

Evolución de genes (codificantes de proteínas)

Evolución Orgánica

B-405

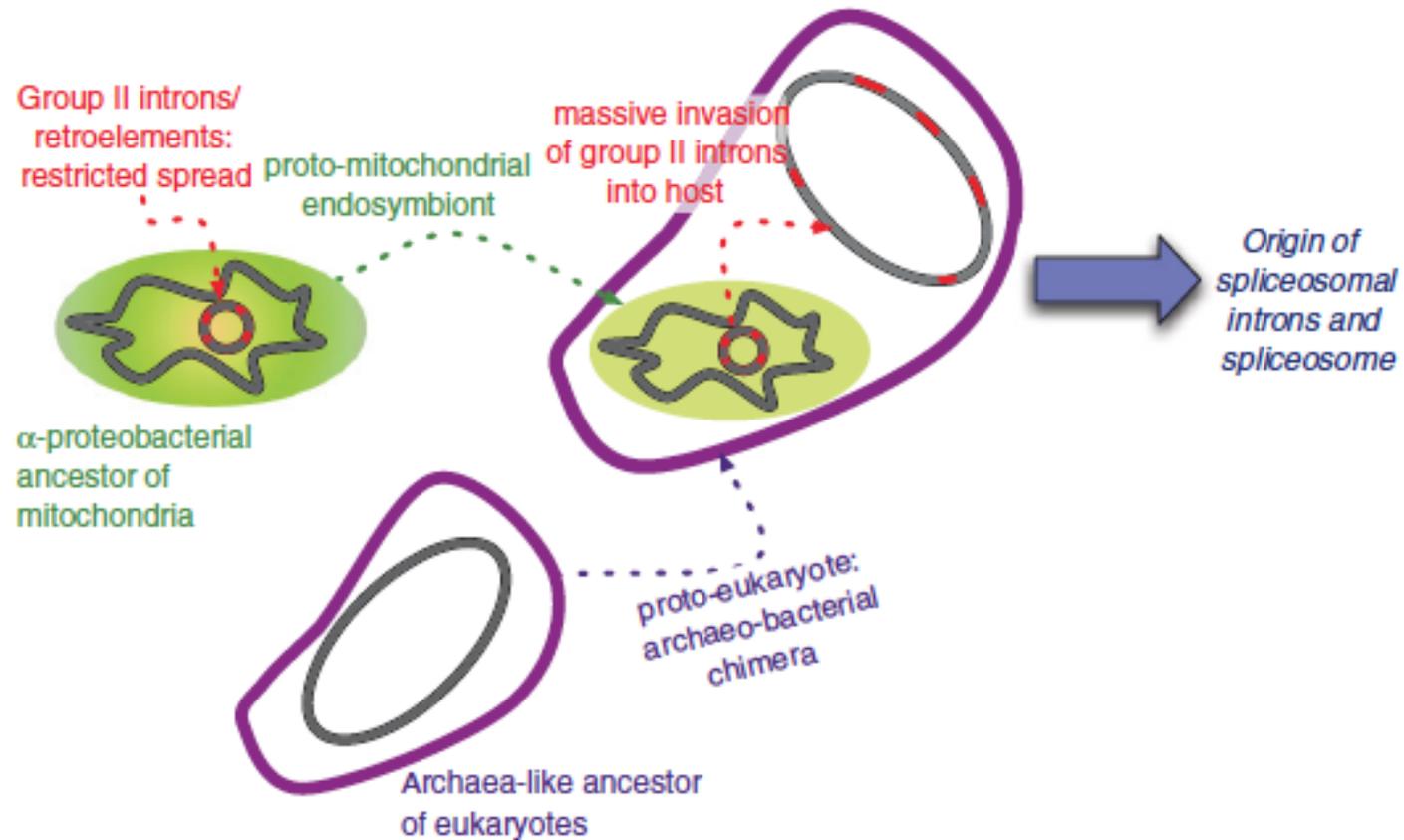
II Semestre 2024

Contenido de tema: Evolución de genes codificantes de proteínas

- Términos
 - Gene *de novo*
 - Duplicación/Intercambio de exones

Origen hipotético de genes eucarióticos

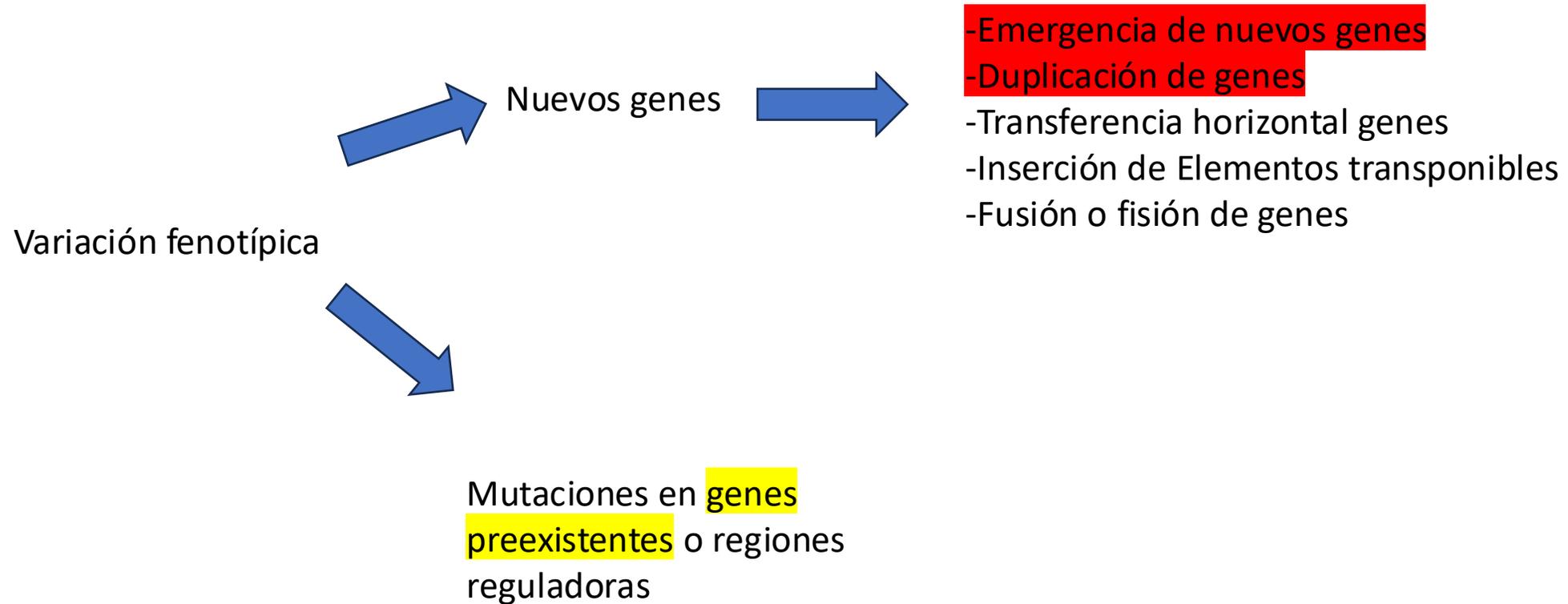
(Koonin, E.V., 2013)



Origen de los genes (dominios)

- Origen de los genes nuevos...generados a partir de secuencias no codificantes
- Hay datos describiendo la duplicación de genes y sus variantes pero no hay evaluaciones sobre los modelos de surgimiento de genes y sus predicciones.
- Cuáles son los mecanismos selectivos y evolutivos involucrados con el surgimiento, mantenimiento y evolución (“destino”) de los genes duplicados?

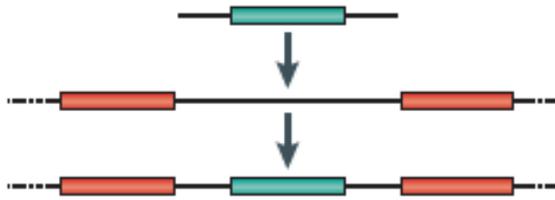
Origen de genes: Variación fenotípica



Mecanismo de origen de nuevos genes

(Chen et al. 2013)

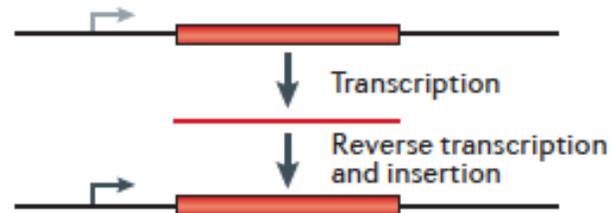
a Exon or domain shuffling



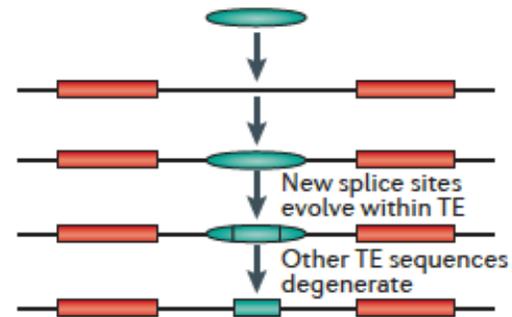
b Gene duplication



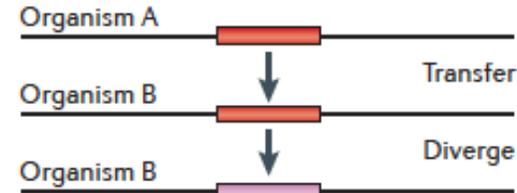
c Retrotransposition



d TE domestication



e Lateral gene transfer



f Gene fission or fusion



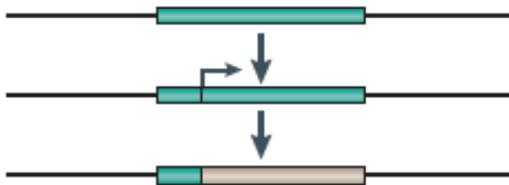
Mecanismo de origen de nuevos genes

(Chen et al. 2013)

