

# ¡Hola!

## Soy Santiago Rentería.

Puedes contactarme en [santiagorenteria25@gmail.com](mailto:santiagorenteria25@gmail.com)





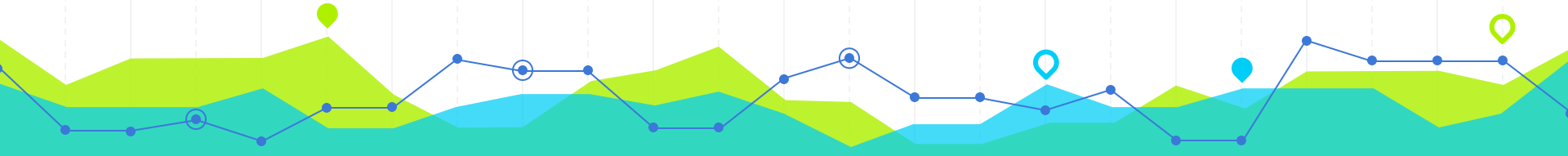
**La vida para la IA  
la IA para la vida**



**La vida para la IA**

“

*¿Hay inteligencia más allá de lo humano?*





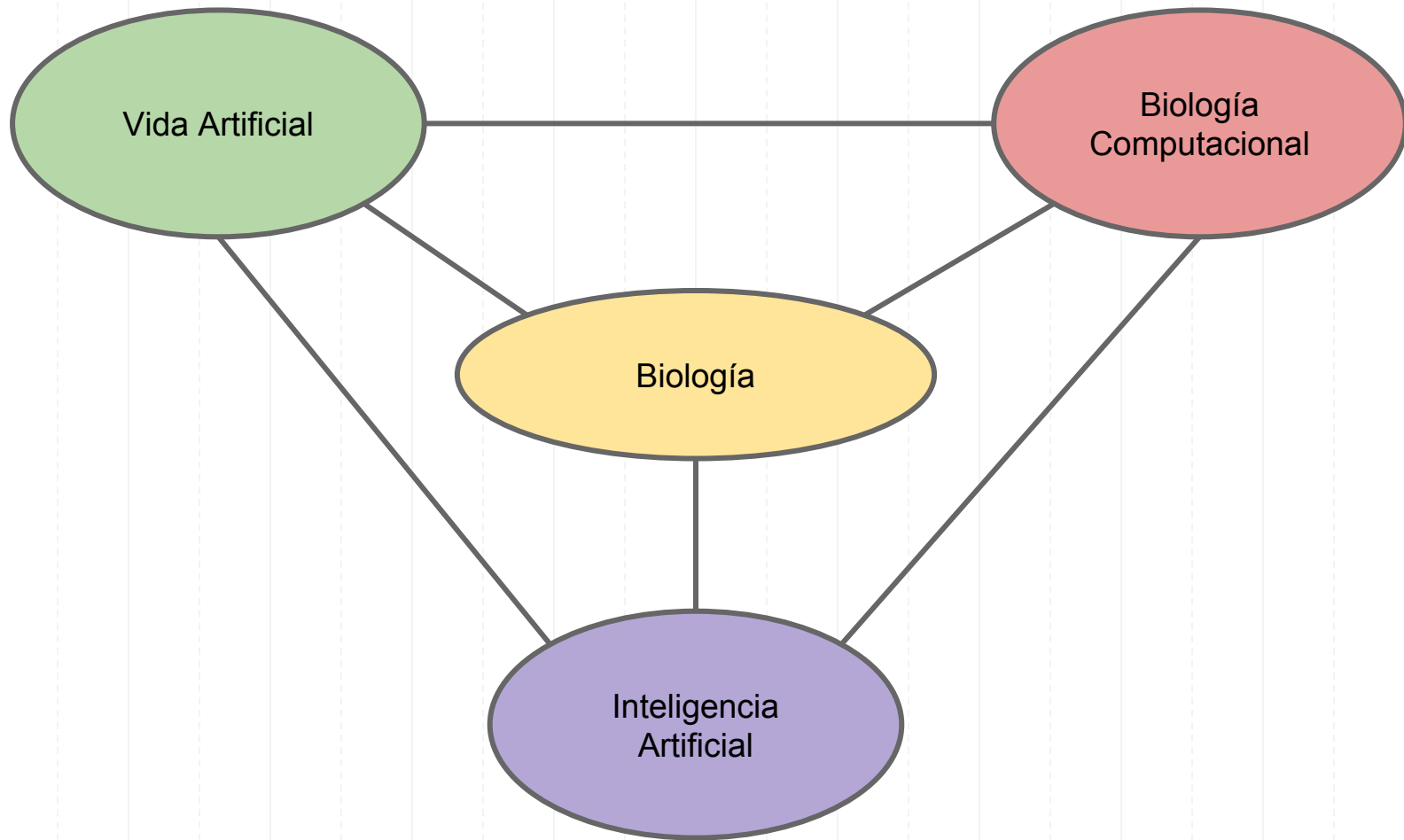
Boston Dynamics



# I.M. Intuition Machine

## The Many Tribes of ARTIFICIAL INTELLIGENCE

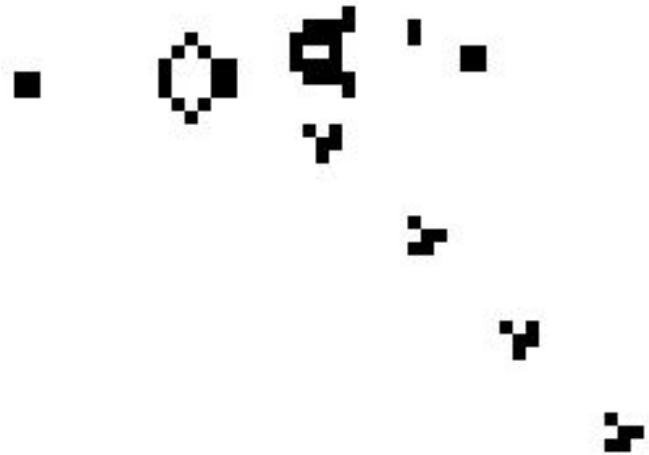




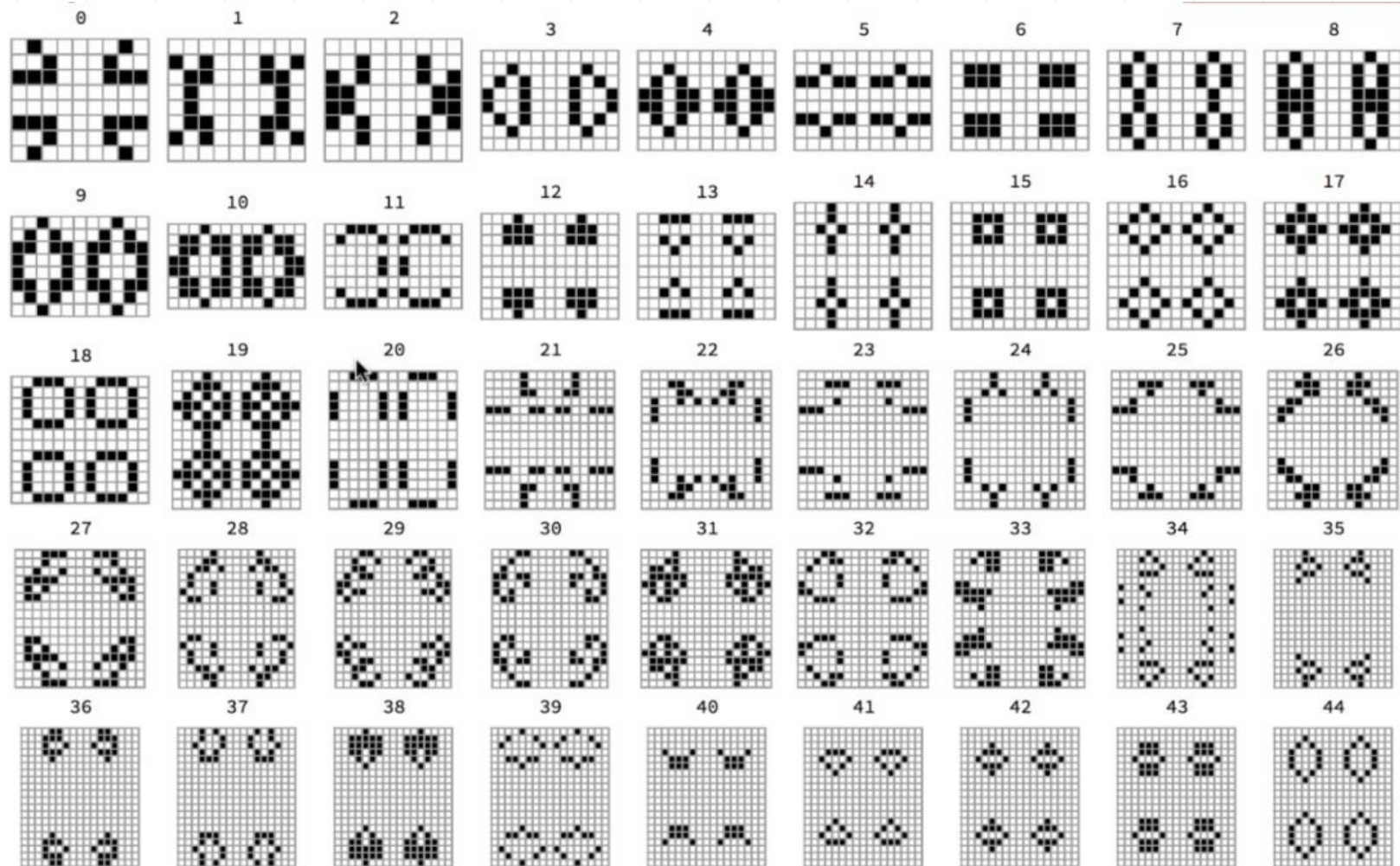
## Vida Artificial: Life as it could be

El estudio de la vida potencial más allá de su forma actual:

- Agregación Limitada por Difusión (Morfogénesis)
- Autómatas Celulares
- Fractales
- Evolución molecular
- Inteligencia de enjambre



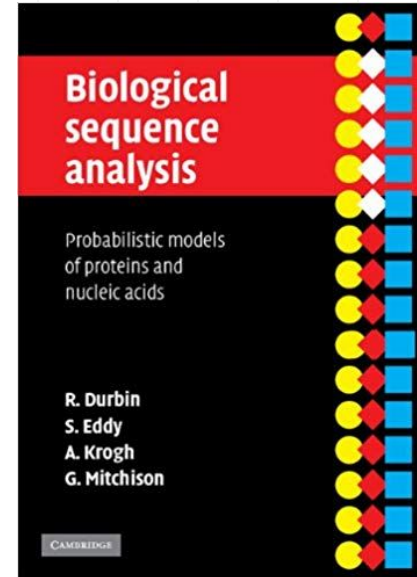




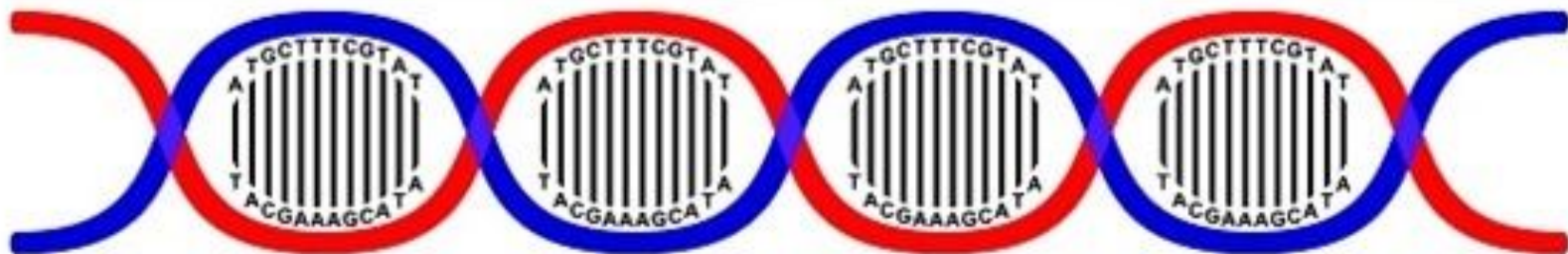
# Biología Computacional

Desarrollo de algoritmos y modelos para responder hipótesis biológicas.

- Bioinformática
  - Proteoma
  - Genoma
- Neurociencias Computacionales
  - Reconstrucción de circuitos neuronales
  - Simulación



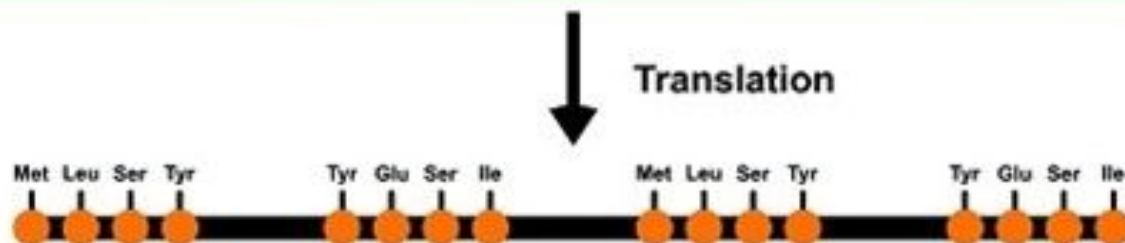
**DNA**



**RNA**



**Protein**



A C T C G C A A T A T G C T A G G C C A G C

A C T \_ \_ \_ T T A T G C T A T G C \_ \_ G C

A C T T G T C T T A T G C

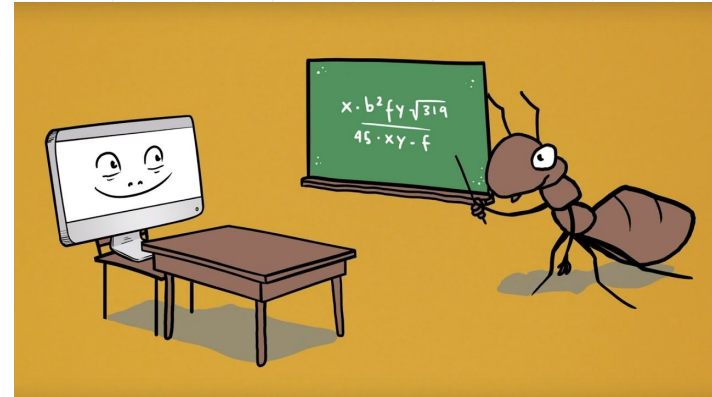
A C T \_ G \_ \_ T T A \_ \_ C



biopython

# Aplicaciones en Inteligencia Artificial

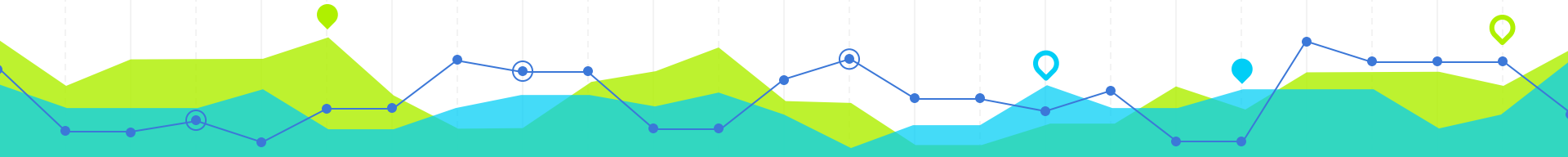
- Computación Evolutiva
- Inteligencia de Enjambre
- Sistema Inmune Artificial
- Robótica basada en el comportamiento
- Sistemas Celulares (CA)

