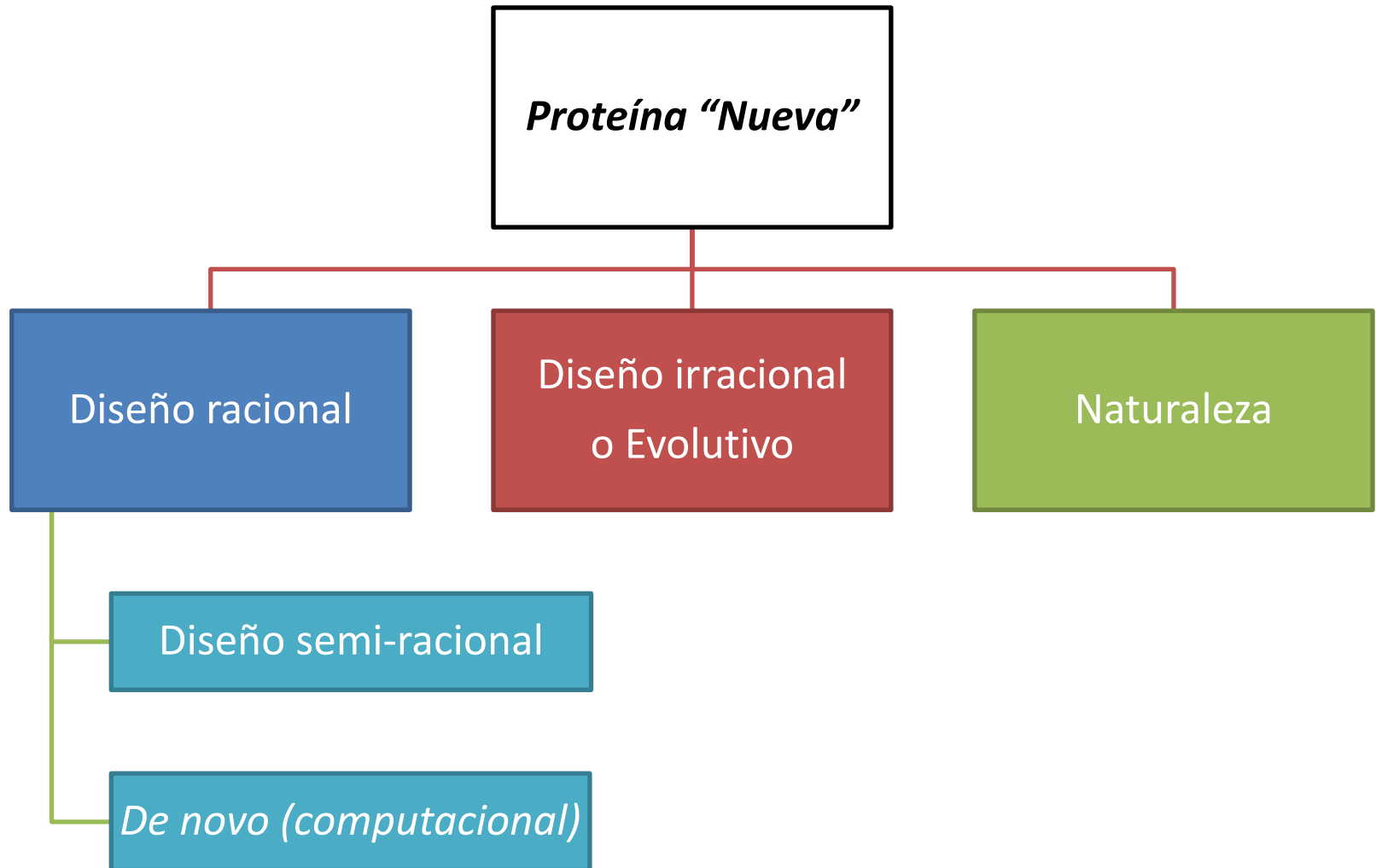


Como?



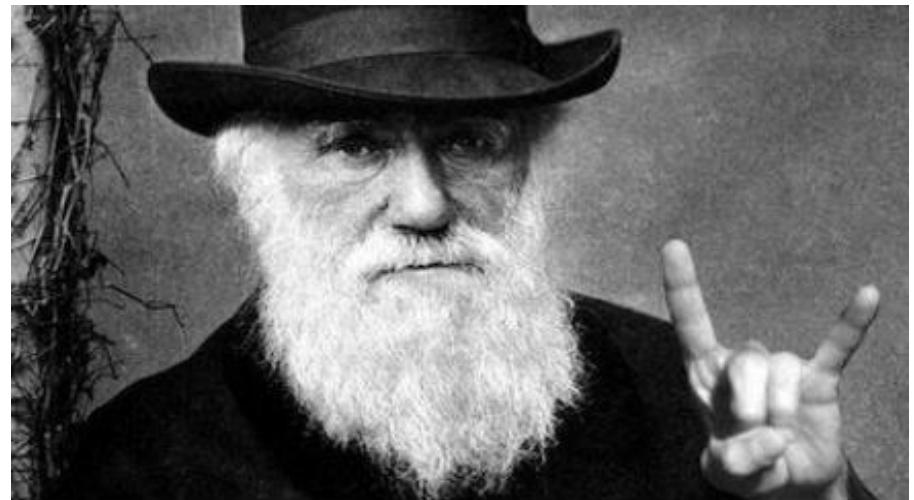
- Darwin => natural selection
- 1970 – John Maynard Smith
 - Evolution is a walk from one functional protein to another in the landscape of all possible sequences

Directed Evolution

Simula el proceso Darwiniano de evolución, mediante la combinación de mutagénesis al azar y/o técnicas de recombinación de DNA con técnicas de *screening* o selección de variantes de proteínas que hayan incorporado las propiedades deseadas

NO es necesario conocer:

- Estructura 3D
- Mecanismo de acción



- Is an **engineering** strategy used to improve protein functionality through **repeated** rounds of **mutation** and **selection**
- Around **1%** of all random mutations estimated to be beneficial
- Based on **natural evolution** processes, but in a much quicker timescale

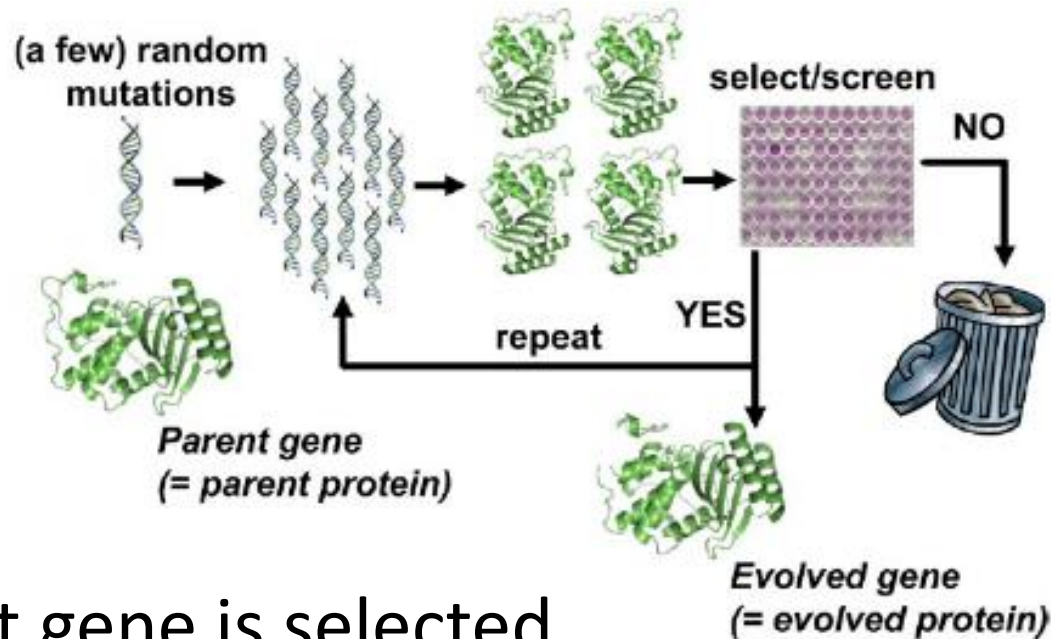
Why is it Needed?

- Biotechnology – increased demand for specific properties that don't necessarily occur naturally
- Can be used to improve existing proteins' functionality
- Can be applied as far as the ideas come – enzymes and catalysts to pharmaceuticals or crops

Successful Directed Evolution

- Desired function should be/have:
 - Physically feasible
 - Biologically or evolutionarily feasible
 - Libraries of mutants complex enough to contain rare beneficial mutations
 - Rapid screen to find desired function
- Increases understanding of protein function and evolution – disconnects protein from natural context

Basic Method



- A parent gene is selected
- Mutations/diversity are induced (mutagenesis or recombination)
- Selection/screening criteria applied
- Repeat with new parent genes selected

Directed Evolution Library

Even a large library $\rightarrow$ (10^8 independent clones)

will not exhaustively encode all possible single point mutations.

Requirements would be:

20^N independent clones $\rightarrow$ to have all possible variations in a library
(+ silent mutations)

