Directed evolution: protein improvement by mutation and selection

Dr. Nick Coleman

Directed evolution

Learning outcomes. Be able to…

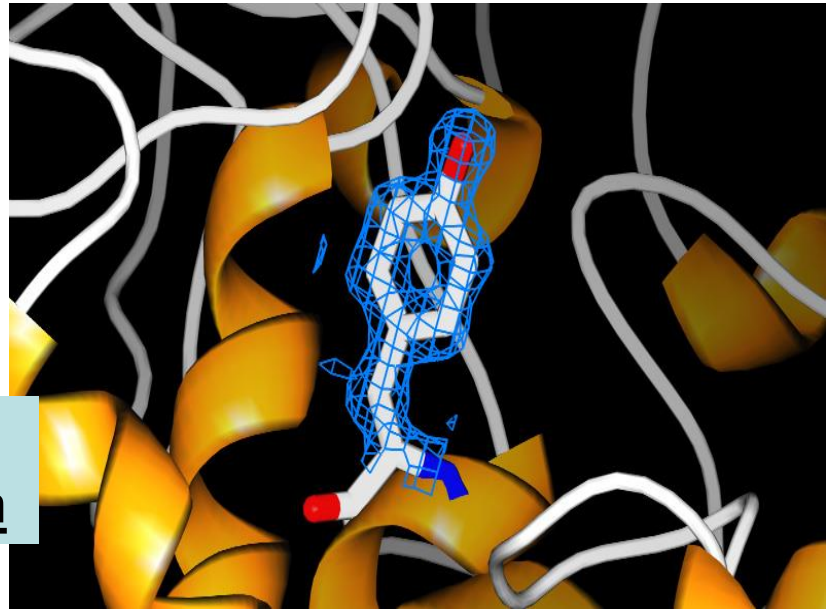
- **Describe the concepts and methods used for directed evolution of DNA sequences**
- **Describe what directed evolution offers to biotechnology**
- **Describe in detail three directed evolution methods for generation of mutant diversity**
- **Explain differences, advantages, disadvantages of screening vs. selection strategies for clone detection**

What is directed evolution?

- Creating desirable enzyme properties through mutation of genes, and selection or screening of clones*.

- Different to natural evolution:
  - desired function known in advance
  - controlled conditions of selection/screening
  - can produce traits with neutral/negative effect on host

- Different to rational design:
  - No need for knowledge of protein structure or function



Cellulase contains tyrosine at the active site: a target for rational design

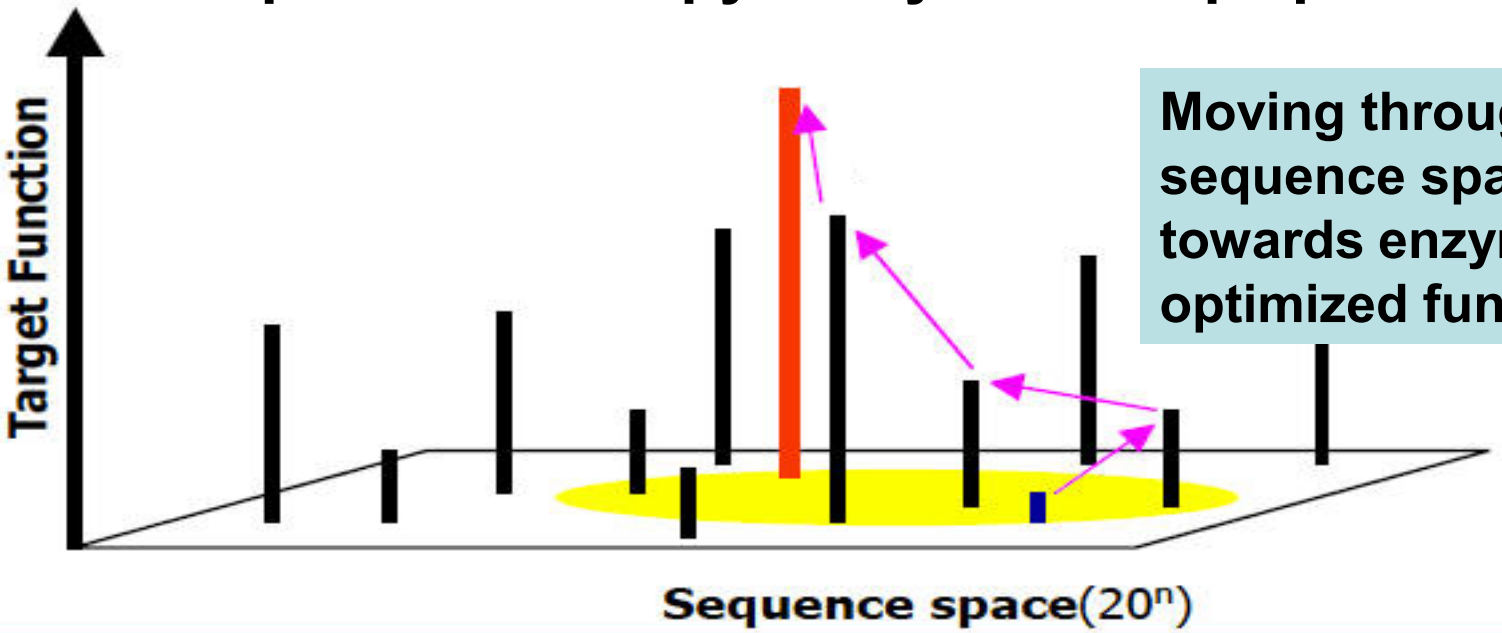
What is directed evolution?

- Searching through 'sequence space' for novel sequences that encode useful phenotypes
- Dimensions : number of different amino acids (20) to the power of protein length (typically 100-500 a.a.)

→ sequence space is HUGE; approx 20^{500} possibilities

…for a single protein !

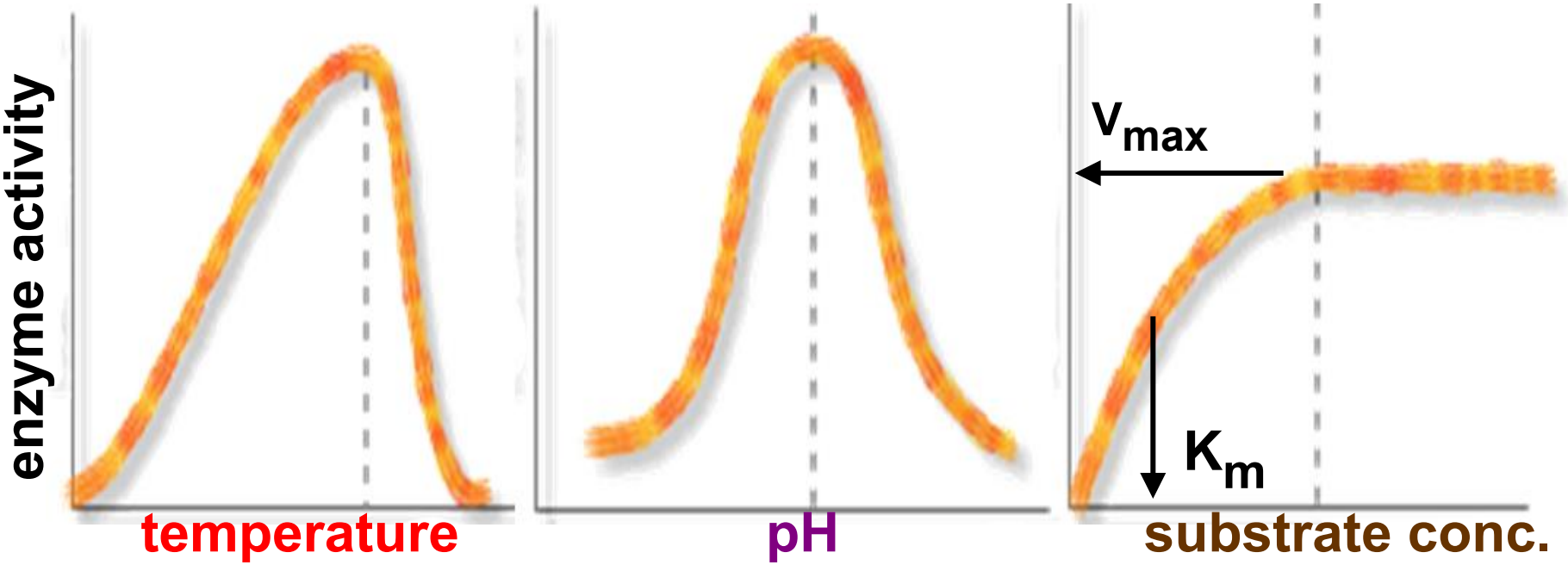
- Natural proteins occupy a tiny % of seq. space



Moving through sequence space towards enzyme with optimized function

Why use directed evolution?

- Proteins are well-adapted to their role in the cell, but may not be well-adapted for biotech applications
- <u>Common targets for protein improvement</u>:
  - thermostability and acid/alkali stability
  - higher reaction rate (V_{max}) or substrate affinity (K_m)
  - expansion of substrate range



Why use directed evolution?