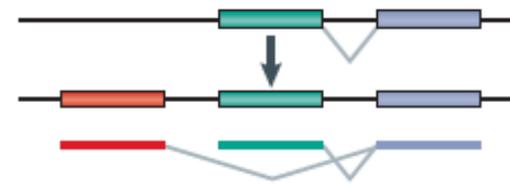
g *De novo origination*



h *Reading-frame shift*



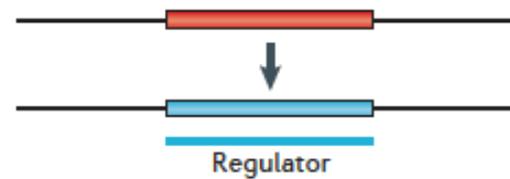
i *Alternative splicing*

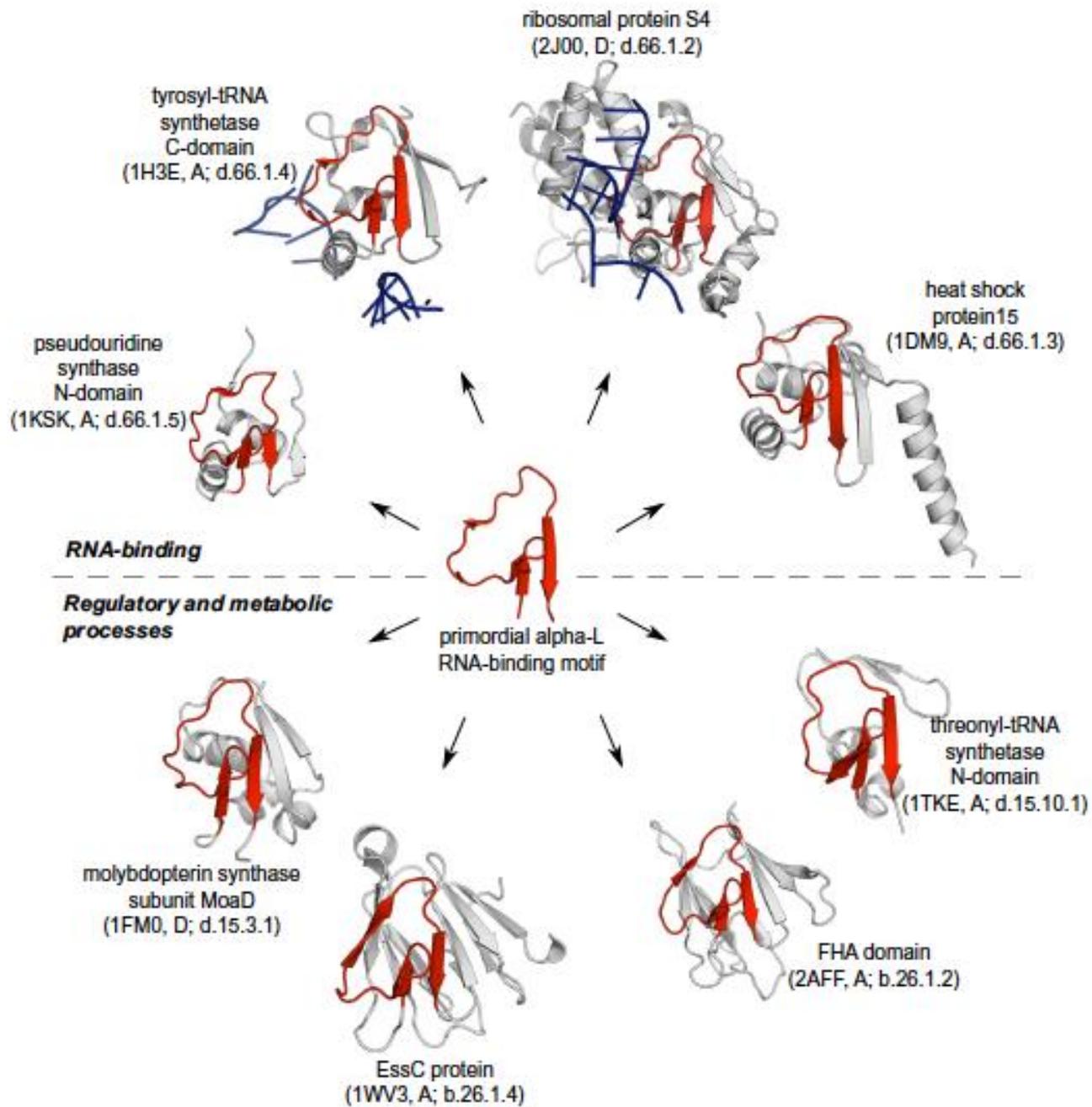


j *Non-coding RNA*

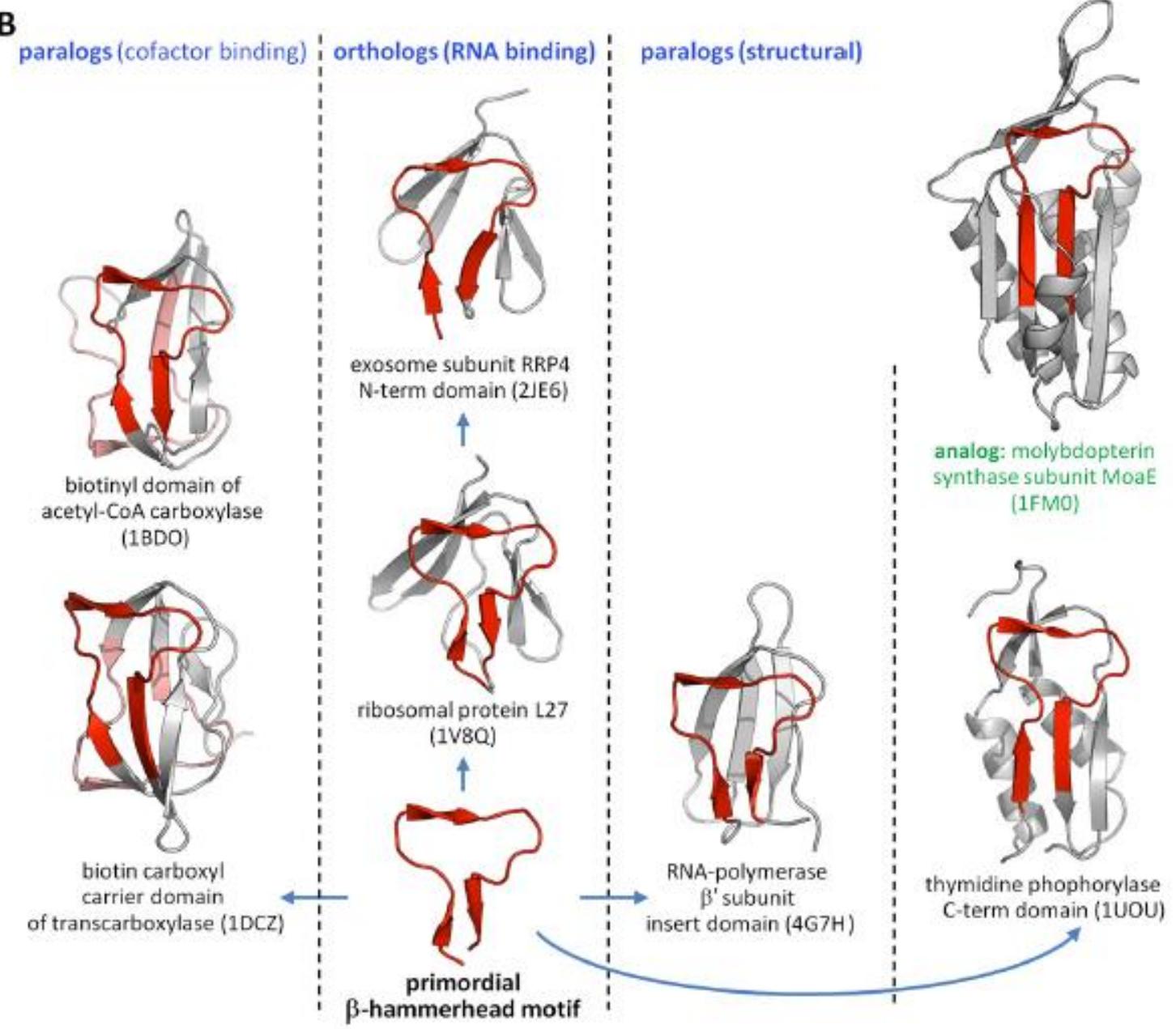


k *Pseudogene as RNA regulator*





Origen de genes a partir de dominios prototipos, se diversificaron las proteínas (Lupas & Alva, 2017)



Dominio primordial “ β -hammerhead” en cuatro plegamientos diferentes, abarca funciones muy diversas (Alva, V. et al., 2015)

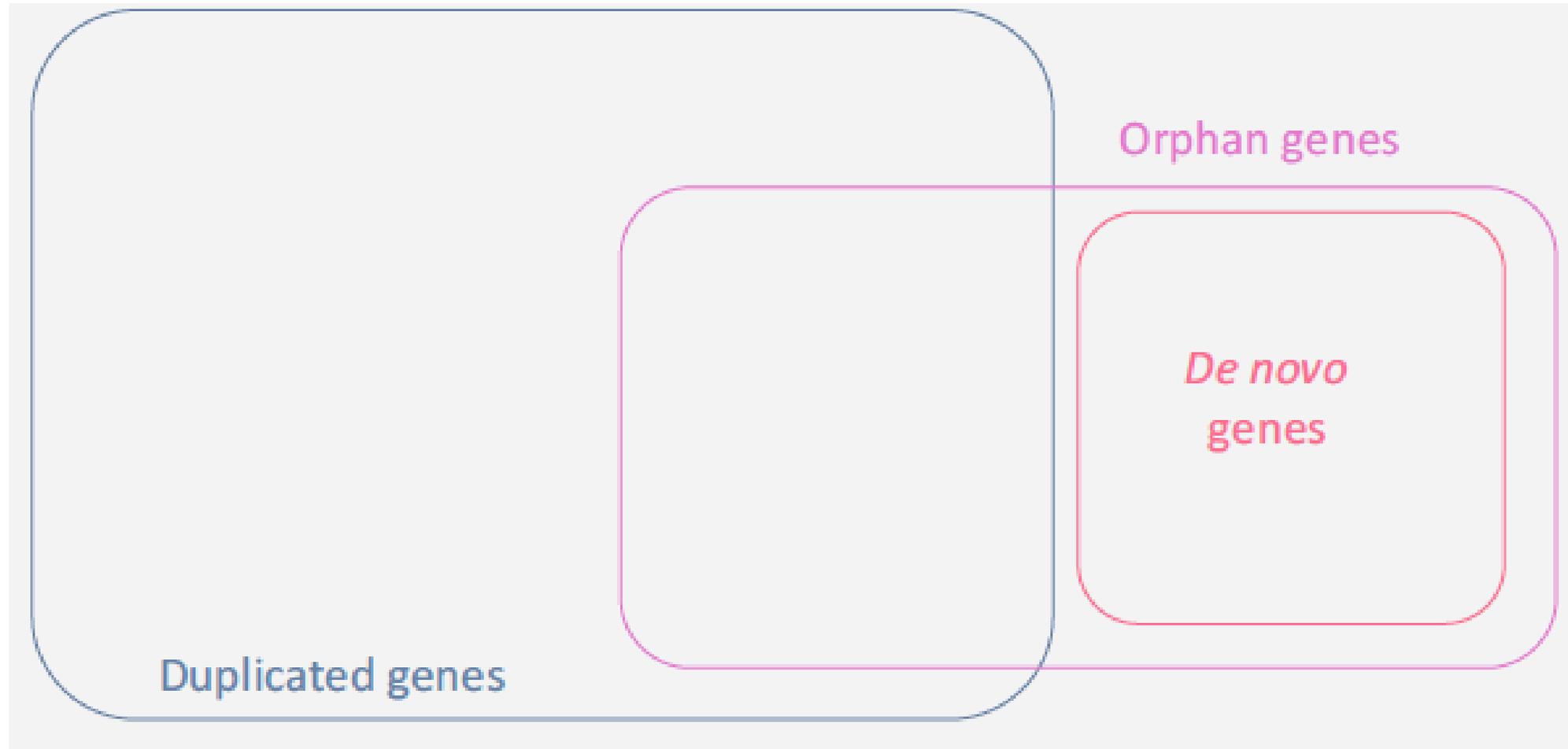
Ortólogo: secuencia homóloga en diferentes especies que se originan de ancestro común. (vía especiación)

Parólogo: secuencia cuyo ancestro común es Diferente (vía duplicación)

Origen de los genes (dominios)

- -**Genes “antiguos”**, son aquellos que han cambiado muy poco comparando grupos filogenéticamente poco derivados (tioredoxina)
- -**Genes nuevos**, son aquellos que han surgido recientemente de un linaje dado.
- **Genes huérfanos “nuevos”**, son genes donde **no** se han encontrado homólogos en otros linajes.