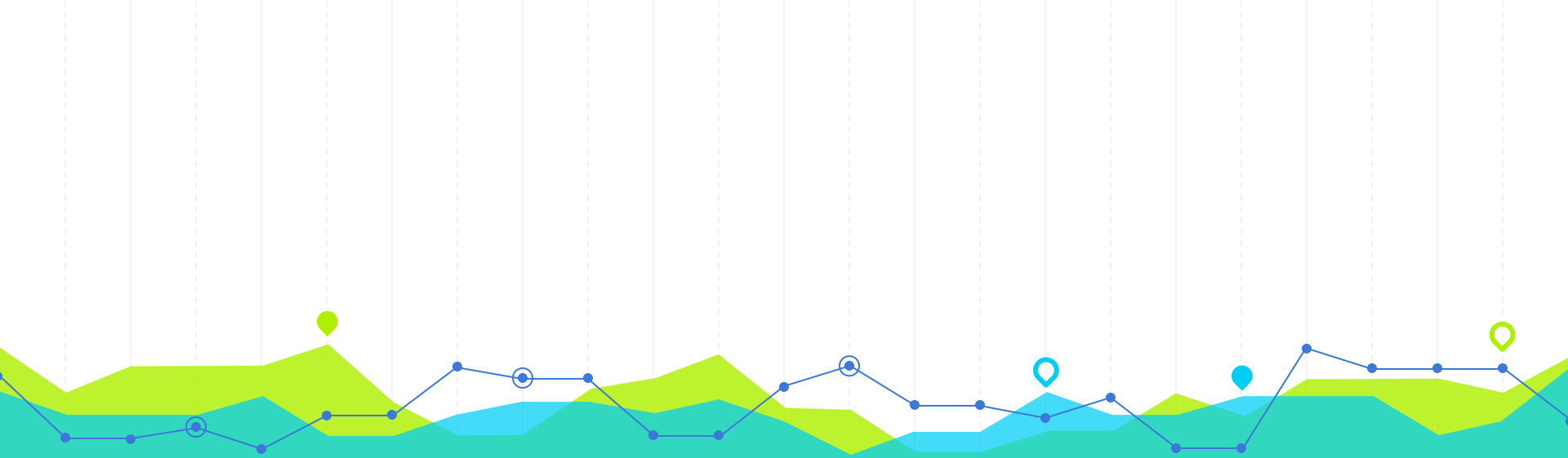


# Agenda



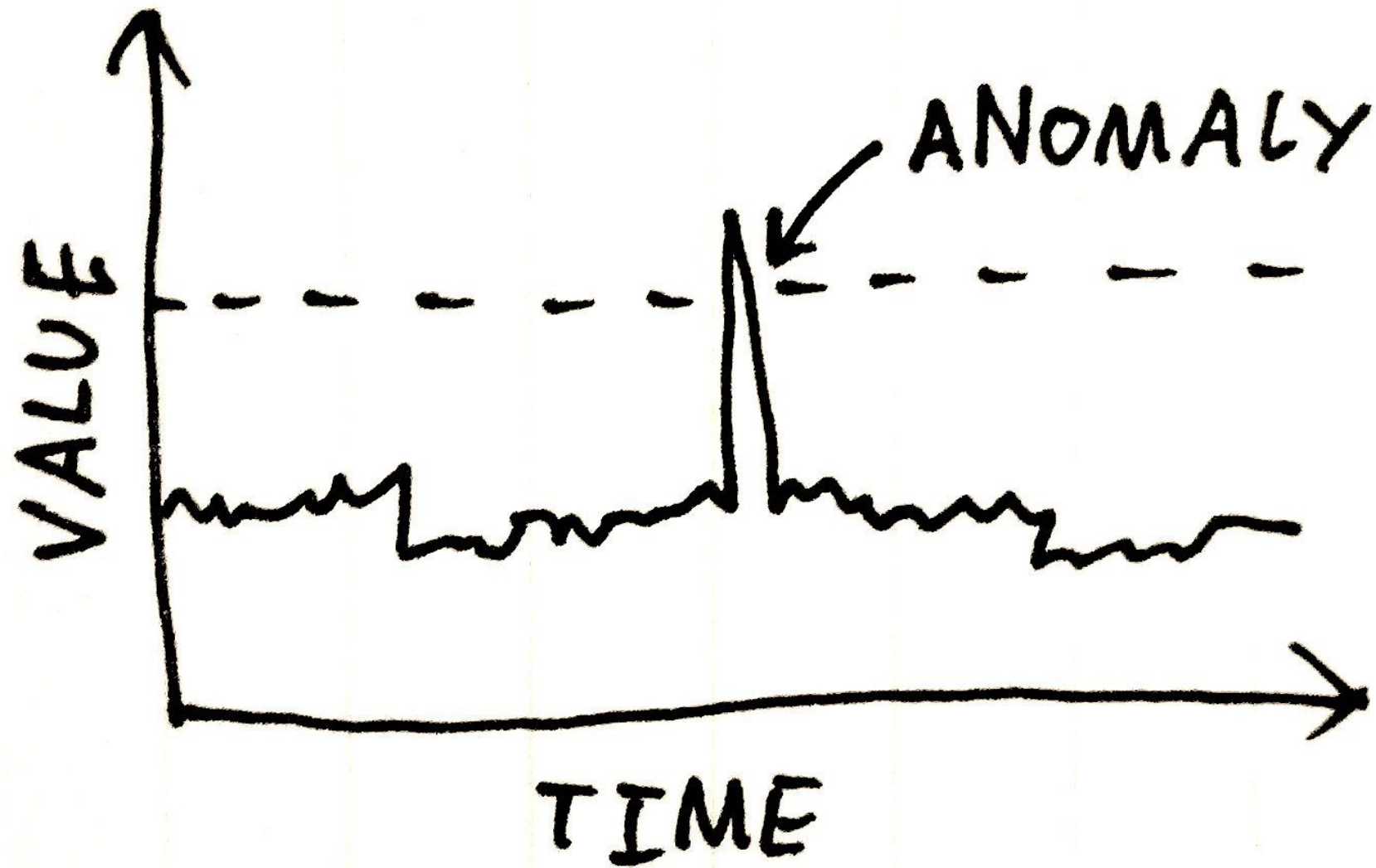
- La Vida para la IA
  - Sistema Inmune Artificial
  - Algoritmo Genético
- Bioinformática
  - Alineamiento Múltiple
  - Aplicaciones
- La IA para la Vida
  - Predicción de mutaciones
  - Clasificación de cáncer de piel





# Sistema Inmune Artificial







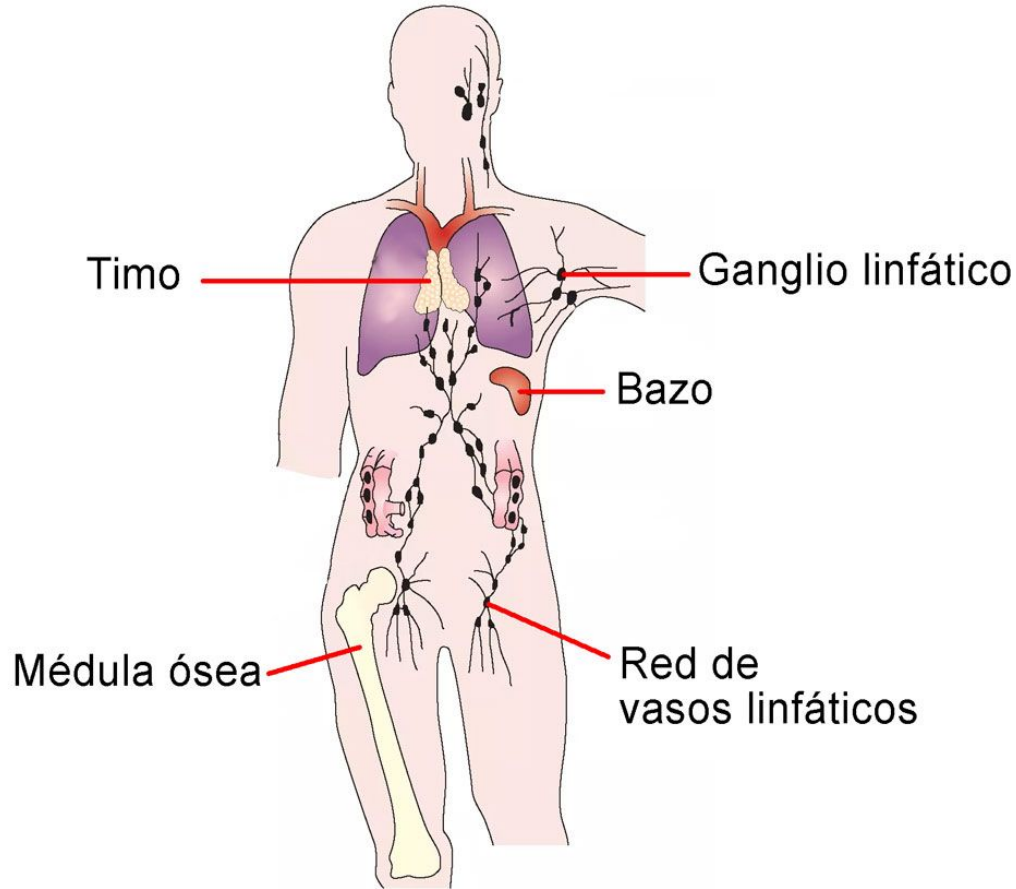
## Algoritmo de selección negativa (Forrest et. al)

### Aplicaciones:

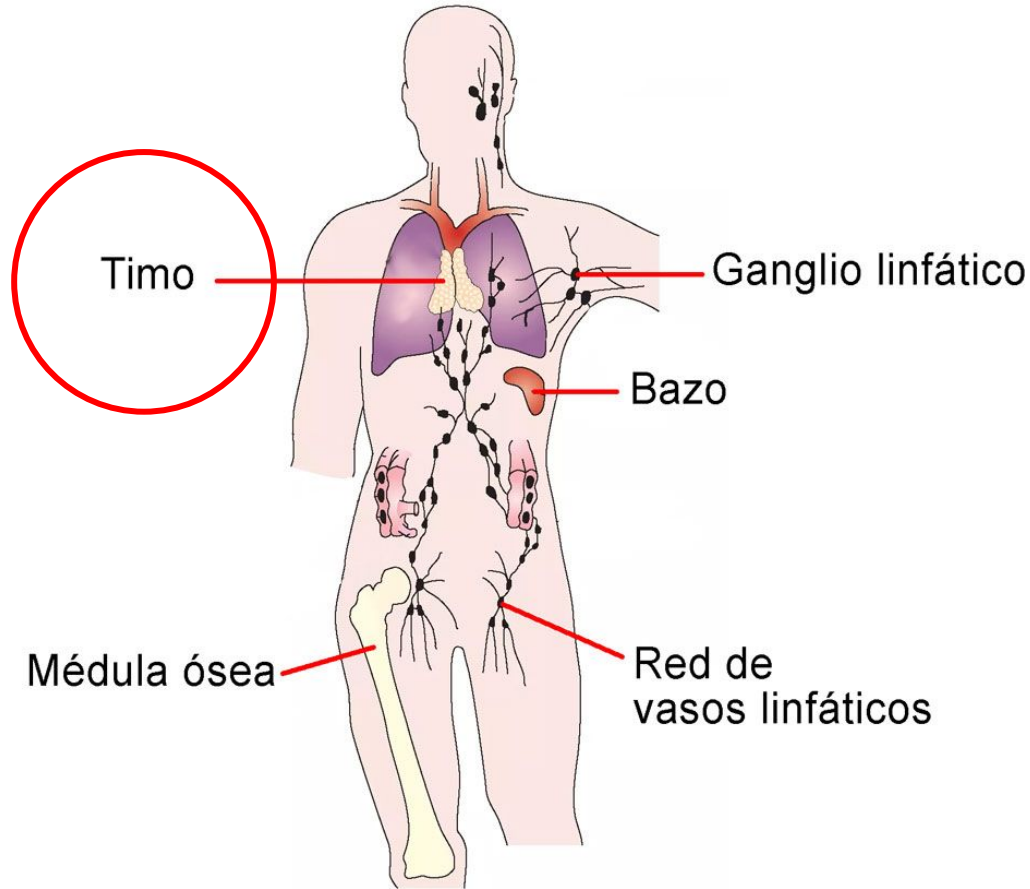
- Detección de anomalías en series de tiempo
- Seguridad Informática
- Monitoreo de procesos de sistema operativo
- Detección de patrones en ventas