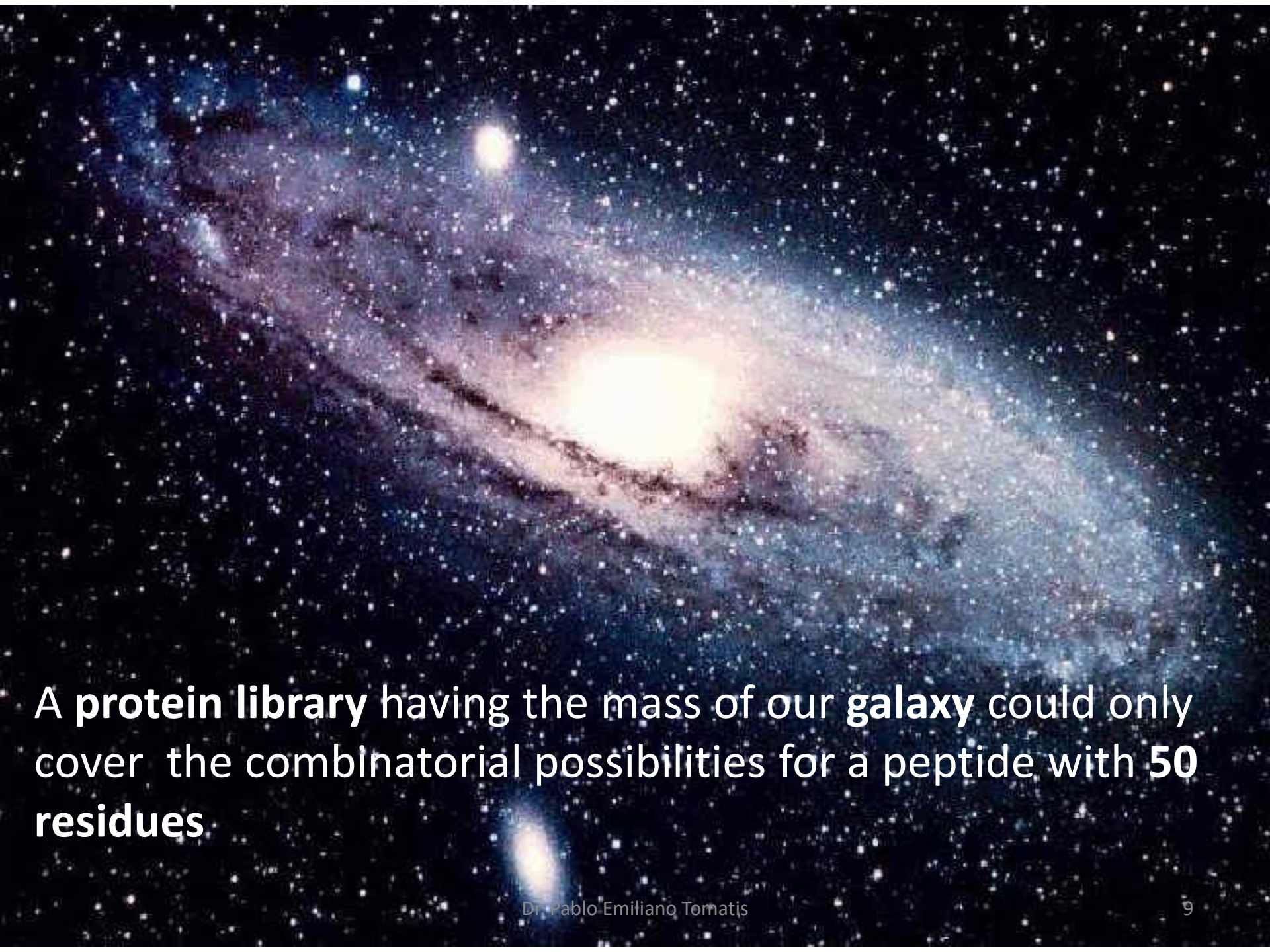
N..... number of amino acids in the protein

For a small protein: $\rightarrow$ Hen egg-white Lysozyme (129 aa; 14.6 kDa)

$\rightarrow$ library with 20^{129} (7×10^{168}) independent clones

Consequence $\rightarrow$ not all modifications possible

$\rightarrow$ modifications just along an evolutionary path !!!!



A **protein library** having the mass of our **galaxy** could only cover the combinatorial possibilities for a peptide with **50 residues**

Limitation of Directed Evolution

1. Evolutionary path must exist - > to be successful
2. Selection/Screening method must be available
 - > You get (exactly) what you ask for!!!
 - > need to be done in -> High throughput !!!

ORGANISMO HOSPEDADOR

- La expresión funcional heteróloga de los genes que codifican la proteína de interés es el primer requerimiento para llevar a cabo un experimento de evolución dirigida
- organismos hospedadores más utilizados
 - *E. coli* y
 - *S. cerevisiae*
 - *Bacillus subtilis*
- Recientemente ;***Pichia pastoris***, (recuperación de los genes mutados y baja eficiencia de transformación limita su aplicación práctica)

ORGANISMO HOSPEDADOR

- ***E. coli* es el organismo hospedador más frecuentemente utilizado.**
 - eficiencia de transformación
 - estabilidad del DNA plasmídico
 - tasa de crecimiento,
- **Expresión bacteriana de genes eucarióticos puede estar limitada** por la diferencia entre los sistemas de expresión bacterianos y los nativos.
 - diferente uso de codones
 - modificaciones postraduccionales
 - procesamiento proteolítico
 - Glicosilaciones

pueden causar bajos niveles de expresión, o un plegamiento inapropiado, con la consiguiente formación de cuerpos de inclusión

- ***S. cerevisiae***
 - glicosilar y secretar las proteínas al medio extracelular
 - elevada eficiencia de transformación

Type of mutation - Fitness of mutants

Type of mutations:

- ⇒ Beneficial mutations (good)
- ⇒ Neutral mutations
- ⇒ Deleterious mutations (bad)

⇒ Beneficial mutations are diluted with neutral and deleterious ones

!!! Keep the number of mutations low per cycle

-> improve fitness of mutants!!!

GENERACIÓN DE DIVERSIDAD

- **Non-recombinative methods (asexual):**
 - > Oligonucleotide Directed Mutagenesis (random saturation mutagenesis)
 - > Chemical Mutagenesis, Bacterial Mutator Strains
 - > Error-prone PCR
- **Recombinative methods (sexual)**
 - > Mimic nature's recombination strategy

Used for: Elimination of neutral and deleterious mutations

In vitro

- > **DNA (family) shuffling**
- > Random priming recombination, Staggered extension process (StEP), RACHITT
- > ITCHY, SCRATCHY, SHIPREC, in vivo Exon Shuffling

In vivo

- > Recombination (Yeast, *S. Cerevisiae*)

Random Mutagenesis

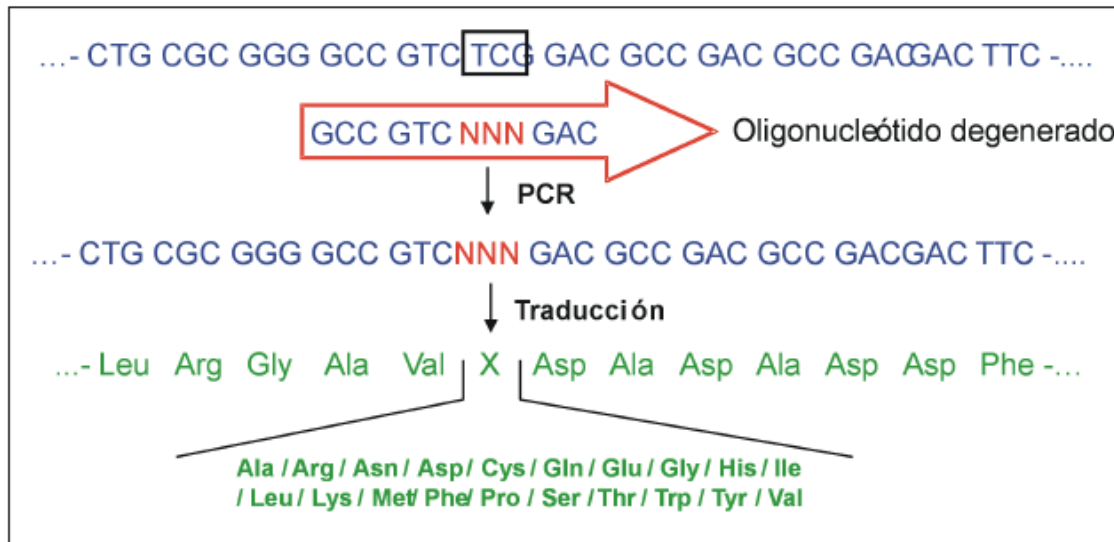
- Old traditional method
- Point mutation based – **error prone PCR**
- Frequency of beneficial mutations very low
- Multiple mutations virtually impossible to come out positive

Random Mutagenesis

- **oligos degenerados**

This approach has two advantages:

- The researcher does not have to know which amino acid is important or how to change it.
- The range of mutants produced is broad with lots of interesting possibilities.



Random Mutagenesis

- **oligos degenerados**

This approach has two advantages:

- The researcher does not have to know which amino acid is important or how to change it.
- The range of mutants produced is broad with lots of interesting possibilities.

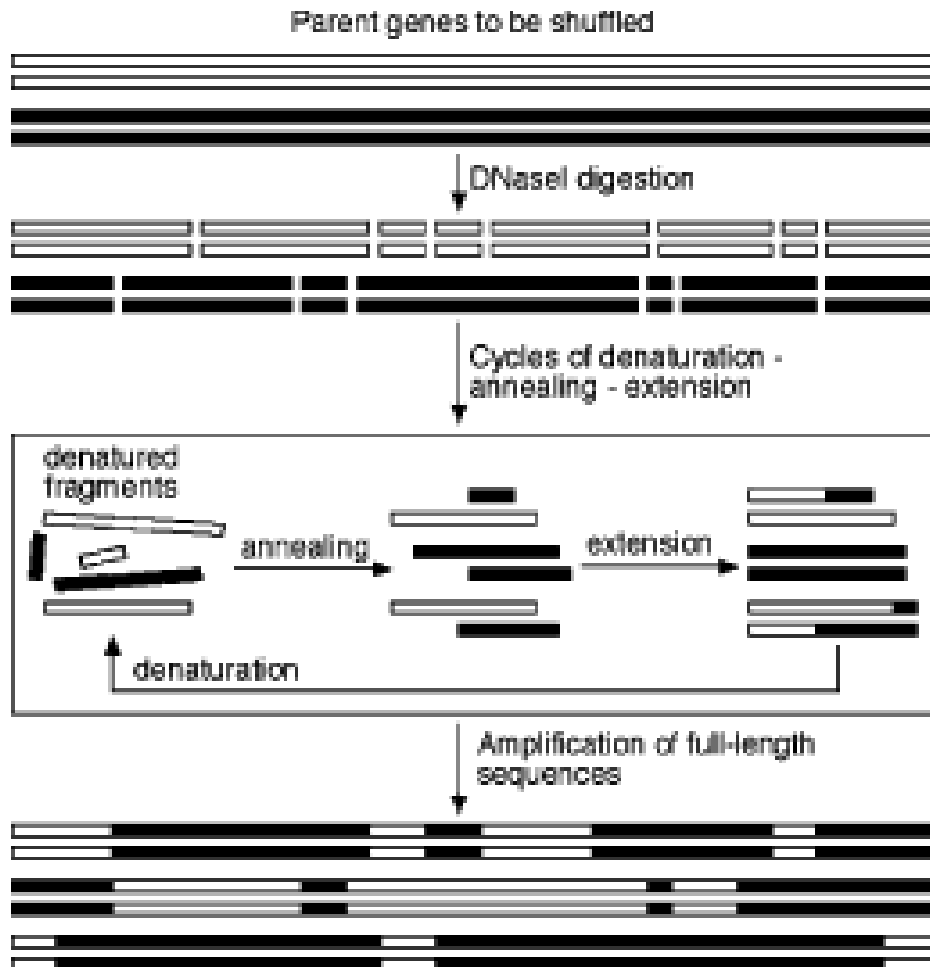
- **Random Insertion/Deletion Mutagenesis**

- Delete a small number of nucleotides at random positions along the gene.
- Insert either specific or random sequences into that position.