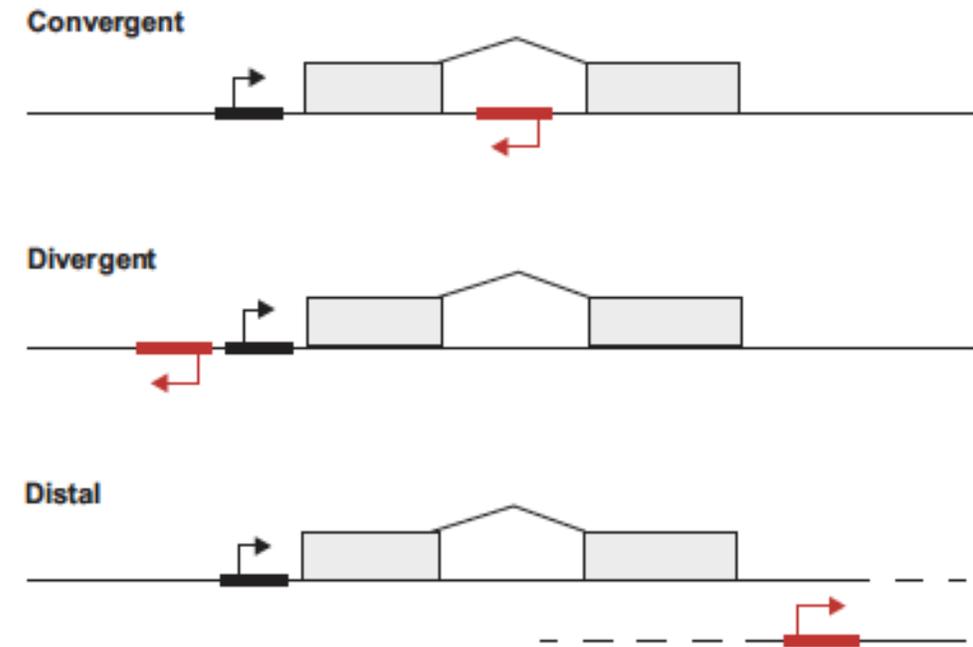
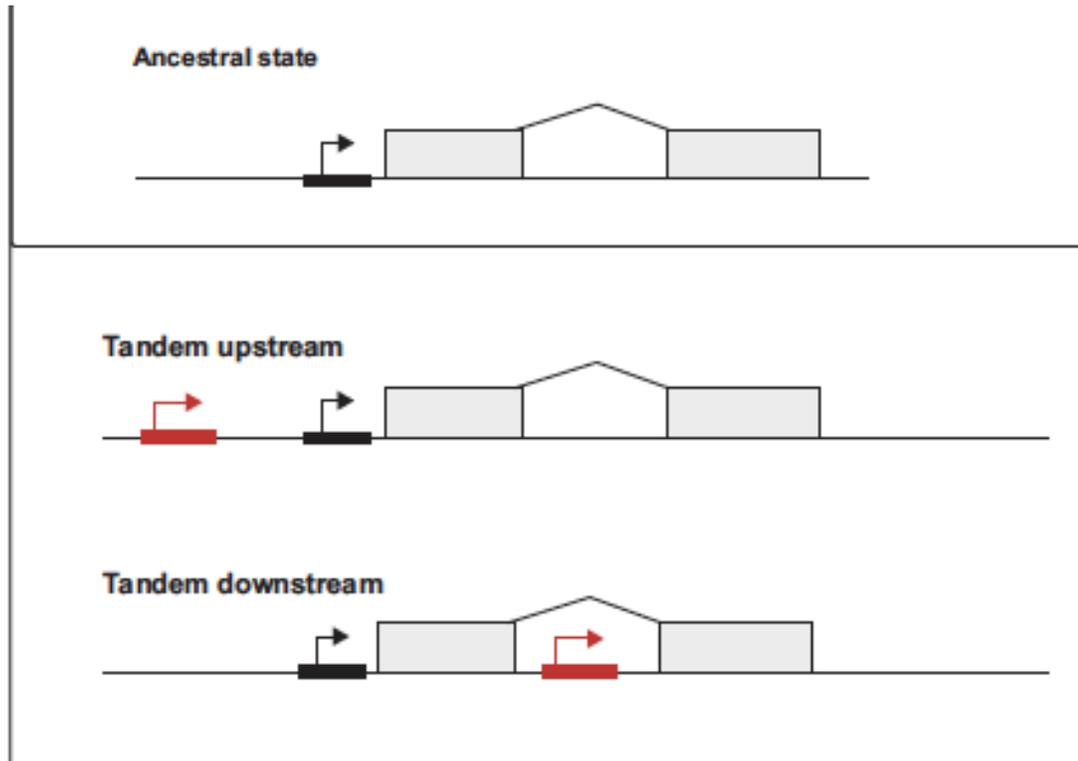
Nuevos genes? (Rödelspergel, Ch. et al. 2019)



Modelos sobre origen de genes de novo (Li et al. 2024)

	Cultivator model	ORF-first model	Transcription-first model
Origination of variation	Preexisting or rapidly evolvable adaptive variation is fixed in response to altered selective conditions.	Neutral variants exist prior to stress-induced transcription/translation, favorable variants increase in population frequency during stress	Neutral and/or selection for regulatory function
Exposure to selection	Regulation of cultivator gene(s) expression exposes de novo gene variants to selection	Induction of stress exposes cryptic protein-coding variants of unknown function to selection via promiscuous transcription and translation	Reduction of deleterious effects of noncoding RNA transcription may expose variants to selection. Alternatively, the appearance of favorable protein products of unknown functions exposes variants to selection.
Gain of transcription	Appears as by-product of <i>cis</i> -regulation of cultivator gene(s), becomes fixed by linkage	Appears due to stress induction, increases in population frequency and potentially becomes fixed based on adaptive value of proto-gene product	Assumed to exist prior to selection as enhancer RNA or bidirectional promoter
Gain of posttranscriptional processing	May appear based on adaptive value of <i>trans</i> -regulation of cultivator gene(s) by de novo transcript product likely via improved regulation of same cultivator gene(s). Potentially occurs through the "U1-PAS axis"	Gain of gene-like properties like gene length and exonic structures continues to evolve based on adaptive value of proto-gene product	Reduces deleterious variation through splicing, gains easily evolvable poly-A processing through chance via the "U1-PAS axis"
Gain of translation	Appears based on the adaptive value of <i>trans</i> -regulation of cultivator gene(s) by de novo protein-coding gene product, potentially via improved regulation of same cultivator gene(s)	Gain of gene-like properties continues to evolve based on adaptive value of proto-gene product	May appear based on the adaptive value of de novo protein-coding gene product

Posibles surgimientos de nuevos promotores (Li et al. 2024)

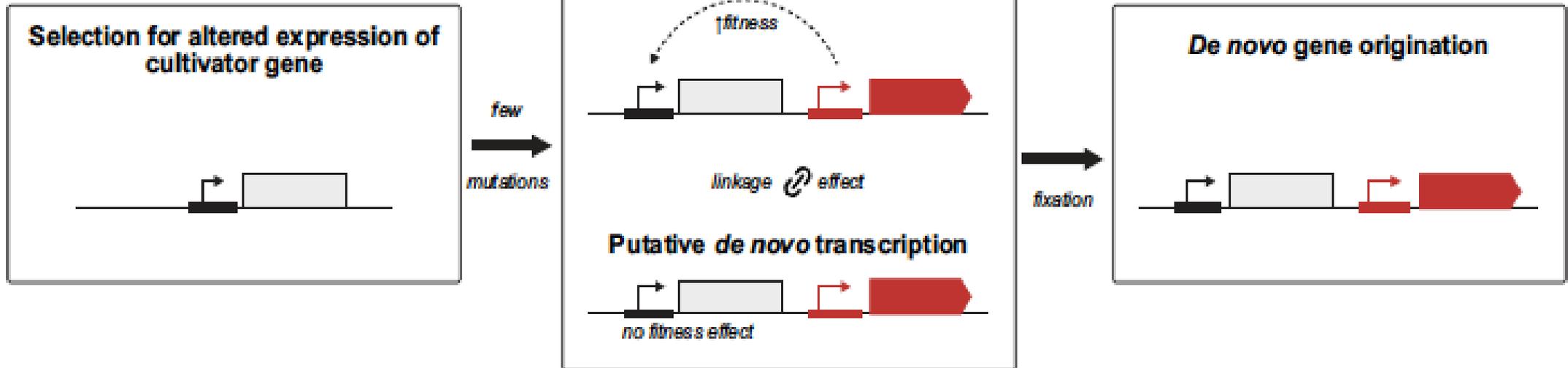


Condiciones selectivas y acumulación de sustituciones en regiones reguladoras/promotoras de un “gen semilla”

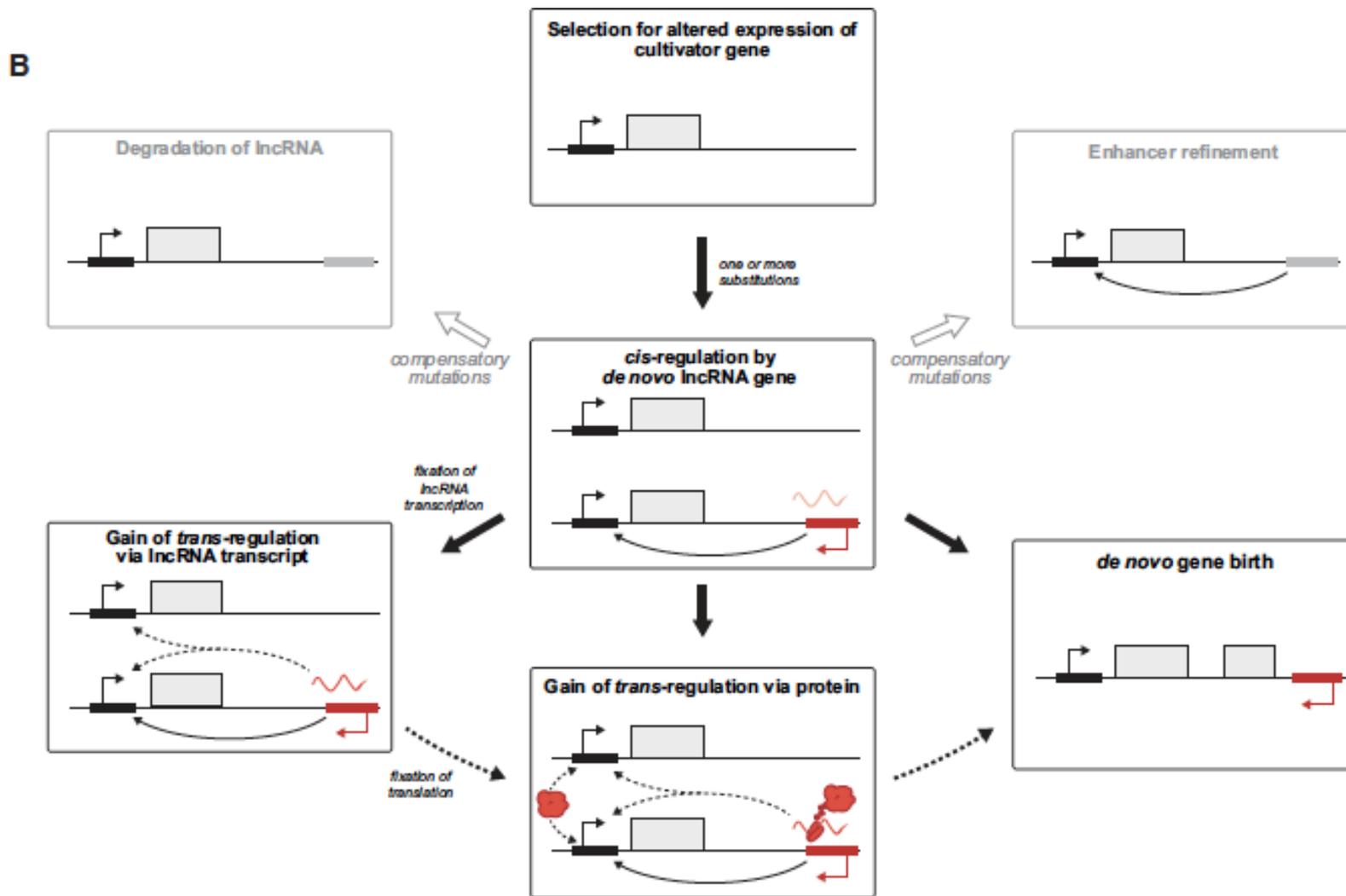
(Li et al. 2024)

A

Cultivator Model:
selection on cultivator gene



Múltiples opciones
potenciales para un
nuevo
cis gen ARNInc
(Li et al. 2024)



