ADVANCED ENZYME TECHNOLOGY

PROTEIN ENGINEERING (PART 2)

by
AIZI NOR MAZILA RAMLI
Faculty of Industrial Sciences & Technology
aizinor@ump.edu.my

Chapter Description

- Expected Outcomes
  - Describe about the general principal of protein engineering
  - Discuss about the strategies of protein engineering
  - Identify the differences between each strategies

CONTENT

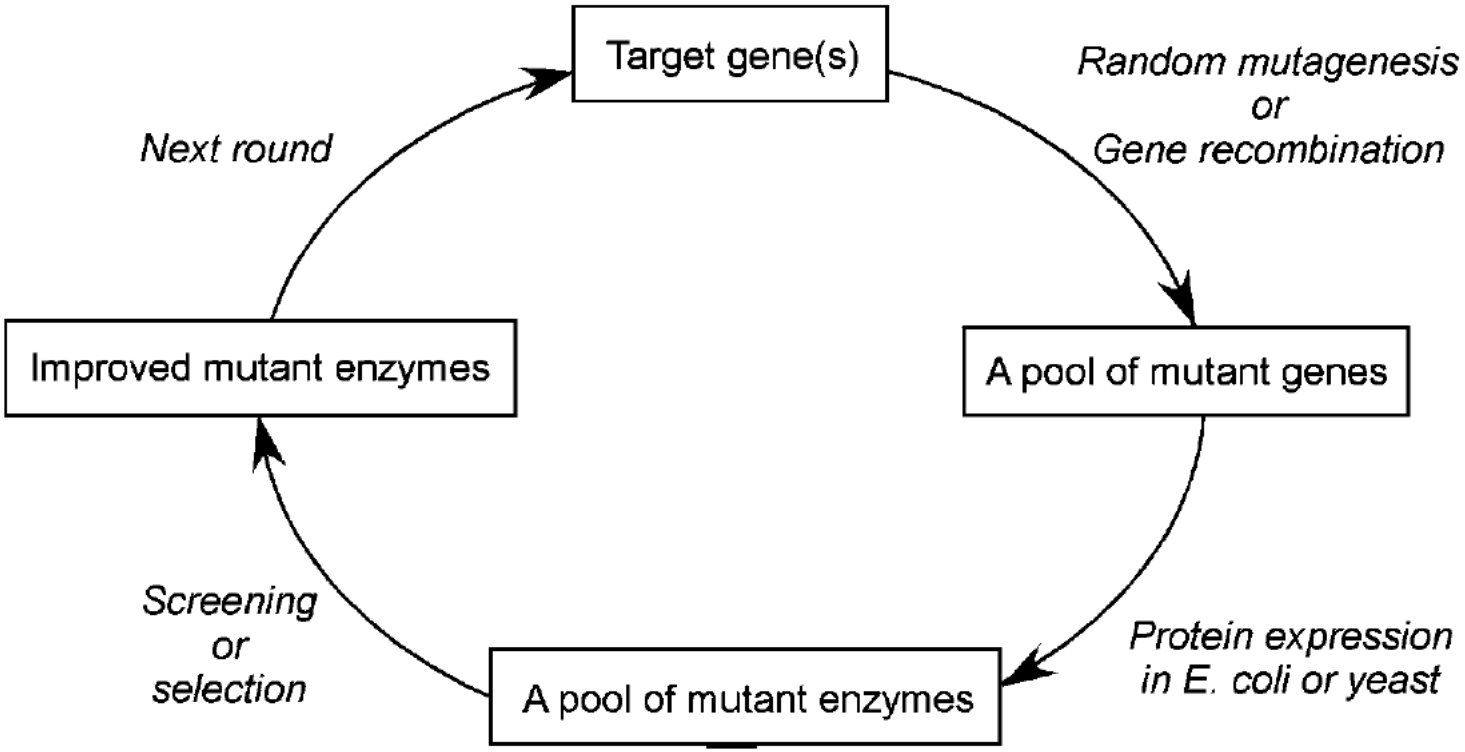
- General introduction
- Strategies of protein engineering
  - *De novo* design
  - Rational design
    - Site directed mutagenesis
      - Overlap extension PCR mutagenesis
  - Directed evolution
    - Random mutagenesis
      - Error-prone PCR