DNA Shuffling

- Recombination used to create chimeric sequences containing multiple beneficial mutations
- “Family shuffling” of homologous genes or “*molecular breeding*”
- “Synthetic shuffling” – oligonucleotides combined to create full-length genes
- Whole-genome shuffling – accelerated phenotypic improvements
- *Drawback – high homology required*

DNA Shuffling In vitro homologous recombination



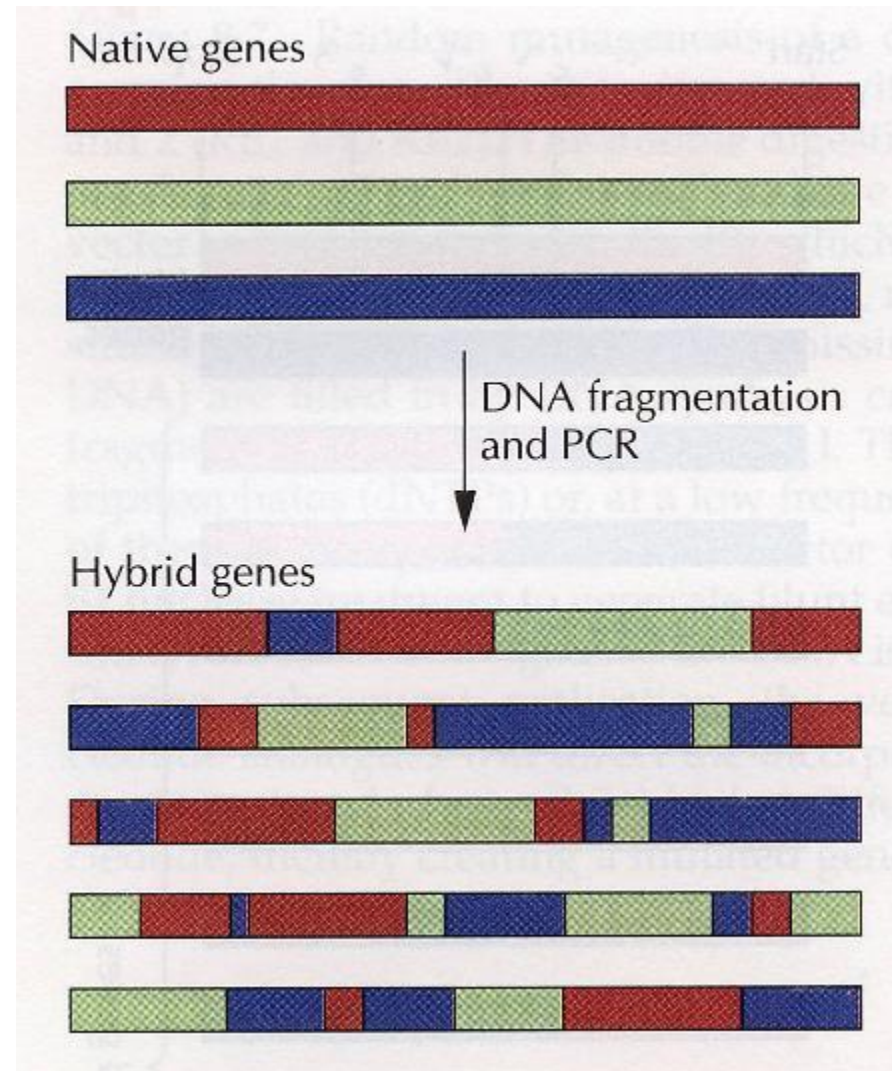
- 3.7 crossovers per 2.1 kb gene (1%) with a low mutagenesis rate (0,01%)
- successfully to recombine parents with only 63% DNA sequence identity

DNA Family Shuffling

Family Shuffling

Genes coming from the same
gene family -> highly
homologous

-> Family shuffling



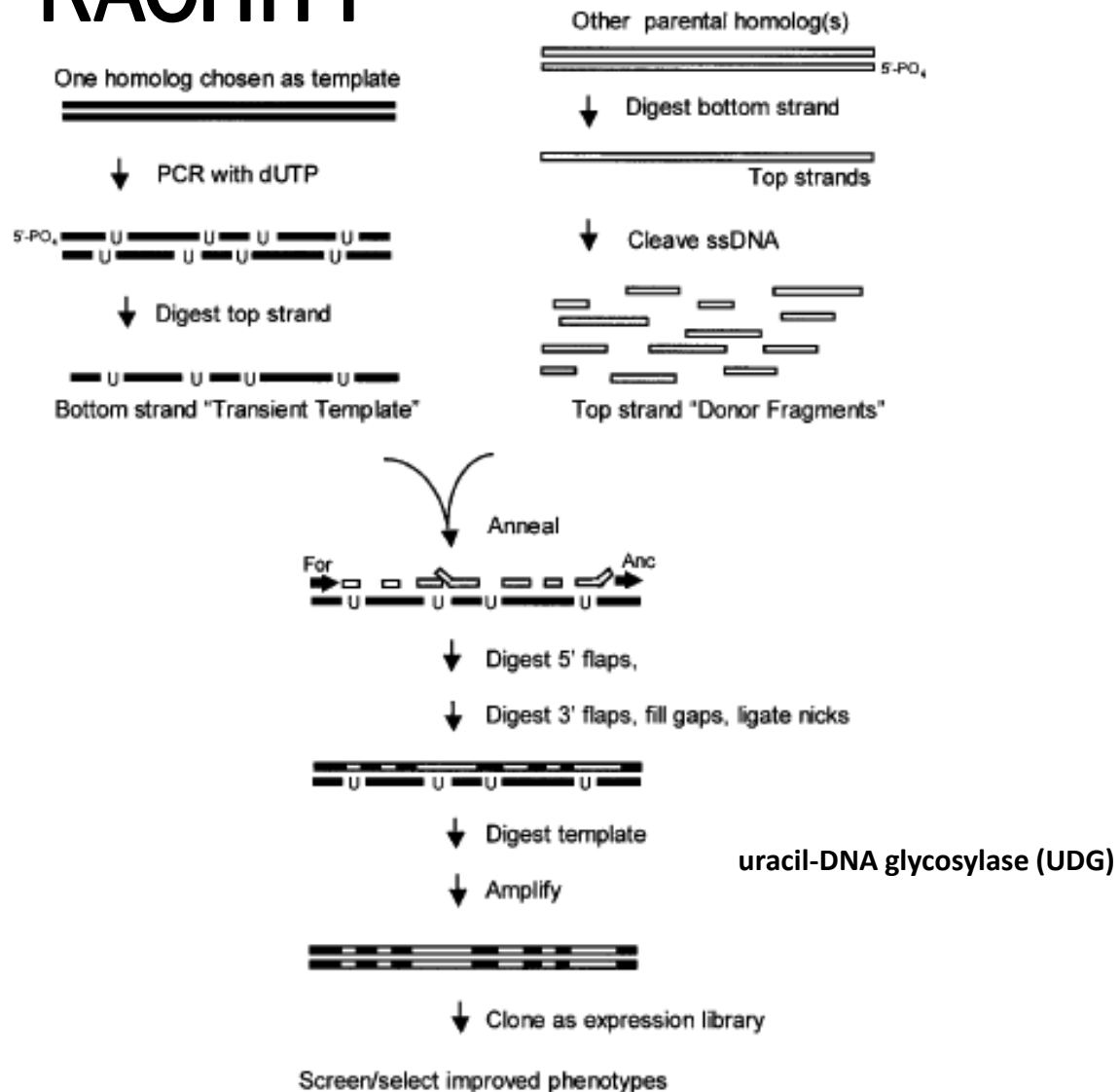
RACHITT

- Random Chimeragenesis on Transient Templates
- Small DNA fragments hybridized on a scaffold to create a chimeric DNA fragment
- Incorporates low-homology segments
 - (generación aleatoria de secuencias quiméricas sobre un templado transitorio)
 - En este método, se generan fragmentos aleatorios provenientes de 1 sola hebra de ADN de las dos posibles de cada uno de los genes homólogos.
 - Estos fragmentos son reensamblados posteriormente usando como templado la otra hebra complementaria que no se fragmentó.
 - Las regiones que no ensamblaron se remueven por digestión con una exonucleasa, y los huecos generados se completan con una DNA polimerasa para finalmente ligarse.
 - El templado se remueve con endonucleasa V de forma que quede solo la hebra que contiene los fragmentos ligados y a partir de esta se genera una doble hebra de ADN.
 - Este método genera un gran número de recombinaciones (Coco, 2003; Farinas et al., 2001). La desventaja del método es el enorme número de pasos requeridos

RACHITT

Gene Family Shuffling by Random Chimeragenesis on Transient Templates

- average 12 crossovers per gene in a single round
- limited due to the complexity



ADO

- Assembly of Designed Oligonucleotides
- Nonconserved regions with conserved parts as linkers
- PCR with dsDNA without primers
- Full length genes in expression vector
- Creates large diversity of active variants without codon bias for parental genes