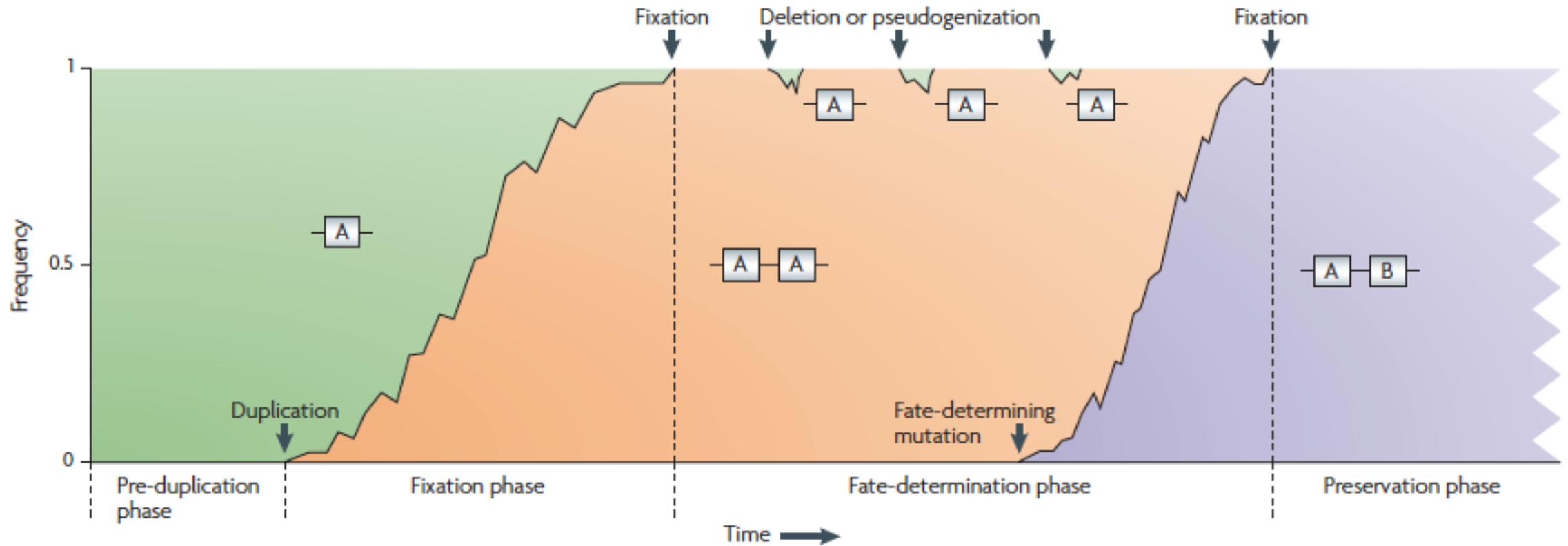
Ejemplos de genes nuevos en *Drosophila* (Chen et al. 2013)

New genes	Age (million years)	Origination mechanism	Expression	Phenotype	Function
<i>Drosophila spp.</i>					
<i>Sdic</i> family	0–3	DNA duplication	Testis	Sperm competition	Cytoskeleton
<i>sphinx</i>	0–3	Retrotransposition	Neuronal and reproductive tissue	Male courtship	ncRNA
<i>jingwei</i>	0–3	Retrotransposition	Testis	Recruitment pheromone and hormone	Alcohol metabolism
<i>p24-2</i>	0–3	DNA duplication	Multiple stages and tissues	Development, male reproduction	Protein trafficking
<i>Xcbp1</i>	3–6	Retrotransposition	Neuronal tissue	Foraging behaviour	Chaperone
<i>Zeus</i>	3–6	Retrotransposition	Male reproductive tissues	Male reproduction	Gene regulation
<i>nsr</i>	3–6	DNA duplication	Imaginal discs, sperm	Sperm maturation	Post-transcriptional regulation
<i>Sflc</i>	6–11	DNA duplication	Wide range of tissues	Male and female fecundity	Protease
<i>ms(3)K81</i>	6–11	DNA duplication	Sperm	Zygote development	Paternal chromosome telomere capping
<i>Umbrea</i>	6–11	DNA duplication	Multiple stages and tissues	Development	Chromatin binding
<i>Pros28.1A</i>	6–11	Retrotransposition	Multiple stages and tissues	Development	Signalling, protein regulation
<i>Desr</i>	11–25	Retrotransposition	Neuronal	Foraging behaviour	Transcription regulator
<i>MstProx</i> (also known as <i>Toll3</i>)	11–25	DNA duplication	Larva and pupa	Development	Immune response
<i>bicoid</i>	~150	Unclear	Early anterior embryo	Embryonic development	Maternal-effect morphogen
Mammals					
<i>FGF4</i>	~0.01	Retrotransposition	Distal humerus	Humerus development	FGF signalling
<i>SRGAP2C</i>	1.0–3.4	Partial DNA duplication	Brain	Predicted to affect cortex development in a mouse model	Unknown

Origen de los genes (dominios)

- Origen de los genes nuevos...generados a partir de secuencias no codificantes.
- **Hay datos describiendo la duplicación de genes y sus variantes pero no hay evaluaciones sobre los modelos de surgimiento de genes y sus predicciones.**
- **Cuáles son los mecanismos selectivos y evolutivos involucrados con el surgimiento, mantenimiento y evolución (“destino”) de los genes duplicados?**

Fases de preservación estable de un gen duplicado (Innan & Kondrashov, 2010)

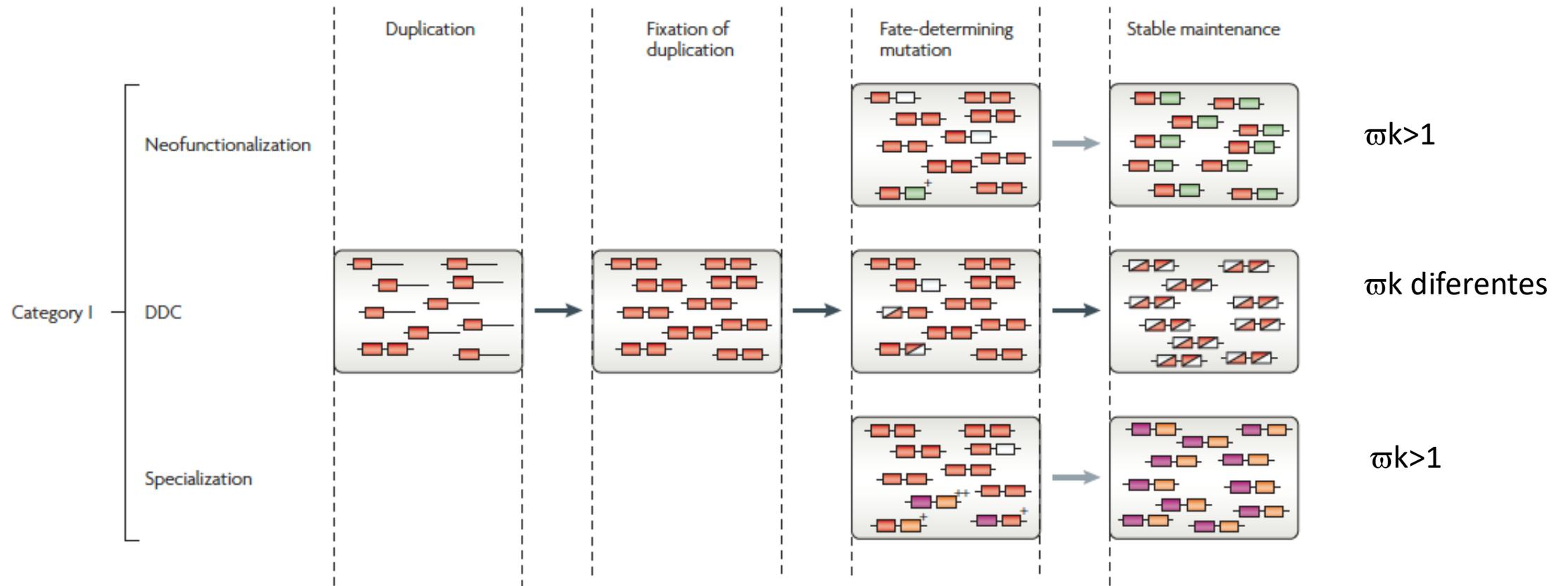


Indices para evaluar la fijación de un gen duplicado (Innan & Kondrashov, 2010)

- π : the average number of pairwise nucleotide differences
- K_S : rate of synonymous evolution, measured as the number of synonymous substitutions divided by the number of synonymous sites
- K_N : rate of non-synonymous evolution, measured as the number of non-synonymous substitutions divided by the number of non-synonymous sites
- π_S : synonymous level of polymorphism, measured as the average number of pairwise synonymous differences divided by the number of synonymous sites
- π_N : non-synonymous level of polymorphism, measured as the average number of pairwise non-synonymous differences divided by the number of non-synonymous sites
- ω_K : equal to K_N/K_S , representing the strength of selection on non-synonymous substitutions
- ω_π : equal to π_N/π_S , representing the strength of selection on non-synonymous polymorphisms

Modelos de genes duplicados por proceso selección neutral

(Innan & Kondrashov, 2010)



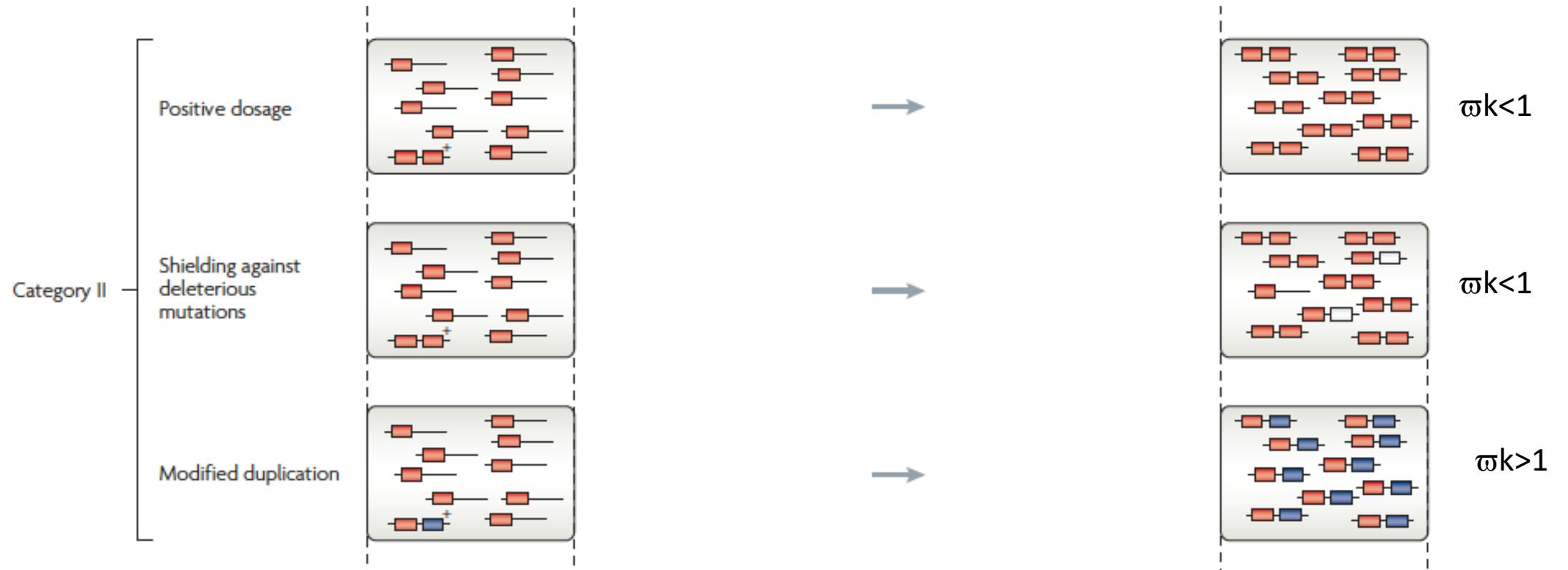
Modelos de genes duplicados por proceso selección neutral

(Innan & Kondrashov, 2010)

Name	Functional evolution			Fixation phase	Fate-determination phase		Preservation phase*	
	Function of original copy	Function of new copy	Fate-determining mutation	Selection on new copy	Selection on original copy	Selection on new copy	Molecular evolution in original copy	Molecular evolution in new copy
<i>Category I</i>								
Neofunctionalization	Kept	Novel	Gain-of-function mutations	Neutral	Purifying selection	Neutral	α	β
DDC	Subfunctionalized	Subfunctionalized	Loss-of-function mutations	Neutral	Relaxed purifying selection	Relaxed purifying selection	β	β
Specialization or EAC	Subfunctionalized	Subfunctionalized	Gain-of-function mutations	Neutral	Relaxed purifying selection	Relaxed purifying selection	β	β