Directed Evolution

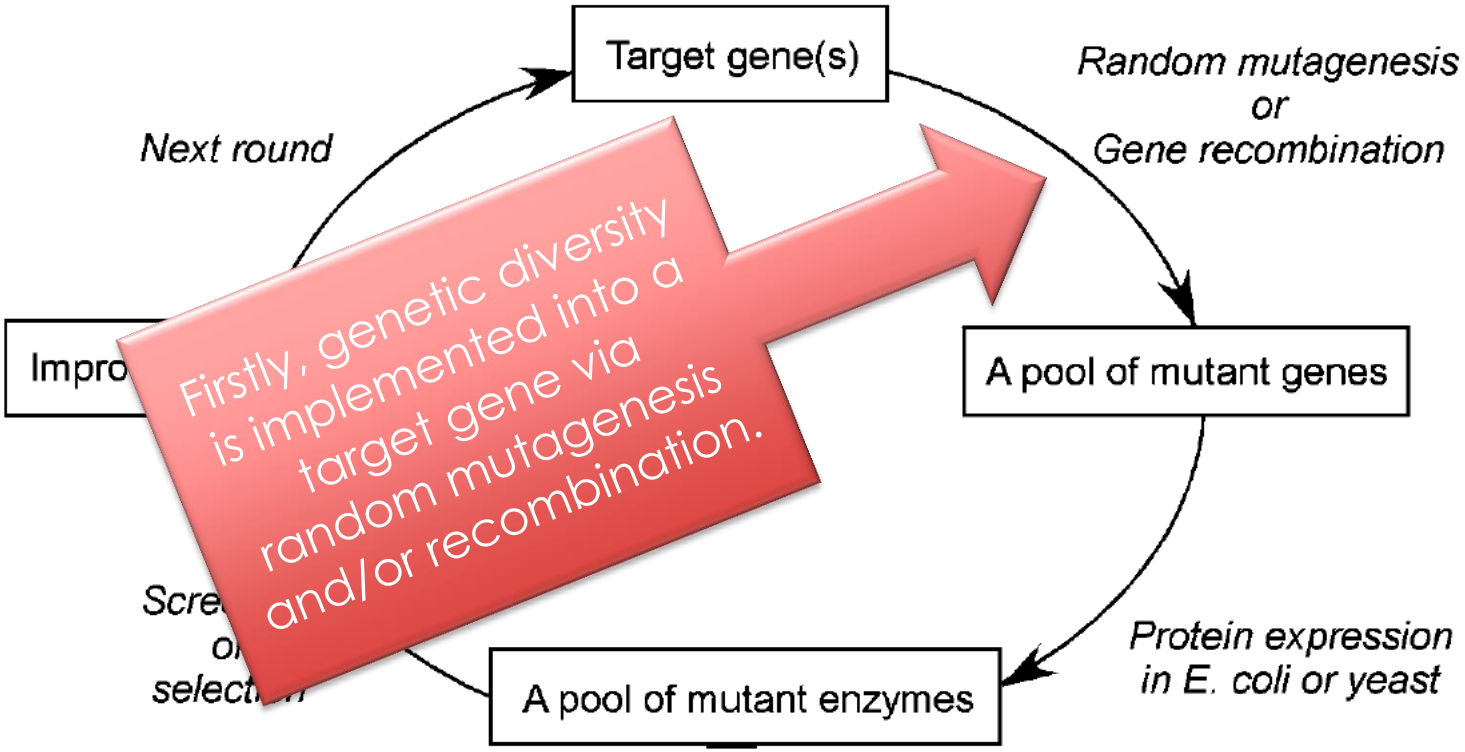
- Although **rational design** strategy has been reported as
  - one of an effective protein design strategy, **percentage of successful findings is still considered as low** owing to the **inadequate understanding regarding protein folding, structure, function, and dynamics**.
  - In addition, this technique relatively **time-consuming** due to the requirement of a high-quality 3D protein structure as an experimental guidance.

- In contrast, **directed evolution** does not rely on any predetermined notions about what is important
  - only **depend on a very simple algorithm**: **diversification together with selection**.
  - In principle, directed evolution resembles natural Darwinian evolution implemented in a laboratory environment.