MURA

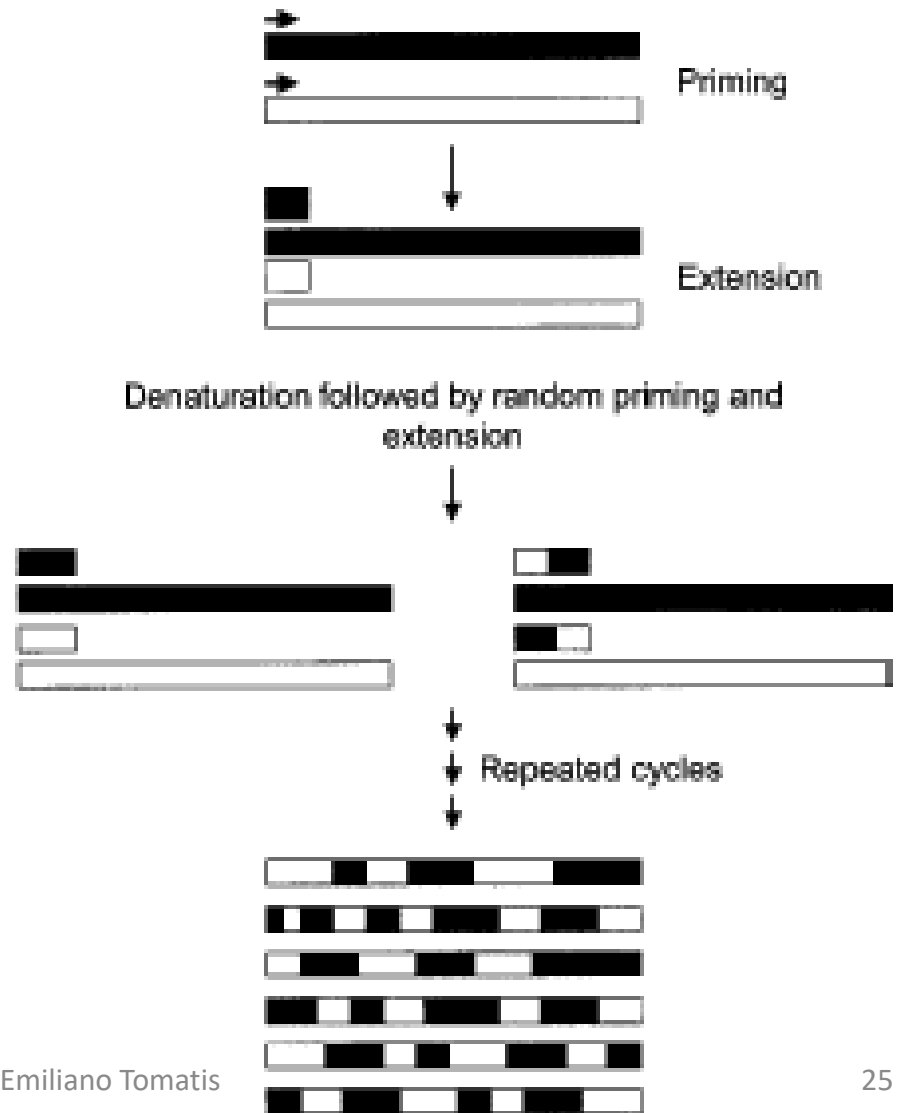
- Random fragmentation of parental gene
- Reassembled with unidirectional primers for specific restriction site
- Generates N-terminally truncated DNA shuffled libraries

Diseño irracional o Evolutivo

Staggered Extension Process (StEP) In Vitro Recombination

based on cross hybridization of growing
gene fragments during polymerase-
catalyzed primer extension

It is much easier than DNA shuffling



Diseño irracional o Evolutivo

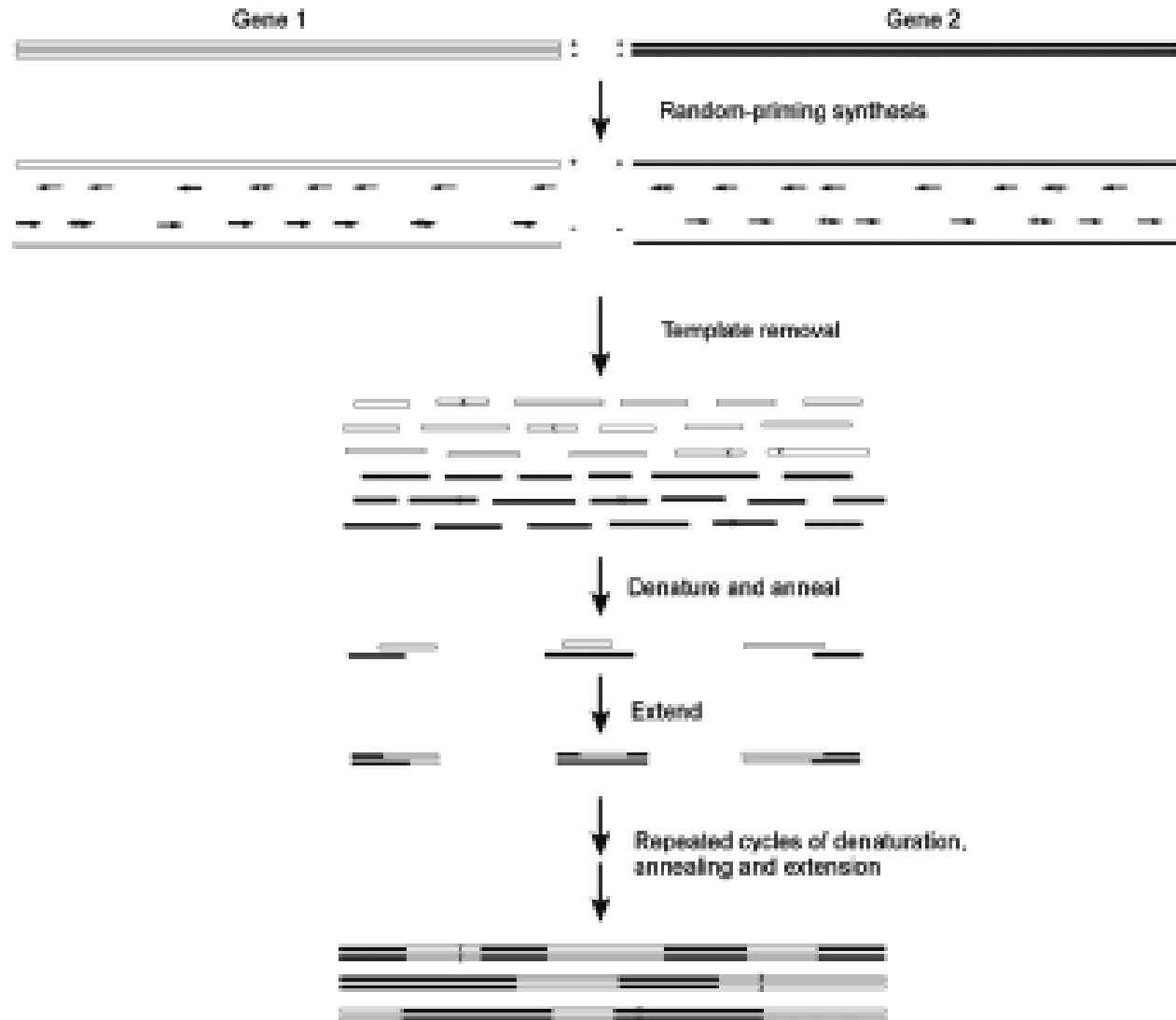
Staggered Extension Process (StEP) In Vitro Recombination

- La StEP se basa en un protocolo de PCR modificado, donde se utilizan dos o más genes homólogos como templados, un grupo de cebadores (específicos para cada gen) y tiempos cortos para las reacciones de hibridación y extensión.
- Los pequeños oligómeros generados en una primera reacción se disocian del templado original y cambian a otro templado aleatoriamente extendiendo otro gen nuevamente en tiempos cortos de hibridación y extensión.
- Varias repeticiones permiten la recombinación y formación de genes completos, con los que se construye una genoteca de variantes y después se someten al proceso de selección

Random priming *in vitro* recombination (RPR)

RPR has several potential advantages vs. *DNA shuffling*:

- no use of DNase
- random-priming DNA synthesis is independent of the length of the DNA template
- It can be used single-stranded DNA or RNA templates
- mutations introduced by misincorporation and mispriming can further increase the sequence diversity



Random priming *in vitro* recombination (RPR)

Recombinación por cebado al azar

- A diferencia de la StEP, en la recombinación por cebado al azar se utilizan cebadores no específicos con introducción de mutaciones puntuales adicionales los cuales pueden hibridar con los genes homólogos de interés para generar fragmentos aleatorios de ADN.
- Estos fragmentos sirven como cebadores de su propio gen y de los otros genes, generando así un gen completo del mismo modo en que ocurre en las técnicas descritas para StEP.
- La frecuencia de mutación se puede modificar variando la concentración y tamaño de los cebadores, el tiempo de reacción y la temperatura de hibridación

Non-homologous Recombination

- Creation of **new protein folds**
- Structures not present in nature – useful for evolution of multifunctional proteins
- Incremental truncation for the creation of hybrid enzyme (**ITCHY**) – two genes in expression vector with unique restriction sites, blunt end digestion, ligated -> **SCRATCHY**
- Nonhomologous random recombination – potentially higher flexibility in fragment size and crossover frequency

The Creation of ITCHY Hybrid Protein Libraries

An ITCHY library created from a single gene consists of genes with internal deletions and duplications.

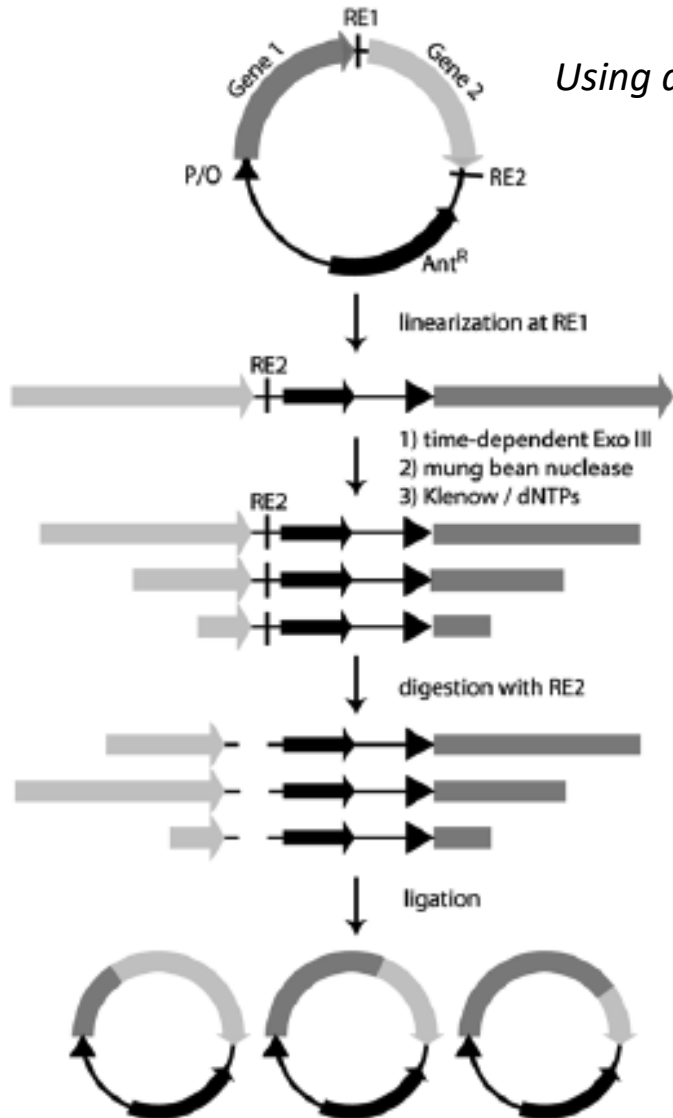
An ITCHY library created between two different genes consists of gene fusions created in a DNA-homology independent fashion.