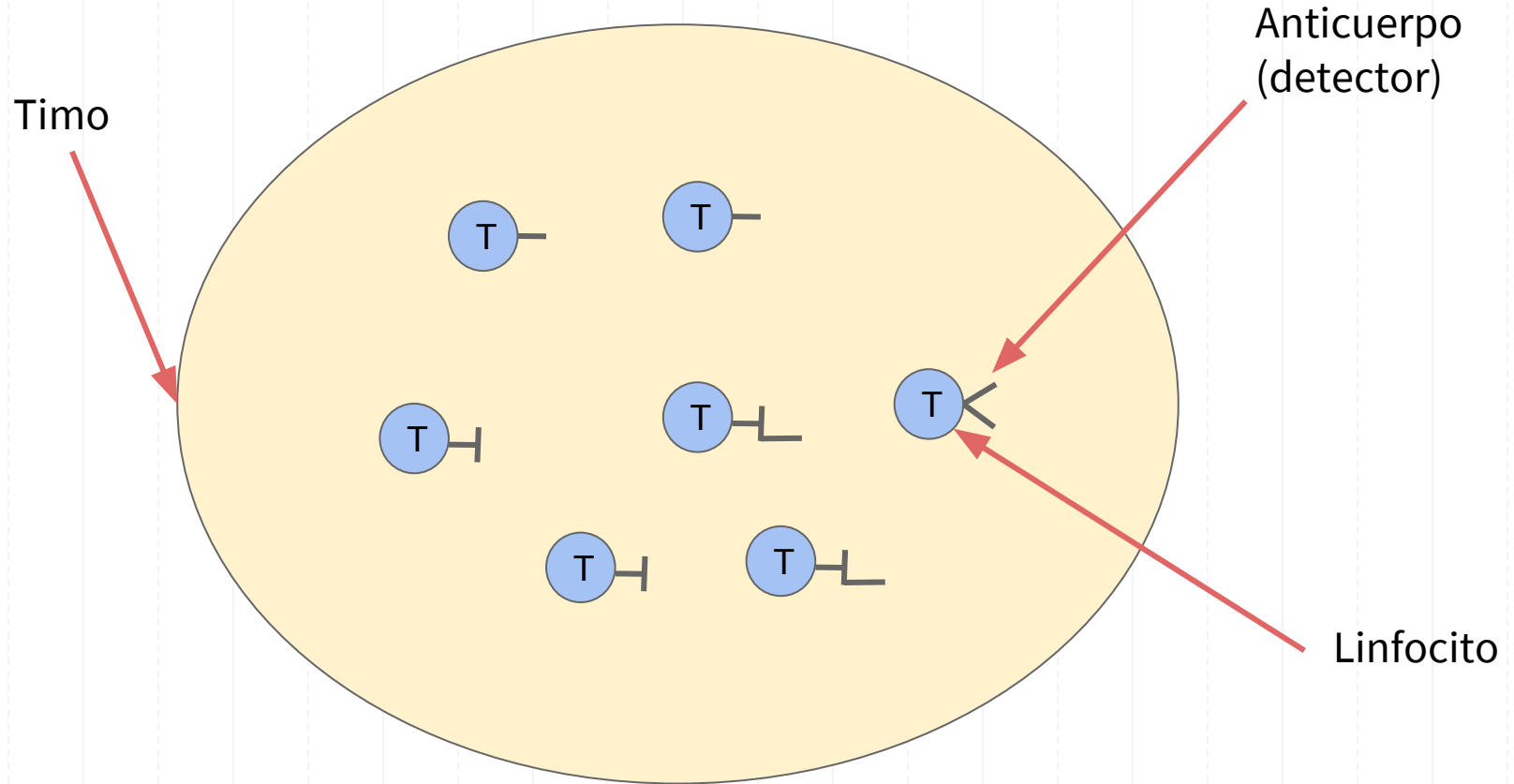
# Sistema inmunológico



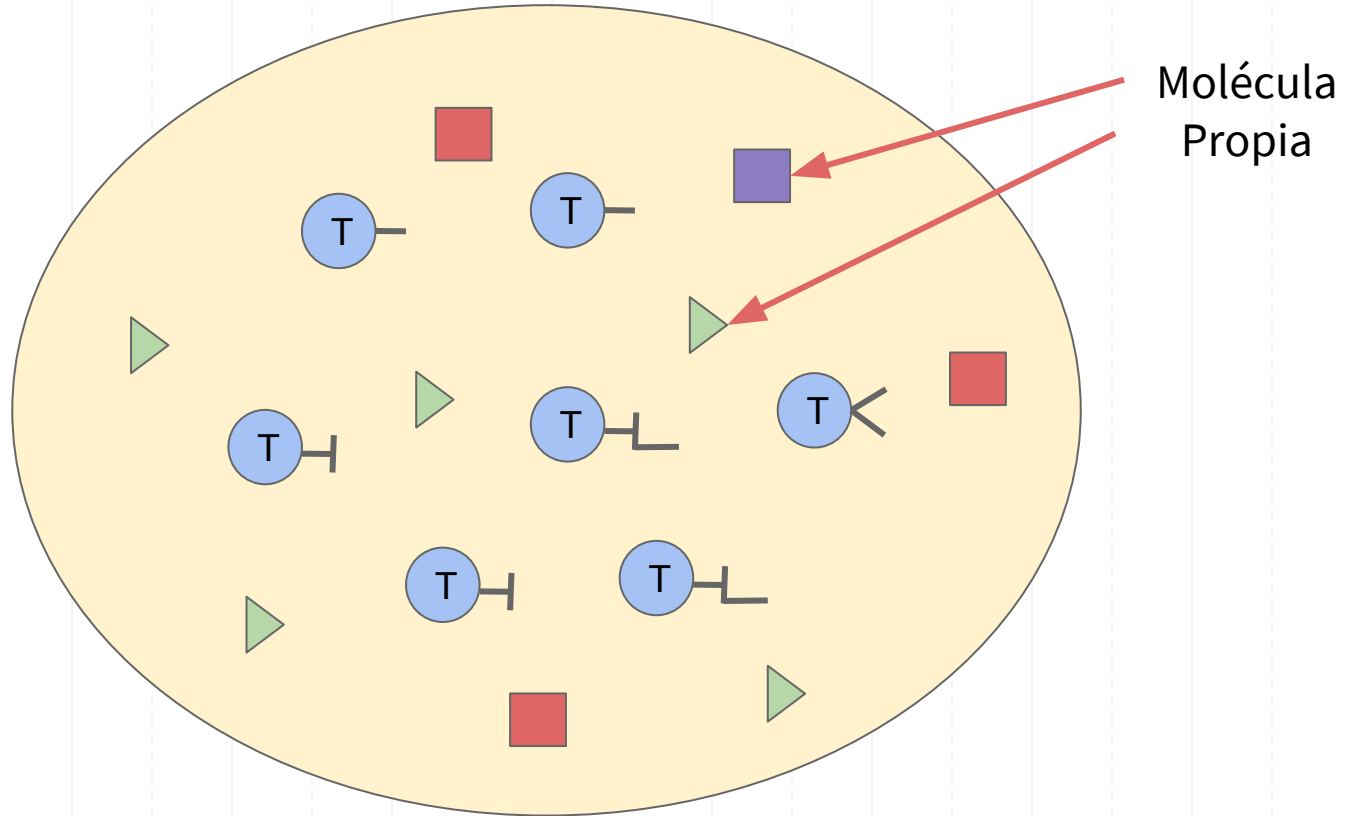
# Sistema inmunológico



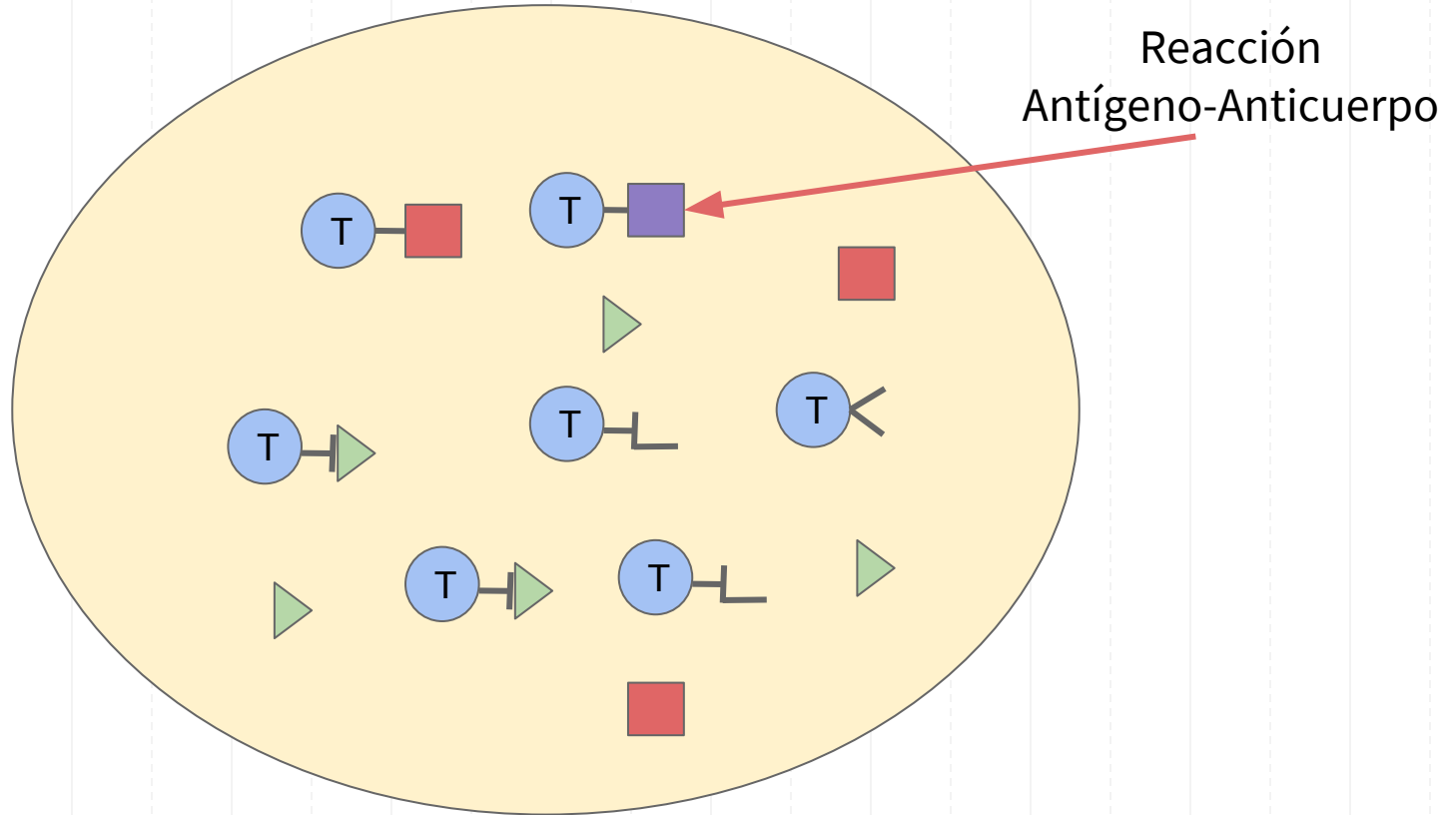
## Maduración de Linfocitos T (Fase de generación)