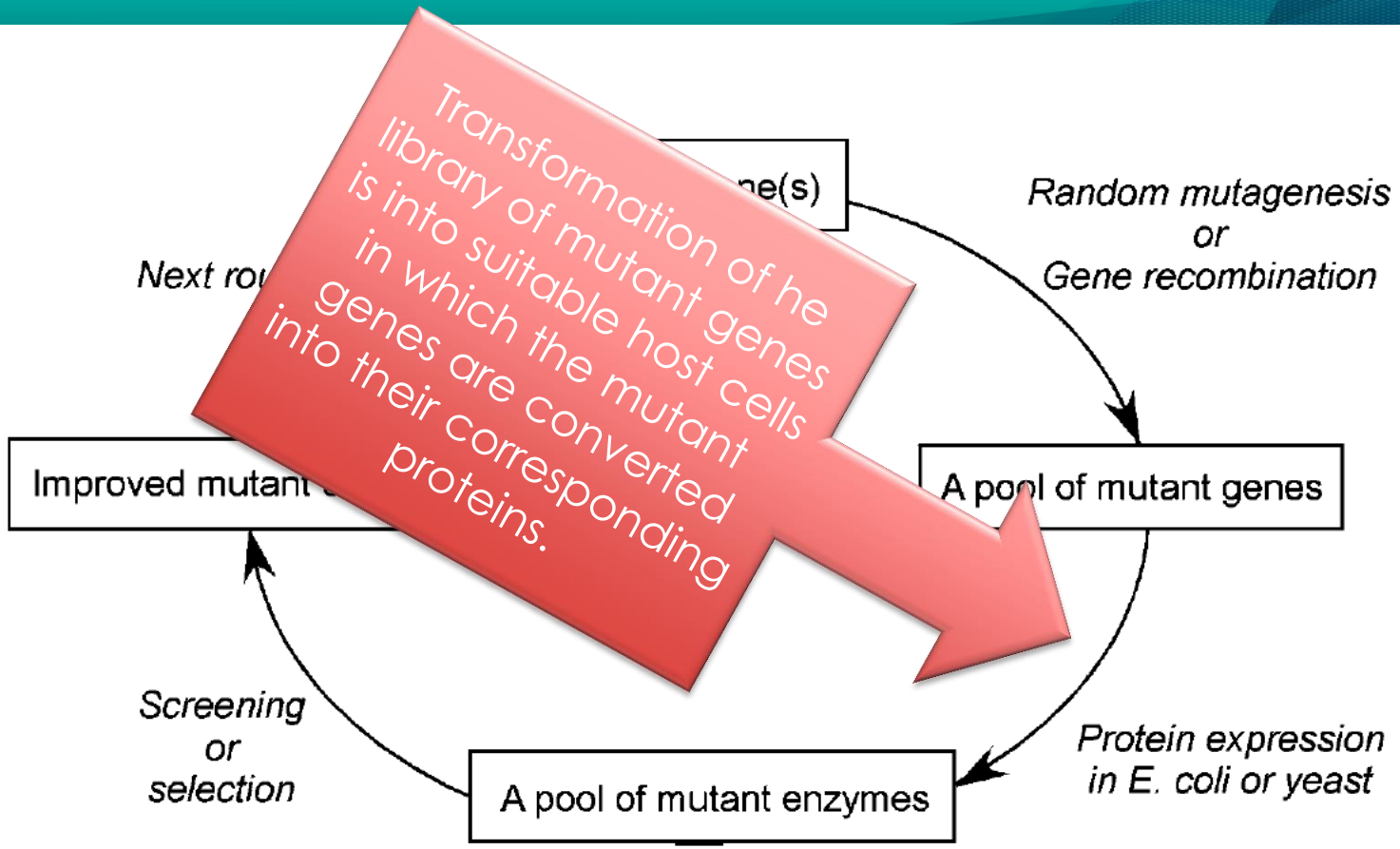
The general scheme of directed evolution