Modelos de genes duplicados con efecto ventajoso

(Innan & Kondrashov, 2010)

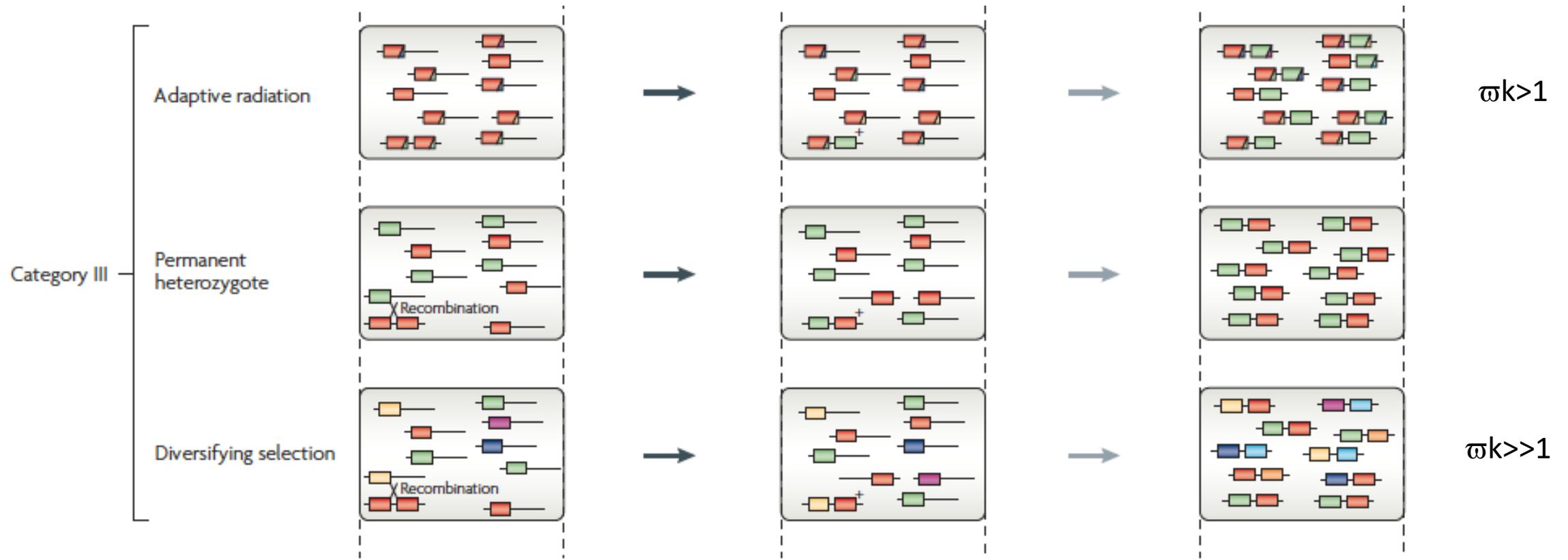


Modelos de genes duplicados con efecto ventajoso (Innan & Kondrashov, 2010)

Name	Functional evolution			Fixation phase	Fate-determination phase		Preservation phase*	
	Function of original copy	Function of new copy	Fate-determining mutation	Selection on new copy	Selection on original copy	Selection on new copy	Molecular evolution in original copy	Molecular evolution in new copy
<i>Category II</i>								
Positive dosage	Kept	Same as original	NA	Positive selection on duplication	NA	NA	α'	α'
Shielding against deleterious mutations	Kept	Same as original	NA	Positive selection on duplication	Relaxed purifying selection	Relaxed purifying selection	NA	NA
Modified duplication	Kept	Novel	Gain-of-function mutations	Positive selection on duplication	NA	NA	α	β

Categoría II a) Dosis, ejemplos: Genes respuesta a estres, ribosomale e histonas

Modelos de genes duplicados con variación génica (Innan & Kondrashov, 2010)



Modelos de genes duplicados con variación génica

(Innan & Kondrashov, 2010)

Name	Functional evolution			Fixation phase	Fate-determination phase		Preservation phase*	
	Function of original copy	Function of new copy	Fate-determining mutation	Selection on new copy	Selection on original copy	Selection on new copy	Molecular evolution in original copy	Molecular evolution in new copy
<i>Category III</i>								
Permanent heterozygote	Subfunctionalized	Subfunctionalized	Gain-of-function mutations	Positive selection on pre-duplicational variation	NA	NA	β	β
Adaptive radiation model	Kept	Novel	Gain-of-function mutations	Positive selection on pre-duplicational variation	NA	NA	α	β
Diversifying selection	Multiple functions	Multiple functions	Gain-of-function mutations	Positive selection on pre-duplicational variation	NA	NA	\circ	\circ

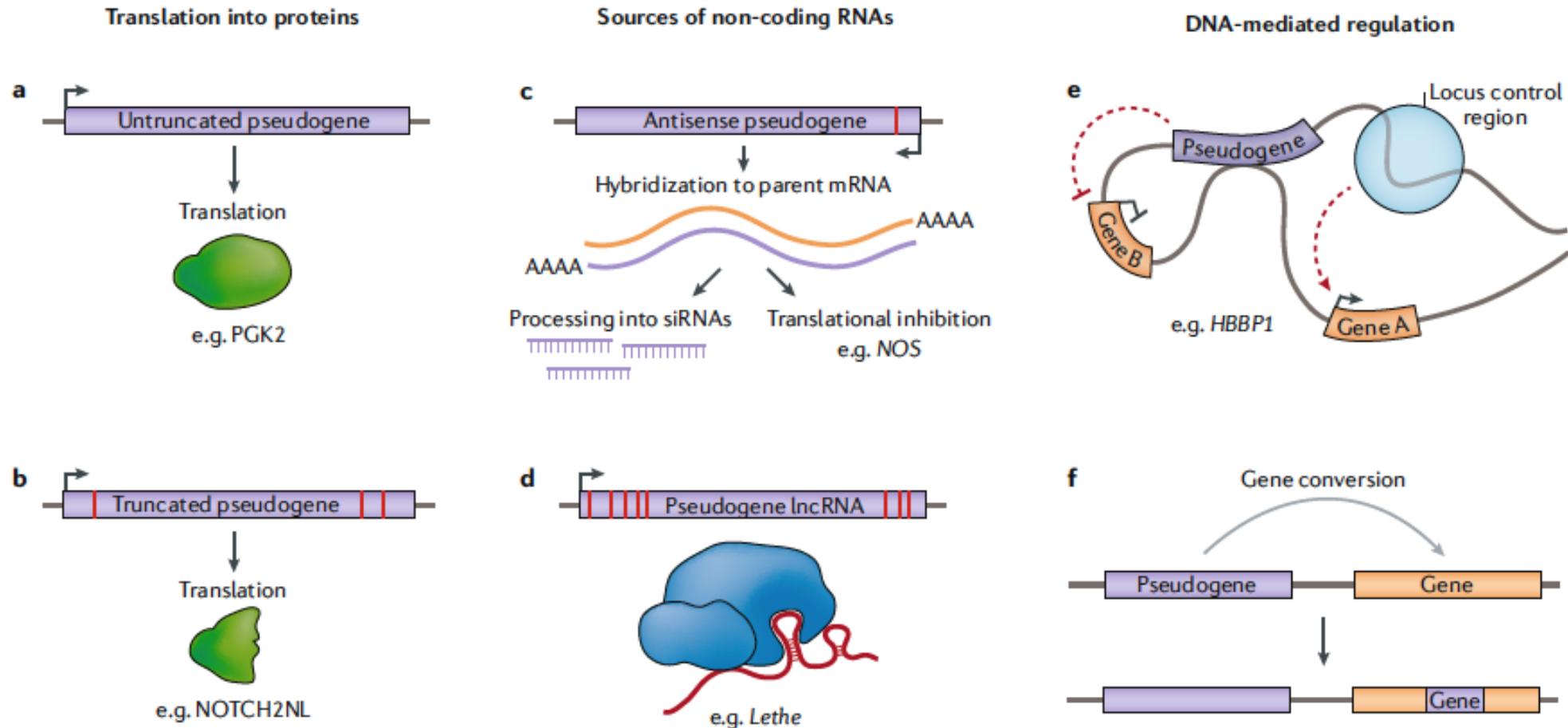
Categoría III con ejemplos como los genes MHC

Los pseudogenes son la transición entre una secuencia defectuosa y una funcional?

- Definición:
 - -Cualquier secuencia genómica que es similar a otro gen y es defectuoso
 - Se han identificado en genomas de metazoos unos 10 000 a 20 000 regiones pseudogénicas.
- Pseudogenes funcionales:
 - Secuencias donde se identificó la región parental funcional y tiene un impacto biológico o función
 - Se han identificado cinasas, nitrato sintasas, receptores de membrana, proteínas ribosomales y globinas.

Pseudogenes funcionales en eucariotas

(Cheetham, S.W. et al., 2019)