- Knowledge of protein structure or function not required – only need a gene and a screening/selection method

Target gene(s)

↓

Create diversity: mutagenesis and/or recombination

↓

Expression of proteins (clone library or original host)

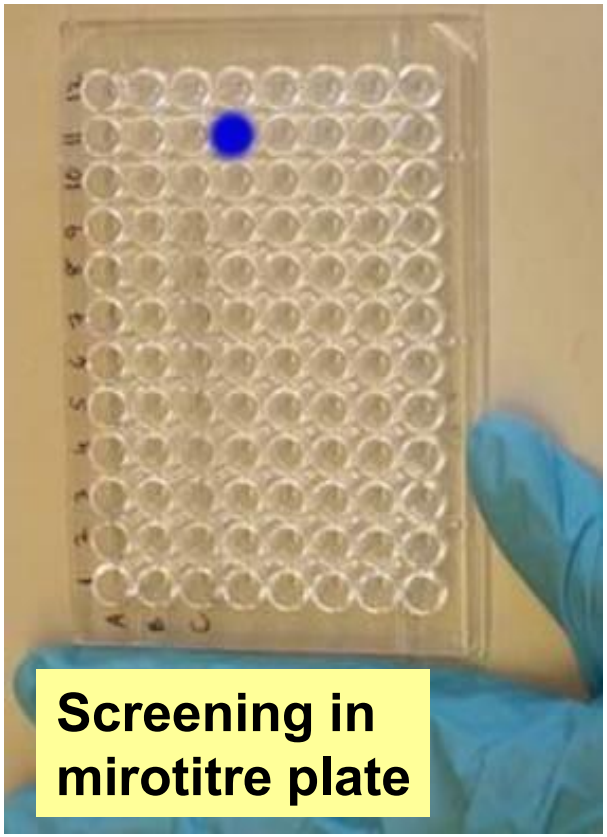
↓

Screening or selection → Improved mutant

↓

Uninteresting mutants

Further iterations lead to multiple beneficial mutations



Screening in mirotitre plate

Generation of mutant diversity

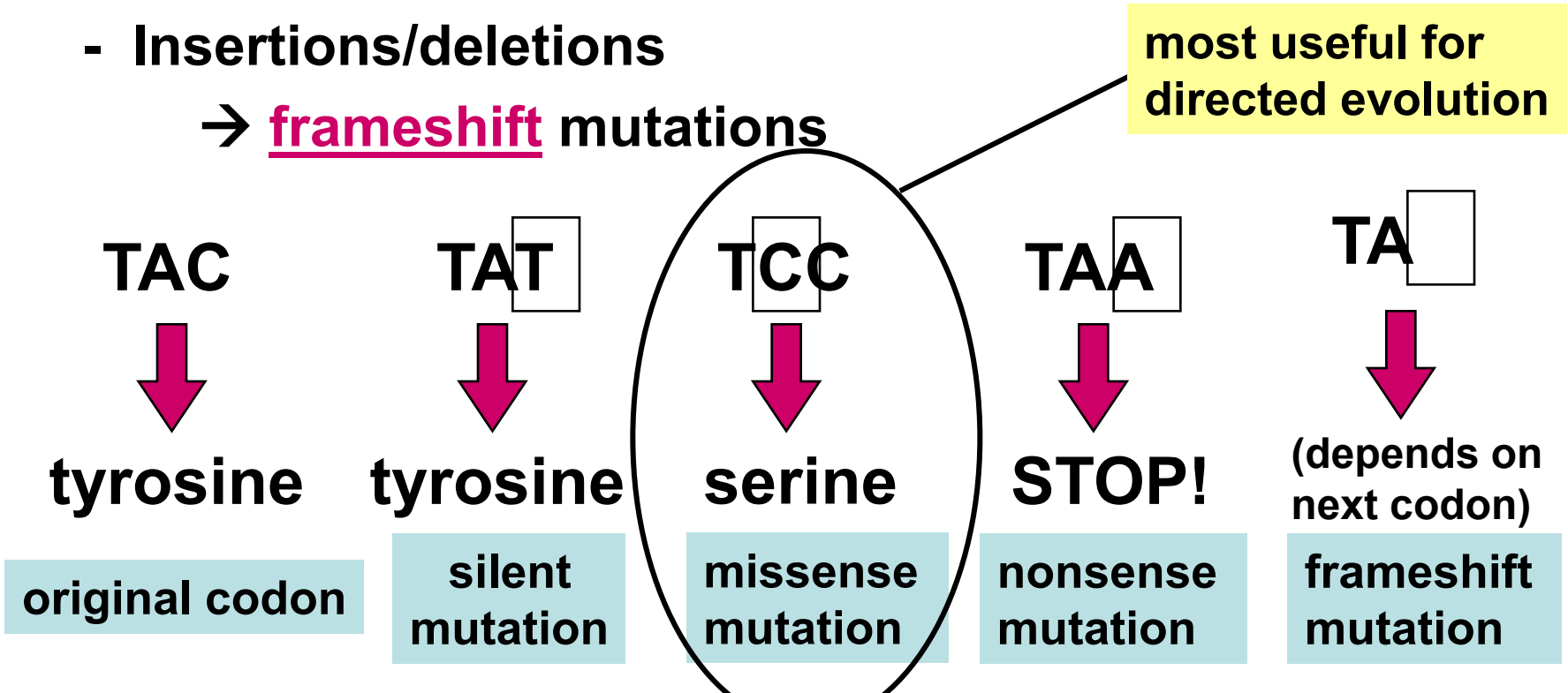
- Creating new protein functions depends on making mutations (heritable changes) in corresponding genes
- Types of mutations:

  - Point mutations

    → silent, missense, nonsense mutations

  - Insertions/deletions

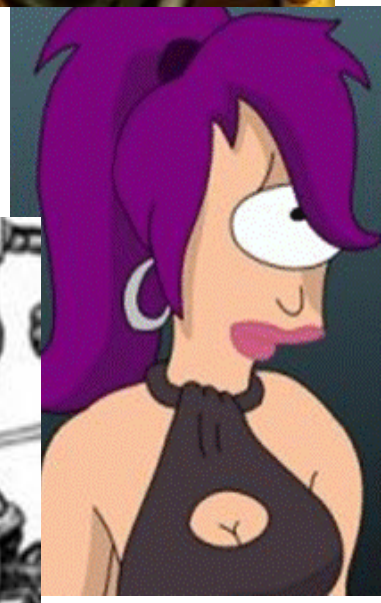
    → frameshift mutations



Generation of mutant diversity – ‘natural’ methods

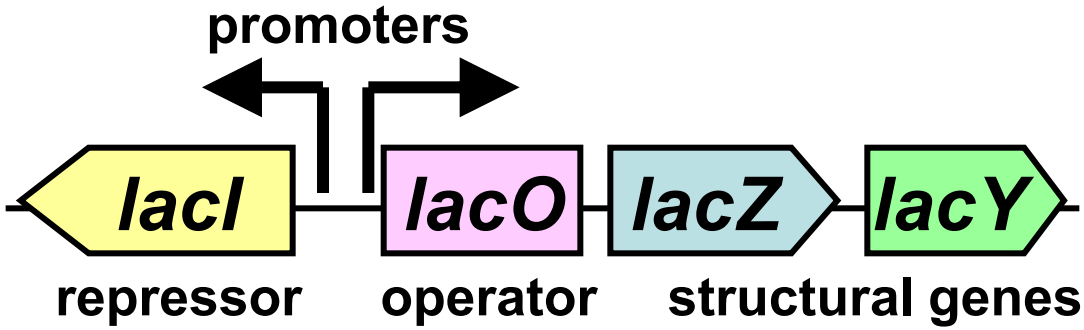
- **Spontaneous**
  - replication & repair errors
- **Chemicals**
  - alkylating agents, intercalators
- **Radiation**
  - UV, X-ray, gamma-ray
- **Transposable elements**
  - Transposons + insertion sequences

Mutants are cool.



Generation of mutant diversity – 'natural' methods

- **Disadvantages of natural methods:**
  - **cause multiple mutations throughout genome**
    → affect non-target genes, reduce cell viability
  - **not targeted at gene(s) of interest**
    → difficult to localise and characterise mutations

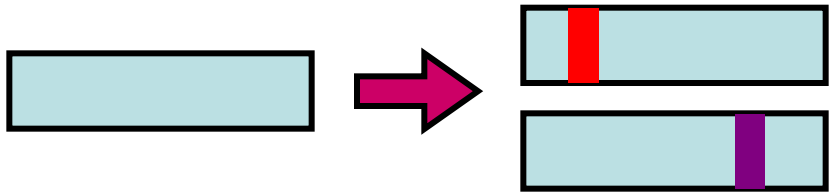


Different mutants of *E.coli* with increased growth on lactose can have mutations in lacI, lacO, lacZ, LacY, or lac promoters

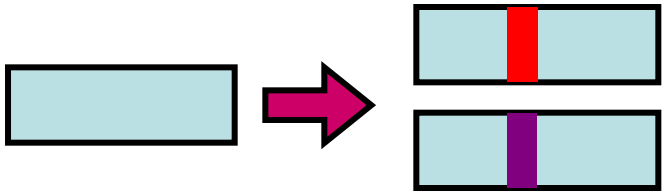
- **Molecular biological tools offer new and better ways to generate mutant diversity ….**

Generation of mutant diversity – directed evolution

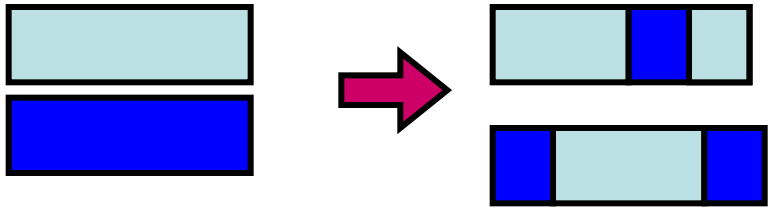
- **Random point mutations:**
  - Error-prone PCR
  - Mutator strains



- **Site-directed mutations:**
  - Strand overlap extension PCR



- **Homologous recombination:**
  - DNA shuffling

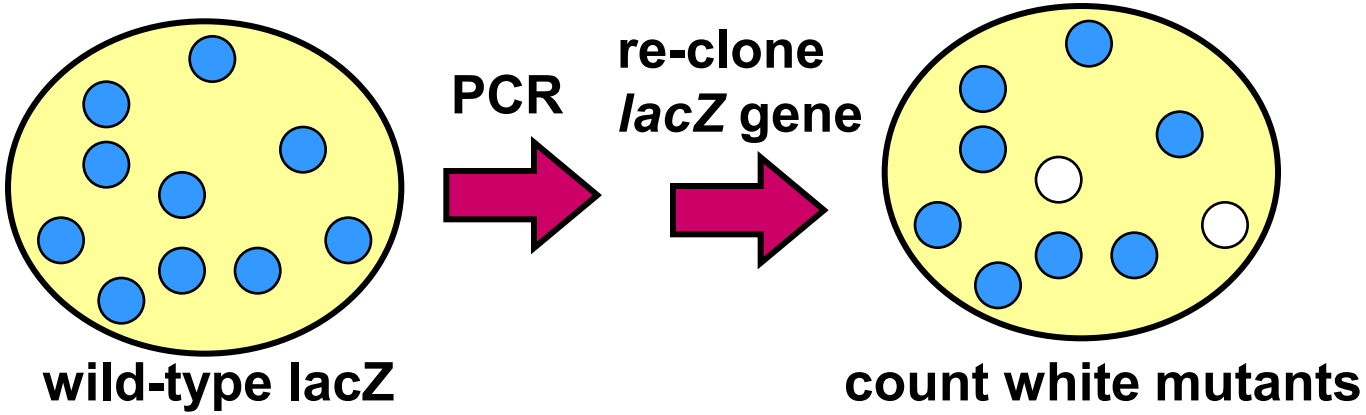


Generation of mutant diversity – directed evolution

- **Error-prone PCR**
- Error rate of thermostable DNA polymerases (eg. Taq / Pfu) can be increased by several methods:
  - add Mn^{2+} (manganese ions)
  - unequal ratio of dNTPs
  - mutant polymerase
- Combination of methods can increase mutation frequency from baseline ~10^{-6} * up to 0.2 *