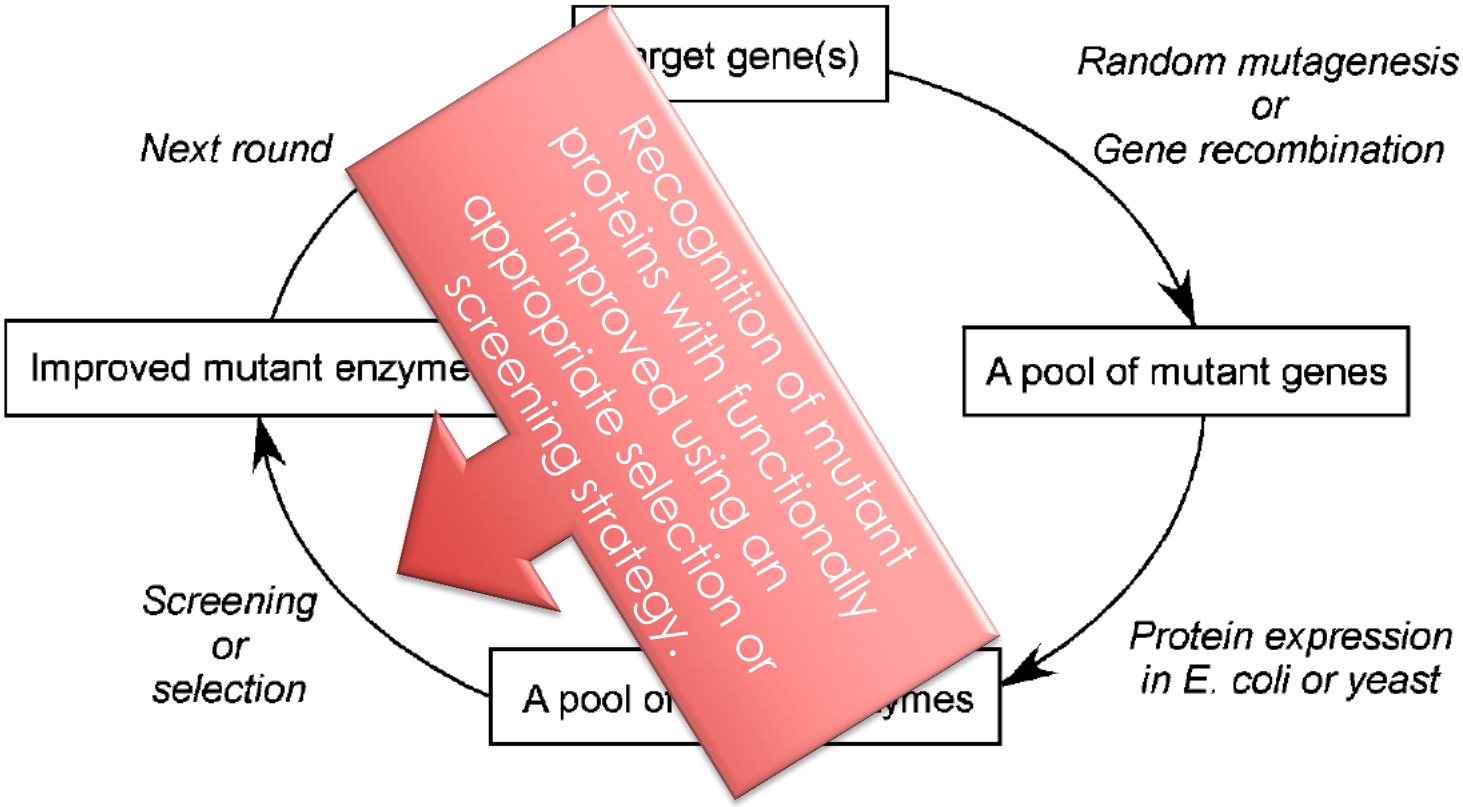
The general scheme of directed evolution