## Maduración de Linfocitos T (Fase de generación)

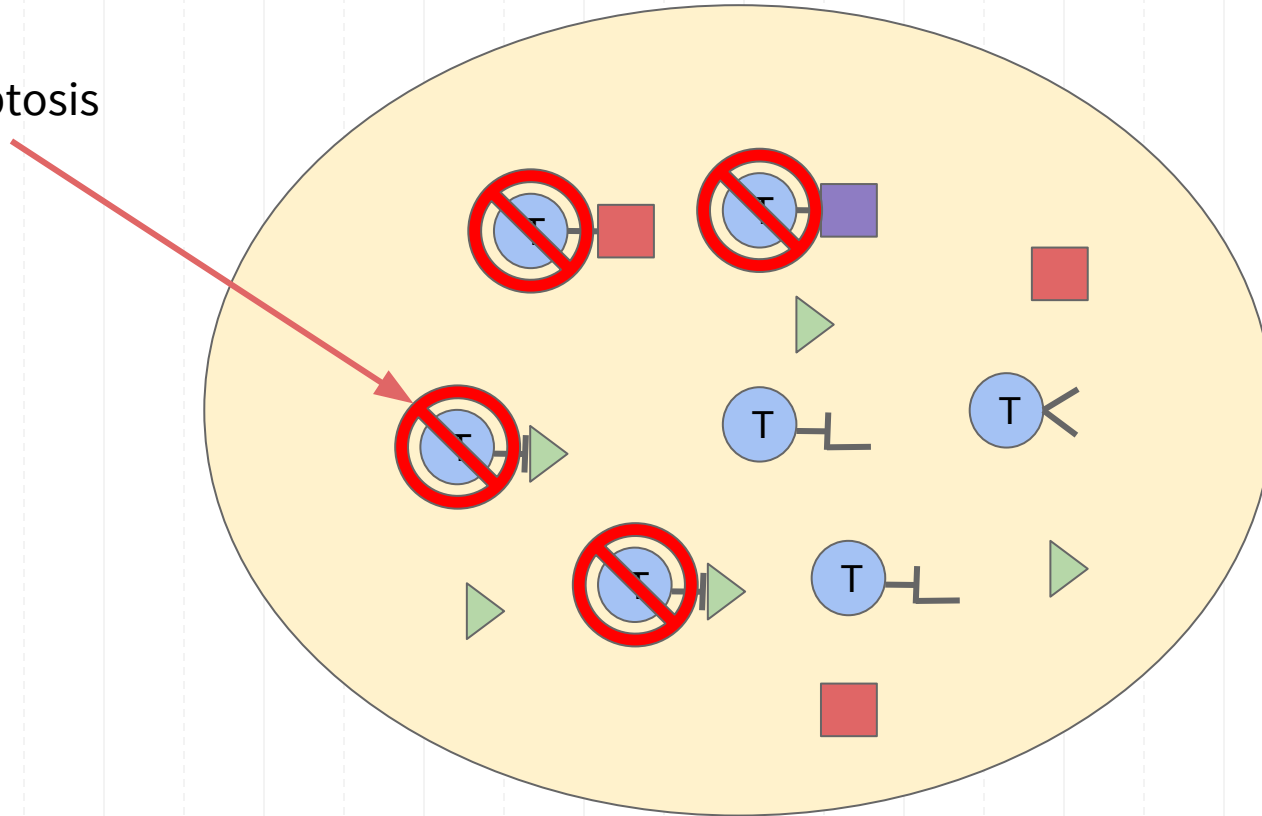


## Maduración de Linfocitos T (Fase de generación)

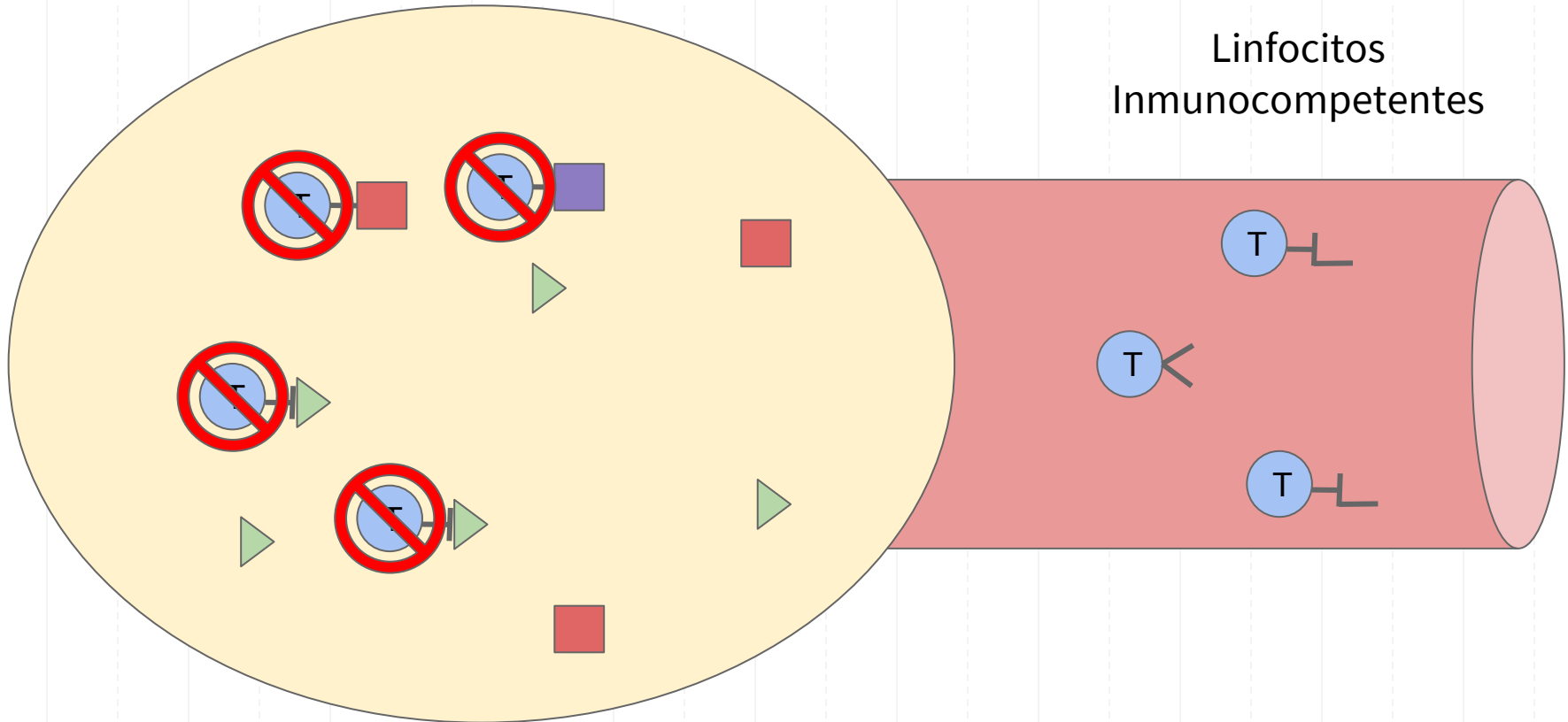


## Maduración de Linfocitos T (Fase de generación)

Apoptosis

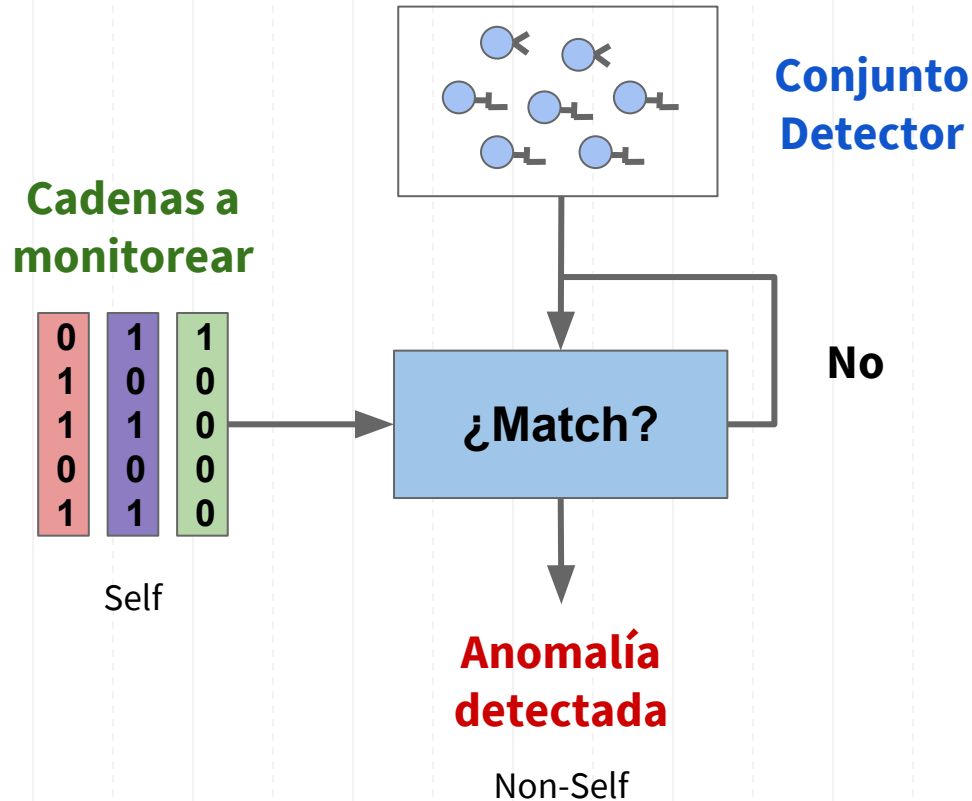


## Maduración de Linfocitos T (Fase de generación)





## Fase de detección



**Autoregulación**

**Extracción de  
características**

**Memoria**

**Reconocimiento**

**Aprendizaje**

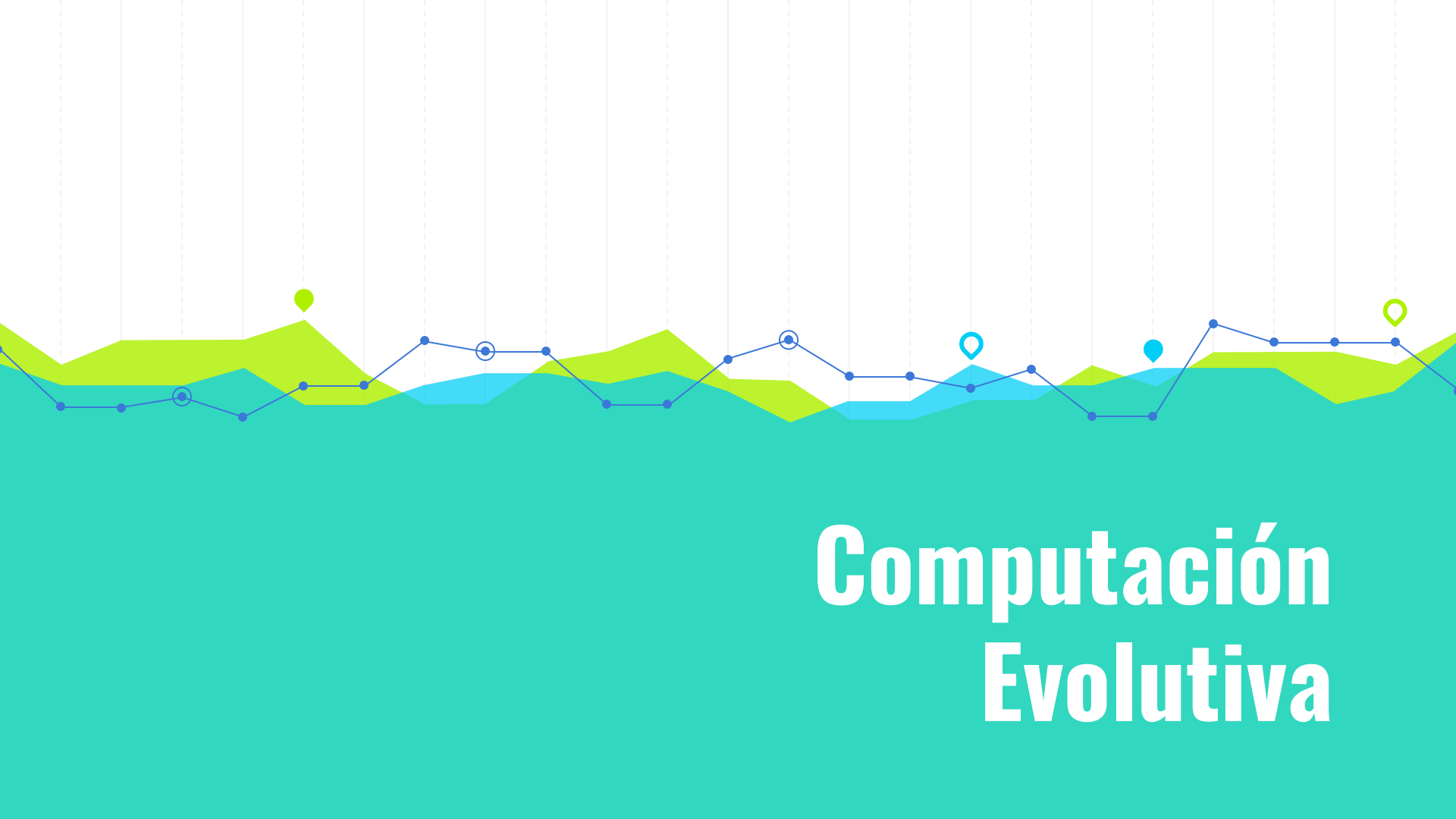
**Detección distribuida**

## Algoritmo de selección negativa (Forrest et. al)

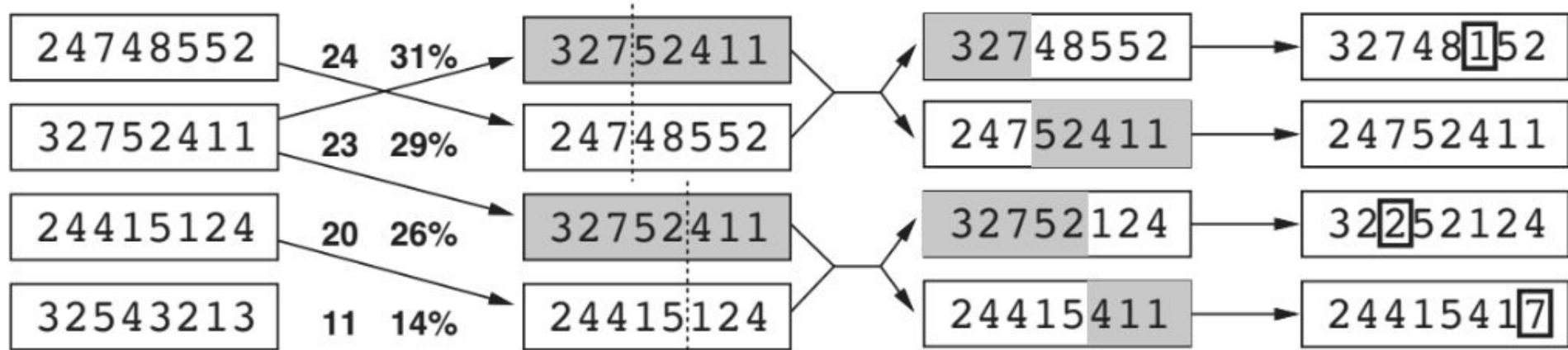
### Características:

- Al igual que nuestro sistema inmunológico, cada **programa** de detección es **único**. Esto contribuye a la seguridad de un sistema distribuido.
- La detección es **probabilística**, por lo tanto no está basada en la búsqueda de patrones específicos.
- Parámetros: Umbral de detección, Criterio de match, #Linfocitos T





# Computación Evolutiva



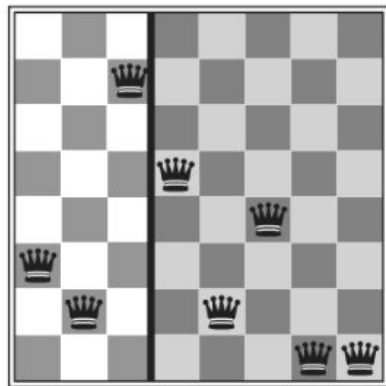
(a)  
Initial Population

(b)  
Fitness Function

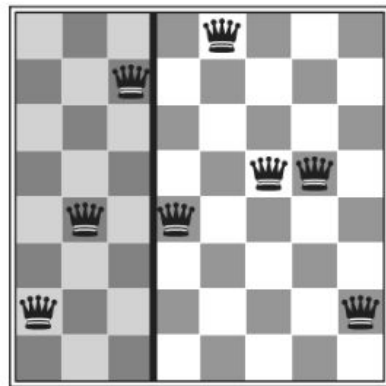
(c)  
Selection

(d)  
Crossover

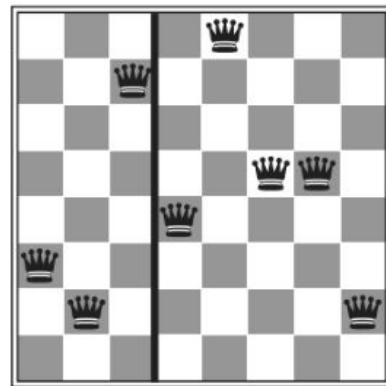
(e)  
Mutation



+

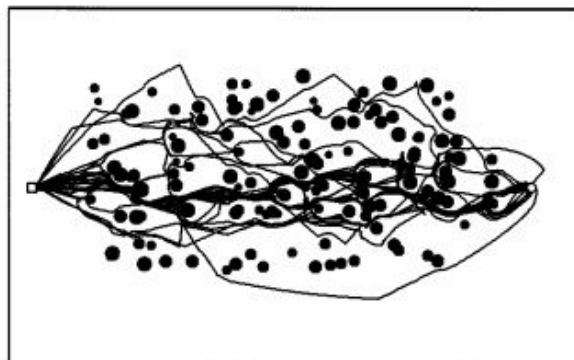
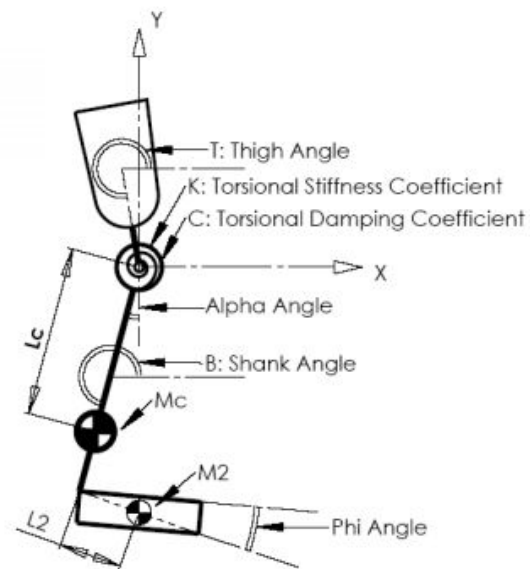
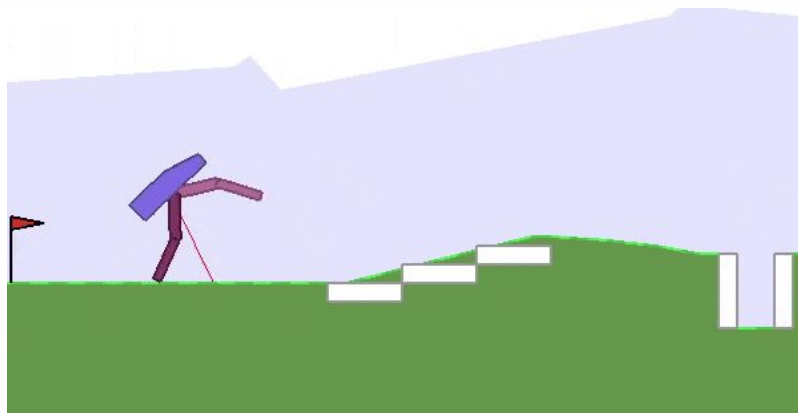