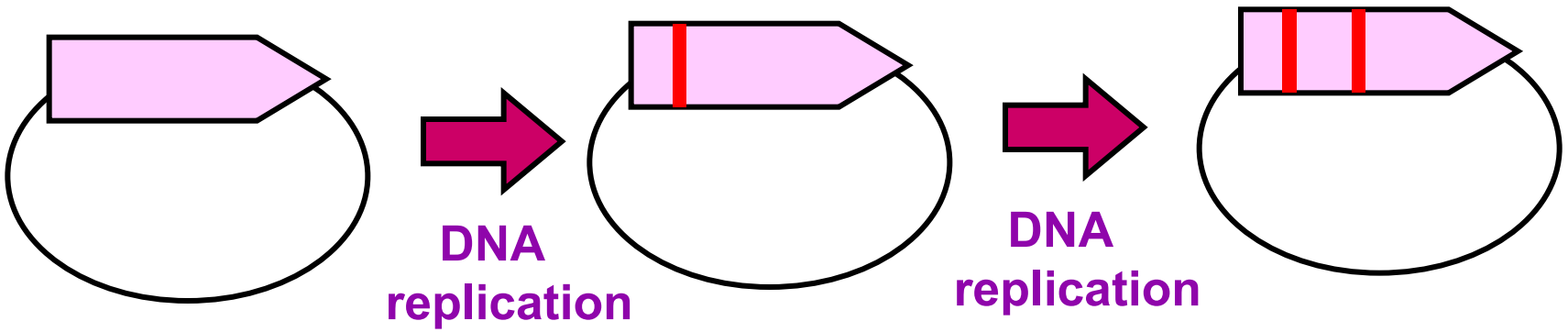
* Units: changes per nucleotide per generation

Estimation of PCR error-rate by amplification of *lacZ* gene



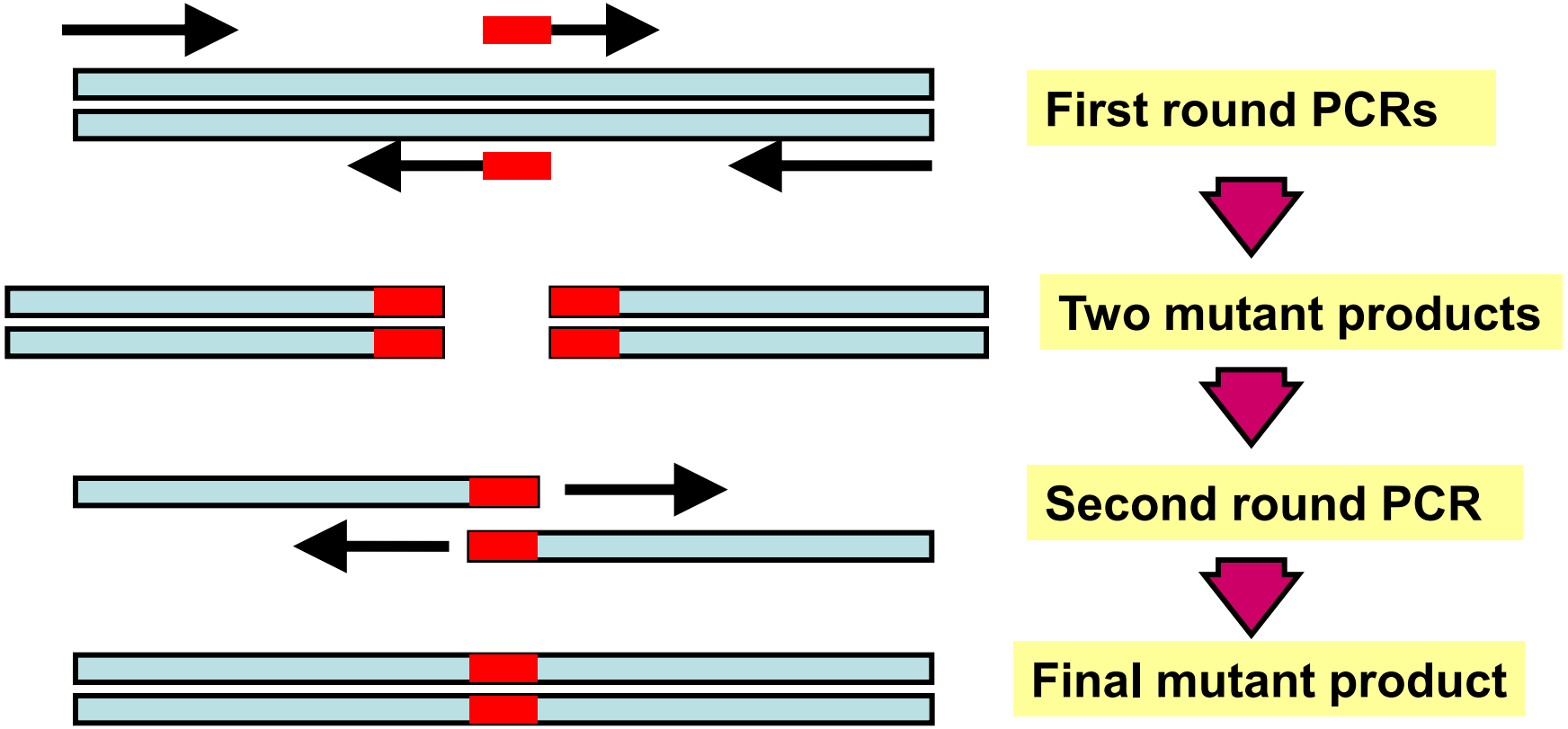
Generation of mutant diversity – directed evolution

- **Mutator strains**
- **Bacteria have multiple DNA repair and error-checking enzymes that keep their mutation rate low**
- **"Mutator strains" have natural or engineered defects in these systems: eg. *E.coli* XL1 Red has mutations in:**
  - ***mutS*: mismatch repair system**
  - ***mutT*: hydrolyzes mutagenic 8-oxo-dGTP**
- **Method: Gene of interest cloned in plasmid, replicated in mutator strain for several generations, then recovered**



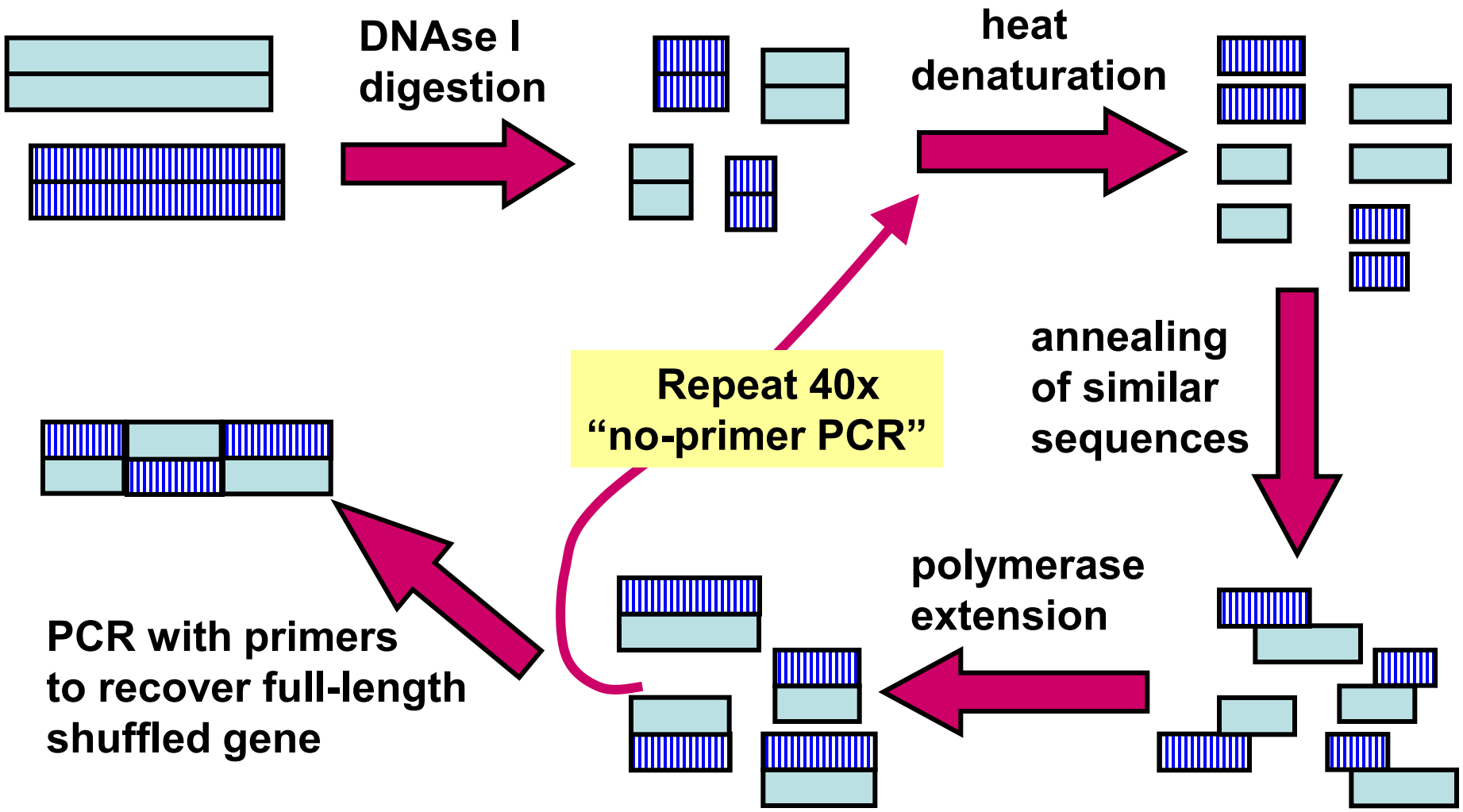
Generation of mutant diversity – directed evolution