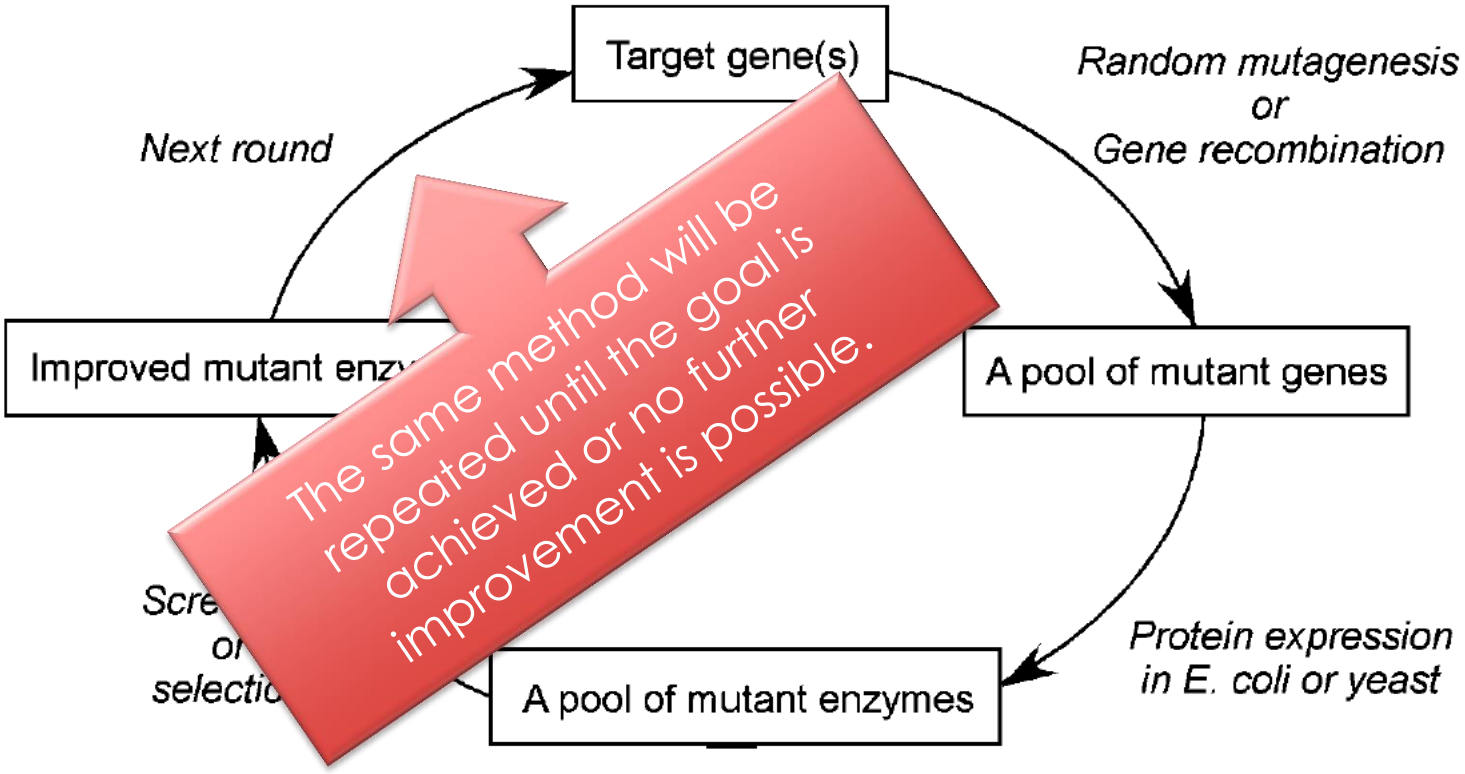
The general scheme of directed evolution



The general scheme of directed evolution



The general scheme of directed evolution



- **NO requirement for structural knowledge**

- **NO requirement for information of mechanism**

Involves 2 General approaches (key component):

1. creating **genetic diversity** via random mutagenesis or gene recombination methods.

2. **Screening and seletion** method development to identify the successful variants.

- Two main categories of genetic diversity:
  - **Random mutagenesis** methods and
  - **Gene recombination** methods .

Differences between random mutagenesis and gene recombination methods.

- **Random mutagenesis**
  - generate a library of variants which contain point mutations (substitution, deletion or insertion) and come from a single parental gene.
  - Example of representative methods such as Error prone PCR, Mutator strains, chemical mutagens and scanning saturation mutagenesis.