- Truncado gradual para crear enzimas híbridas (“Incremental truncation for the creation of hybrid enzymes”, ITCHY).
- La ventaja de este procedimiento es que **no requiere que los genes sean homólogos** para realizar la recombinación.
- En esta técnica, dos genes seleccionados se cortan de forma gradual con exonucleasa III en condiciones controladas que permitan obtener genes truncados con diferencia de una sola base. Los extremos 5' de un gen y los extremos 3' del otro gen se ligan de forma aleatoria para generar una genoteca de secuencias quiméricas que incluye todas las posibles combinaciones de fragmentos de distintos tamaños de uno y otro gen

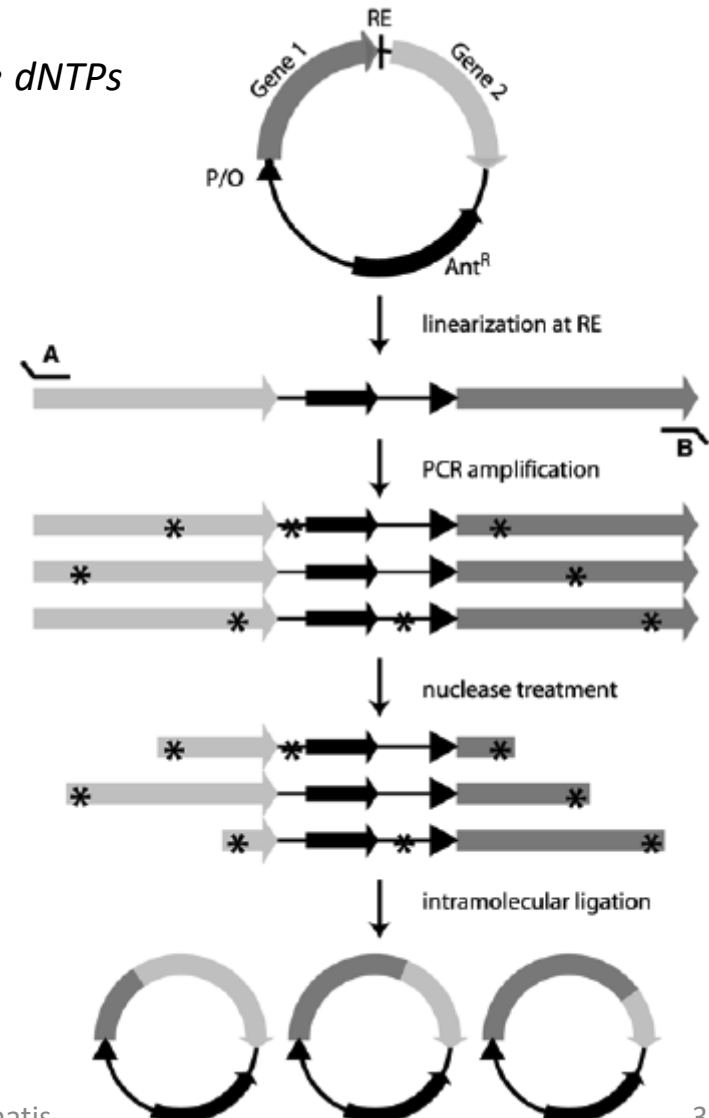
Diseño irracional o Evolutivo

The Creation of ITCHY Hybrid Protein Libraries

Using α -phosphorothioate dNTPs



*Blunt ends
generation*

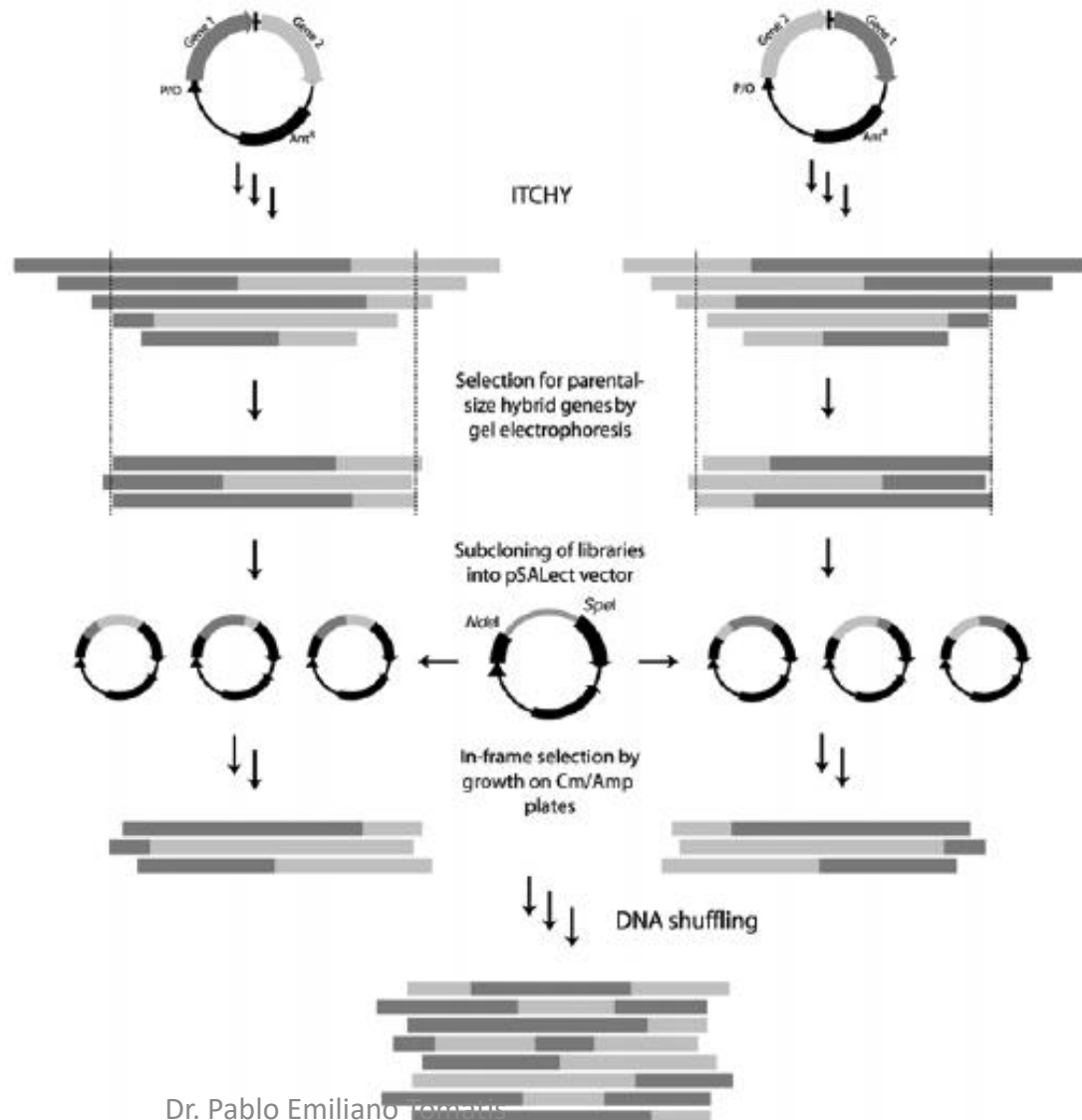


Diseño irracional o Evolutivo

Preparation of SCRATCHY Hybrid Protein Libraries

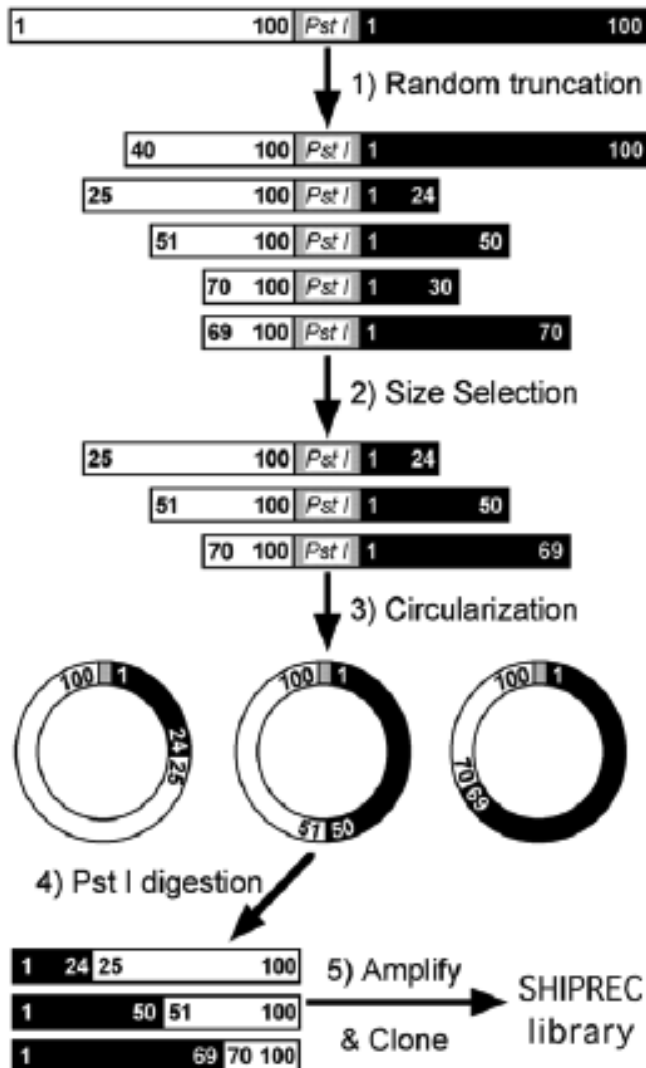
Size- and In-Frame Selection of Nucleic Acid Sequences

*A combination between
ITCHY and DNA shuffling*



Diseño irracional o Evolutivo

Sequence Homology-Independent Protein Recombination (SHIPREC)



chimeras are fused via a linker sequence containing several restriction sites

DNase I and S1 nuclease

DNA corresponding to the length of the parental genes is isolated and subsequently circularized

crossovers occur at structurally related sites

to generate all possible single-crossover chimeras, SHIPREC must be performed twice starting with both possible parental gene fusions, e.g., A-B and B-A.

Exon Shuffling

- Similar to natural splicing of exons
- Chimeric oligos mixed together, controlling combination of which exons to be spliced
- Protein pharmaceuticals based on natural human genes – less immune response

Generación de células artificiales

- “*In Vitro* Compartmentalization”, IVC.
- Se genera previamente una genoteca de variantes y las células artificiales son generadas con una emulsión agua-aceite, cada gota contiene un gen con toda la maquinaria molecular necesaria para expresarlo.

