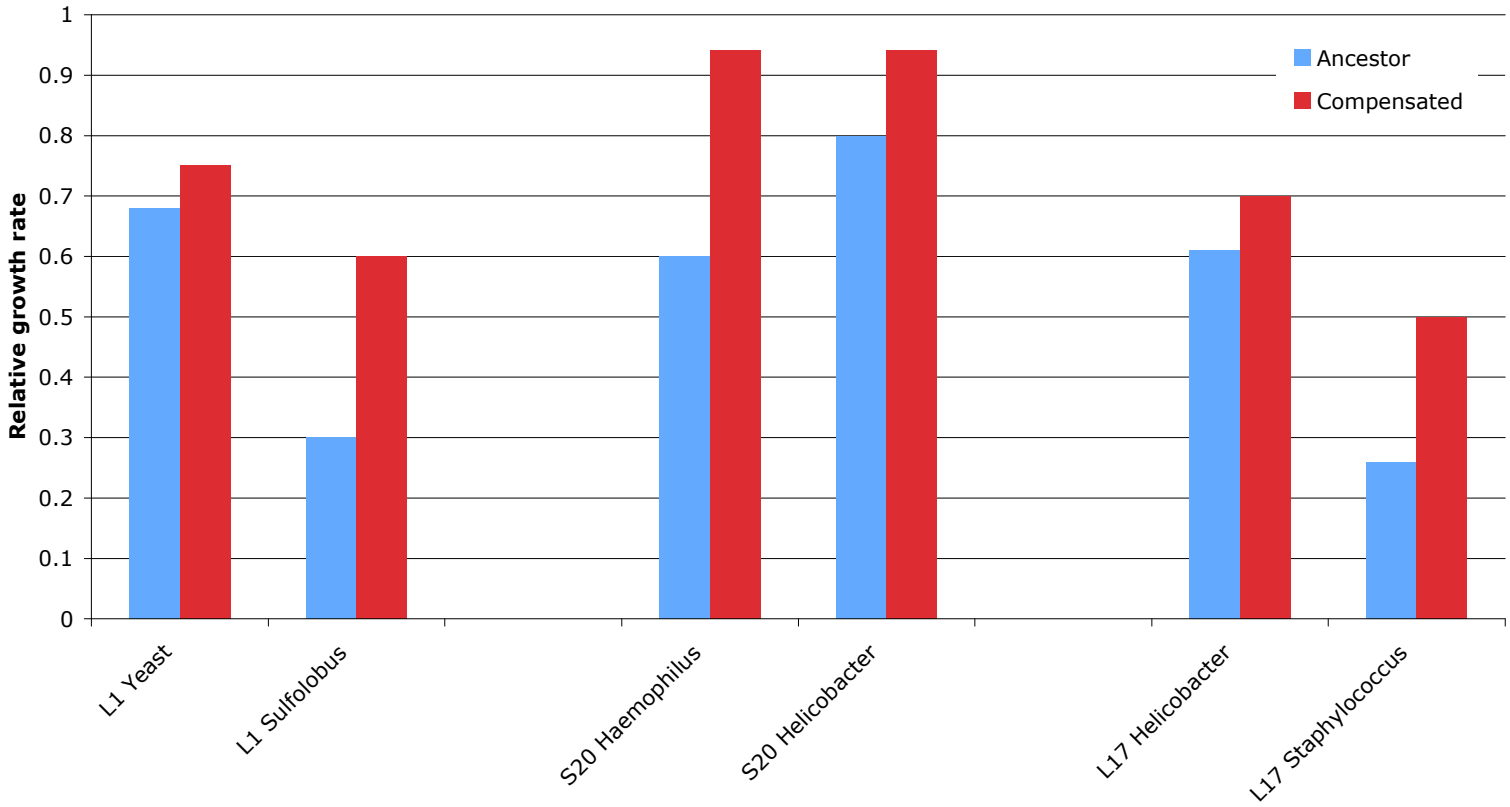
  A counter-selected import is essentially unable fix in a bacterial population

  For example, s= -0.01, N_e=N= 10^5, P_{fix}= 10^{-436}

Amelioration of fitness costs associated with HGT

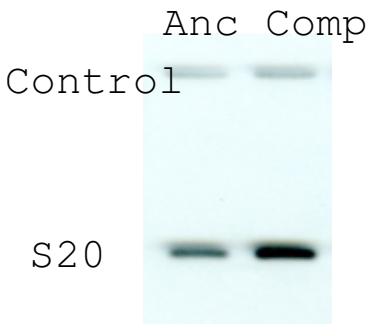
Serial passage of HGT mutants with large fitness costs,
10^6 cells transferred to 1 ml LB --> grow to 10^9 cells--> re-inoculate-->
10 generations of growth per passage

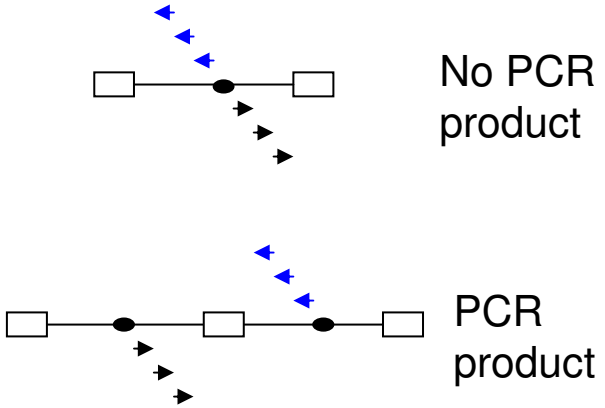
Substantial increases in fitness for all mutants after <300 generations