- **gene recombination**
  - generally begins by creating blockwise sequence information exchange within the parental genes which can be either single or a pool of homologous genes.
  - Involve breaking of DNA fragments followed by rejoining to form a new DNA combination.
  - Example of representative methods such as DNA shuffling, family shuffling, StEP, ITCHY, SHIPREC

Random Mutagenesis

- Point mutations introduced in all areas within DNA of interest
- Resulted into a library of variant containing both wild-type and random mutated DNA.
- Generally for a real library -> contain many variants -> screening !!!
- If methods efficient -> mostly mutated DNA will be obtained

General strategy for random mutagenesis

Requirements:

- DNA of interest (gene or promoter) must be cloned

- Expression system must be available -> for testing phenotypic change

Error-prone PCR

- The most recognized method of random mutagenesis – simple and efficient!
- Involve a series of the standard PCR which their reaction conditions have been slightly modified.
- Generally, In a standard procedure of DNA amplification via PCR, the possibility of Taq DNA polymerase to introduce incorrect nucleotides into the progeny genes is very limited .

- However, a very minor error rate can be intensely enlarged by altering the reaction buffer that during amplification process will influence the Taq DNA polymerase to introduce more incorrect nucleotides.

- The modification will consist of:
  1. consumption of imbalanced concentrations of dNTPs;
  2. Usage of MnCl2 in addition to MgCl2; and
  3. Increase the concentration of Mg2Cl2 (up to 7 mM).

- The major difficulty of error prone PCR is that it
  - can only get into six amino acid substitutions approximately at a given residue position due to the degeneracy of the genetic code, thus reducing the potential diversity significantly.
  - Mutations occurred are not totally random. For instance, a typical preference of error-prone PCR is the high exchange of AG nucleotide.

DNA shuffling

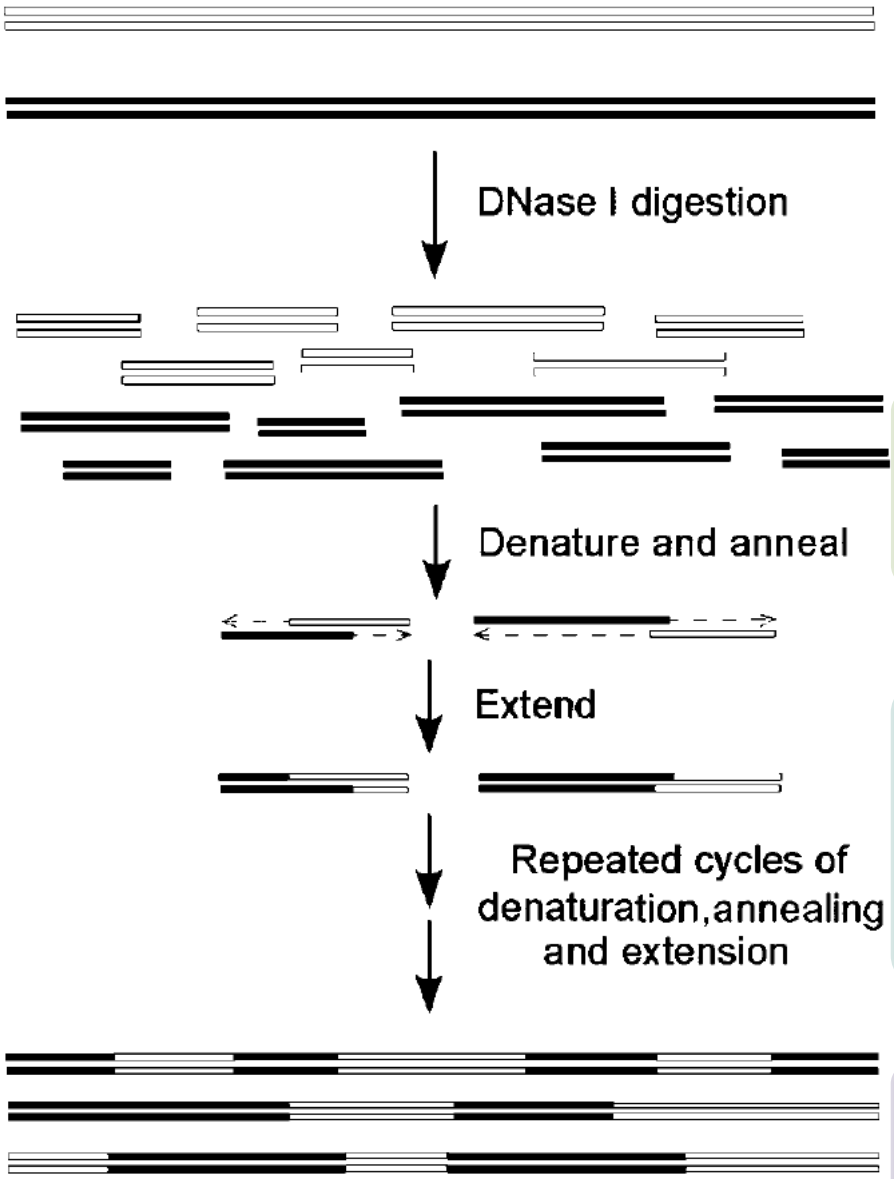
- Also known as "sexual PCR"
- The first and most established method of gene recombination method
- Begin with the formation of a small double-stranded DNA fragments with the read length between 20 to 50 bp. In this case, DNase I will be used as an enzyme to digest randomly a pool of closely related genes containing point mutations.
- A PCR-like reaction without any primers will be performed to purify and reassembled a small double-stranded DNA fragments produced into a full-length gene.