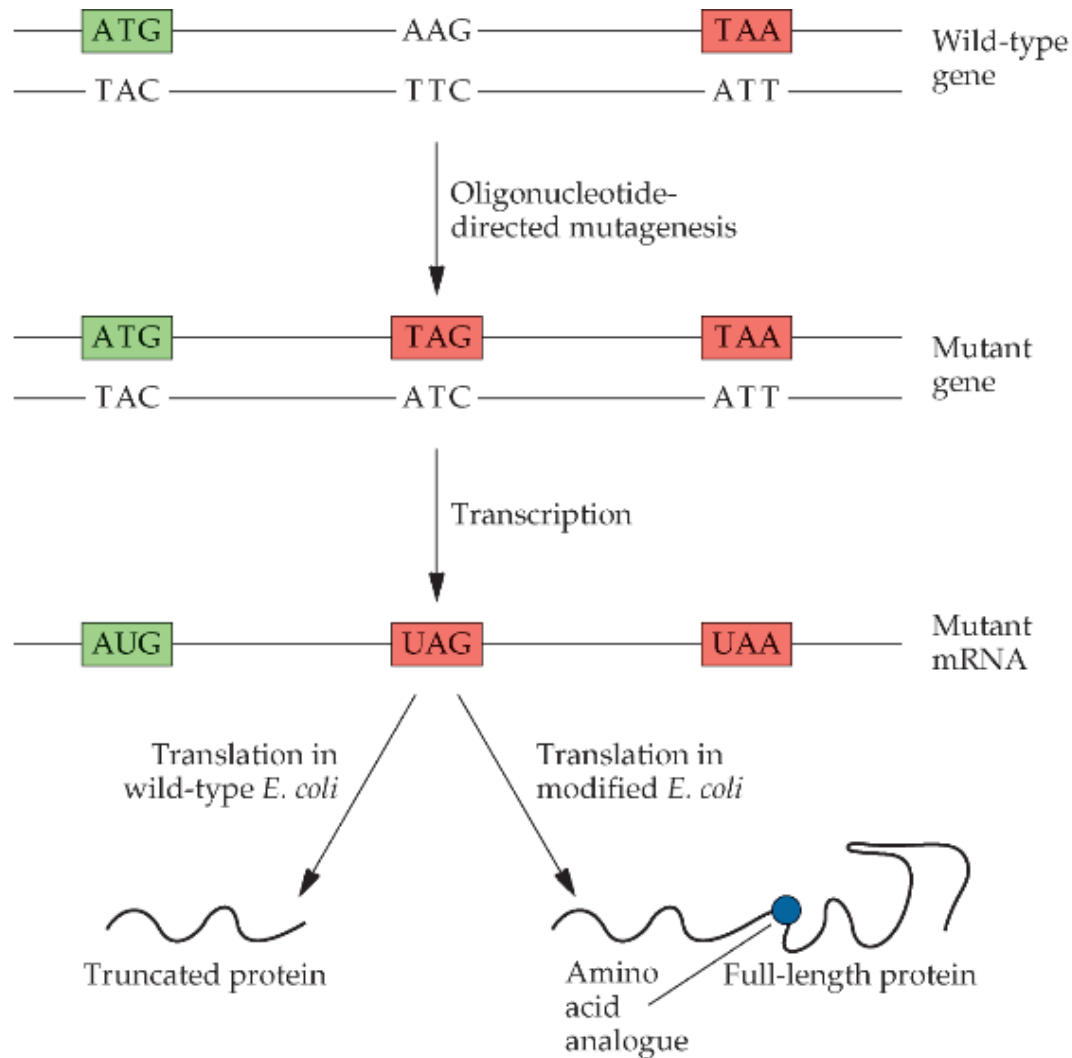
Nuevas metodologías

- Se siguen desarrollando nuevas metodologías considerando las limitaciones que presentan cada una y la necesidad de desarrollar mejores técnicas de recombinación.
- Muchas son derivadas o combinan las anteriormente descritas.

Unusual Amino Acids

- Any protein can be altered by substituting one amino acid for another using directed mutagenesis.
- However, it is limited to the 20 amino acids.
- One way to increase the diversity of the proteins is to introduced synthetic amino acids with unique side chains at specific sites.
- It is possible using modified tRNAs and tyrosine-tRNA synthetase to introduce unusual amino acid into mutagenesis amber stop codon.
- An attempt to alter the final protein activity in ways that aren't normally possible.

Mutant Proteins with Unusual Amino Acids



Metodos de aislamiento y selección de mutantes

- Uno de los mayores retos en evolución dirigida es el contar con un sistema de selección adecuado y eficiente.
- Las características más importantes del método son:
la rapidez, exactitud y que esté dirigido a la identificación del biocatalizador esperado dentro de la numerosa genoteca de mutantes
- 2 enfoques:
 - Selección
 - Tamizado o screening,

pueden combinarse o usarse consecutivamente.

Selección: solo los clones deseados aparecen

Screening: todos los miembros de la librería están presentes, pero se eligen solo los mejores

El mayor inconveniente del screening es que el tamaño de la librería que puede explorarse es limitado

Selección:

utilizado tradicionalmente para:

- enriquecer ciertos microorganismos,
- identificación de los productos de clonación deseados.

se puede ver a simple vista la característica de interés que se desea seleccionar.

Permite análisis más rápido de una genoteca numerosa

Aplicación limitada debido a la falta de sistemas de selección cuantitativos y aplicables.

Screening:

Tamizado o Exploracion:

permiten separar las variantes deseadas de acuerdo a sus propiedades.

permiten cuantificar la propiedad deseada con técnicas analíticas reproducibles, automatizadas y en escala de volúmenes reducidos.

Métodos colorimétricos y fluorométricos en formatos de microplaca (volúmenes de μL)

sistemas automatizados que pueden tomar cientos de muestras simultáneamente y someterlas a un tratamiento que nos permita cuantificar la actividad de la enzima que nos interesa.

reduce enormemente el tiempo dedicado a la exploración y selección de mutantes.

mejora la reproducibilidad de las mediciones y puede hacerse en escalas que resulten económicamente atractivas

Screening:



- a) Robot picador de colonias.
- b) Lector de microplacas.
- c) Robot manipulador de líquidos.
- d) Inyector automático HTP para HPLC.

Screening:

Desarrollos de nuevos sistemas

- electroforesis capilar en paralelo
- arreglos de termistores (componentes electrónicos cuya resistencia cambia con la temperatura)
- termografía de infrarrojo

Recientemente se han automatizado también tecnologías de espectrometría de masas

“Electrospray Ionization Mass Spectrometry” (ESI-MS) y

“ Matriz Assisted Laser desorption-ionization Mass Spectrometry” (MALDI-MS)

para la identificación de intermediarios, sustratos y productos de bajo peso molecular en una mezcla de reacción

Screening:

ensayos cuantitativos rápidos basados en **indicadores de pH**.

Ej: lipasas y esterases; hidrólisis de un éster produce un equivalente químico del ácido correspondiente que a su vez produce una disminución del pH.

detectada espectrofluorométricamente con un compuesto sensible a cambios de pH.

es posible inmovilizar el colorante para seguir una reacción catalizada

Otra alternativa para los ensayos de aislamiento es el uso de la tecnología de separación de células activadas por fluorescencia (**FACS**).

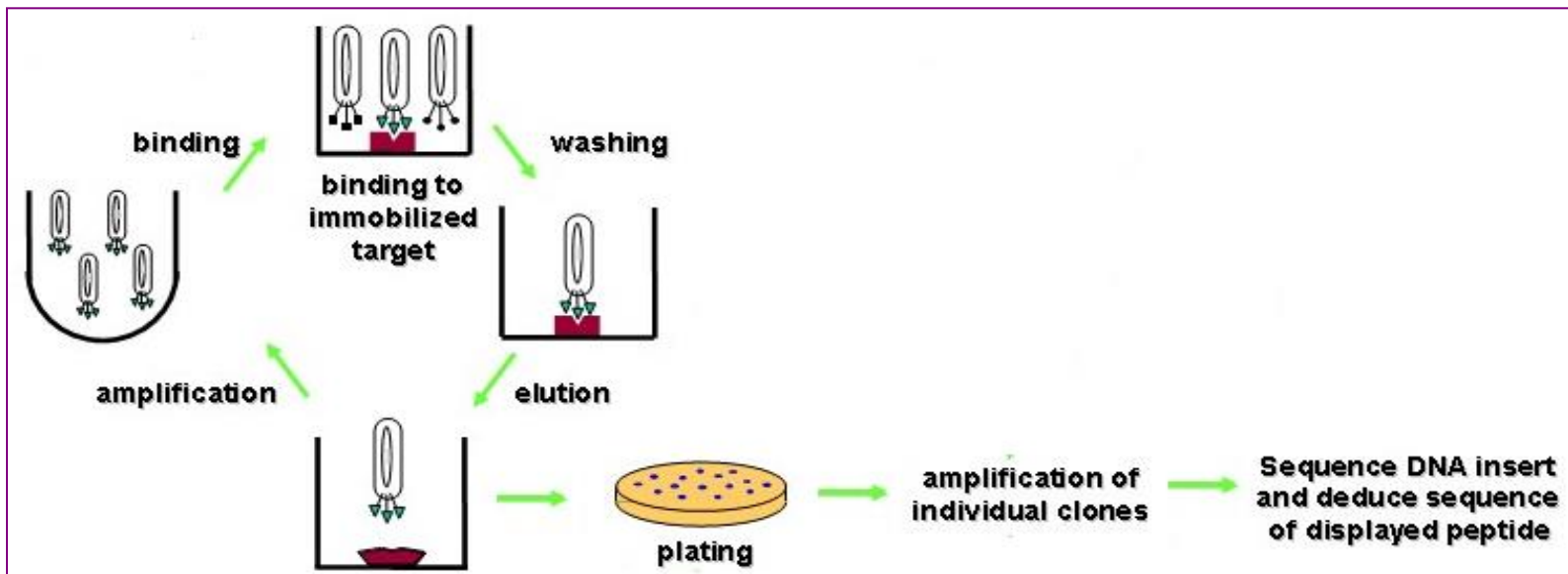
La tecnología FACS se basa en la separación física de poblaciones de células previamente reconocidas por la fluorescencia de sus marcadores.

Screening:

Phage display o presentación en fagos

fusiona la enzima de interés con una de las proteínas de la cápside de un bacteriófago, de modo que la molécula queda expuesta en la superficie del bacteriófago.

- Affinity Selection
- Enzymes Displayed
 - Trypsin, Prostate Specific Antigen, β -lactamase, Lysozyme



Screening:

Phage display o presentación en fagos (esquema en prox.)