- **Strand overlap extension (SOE) PCR**
- **Mutations produced by PCR, using mutagenic primers**
- **Creates site-specific mutations in final PCR product**
- **First PCR makes mutations, second PCR joins products**



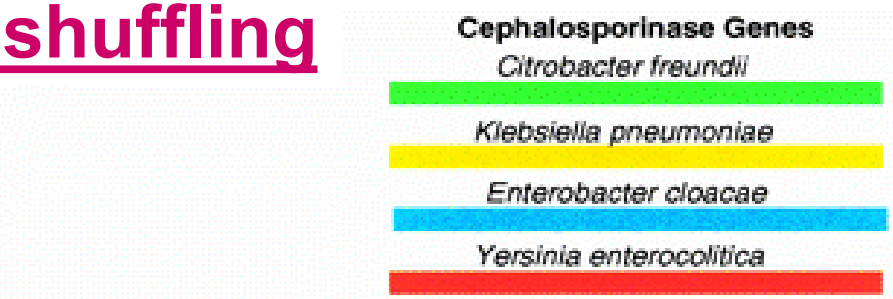
Generation of mutant diversity – directed evolution

- **DNA shuffling**
- **Relies on in-vitro recombination to join fragments of closely-related genes to give a hybrid gene**

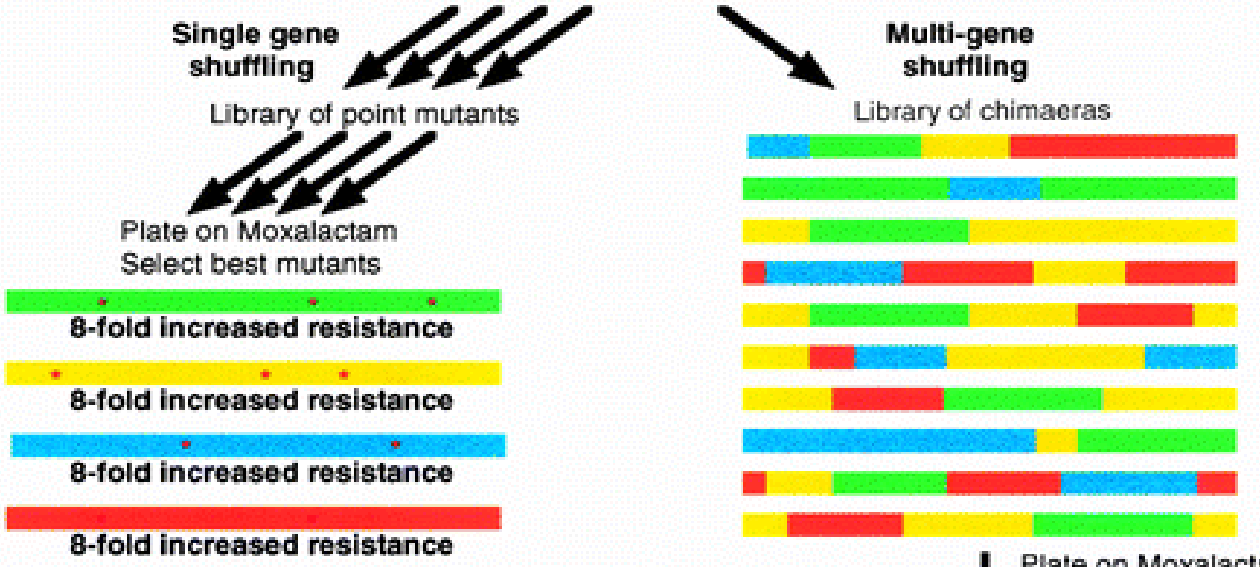


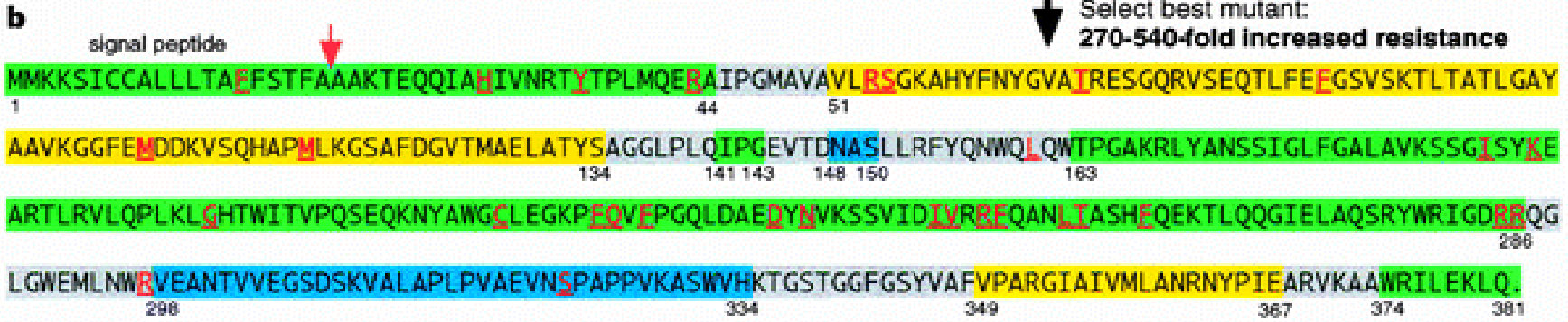
Generation of mutant diversity – directed evolution

DNA shuffling



Crameri A, et al. (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* 391 (6664): 288-291