DNA shuffling and family shuffling

- Fragments originated from dissimilar parental genes will prime one another in a process of Recombinogenic events.

- A standard PCR reaction will then be implemented in which reassembly mixture is used as a template with primers flanking the gene of interest.

- The final amplified product will consist of a library of full-length genes containing recombined mutations from different parental genes.



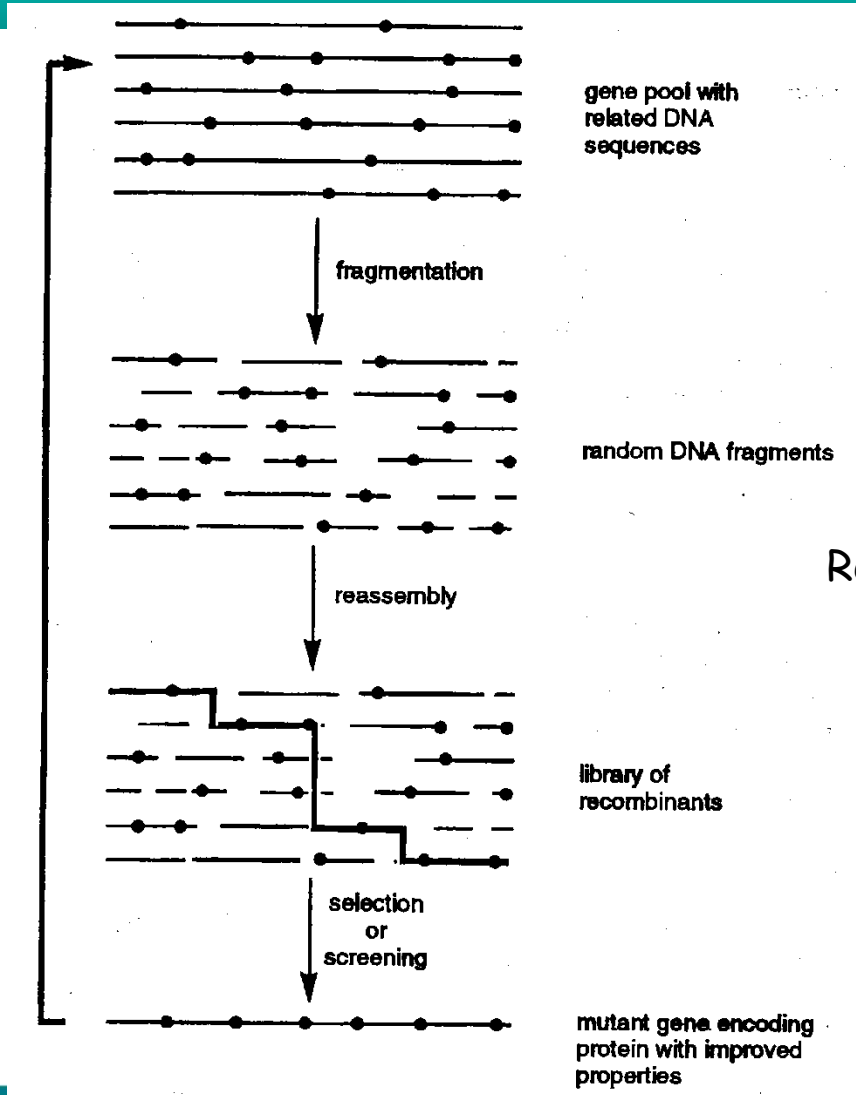
Begin with the formation of a small double-stranded DNA fragments (20 to 50 bp). DNase I will be used as an enzyme to digest randomly a pool of closely related genes containing point mutations.