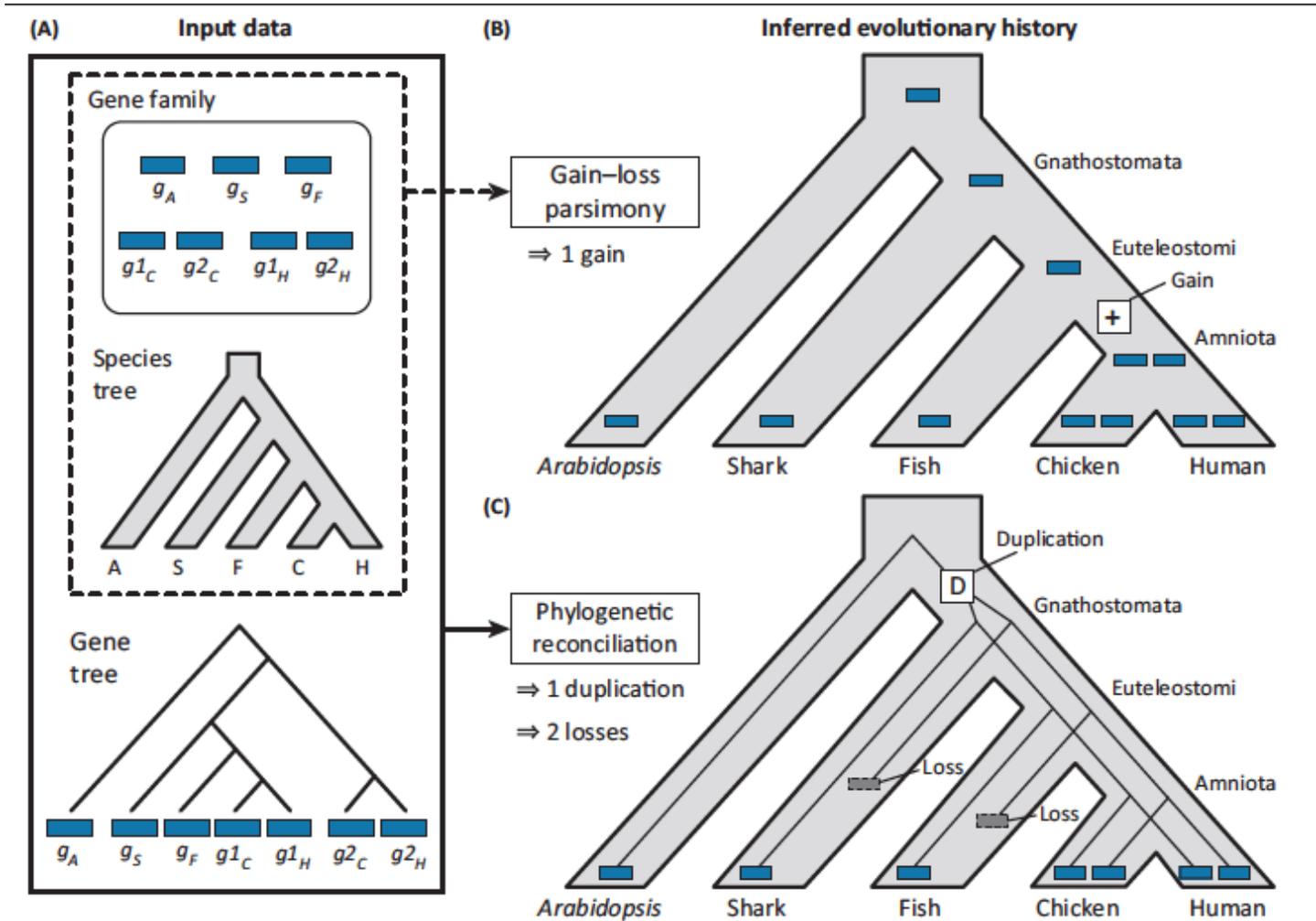


Estudio de Caso: Familia g2 de la Fosfolipasa A2

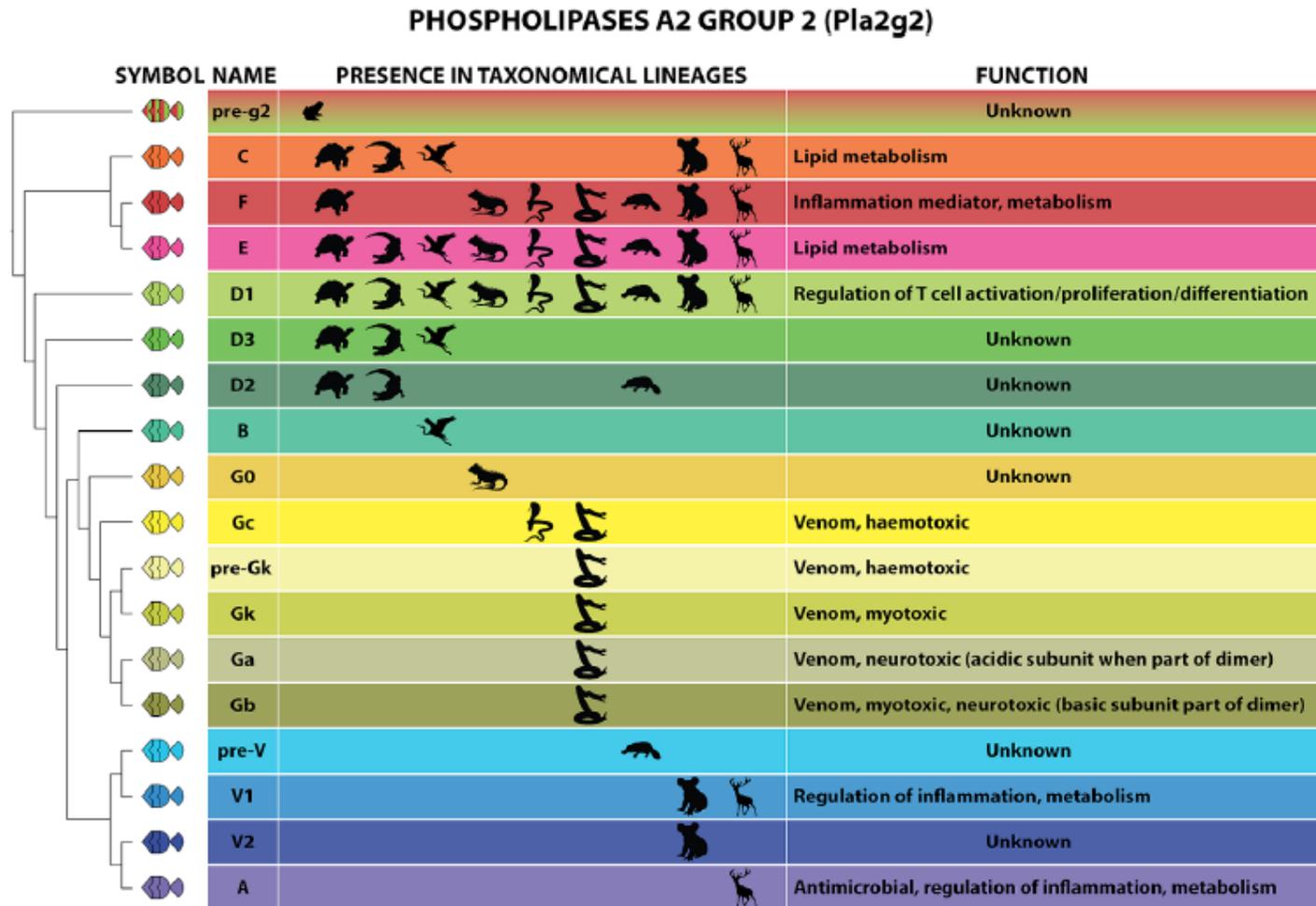
(Koluradov, I. et al., 2019)

- Se analizaron 90 genomas
- Posee múltiples roles funcionales:
 - Inmunidad en mamíferos
 - Venomas en serpientes
- *Los venenos son rasgos funcionales y ecológicos*
- El cluster de genes ubicado entre dos genes *otuD3* y *ubxN10*; en una región sinténica de 10 genes flanqueantes
- La expansión de la familia de genes (aprox. 14 miembros) son extracelulares y "exoquímicas"

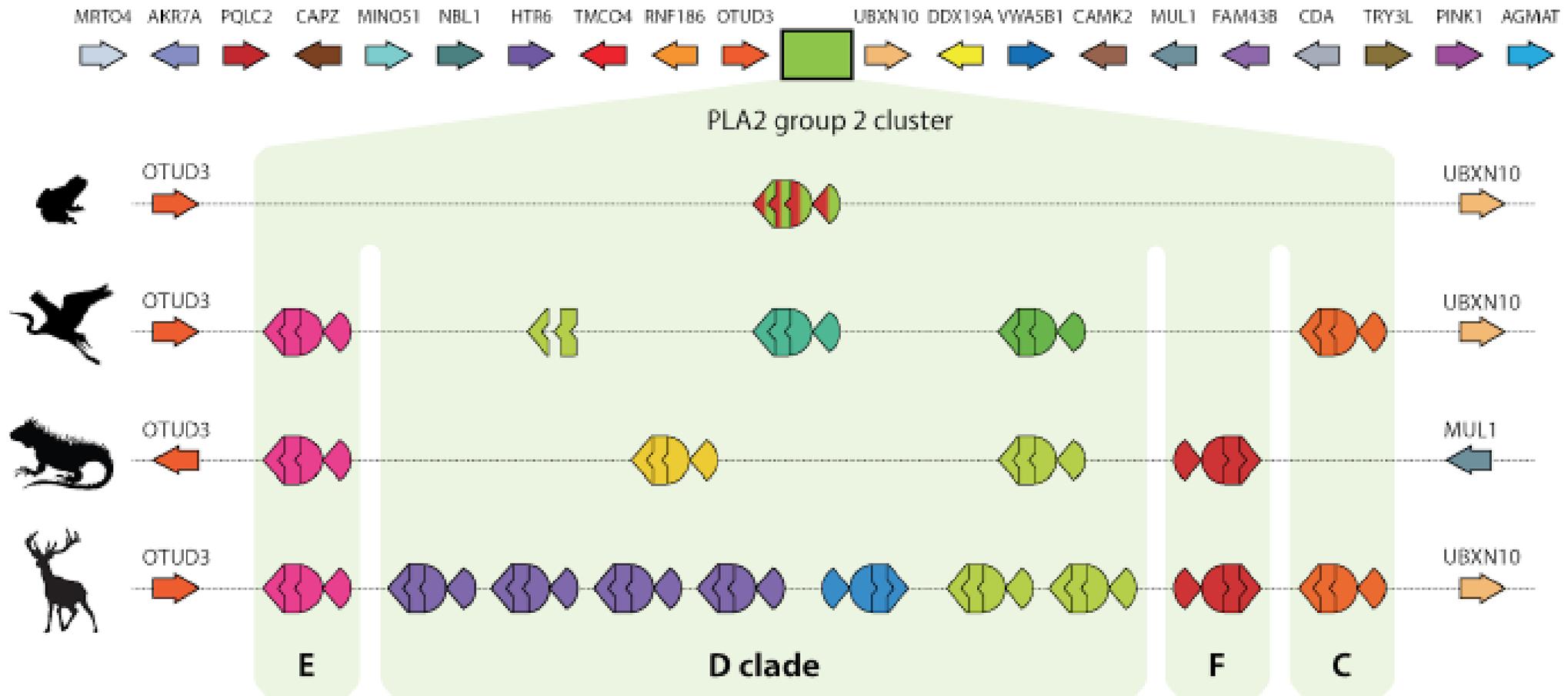
La reconciliación filogenética para evitar error de método ganancia-pérdida (Capra, J.A. et al., 2013)



Familia g2 con sus relaciones y funciones de la fosfolipasa A2 (Koluradov, I. et al., 2019)



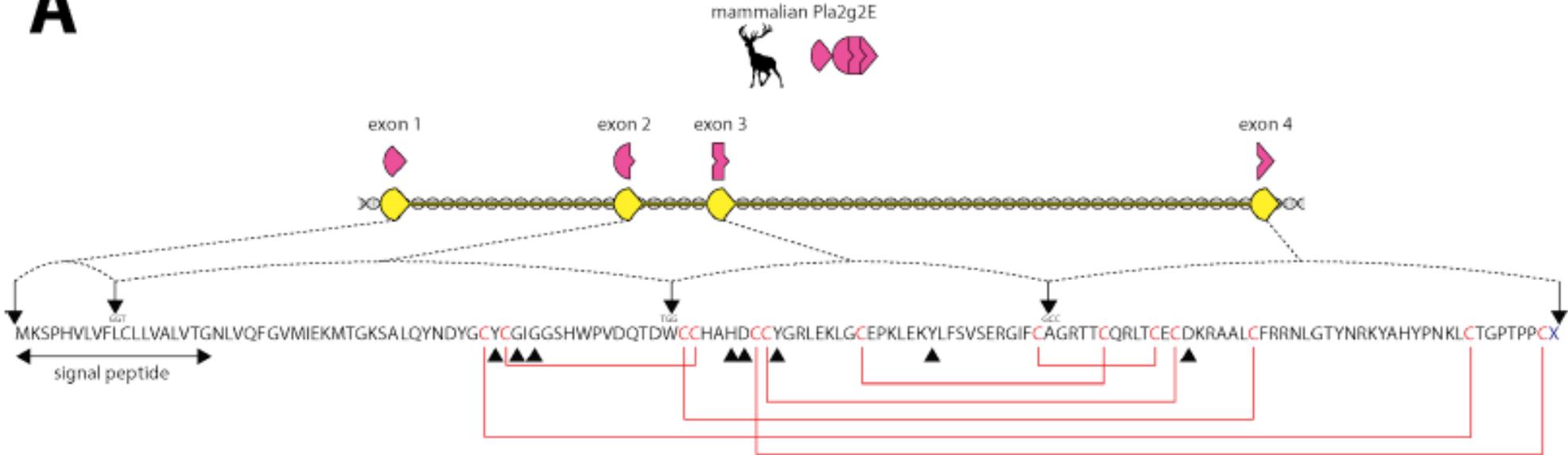
Localización del cluster Pla2g2 en anfibios, aves, esquamata y mamíferos (Koluradov, I. et al., 2019)