Los bacteriófagos aislados de un medio de cultivo son posteriormente seleccionados en base a su capacidad de unirse a una fase sólida que tiene unido covalentemente el ligando de la proteína, comportándose como una matriz cromatográfica de afinidad

Las ventajas de éste método son: su eficiencia, ya que se pueden probar simultáneamente hasta 10^{12} variantes de un gen, permite seleccionar en un solo paso cromatográfico a aquellas variantes que cumplen con el requisito de afinidad por un ligando y puede ser modificado para detectar catálisis en lugar de afinidad.

Sin embargo, proteínas multiméricas no pueden organizarse en multímeros y es posible que la proteína fusionada a la proteína del fago no se pliegue adecuadamente perdiendo sus propiedades de unión y catálisis

Screening:

Cell free display systems

exploit in vitro protein translation or cell free translation.

These methods include

- **mRNA display,**
- **ribosome display,**
- **covalent and non covalent DNA display,**
- **in vitro compartmentalization**

Diseño irracional o Evolutivo

Resumen

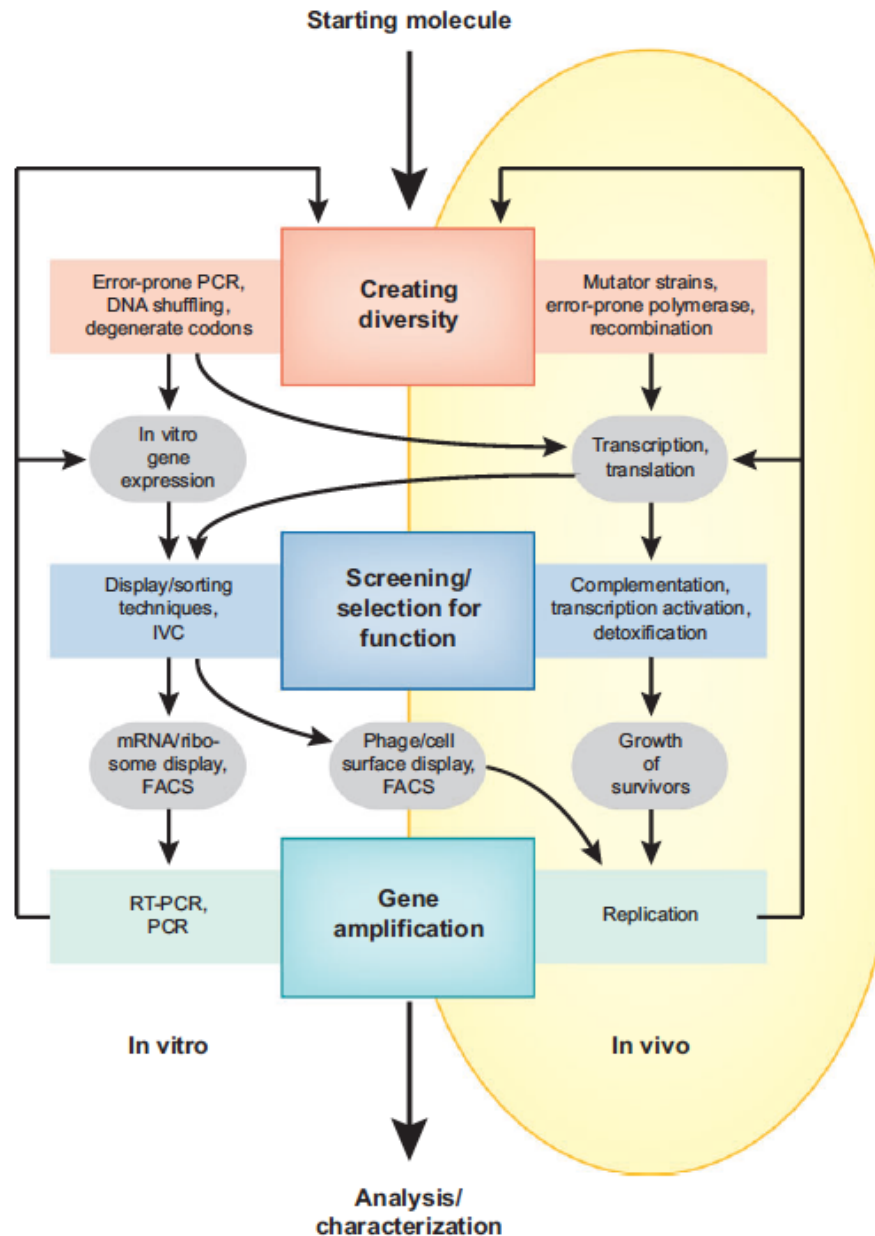


Figure 1

Schematic overview of the principal processes, strategies, and techniques of directed evolution. Today, numerous experimental methods are available to perform the fundamental processes of true Darwinian evolution (*central boxes*) in the laboratory, either in vivo within microorganisms or entirely in vitro in the test tube. Arrows indicate possible routes for connecting individual evolutionary steps. Abbreviations: PCR, polymerase chain reaction; RT-PCR, reverse transcription PCR; IVC, in vitro compartmentalization; FACS, fluorescence-activated cell sorting.

Ejemplos: Directed Evolution in Action

- improve polymerases, nucleases, transposases, integrases, recombinases
- Applications in genetic engineering, functional genomics, and gene therapy
- Optimized fluorescent proteins and small-molecule probes for imaging and techniques like FRET

Diseño irracional o Evolutivo

Enzima	Modificación	Método de mutagénesis	Selección y/o aislamiento	Resultado	Observaciones	Referencia
Esterasa de <i>Pseudomonas fluorescens</i>	Especificidad de sustrato	Cepas mutagénicas	Crecimiento con indicador de pH	Mutante con el doble de actividad	Enantioselectividad baja	Bornscheuer, <i>et al.</i> , 1999
Lipasa de <i>Pseudomonas aeruginosa</i>	Incrementar enantioselectividad	ep-PCR y cepas mutagénicas	Ensayos microplacas con <i>p</i> -nitrofenil ésteres	Incremento de enantioselectividad (E) de 1 a 26.	5 mutaciones, ninguna en el sitio de unión del sustrato	Liebeton <i>et al.</i> , 2000; Reetz <i>et al.</i> , 1999; Reetz <i>et al.</i> , 2000
Esterasa de <i>Bacillus subtilis</i>	Actividad en presencia de dimetilformamida (DMF)	ep-PCR, cepas mutagénicas y barajado de ADN	Ensayos microplacas con <i>p</i> -nitrofenil ésteres	Incremento 150 veces la estabilidad en DMF	Mayor actividad sobre <i>p</i> -nitrofenil éster de palmitato que sobre éster de butirato	Arnold & Moore, 1997
Hemoperoxidasa de <i>Coprinus cinereus</i>	Estabilidad y actividad bajo condiciones de lavado	ep-PCR, cepas mutagénicas y barajado de ADN	Incubación microplacas a pH alcalino, con H ₂ O ₂ , Actividad residual	Aumento 100 veces la estabilidad en H ₂ O ₂ y 170 veces la termoestabilidad		Cherry <i>et al.</i> , 1999
Hidantoinasa de <i>Arthrobacter sp.</i>	Enantioselectividad reversa (isómero L)	ep-PCR, cepas mutagénicas, barajado de ADN	Ensayos microplacas con D y L-hidantoínas	Cambio de actividad sobre el isómero L	Incremento adicional de la actividad específica	May <i>et al.</i> , 2000
Subtilisina S41	Incrementar termoestabilidad	ep-PCR, cepas mutagénicas, barajado de ADN	Crecimiento en placa a temperaturas altas	Variante activa en el rango de 10-60°C		Miyazaki <i>et al.</i> , 2000
Esterasa de <i>Bacillus subtilis</i>	Incrementar termoestabilidad	ep-PCR, cepas mutagénicas, barajado de ADN	Crecimiento en placa a temperaturas altas	Incremento de 14°C en la termoestabilidad		Giver <i>et al.</i> , 1998
Glutarilacilasa de <i>Pseudomonas SY77</i>	Incrementar la actividad de adipil-acilasa	ep-PCR sobre la subunidad α y cepas mutagénicas	Crecimiento en placa en adipil-leucina	Incremento de la relación V_{max}/K_m para adipil-7-ADCA		Sio <i>et al.</i> , 2002
Varias subtilisinas	Varios parámetros simultáneamente	Barajado de genes	Ensayos microplacas	Se obtienen variantes quiméricas muy activas	Se genera mucha diversidad	Ness <i>et al.</i> , 1999
Varias fitoeno desaturasas y licopenoclasas	Generación de nuevos carotenoides	ep-PCR, cepas mutagénicas, barajado de ADN	Crecimiento en placa, identificación visual	Carotenoides nuevos (torolueno)	Se evolucionó la vía de biosíntesis completa	Schmidt-Dannert <i>et al.</i> , 2000

Mis ejemplos

Directed evolution of Metallo beta Lactamases (UNR)
GPCR in *E coli* and in Yeast (UZH)