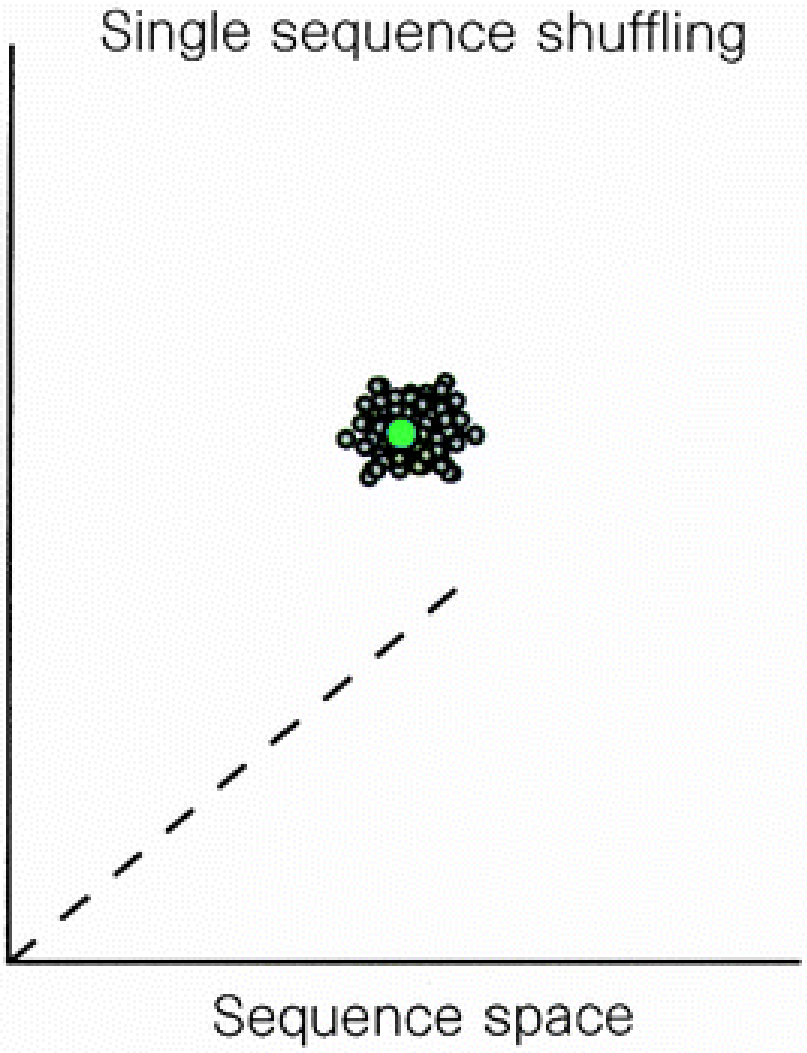


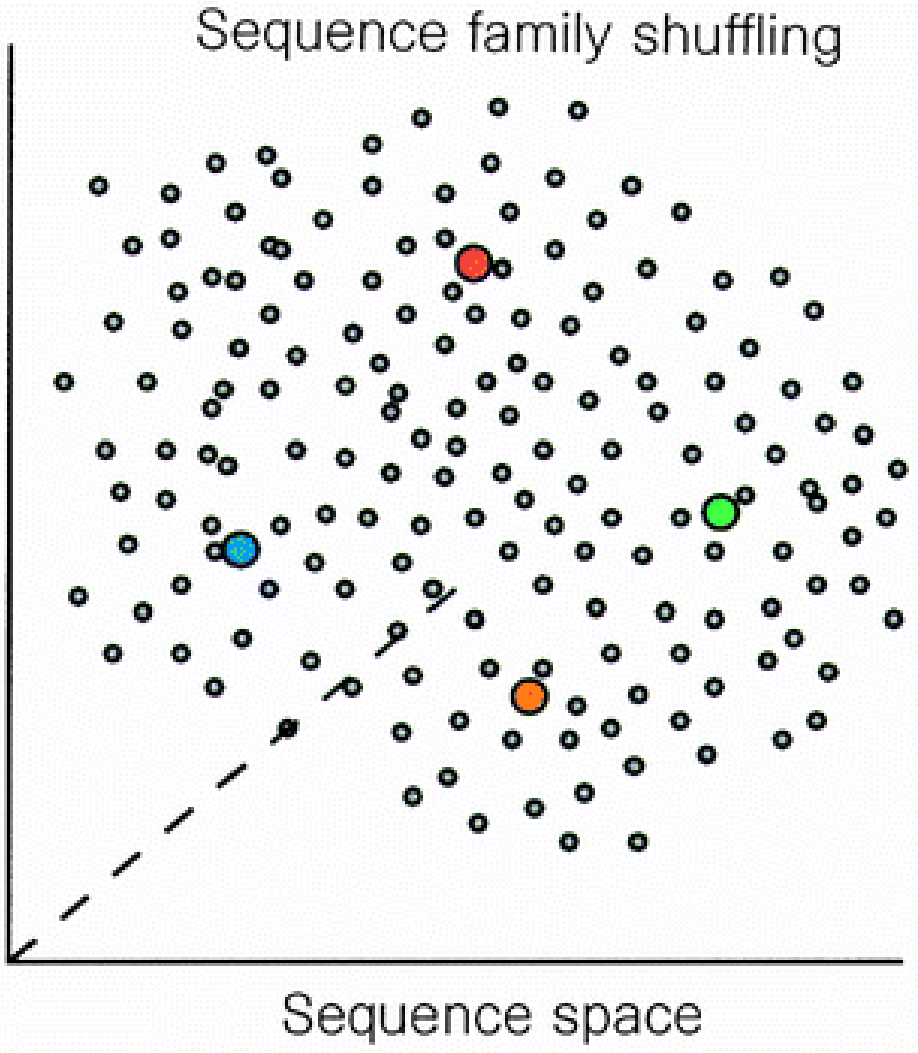


Generation of mutant diversity – directed evolution

DNA shuffling

Crameri et al. (1998) Nature 391:288-291



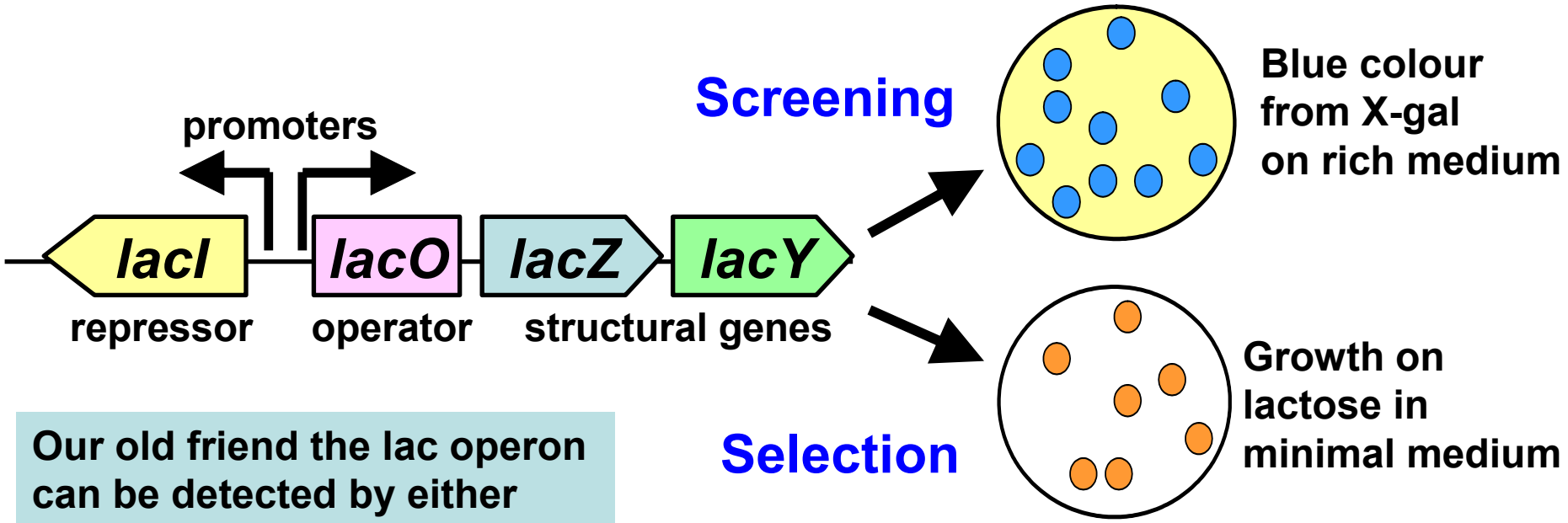


Generation of mutant diversity – comparison of directed evolution methods

- Error prone PCR, Mutator strains: Good control of mutation rate, limited to point mutations
- SOE PCR: Can intensively target regions of known significance, limited to small mutations in single gene
- DNA shuffling: Can include several parent genes, but recombination limited to areas of sequence homology

Screening and selection

- The key to success in directed evolution is appropriate screening and selection strategies
- **Screening:** phenotype **does not benefit** cells; more versatile method, but limited to ~ 10^5 clones
- **Selection:** phenotype **provides benefit** to cells (**in vivo**) **or** gene is **physically linked** to enzyme (**in vitro**)



Our old friend the lac operon can be detected by either screening OR selection

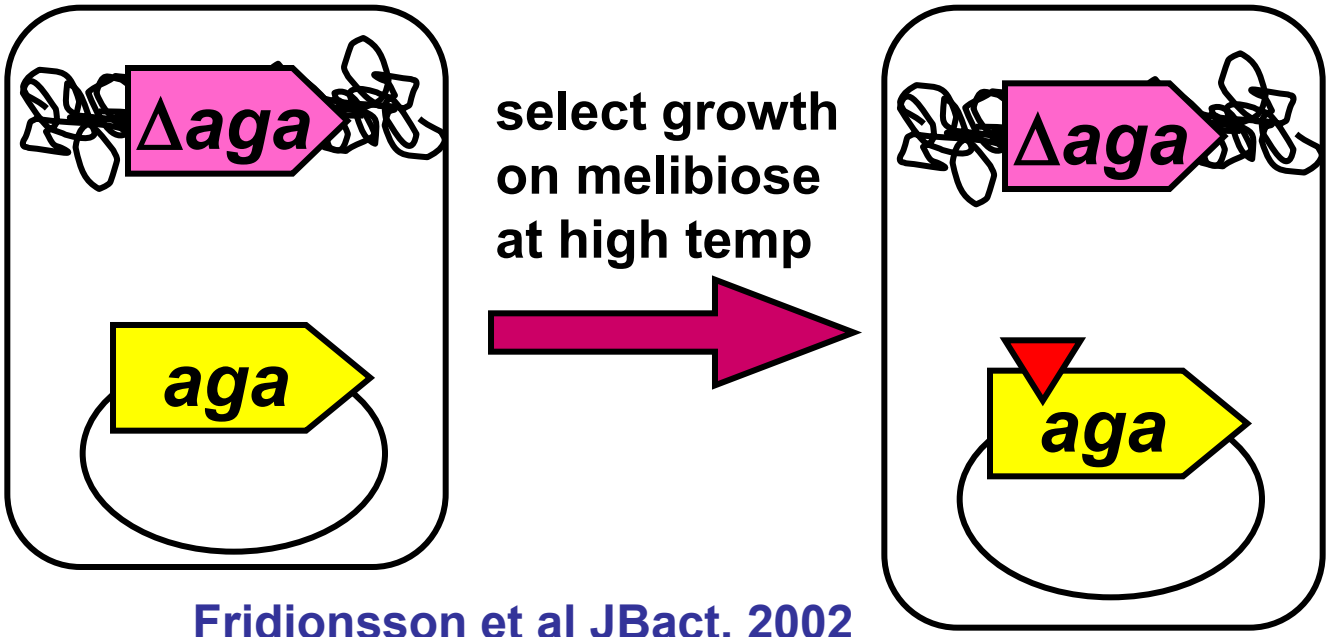
In vivo selection methods

- Require careful management of host and environment
  eg. use of minimal medium and auxotrophic host
- Most useful for resistance and catabolic enzymes, but can be applied to essential enzymes of other types….

eg. thermostable mutants of *Bacillus* alpha-galactosidase (*aga*) selected in a melibiose-negative *Thermus* host:

Deleted *Thermus aga* gene in chromosome

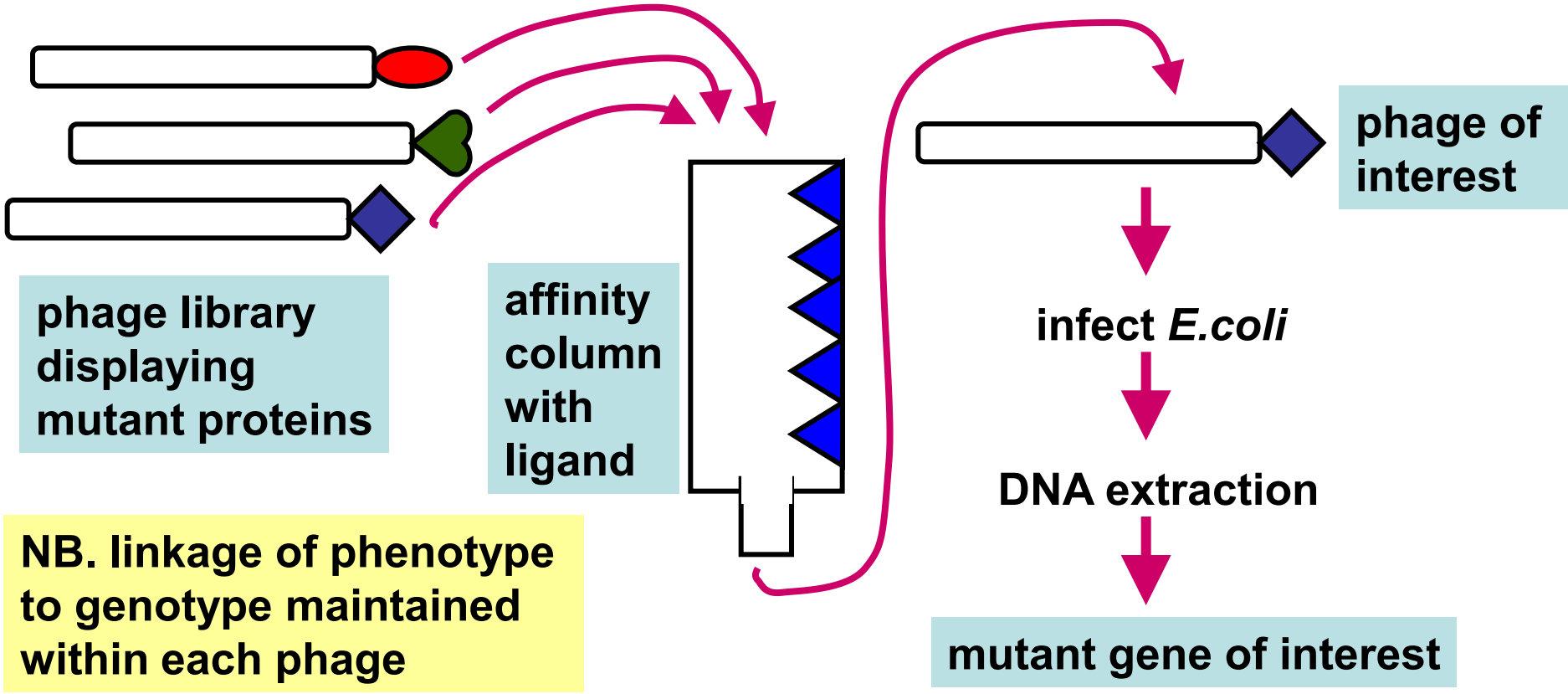
Functional, but less thermostable *Bacillus aga* gene on plasmid