=

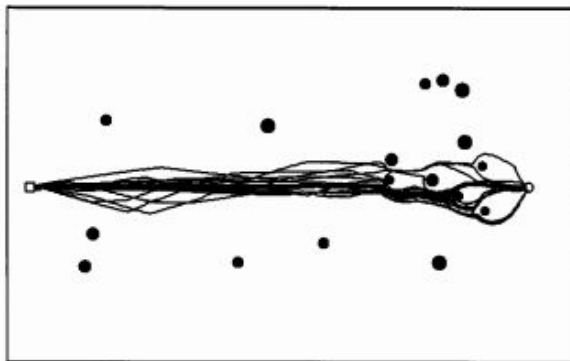




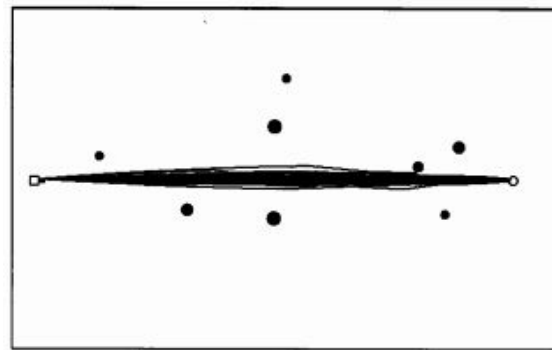
$$\begin{aligned} \text{raw\_fitness} = & \text{collision\_weight} * \text{number\_of\_collisions} \\ & + \text{time\_weight} * \text{number\_of\_steps} \\ & + \text{distance\_weight} * \text{distance\_traveled} \end{aligned}$$



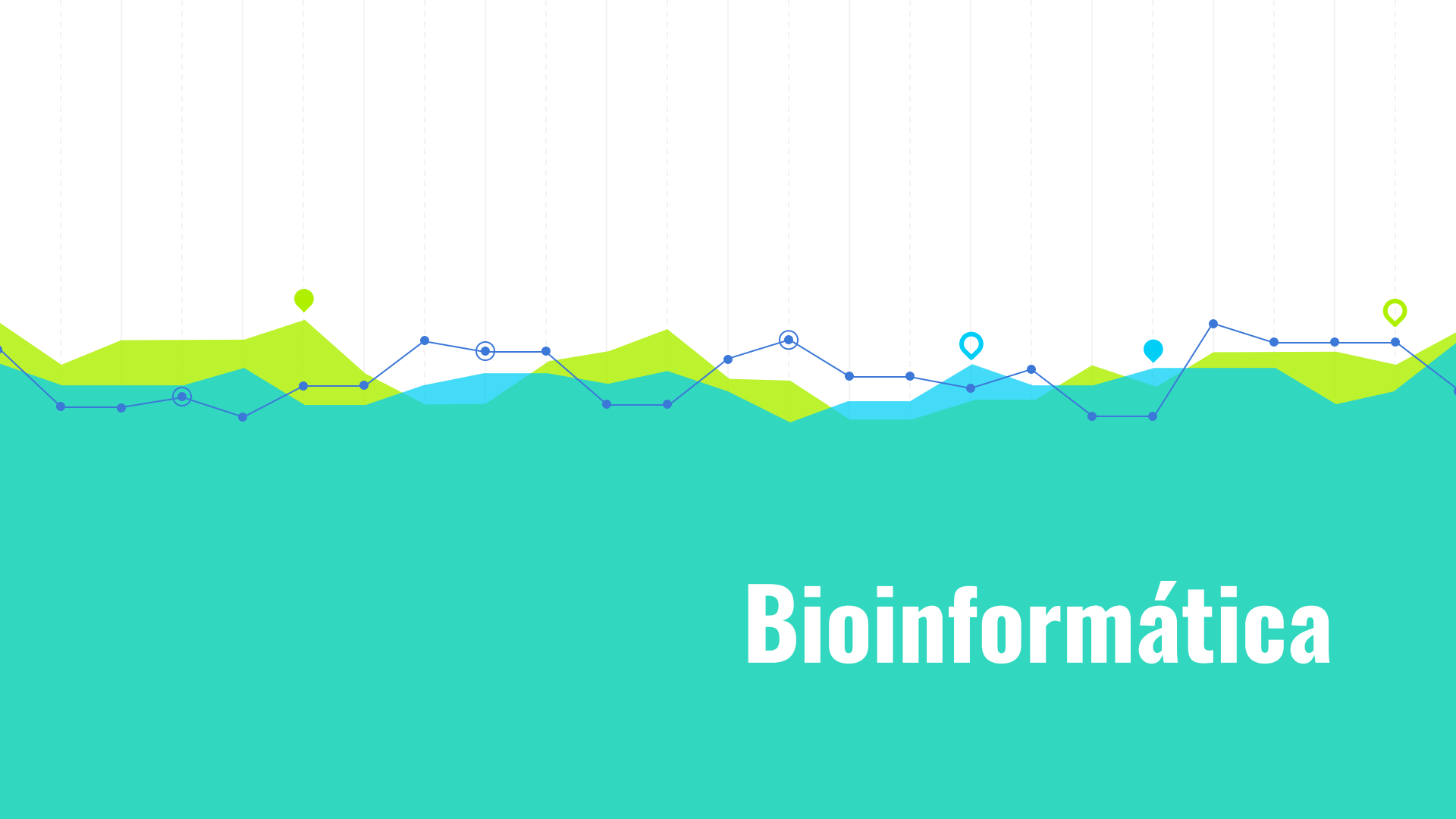
(C)



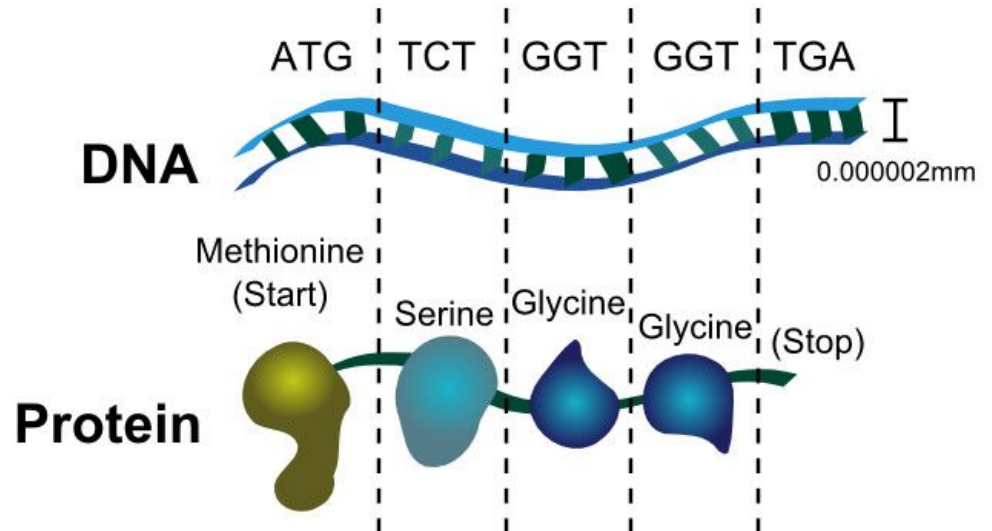
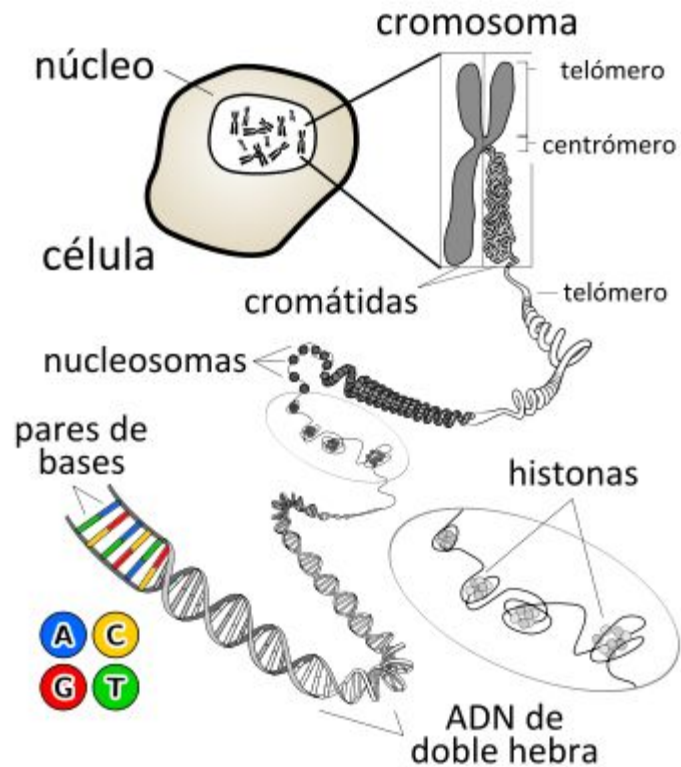
(B)



(A)



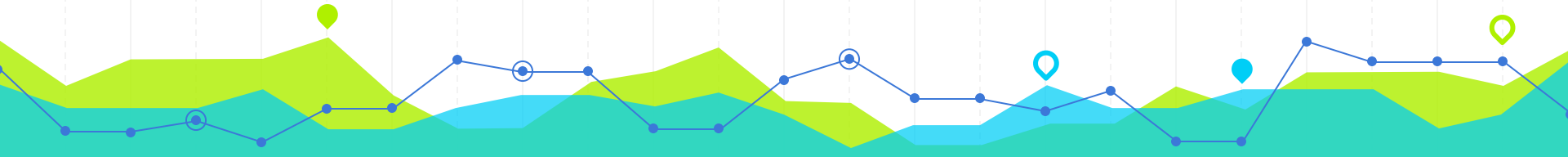
# Bioinformática

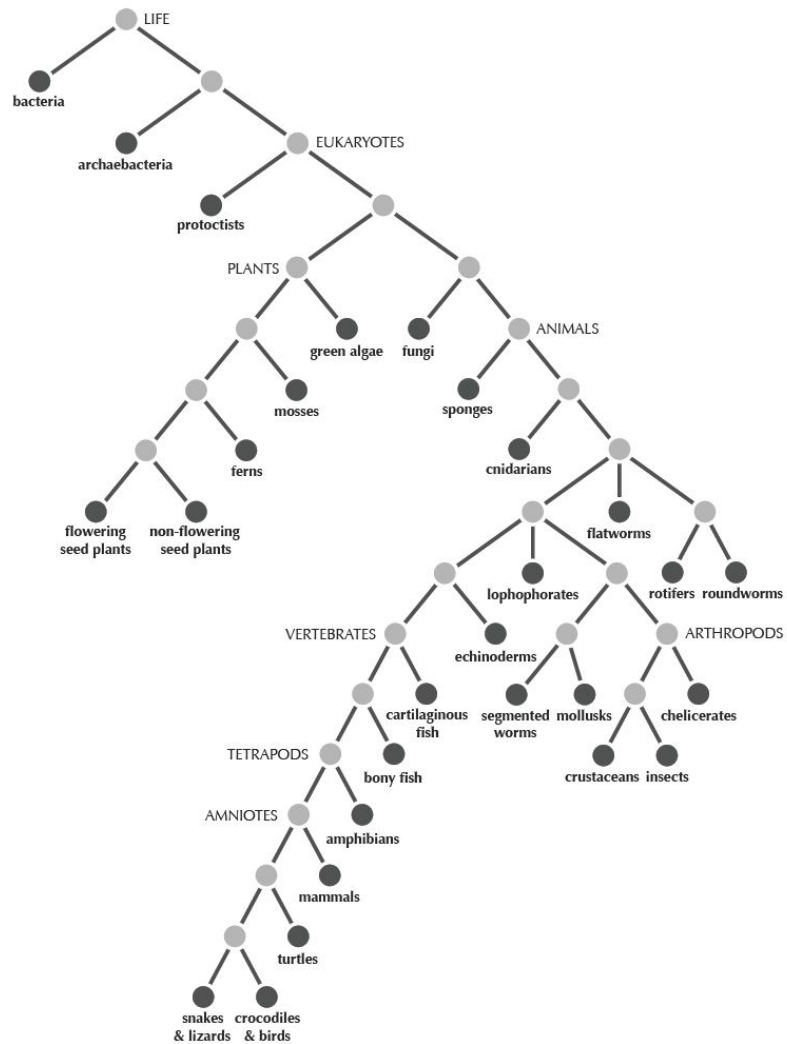




## Similitud de secuencias

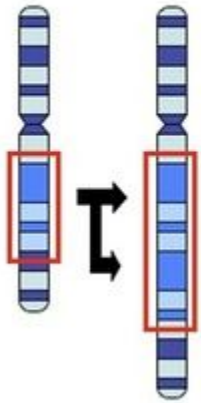
- Secuencias **similares** comparten un ancestro (Trans. Vert.)
- Resultado de **duplicación** (Mitosis)
- La similaridad depende de los tiempos de divergencia
- En aminoácidos ~25% de similitud es buena evidencia de ancestro común.
- Algunas regiones se conservan más que otras debido a su función.