Videos cortos

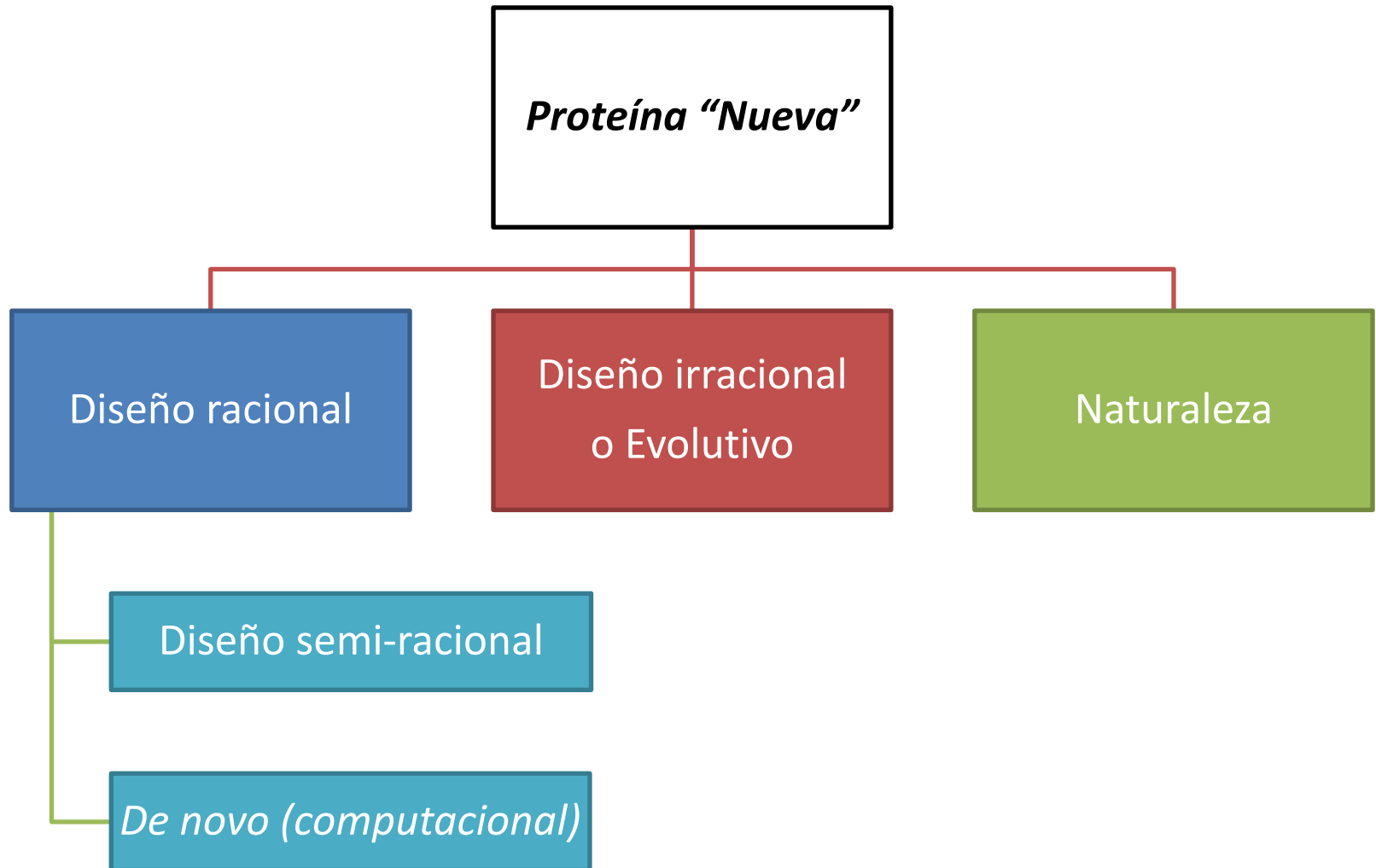
TED talk

https://www.youtube.com/watch?v=BTB_HPL2r88

Frances Arnold: New enzymes by evolution

<https://www.youtube.com/watch?v=05r-FLGtsEQ>

Como?



Semi-rational: a combination

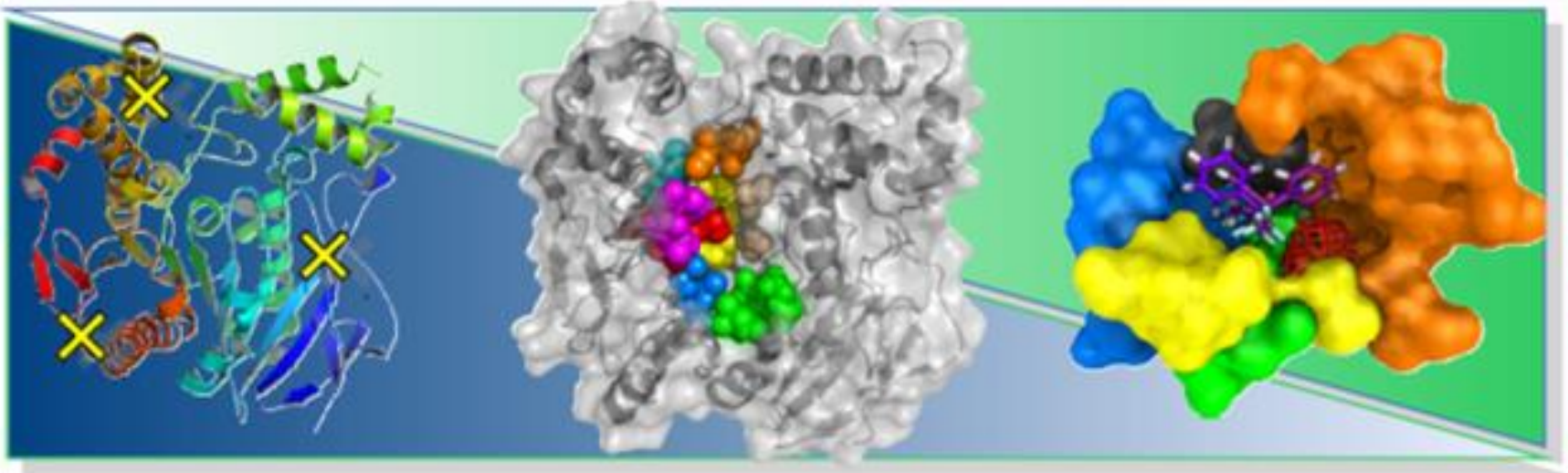
- Rational design with directed evolution
- Success depends on ability to predict fitness of a sequence
- Computationally demanding
- Focuses library diversity for directed evolution

Directed Evolution (Random mutagenesis)

Combined Approaches

Rational Design (Distinct point mutations)

Screening effort



Information intensity

epPCR
gene shuffling
Insertion/Deletion
circular permutation

CASTing/ISM
ProSAR
consensus approach
3DM

Homology model
+ docked substrates
Chrystal structures
+ transition state analogues

Semi-rational protein engineering and design

Semi-rational design uses information about a proteins sequence, structure and function, in tandem with predictive algorithms.

to identify target amino acid residues which are most likely to influence protein function.

Mutations of these key amino acid residues **create focused libraries** of mutant proteins that are more likely to have enhanced properties

- **SEQUENCE-BASED ENZYME REDESIGN**
- **STRUCTURE-BASED ENZYME REDESIGN**
- **COMPUTATIONAL ENZYME REDESIGN**
- **COMPUTATIONAL DE NOVO ENZYME DESIGN**

Semi-rational protein engineering and design

- **STRUCTURE-BASED ENZYME REDESIGN**

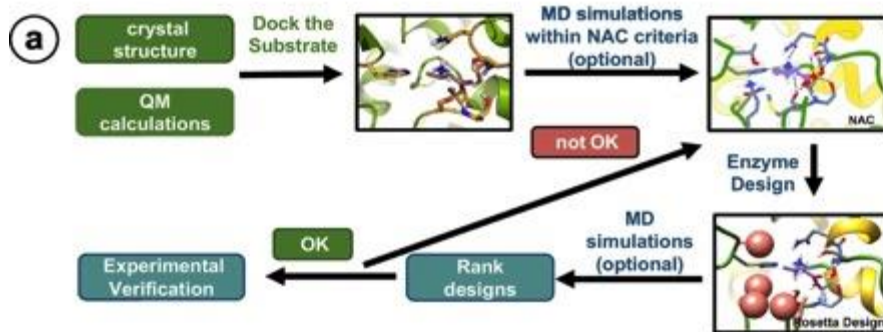
ab initio (AMBER, GROMOS)

fragment based methods (I-TASSER, ROSETTA)

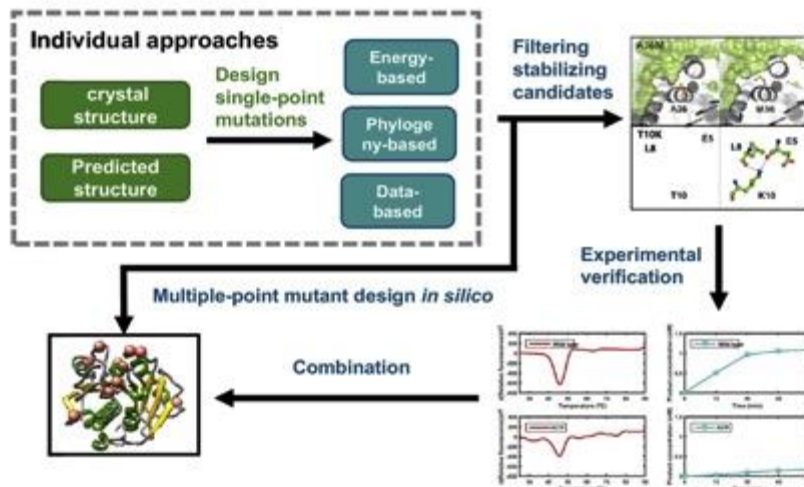
homology modeling (=, SWISS MODEL, MODELLER)

protein threading

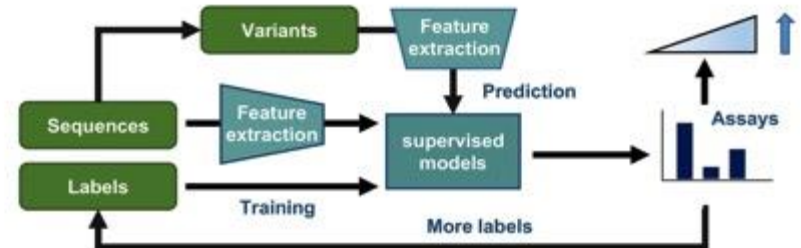
(A) Model-based Enzyme Redesign



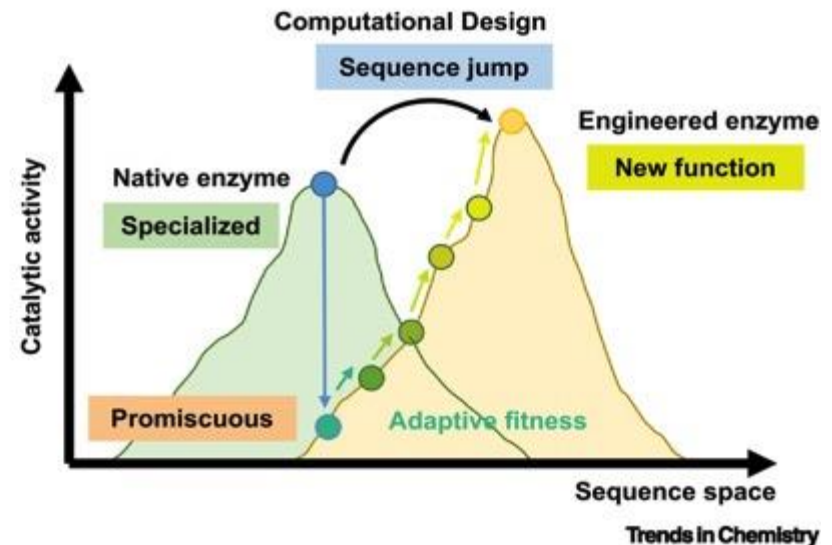
(b) Hybrid Approaches



(B) Data-based Enzyme Redesign



(C)



Para que?

Mejorar la estabilidad y/o función bajo nuevas condiciones

Temperatura, pH, disolvente (orgánico/acuoso), [sal]

Modificación de la especificidad del sustrato enzimático

Mejorar la propiedades cinéticas

Km, Vmax, Allosteric regulation, Cofactor requirements

Alterar las propiedades de unión