Fridjonsson et al JBact, 2002

In vitro selection methods – phage display

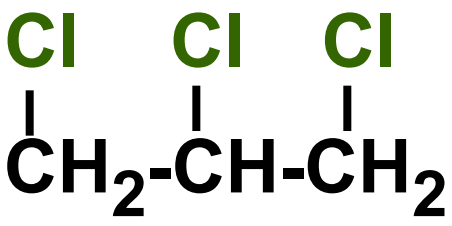
- Library of mutant genes cloned into phage, as translational fusion with coat protein
- Mutant proteins expressed on phage surface; allows capture of a specific phage by affinity chromatography
- Depends on availability of specific ligand

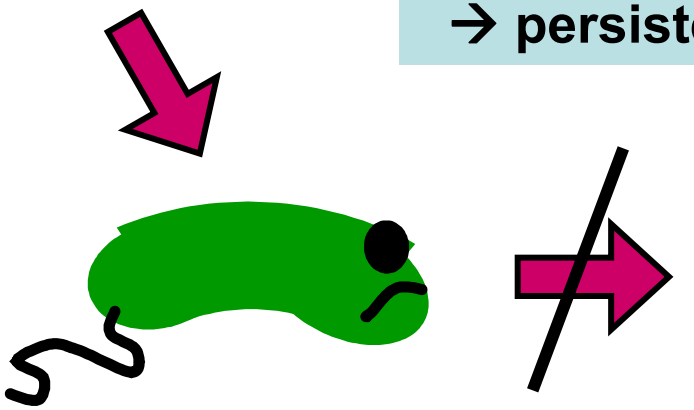


NB. linkage of phenotype to genotype maintained within each phage

Directed evolution of a trichloropropane degrader

- **Trichloropropane (TCP): toxic xenobiotic compound used as a solvent and industrial synthetic intermediate**
- **TCP is resistant to biodegradation, therefore compound accumulates in environment, becomes a pollutant**
- **Can directed evolution be used to make a bacterium that effectively biodegrades TCP? (Bosma et al, AEM 2002)**

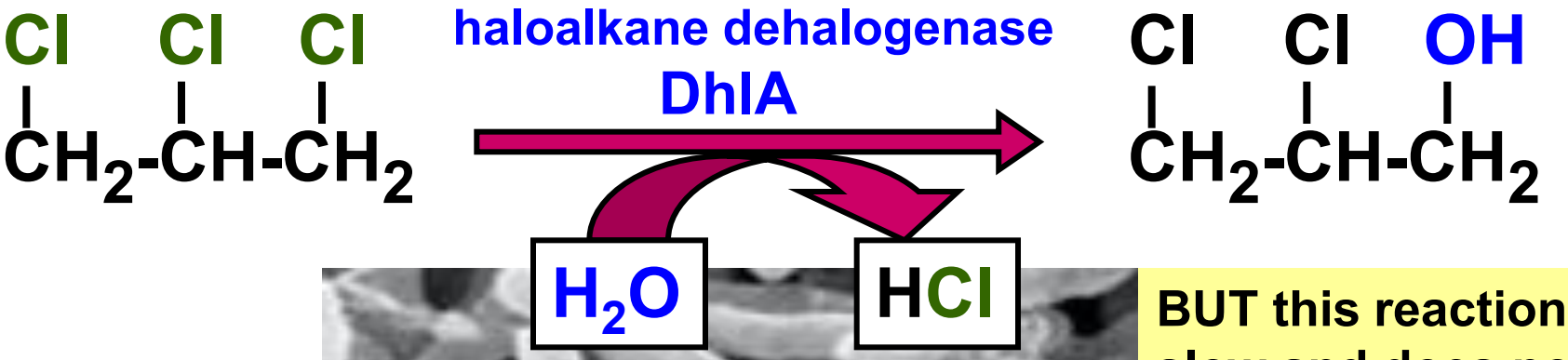


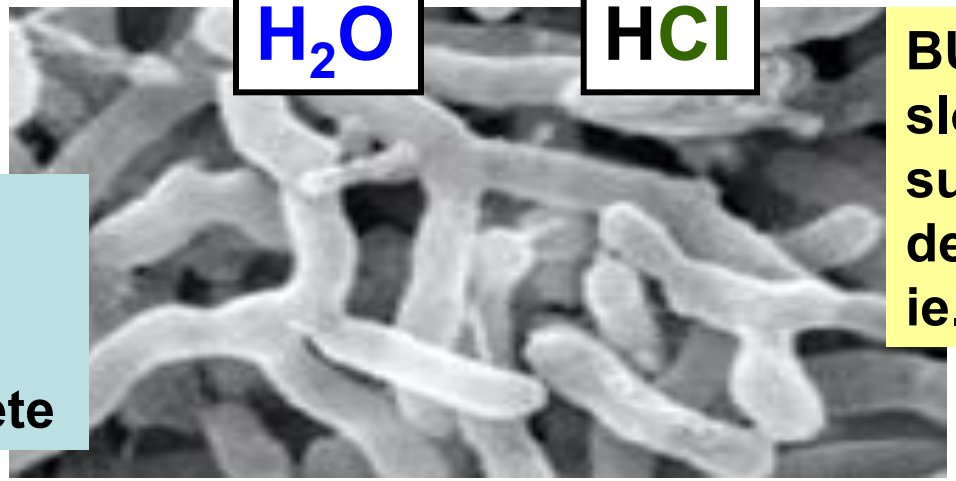


Xenobiotics: not made by biosynthesis, often not recognised by microbial enyzmes → persistence

Directed evolution of a trichloropropane degrader

- **Step 1. Bioprospecting: Where can we find the raw materials for directed evolution?**
- ***Rhodococcus* M15-3 grows on chlorobutane, expresses a haloalkane dehalogenase enzyme that hydrolyses carbon-chlorine bonds in many compounds, including TCP**





***Rhodococcus*: Gram positive branching rods - an actinomycete**

BUT this reaction is slow and does not support growth of degrading bacteria, ie. 'cometabolism'

Directed evolution of a trichloropropane degrader

- **Step 2. Isolate gene, create mutants**
- **Dehalogenase gene *dhlA* amplified by PCR, subjected to single-gene DNA shuffling (DNAse I fragmentation, annealing, extension, re-amplification)**
- **Initial library of *dhlA* gene mutants ligated into expression plasmid pGEF, introduced into *E.coli* BL21**

PCR

dhlA

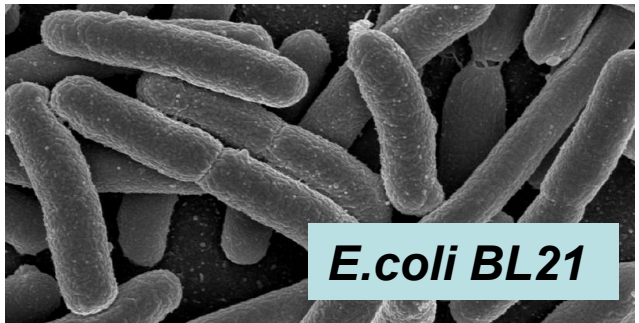
DNA shuffling

dhlA *dhlA* *dhlA* *dhlA*

ligation to vector

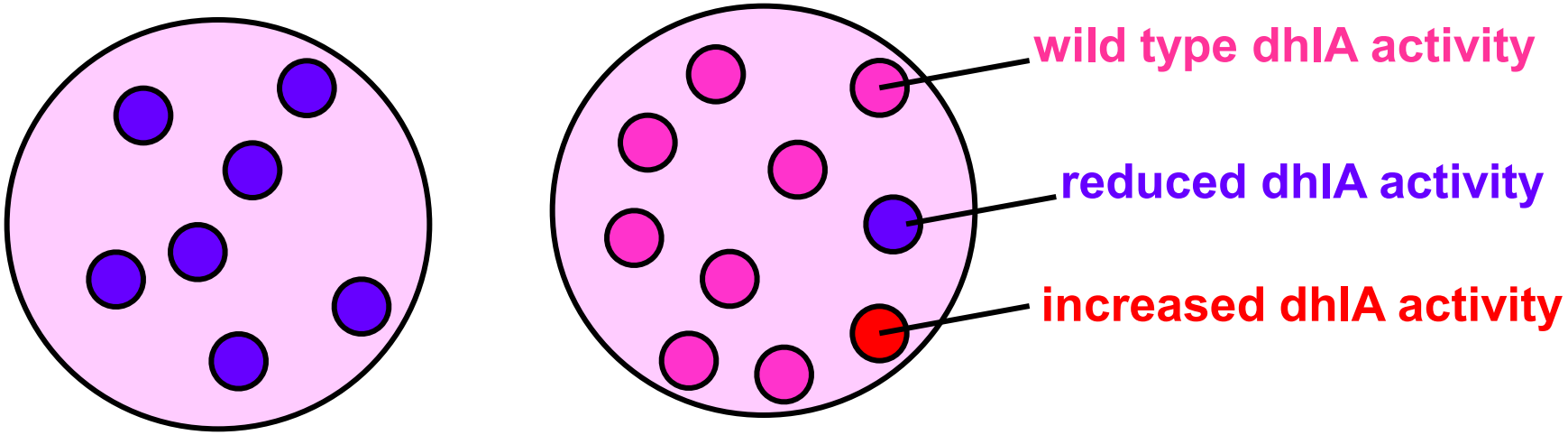
dhlA *dhlA* *dhlA* *dhlA*

transformation into *E.coli*



Directed evolution of a trichloropropane degrader

- **Step 3. Screen mutants for activity**
- **Library of 10,000 recombinant *E.coli* colonies screened for increased dehalogenation activity against TCP**
- **Screening assay: HCl acid production from TCP gives a red colour on eosin-methylene blue (EMB) agar, which contains pH indicator dyes**
- **One improved enzyme mutant gave stronger red colour**