Estructura del gen Pla2g2E en vertebrados

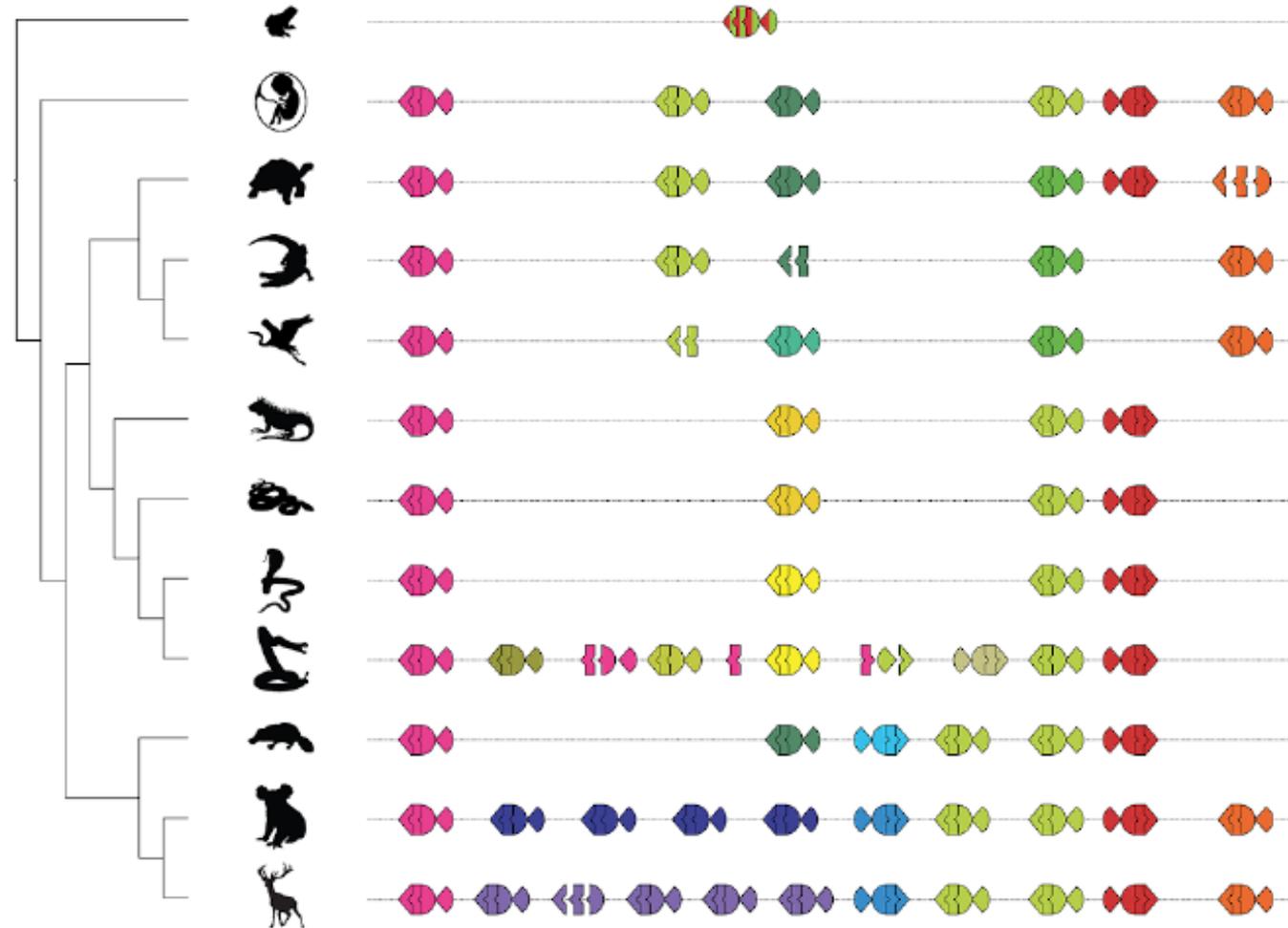
(Koluradov, I. et al., 2019)

A

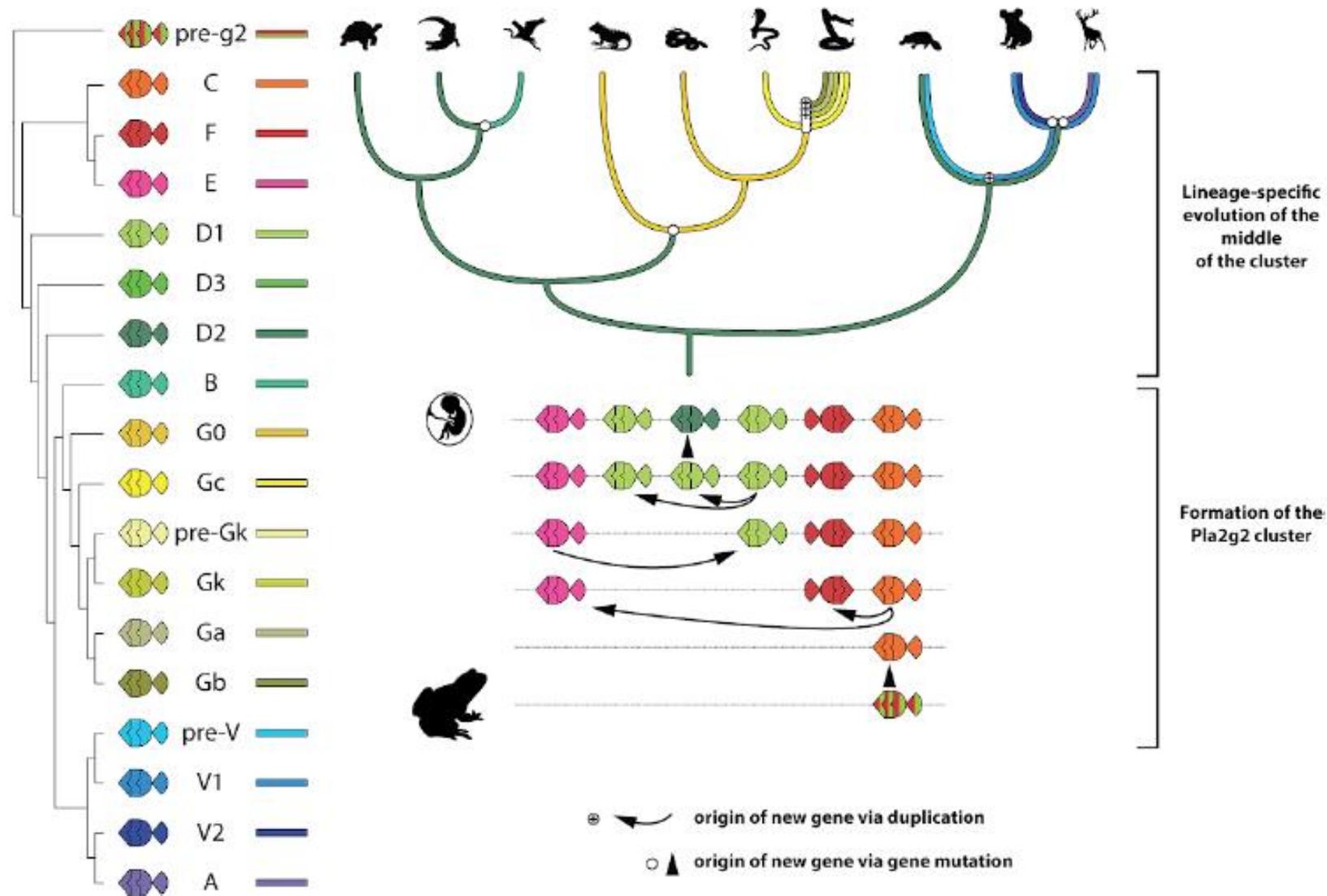


Expansión del cluster Pla2g2 en tetrápodos

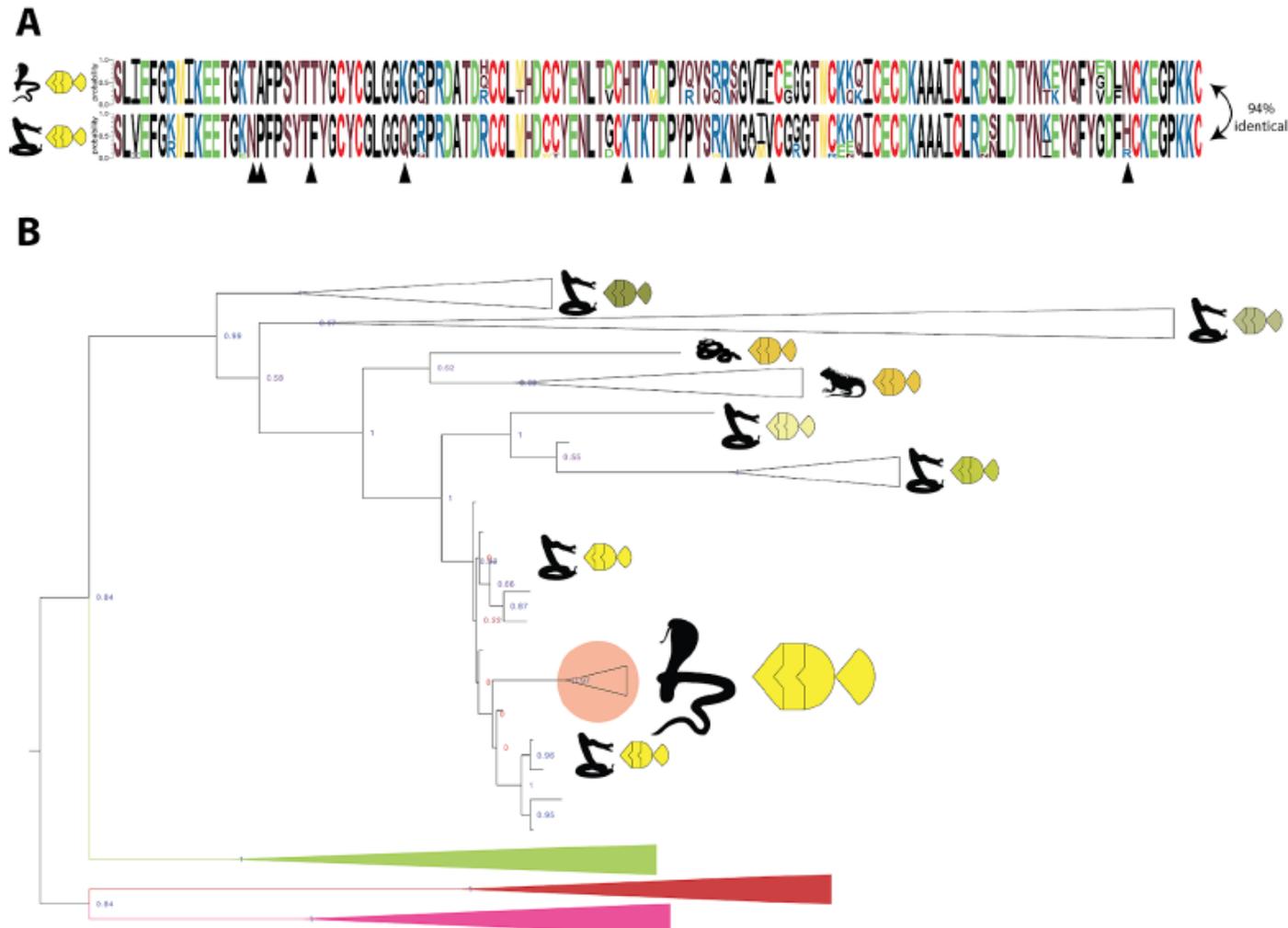
(Koluradov, I. et al., 2019)