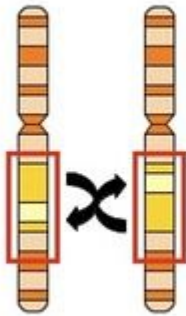


# Mutaciones

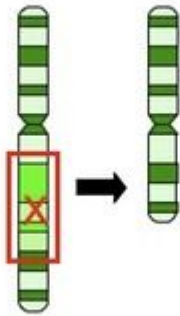
**Duplication**



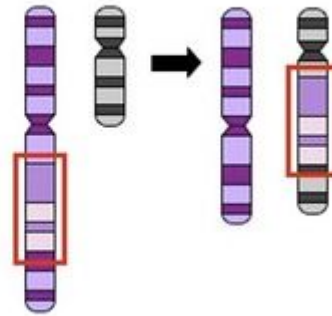
**Inversion**



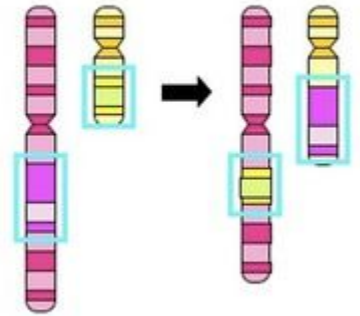
**Deletion**



**Insertion**



**Translocation**



Bioninja

## Alineación (Needleman-Wunsch)



**Figure 3.1.** Distinction between global and local alignments of two sequences.

# Alineación Múltiple

[illegible]



# NIP H A R D !

```

Q5E940 BOVIN -----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLGMGKNTMMRKAIRGHLENN--PALE 76
RLA0_HUMAN -----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLGMGKNTMMRKAIRGHLENN--PALE 76
RLA0_MOUSE -----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLGMGKNTMMRKAIRGHLENN--PALE 76
RLA0_RAT -----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLGMGKNTMMRKAIRGHLENN--PALE 76
RLA0_CHICK -----MPREDRATWKSNYFMKIIQLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLGMGKNTMMRKAIRGHLENN--PALE 76
RLA0_RANSY -----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKQMQQIRMSLRGK-AVVLGMGKNTMMRKAIRGHLENN--PALE 76
Q7ZUG3 BRARE -----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKQMOTIRLSLRGK-AVVLGMGKNTMMRKAIRGHLENN--PALE 76
RLA0 ICTPU -----MPREDRATWKSNYFLKIIQLDDYPKCFIVGADNVGSKQMOTIRLSLRGK-AIVLMGKNTMMRKAIRGHLENN--PALE 76
RLA0_DROME -----MYRENKAAWKAQYFIKVLEFDEFKPCFIVGADNVGSKQMONTIRSLRGL-AVVLGMGKNTMMRKAIRGHLENN--PQLE 76
RLA0_DICDI -----MSGAG-SKRKKLFIEKATKLTFTYDKMIVAADFVGSQLOKIRKSIIRGI-CAVLMGKKTMRKIVIRDLADSK--PELD 75
Q54LP0 DICDI -----MSGAG-SKRKNVTEKATKLTFTYDKMIVAADFVGSQLOKIRKSIIRGI-CAVLMGKKTMRKIVIRDLADSK--PELD 75
RLA0_PLAF8 -----MLLSKQKKMYILQLSSLIQCSKILVHVDVSNQM SVRA LRK AALMGMGKNTIRTRITALKKNLQAV--PQIE 76
RLA0_SULAC -----MIGLAVTITKIAWKDEVALTEKIKTKTIIANIEFIADKILRIRK-ADAVTNNLFNLTAKNAG----YDTK 79
RLA0_SULTO -----MRIMAVITETKIAWKDEVALTEKIKTKTIIANIEFIADKILRIRK-ADAVTNNLFNLTAKNAG----YDTK 79
RLA0_SULSO -----MKRLALALQRKIVSWKEVEKELTELSKNTLGNISFIPQIKIEIKYLRK-ATAVTNTLFKIAAKNAG----IDIE 80
RLA0_AERPE MSVSVLVGQMYK EKPISWK LMLRELEELFSKRVVLADFGTPTVVRVRVLWKYVMVA KRUILRAMKAAGLE----LDDN 86
RLA0_PYRAE -MMLAIGKRRYVTRQYIPARKVKIVSEATELQKYPYVFLFDLHGLSSRIHLEYRYKLRRYGVIKIIRPTLFKTAFTKVYGG--IPAE 85
RLA0_METAC -----MAEERHHTEHIPQWKDEIENIKELTQSHKVFQGMVIEGILATKMQKIRRDLDKV-AVLKVSRTNLTERALNQLG----ETIP 78
RLA0_METMA -----MAEERHHTEHIPQWKDEIENIKELTQSHKVFQGMVIEGILATKMQKIRRDLDKV-AVLKVSRTNLTERALNQLG----ESIP 78
RLA0_ARCFU -----MAAVRGS--PPEYKVRAVEEIKRMISSKPVVAIVSFRNVPAGQMOKIRREFRGK-AEIKVVKNTLLERALDALG----GDYL 75
RLA0_METKA MAVKAKGQPPSGYEPKVAEWRKREVEKELKELMDEYENGLVDLIGIPAPQLQEIIRAKLRERTIIRMSRNTLMRTALEEKLDER--PELE 88
RLA0_METTH -----MAHVAEWKKKEVQELHDLIKGYEVGCIANADIPARQLQKMRQTLRDS-ALIRMSKKTLLISLAEKAGREL--ENVD 74
RLA0_METTL -----MITAESEHIAPIWKIEEVNKLKELKNQVILVDMVEVPARQLQEIIRDKIR-GTMTLKMSRNTLIERAIKEVAEETGNPEFA 82
RLA0_METVA -----MIDAKSEHKIAPWKIEEVNALKELKSNANVILIMMEVPAVQLQEIIRDKIR-DQMTLKMSRNTLIERAEEVAEETGNPEFA 82
RLA0_METJA -----METKVAHVAPWKIEEVKTLKGLIKSPVVAIVMDVPAPQLQEIIRDKIR-DKVKLMSRNTLIERALKEAAEELNPKLA 81
RLA0_PYRAB -----MAHVAEWKKKEVEELANLTKSPVVALVIVSSMPAYPLSQMRRLIRENGGLLRVSNTLIELAIKKAQELGKPELE 77
RLA0_PYRHO -----MAHVAEWKKKEVEELAKLTKSPVVALVIVSSMPAYPLSQMRRLIRENGGLLRVSNTLIELAIKKAQELGKPELE 77
RLA0_PYRFU -----MAHVAEWKKKEVEELANLTKSPVVALVIVSSMPAYPLSQMRRLIRENGGLLRVSNTLIELAIKKAQELGKPELE 77
RLA0_PYRKO -----MAHVAEWKKKEVEELANLTKSPVVALVIVSSMPAYPLSQMRRLIRENGGLLRVSNTLIELAIKKAQELGKPELE 77
RLA0_HALMA -----MSAESERKTETIPEWKQEEVDIAIVEMIESYESVGVVNIAGIPSRQLQDMRRDLHGT-AELRVSNTLLEALDDVD--DGLE 79
RLA0_HALVO -----MSESEVRQTEVIPQWKREEVDLVDFTESYESVGVVGVAGIPSRQLQSMRRELHGS-AAVRMSRNTLVNRLALDEVN--DGFE 79
RLA0_HALSA -----MSAEERQTTEVPPEWKQEEVDAIVDLDTLSDSVGVVNTGIPSKQLQDMRRDLHGT-AALRMSRNTLVNRLALEAG--DGLD 79
RLA0_THEAC -----MKVESQKKELVNEITQRKASRSVAIVDTAGITRQIQDIRGNRGK-INLKVIKKTLLFLKALENGD--EKLS 72
RLA0_THEVO -----MRKINPKKKEIVSELAQDITKSVAIVDIKGVTRTQMODIRAKNRDK-VKIKVVKTTLLFKALDSIND--EKLT 72
RLA0_PICTO -----MTEPAQWKIDFYKNLENIINSRKVAAIVSIKGLRNNEFQKIRNSIRDK-ARIKVSRAALLRLAIENTGK--NNIV 72
ruler 1. 10. 20. 30. 40. 50. 60. 70. 80. 90.

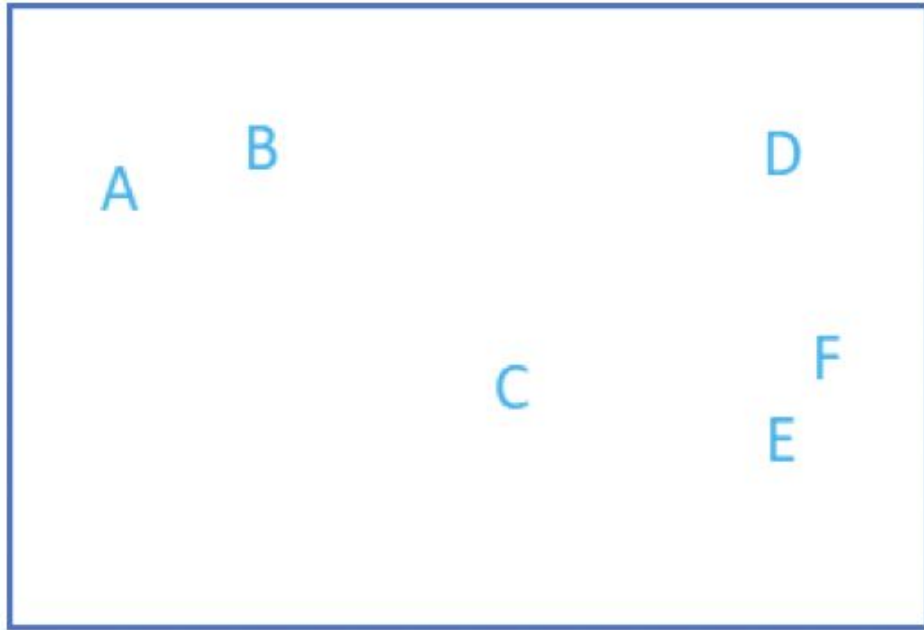
```

Alineamiento Progresivo (Heurística)

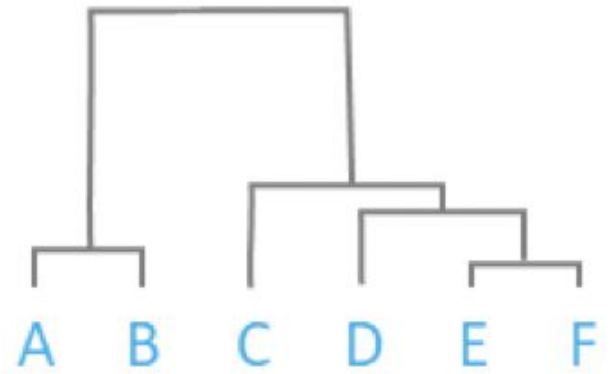
# SOLUCIÓN HEURÍSTICA

No garantiza solución óptima pero  
sí un buen alineamiento

## Alineamiento Progresivo (Heurística)

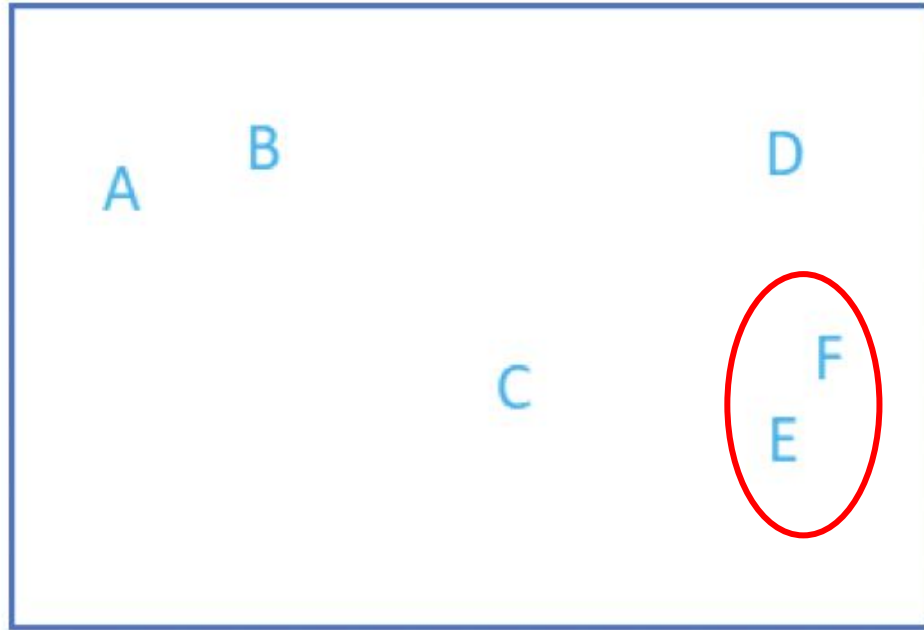


Dendrogram

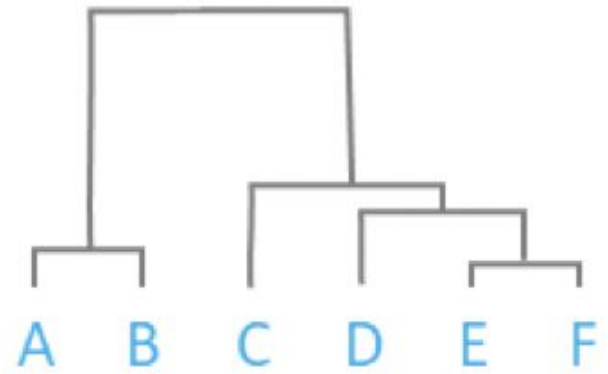




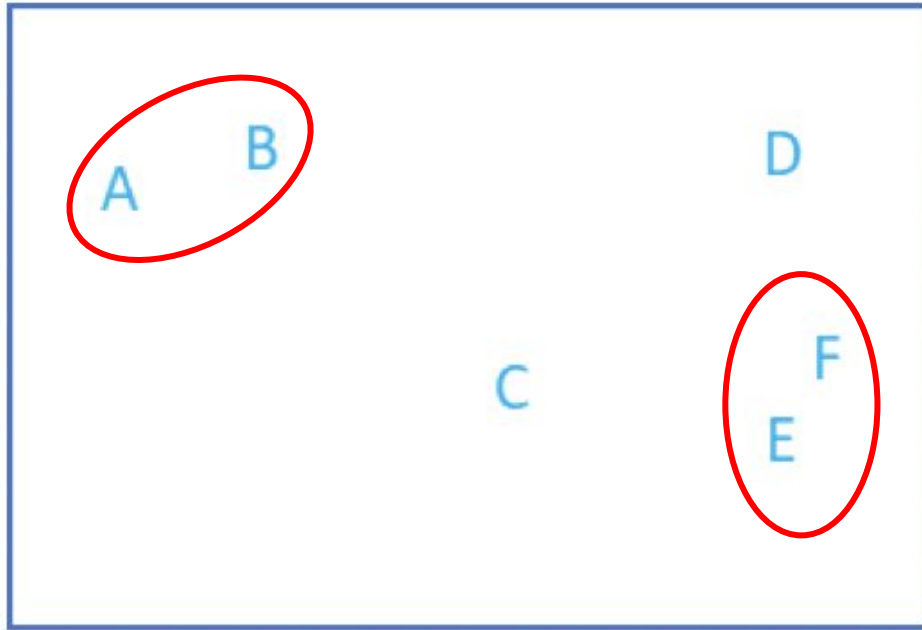
## Alineamiento Progresivo (Heurística)