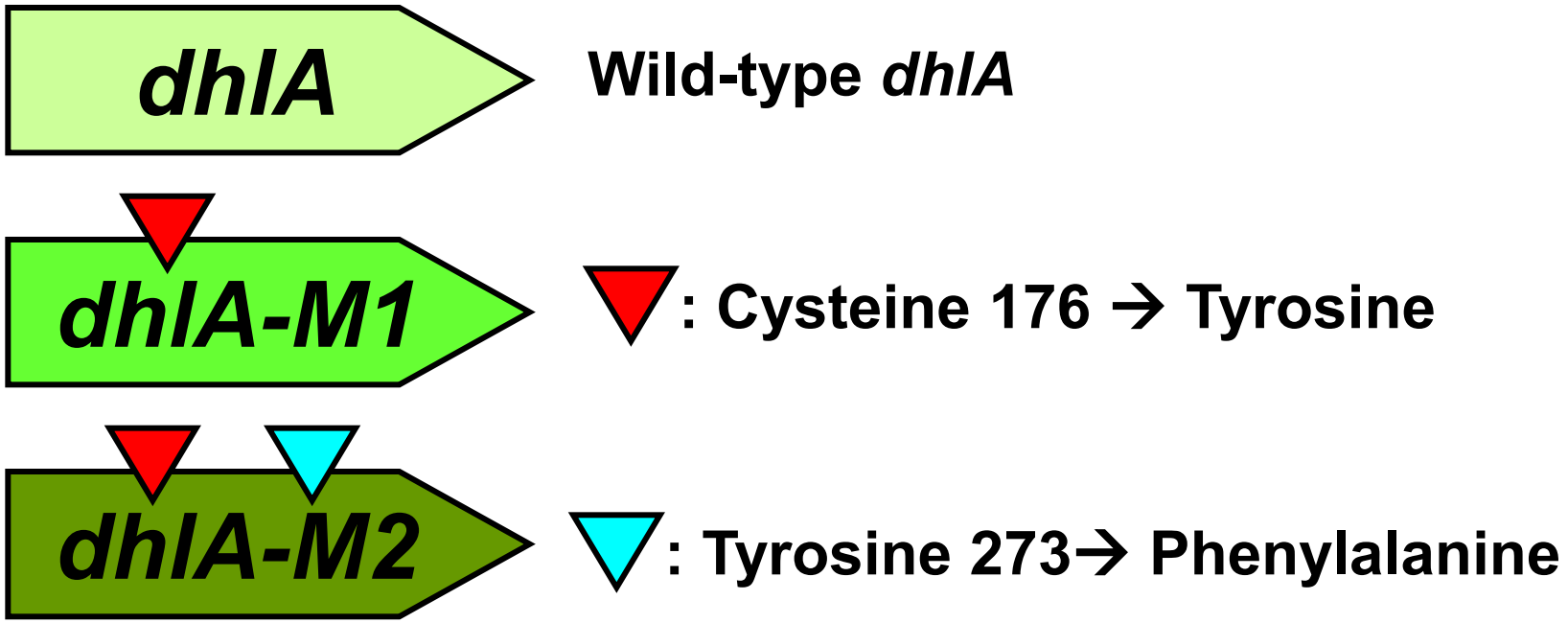


***E.coli* host on EMB-TCP agar**

***E.coli* pGEF-*dhlA* clones on EMB-TCP agar**

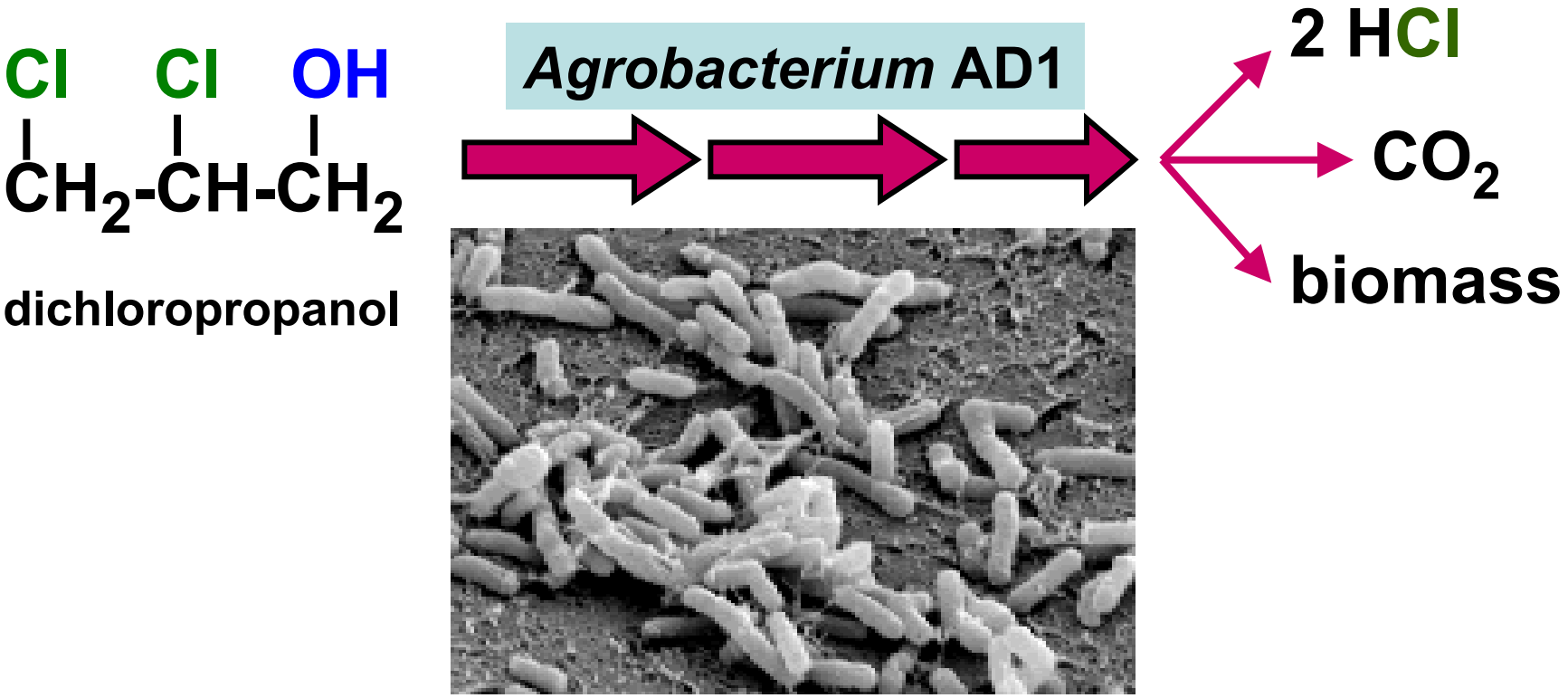
Directed evolution of a trichloropropane degrader

- **Step 4. Further mutagenesis, sequence analysis**
- **Best mutant from DNA shuffling (M1) used as starting point for error-prone PCR mutagenesis**
- **Screening of 10,000 clones on EMB agar repeated → 14 gave increased HCl production from TCP, best mutant (M2) retained for analysis**



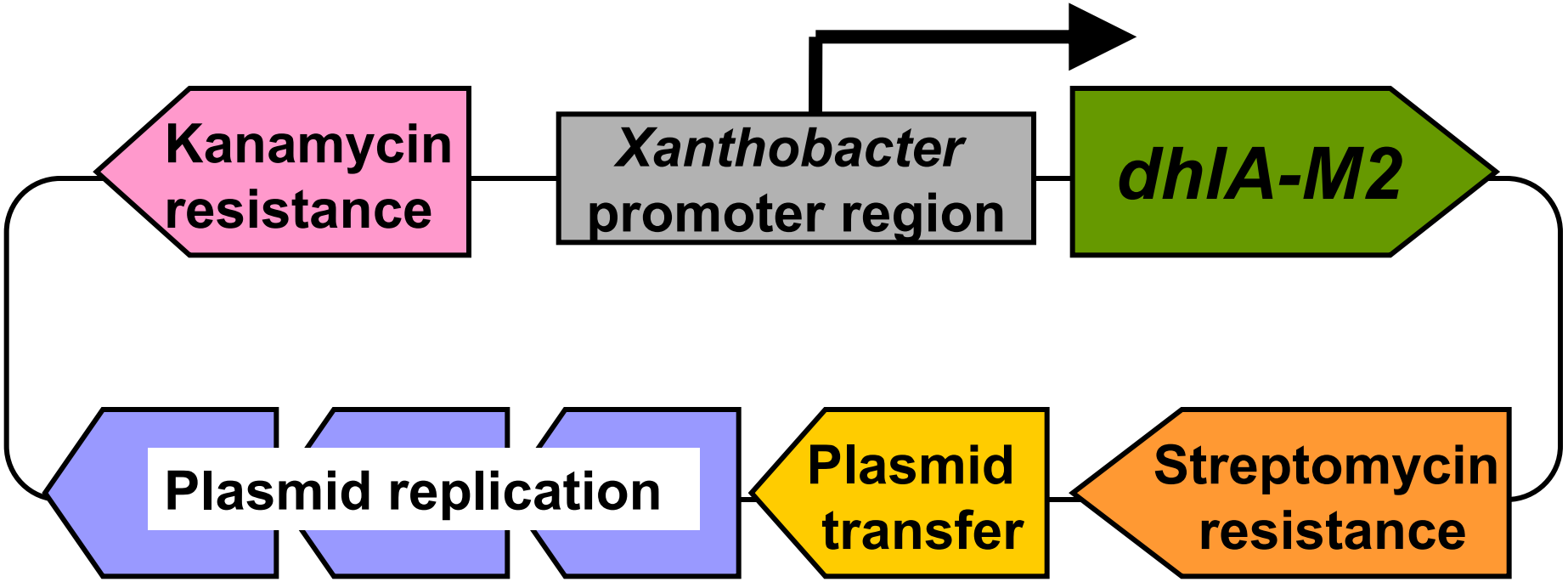
Directed evolution of a trichloropropane degrader