A PCR-like reaction (without primers) will be performed to purify and reassembled DNA fragments into a full-length gene.

Fragments originated from dissimilar parental genes will prime one another in a process of Recombinogenic events. A standard PCR reaction will then be implemented in which reassembly mixture is used as a template with primers flanking the gene of interest.

The final amplified product will consist of a library of full-length genes containing recombined mutations from different parental genes.

DNA Shuffling

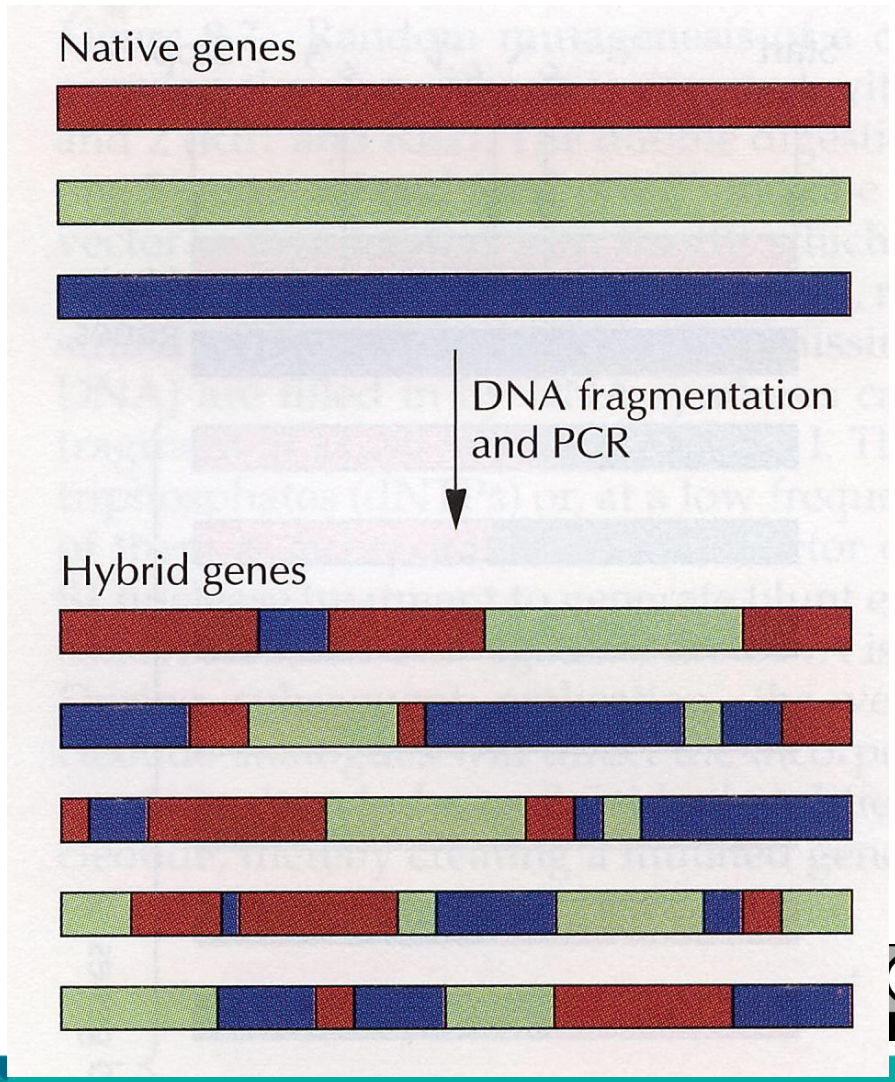


DNase I treatment (Fragmentation, 10-50 bp, Mn^{2+})

Reassembly (PCR without primers, Extension and Recombination)

PCR amplification

Family Shuffling



Several parental genes involves which come from the similar family (highly homologous) -> Family shuffling

Limitation of Directed Evolution

To be succeed:

1. Evolutionary track must be available
2. Screening method must be accessible