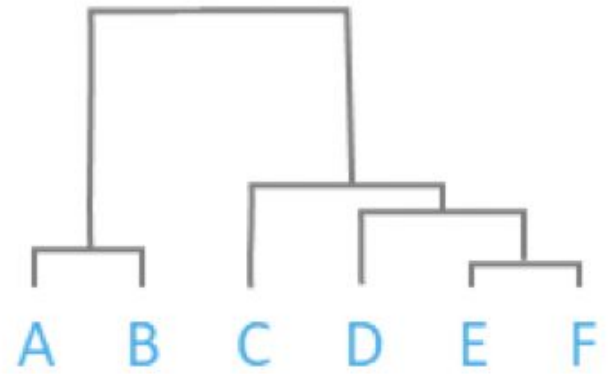
## Dendrogram



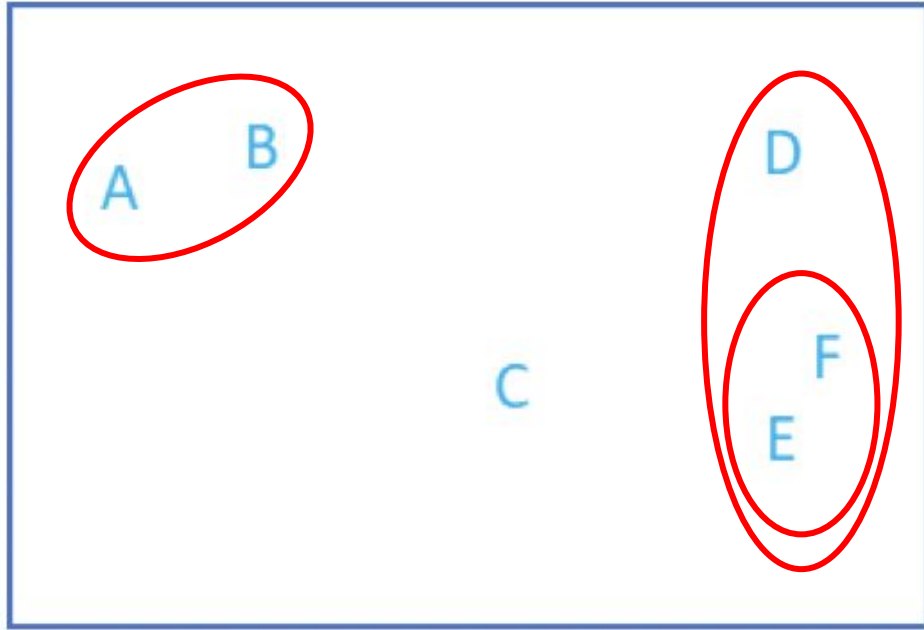
## Alineamiento Progresivo (Heurística)



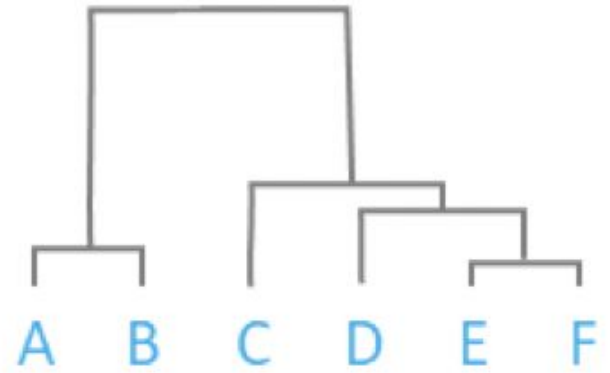
Dendrogram



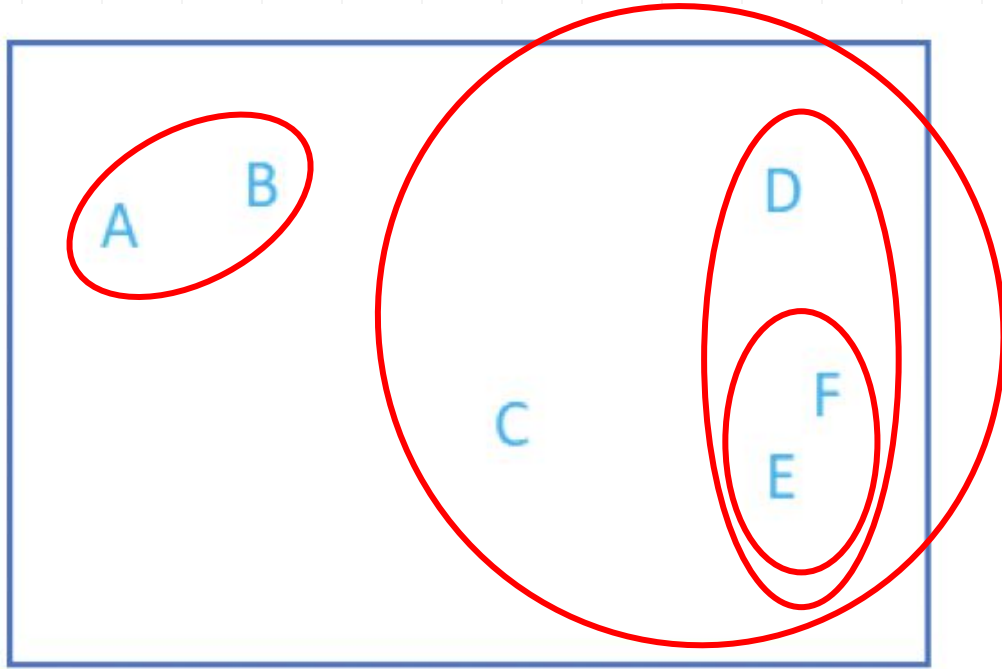
## Alineamiento Progresivo (Heurística)



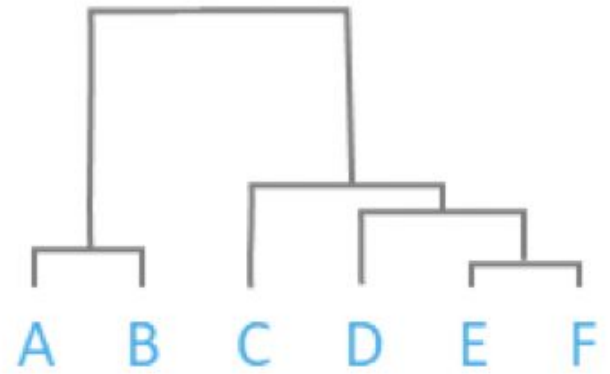
Dendrogram



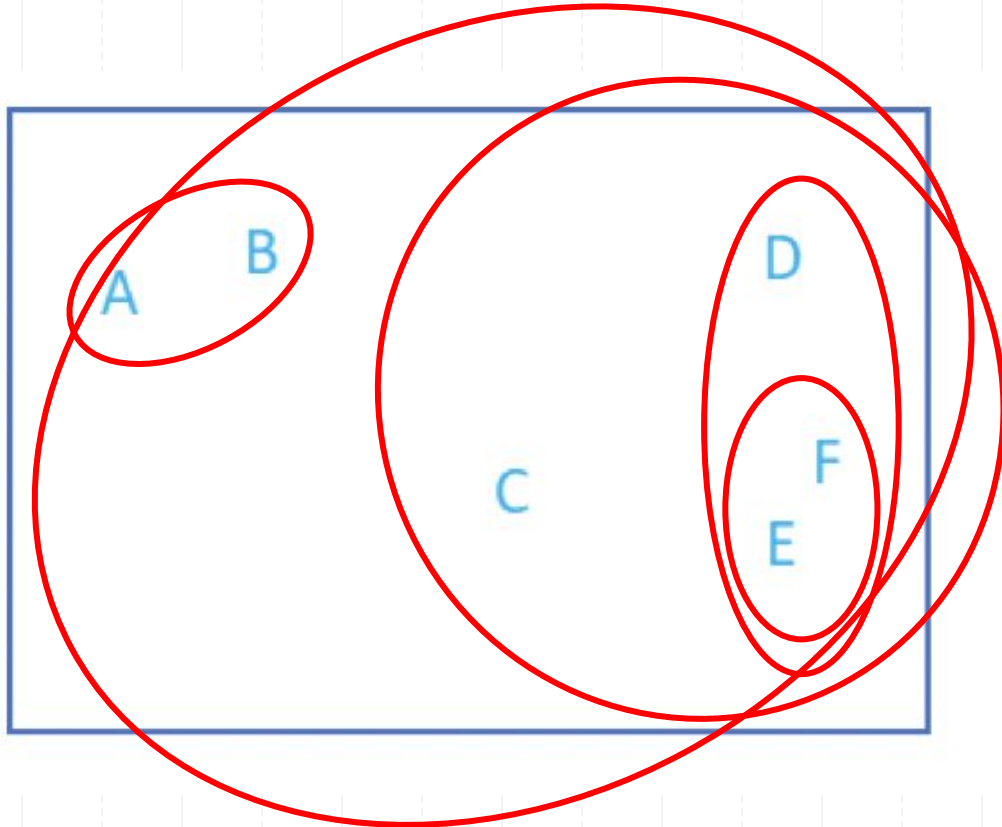
## Alineamiento Progresivo (Heurística)



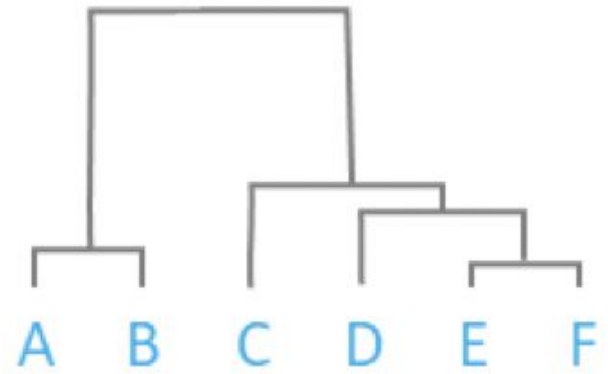
Dendrogram



## Alineamiento Progresivo (Heurística)



## Dendrogram



### Grupo 1:

HFGLACW

MGLACWH

GACHLWL

FGLMCHF

### Grupo 2:

GRVRDIR

QRNSEGL

PRERVGE

DNRSEGA

### Grupo 3:

FGKGC FY

CYFKKKC

YVFGIGT

CYFVGIG

### Grupo 4:

GTWNYFD

SYGSSYY

RPYGN YF

WDWYFDV



### Grupo 1:

HFGLAC--W-

MGLACWH---

-GA-CHLW-L

FGLMC-HF--

Score: 47

### Grupo 2:

GRVRDI-R

QRN-SEGL

PRER-VGE

D-NRSEGA

Score: 29

### Grupo 3:

--FGKGCFY

CYFKKKC--

YVFGIGT--

CYFVG-IG-

Score: 20

### Grupo 4:

GTWN-YFD-

S-YGSSYY-

RPYGN-YF-

--WDWYFDV

Score: 36



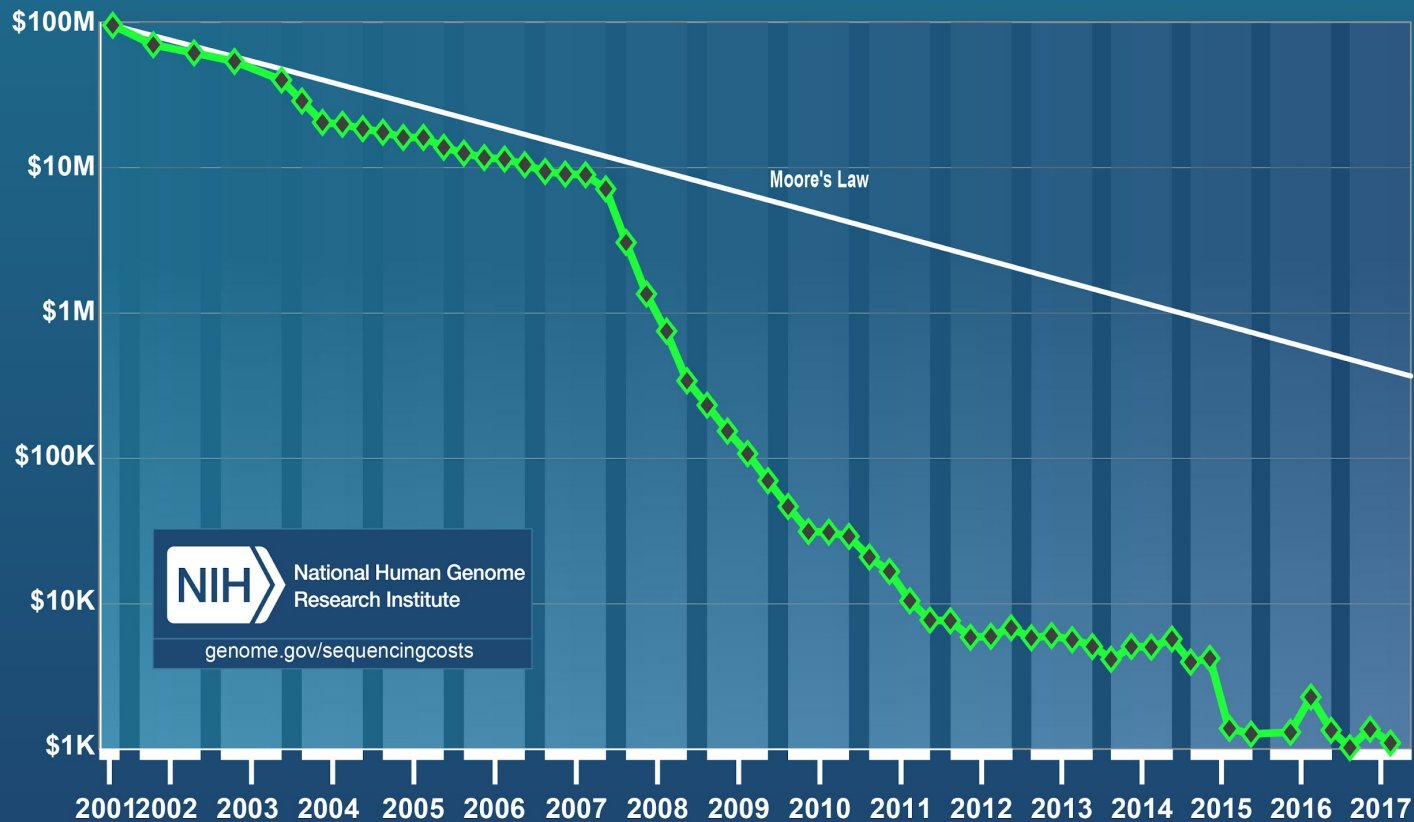
## Aplicaciones

- Análisis de árboles filogenéticos (Biología Evolutiva)
- Lingüística Comparativa
- Análisis de Series de Tiempo (DTW)
- Búsqueda de genes y predicción de función
- Caracterización de proteínas (Bases de datos)