Characteristics of compensated mutants

HGT donor	Fitness	Generations of growth	Copy #	Size of amplification
L1 Yeast	0.68-->0.75	250	2-3	44 kbp (rRNA)
S20 Haemophilus	0.60-->0.93	80	2-4	15-90 kbp
L17 Helicobacter	0.60-->0.70	200	ND	ND





Compensation of suboptimal functions by gene amplification is common

1. Loss of formyl transferase (*fmt* gene) compensated by $tRNA^{Met}$ gene amplification (Nilsson et al, PNAS 2006)

2. Loss of tRNA synthetase activity compensated by *ileS* gene amplification (Paulander et al. Mol Micro 2007)

3. Loss of HemC activity compensated by *hemC* amplification (Paulander et al, in preparation)

4. Low B-lactamase activity compensated by amplification of the bla_{TEM-1} gene (Sun et al, in preparation)

Why are gene amplifications so commonly found as suppressors?

1. Most sub-optimal functions can be improved by increased dosage of the the cognate or non-cognate gene (Patrick WM et al Multicopy suppression underpins metabolic evolvability.Mol Biol Evol. 2007)

2. Gene duplications are several orders of magnitude more common than point mutation (10^{-5} to 10^{-2} versus 10^{-10})

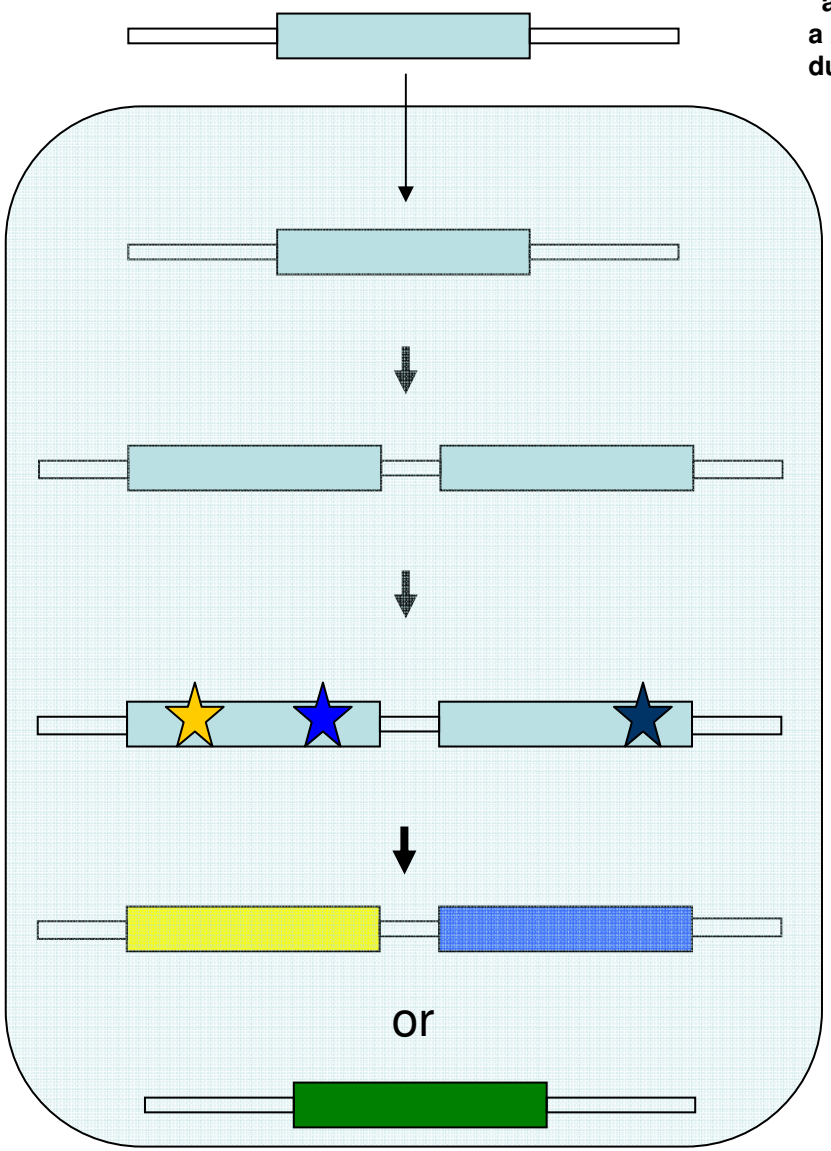
Connecting HGT and the duplication-divergence model for evolution of new genes

(Hooper and Berg, Genome Biology 2003, "a horizontally transferred bacterial gene has a 2- to 10-fold higher probability of being duplicated as compared to an indigenous gene")

HGT

Gene amplification

Divergence by mutation



Beneficial but suboptimal function

Increased gene dosage

Two paralogs

Evolved gene

Conclusions:

Surprisingly small fitness effects when replacing phylogenetically distant ribosomal proteins

However, these small fitness effects (s values around -0.01) are still sufficiently large to effectively prevent any HGT in a bacterial population

Partial amelioration of the fitness costs of foreign genes via increased gene dosage (amplification)

HGT might promote evolution of new genes via the duplication-divergence process