- **Step 5. Further bioprospecting for new host**
- **Improved *dhlA*-M2 gene rapidly attacks TCP with high affinity, but *E.coli* host cannot grow on TCP**
- **Need a host bacterium that can use the reaction end product (dichloropropanol) as carbon and energy source**



Directed evolution of a trichloropropane degrader

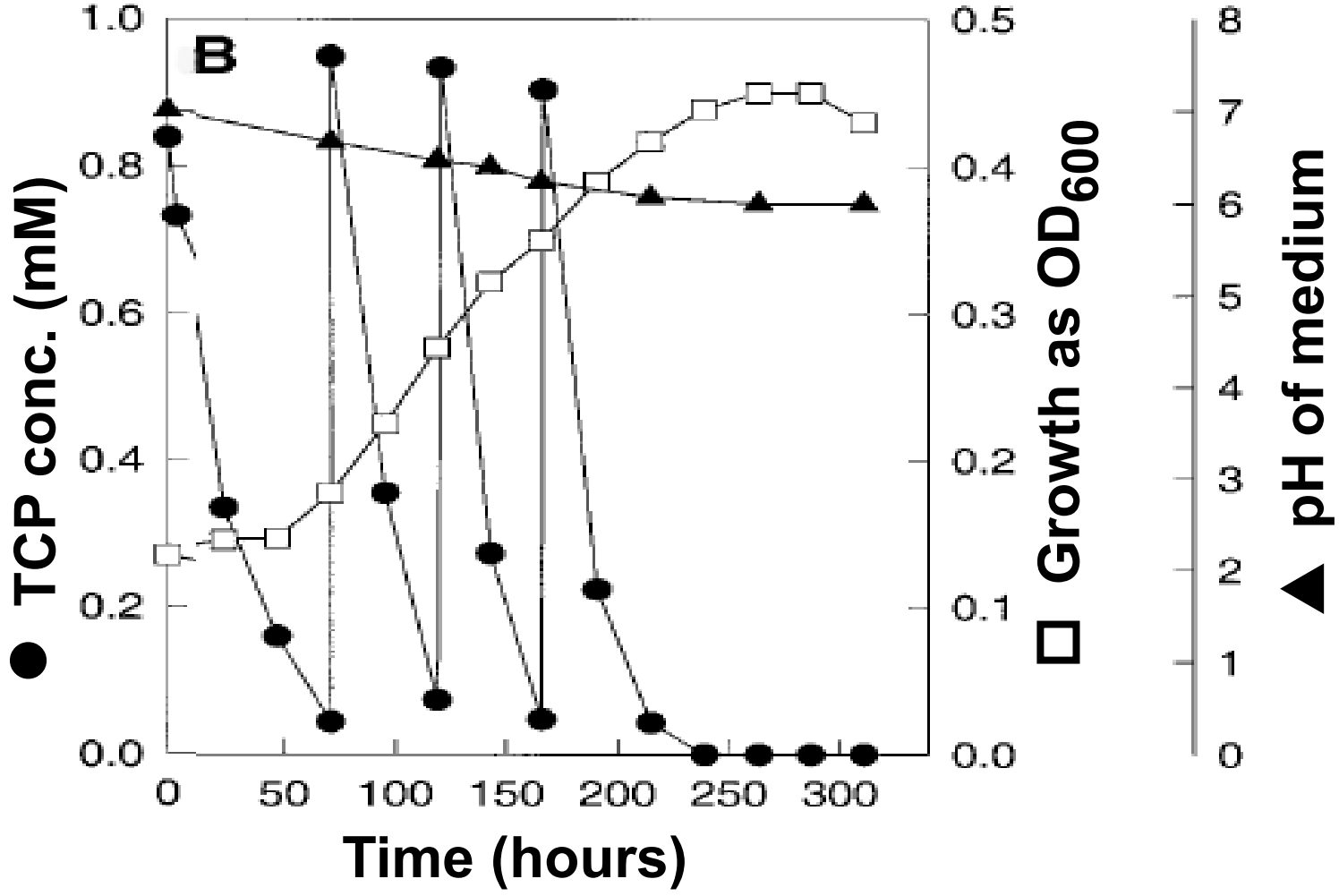
- **Step 6. Transfer of evolved gene into new host**
- ***dhlA-M2*** **gene ligated to broad host range plasmid, under control of constitutive promoter from *Xanthobacter* sp.**
- **plasmid conjugated into *Agrobacterium*, recombinants selected on trichloropropane as sole C source**



structure of broad host range dehalogenase plasmid

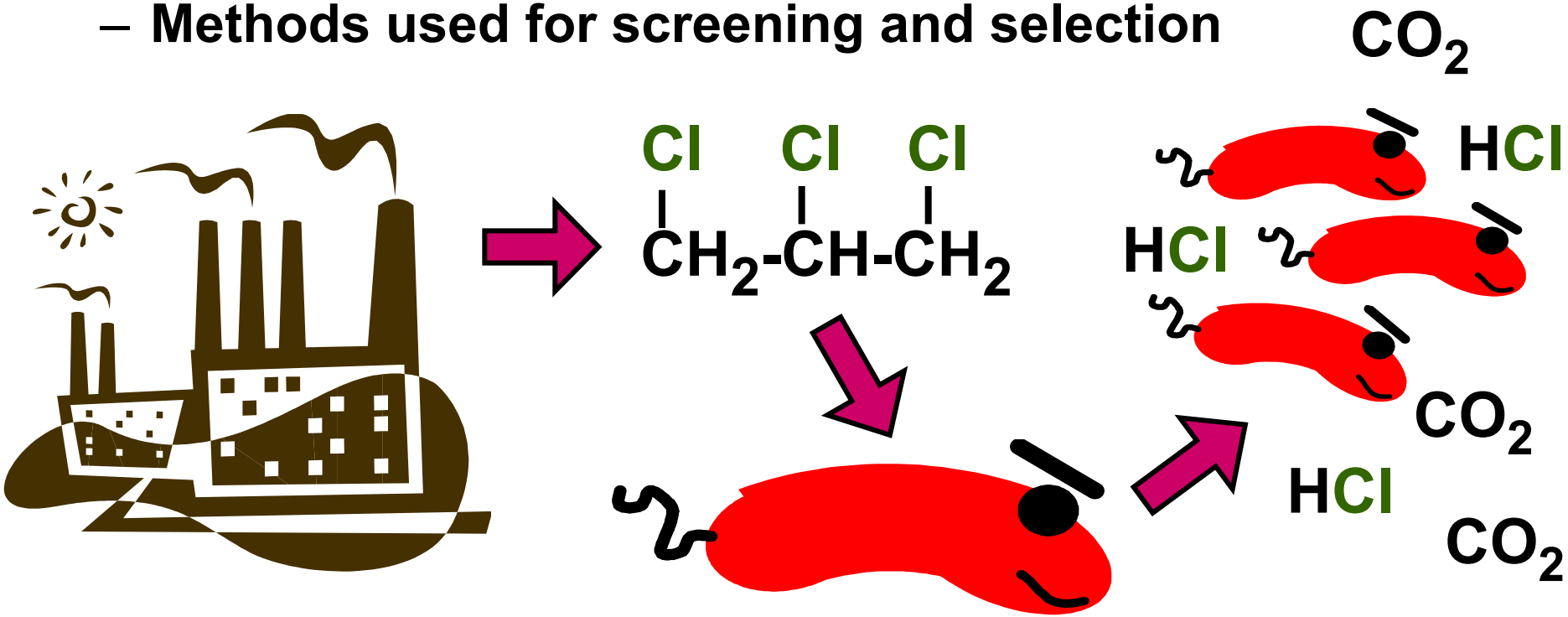
Directed evolution of a trichloropropane degrader

- **Step 7. Confirmation of growth on TCP**
- **Recombinant *Agrobacterium* containing dehalogenase plasmid biodegrades TCP as a sole carbon source**



Directed evolution of a trichloropropane degrader

- **Conclusions - key elements of strategy**
  - Source organisms for enzymes
  - Host organisms for evolution and expression
  - Appropriate vectors for cloning in each host
  - Methods used for directed evolution
  - Methods used for screening and selection



Further reading

-C. Neylon (2004) Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. Nucleic Acids Research 32:1448-1459

-L. Yuan et al (2005) Laboratory-directed protein evolution. Microbiology and Molecular Biology Reviews 69: 373-392

-T.Bosma et al (2002) Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene. Applied and Environmental Microbiology 68:3852-3857

Picture credits

http://lhs.lps.org/staff/sputnam/Biology/U6Evolution/finches.png
http://www.zbh.uni-hamburg.de/wurst/protspace/
http://unit.aist.go.jp/brf/images/iwakura/iwakur1.jpg
www.blobs.org/science/enzyme/index.shtml
http://en.wikipedia.org/wiki/ELISA
http://www.chm.bris.ac.uk/motm/dna/dna1.gif
http://www.europhysicsnews.com/full/18/article7/illenberger_fig2.jpg
http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/A/AmesTest.gif
http://www.agen.ufl.edu/~chyn/age4660/lect/lect_07/FG06_055.GIF

Picture credits

http://fvl.vfu.cz/sekce/sek_patob/1240/obrazkova_priloha/mikrob/img/13-02.jpg
http://upload.wikimedia.org/wikipedia/fr/3/3c/Onpg.png
http://meds.queensu.ca/~mbio318/FACs.gif
http://nar.oxfordjournals.org/content/vol31/issue19/images/large/gng119f4.jpeg
http://www.rhodococcus.ca/
http://www3.niaid.nih.gov/biodefense/Public/Images.htm
http://home.earthlink.net/~mountfuturama/Leela.html
http://upload.wikimedia.org/wikipedia/en/thumb/2/22/Pg02.jpg/300px-Pg02.jpg
http://imagesource.allposters.com/images/pic/
LASR/125-LM~X-Men-Storm-Limited-Edition-Transparency-Posters.jpg