## Cost per Genome





“


*¿Cómo la IA puede salvar vidas?*




Article | Published: 23 July 2018

# Predicting the clinical impact of human mutation with deep neural networks

Laksshman Sundaram, Hong Gao, Samskruthi Reddy Padigepati, Jeremy F. McRae, Yanjun Li, Jack A. Kosmicki, Nondas Fritzilas, Jörg Hakenberg, Anindita Dutta, John Shon, Jinbo Xu, Serafim Batzoglou, Xiaolin Li & Kyle Kai-How Farh 

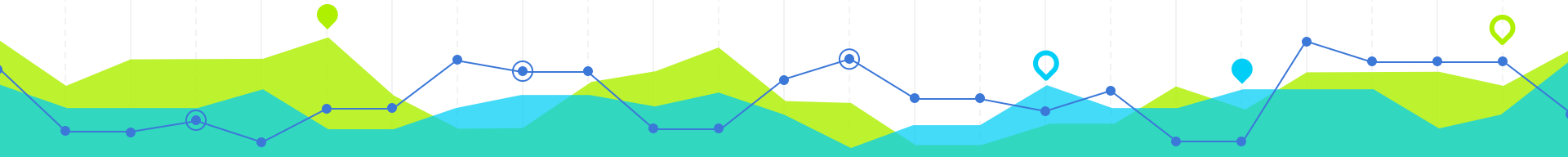
*Nature Genetics* **50**, 1161–1170 (2018) | [Download Citation](#) 

 An [Author Correction](#) to this article was published on 17 December 2018

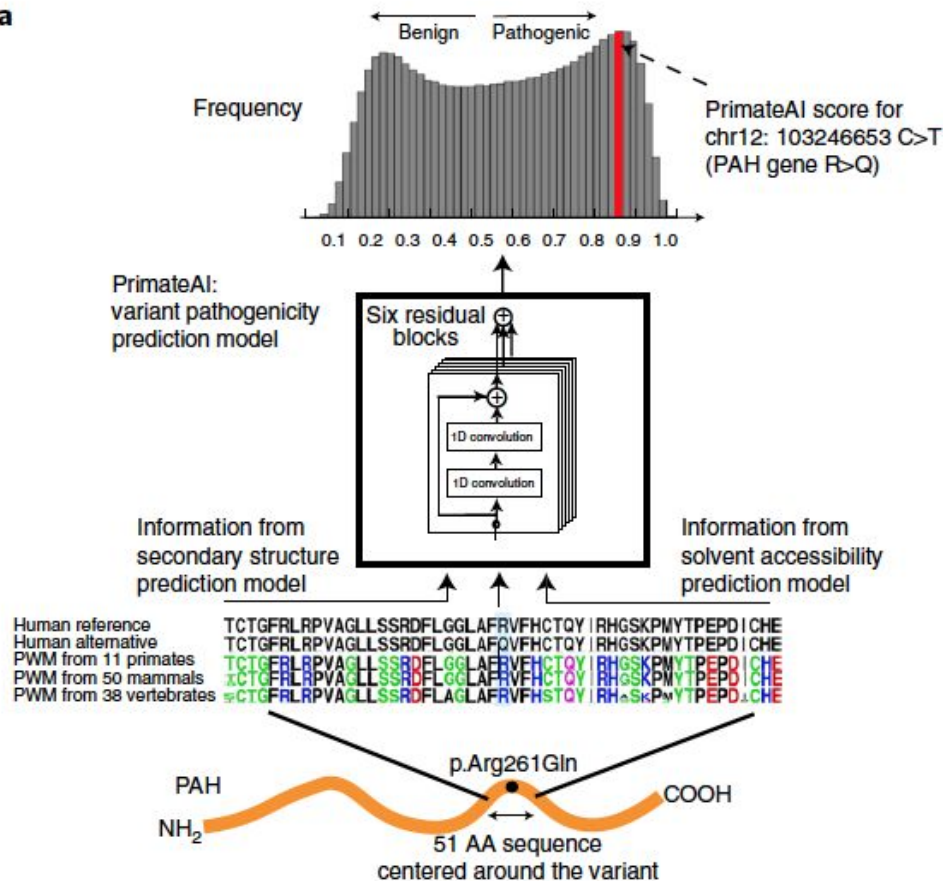
 This article has been updated

## Predicción de mutaciones

- Utilizar **genoma no-humano** (primates) para identificar **variantes humanas** relacionadas a enfermedades.
- Especies: Gorilas, Orangutanes, Chimpancé (~99%).
- Precisión: 88% (Aprendizaje Supervisado: ClinVar)
- Descubrimiento de 14 genes relacionados a discapacidad intelectual



a



## Red de predicción de patogenicidad

### Arquitectura: Deep Residual Network

#### Input:

Secuencias de Aminoácidos (Humano)

Referencia

Centrada en variante

Matriz de pesos

Patrones de conservación por posición

Entre más raro más informativo

Predicción de estructura secundaria

Relacionada a forma 3D de proteína

Predicción de accesibilidad de solvente

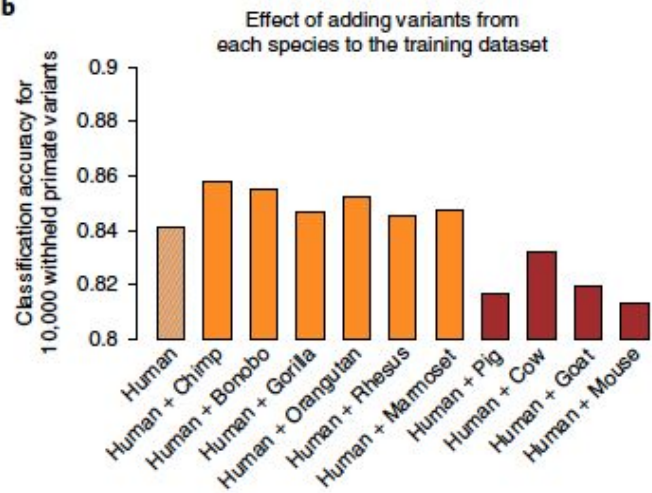
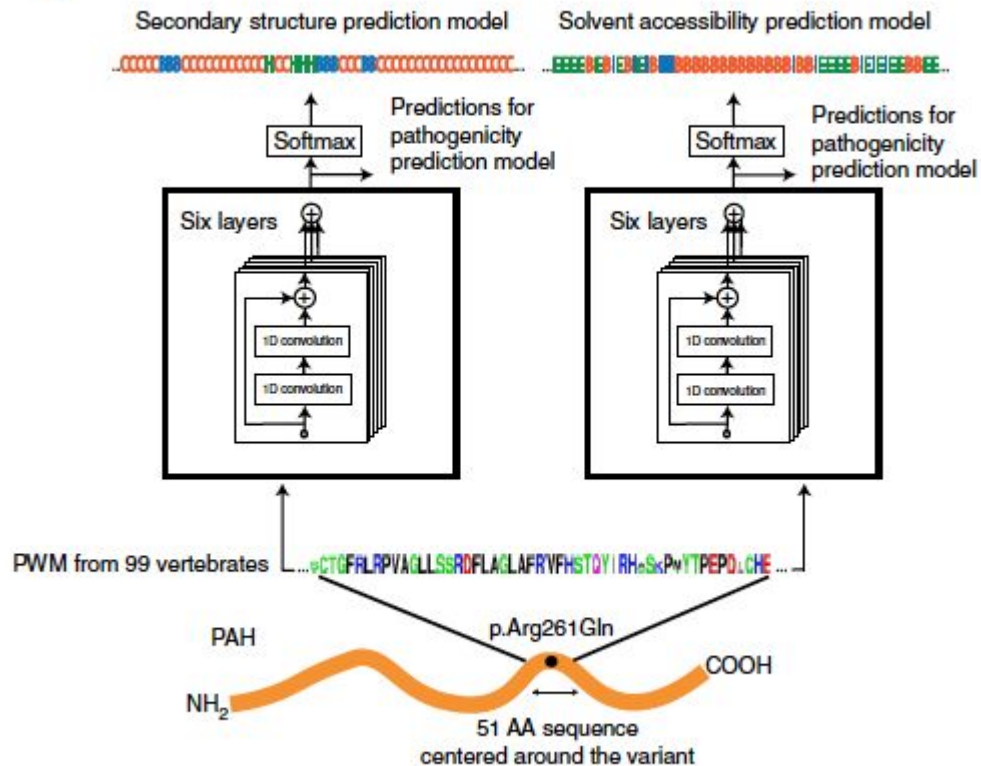
Área de la proteína expuesta solvente

#### Output:

Rango continuo:

0: Benigno

1: Patogénico

**b****b**

Letter | Published: 25 January 2017

# Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva , Brett Kuprel , Roberto A. Novoa , Justin Ko, Susan M. Swetter, Helen M. Blau & Sebastian Thrun 

*Nature* **542**, 115–118 (02 February 2017) | [Download Citation](#) 

 A [Corrigendum](#) to this article was published on 28 June 2017



## Clasificación de cáncer de piel

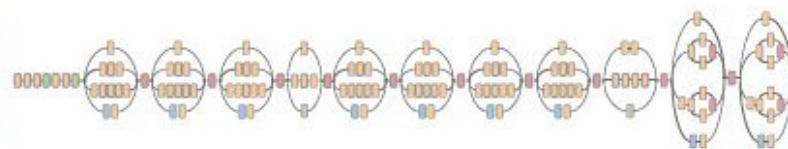
- Uno de los tipos de **cáncer** más comunes
- Diagnóstico **visual**
- End-to-end learning
- **Transfer Learning** (Google Inception V3)
- Entrenada en 757 clases
- Prueba de desempeño con 21 dermatólogos
- Incrementar el alcance del diagnóstico a través de dispositivos móviles. (Acceso universal y de bajo costo)



Skin lesion image



Deep convolutional neural network (Inception v3)



- Convolution
- AvgPool
- MaxPool
- Concat
- Dropout
- Fully connected
- Softmax

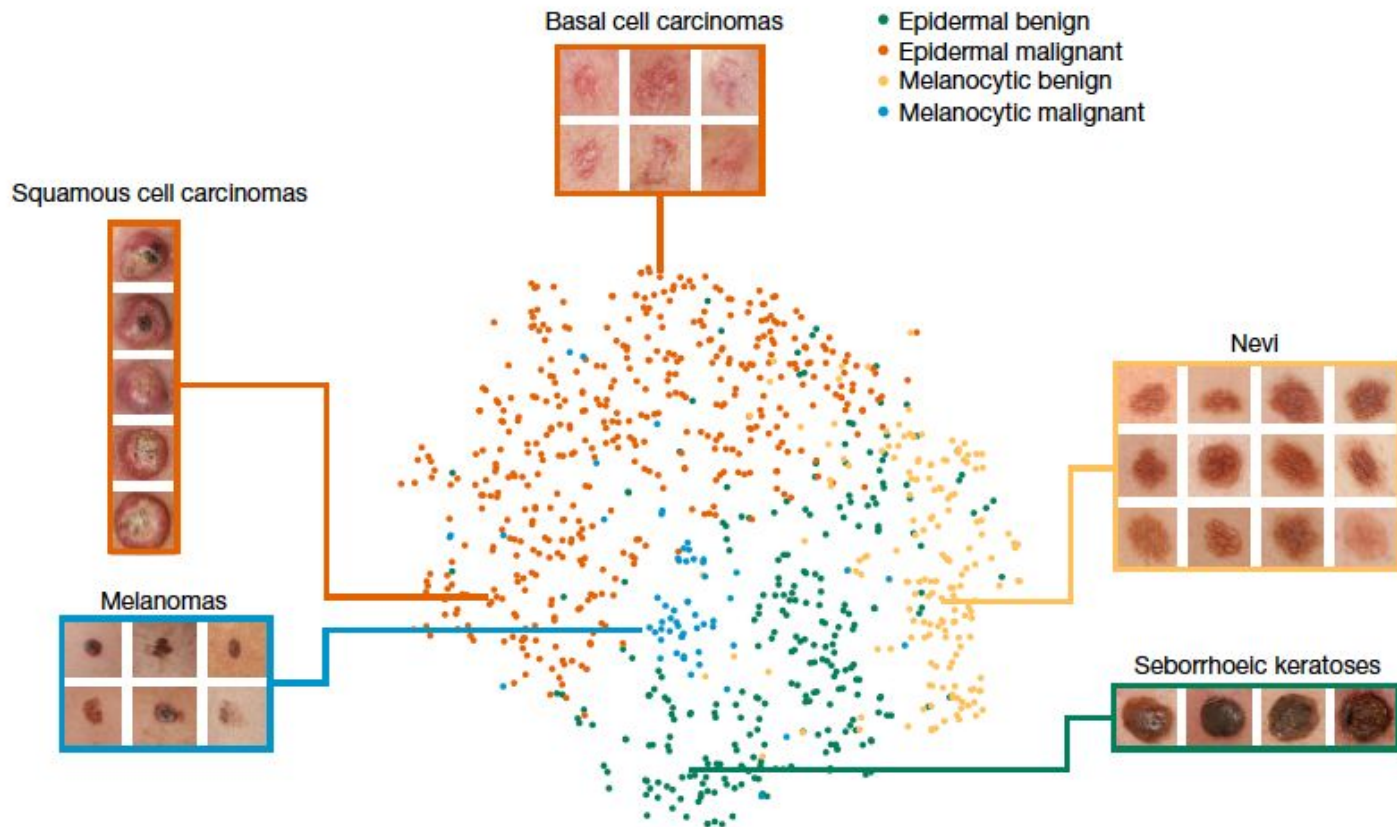
Training classes (757)

- Acrall-lentiginous melanoma
- Amelanotic melanoma
- Lentigo melanoma
- ...
- Blue nevus
- Halo nevus
- Mongolian spot
- ...
- ...
- ...

Inference classes (varies by task)

- 92% malignant melanocytic lesion
- 8% benign melanocytic lesion

# Visualización de capas ocultas con t-SNE



# ¡Gracias!

## Preguntas

[santiagorenteria25@gmail.com](mailto:santiagorenteria25@gmail.com) / [santiagorenterialab.com](http://santiagorenterialab.com)