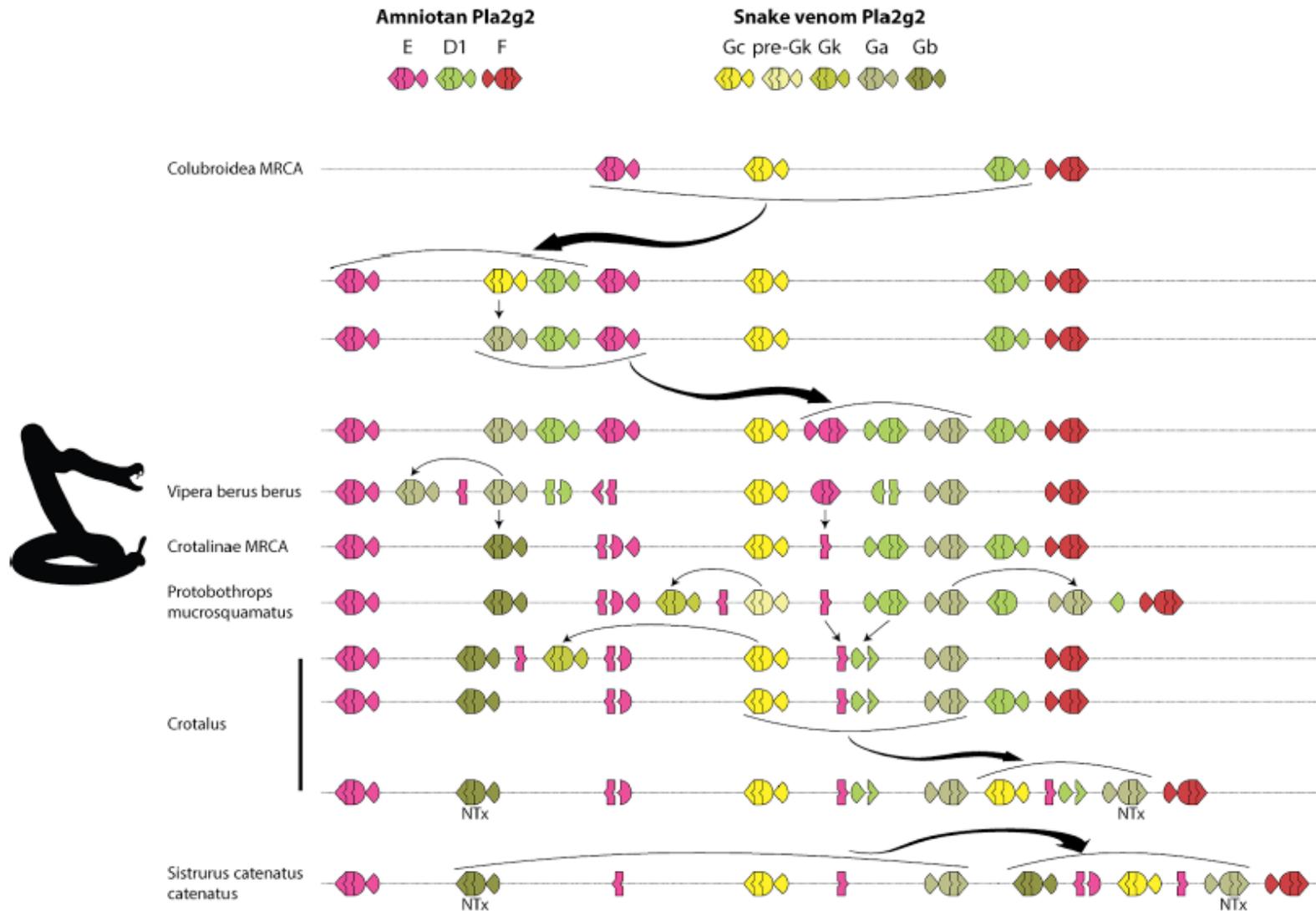
Evolución del cluster Pla2g2 (Koluradov, I. et al., 2019)



Secuencias consenso entre elápidos y vipéridos g2Gc (Koluradov, I. et al., 2019)



Evolución del cluster Pla2g2 (Koluradov, I. et al., 2019)

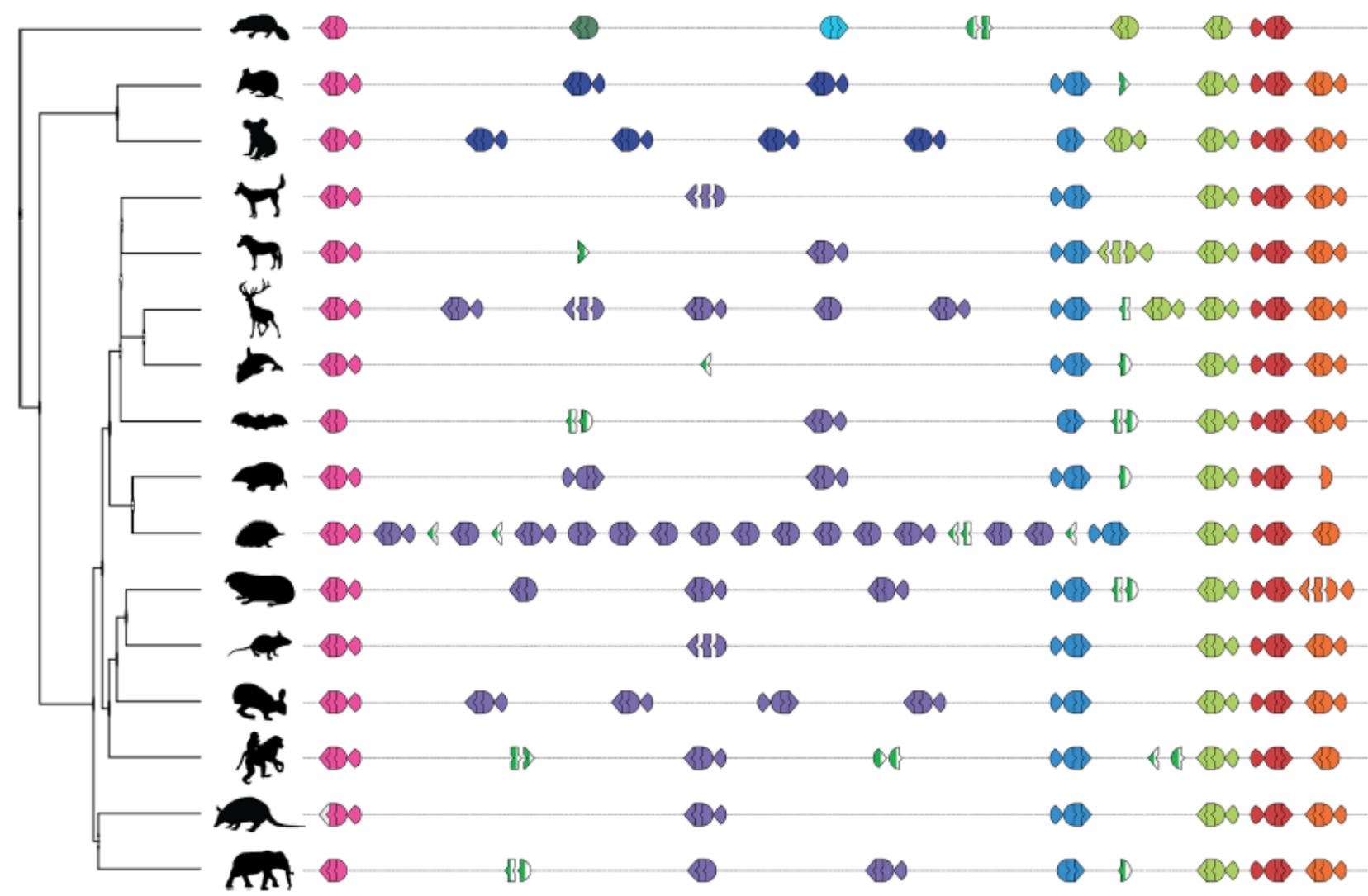


Amniotan Pla2g2

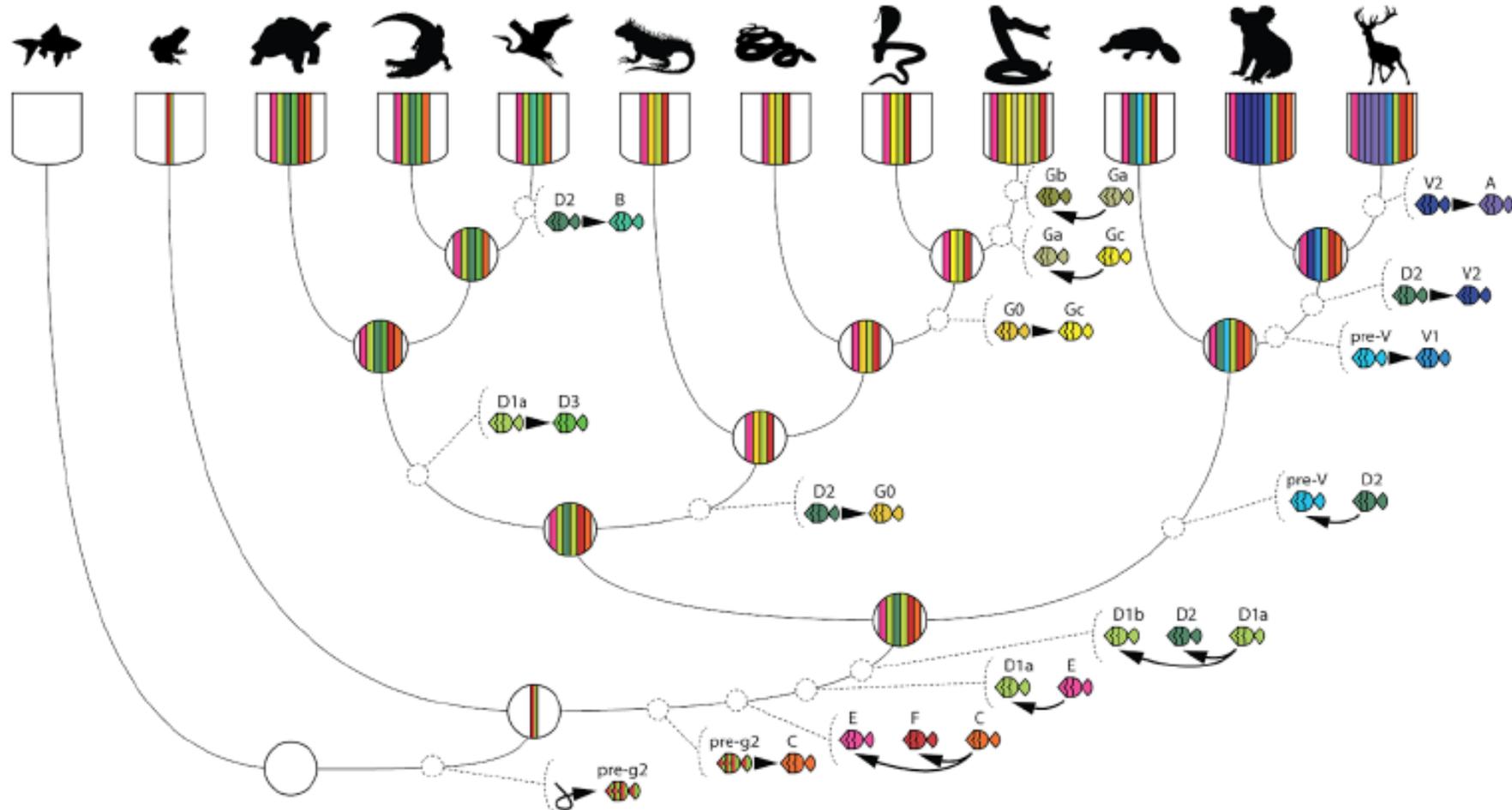
E D2 D1 F C

Mammalian Pla2g2

pre-V V2 A V1

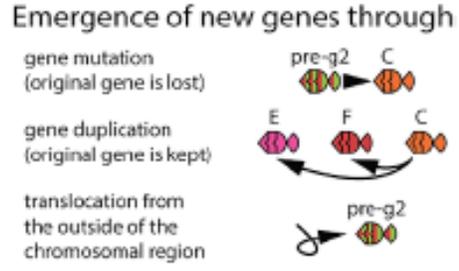
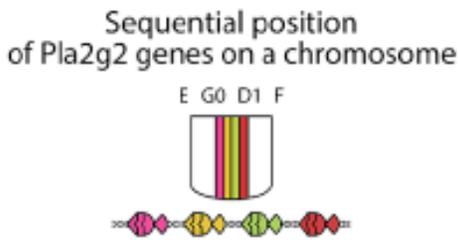
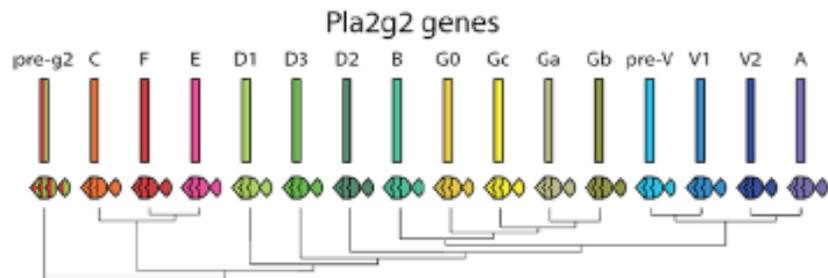


Historia evolutiva reciente del clado g2v en mamíferos (Koluradov, I. et al., 2019)



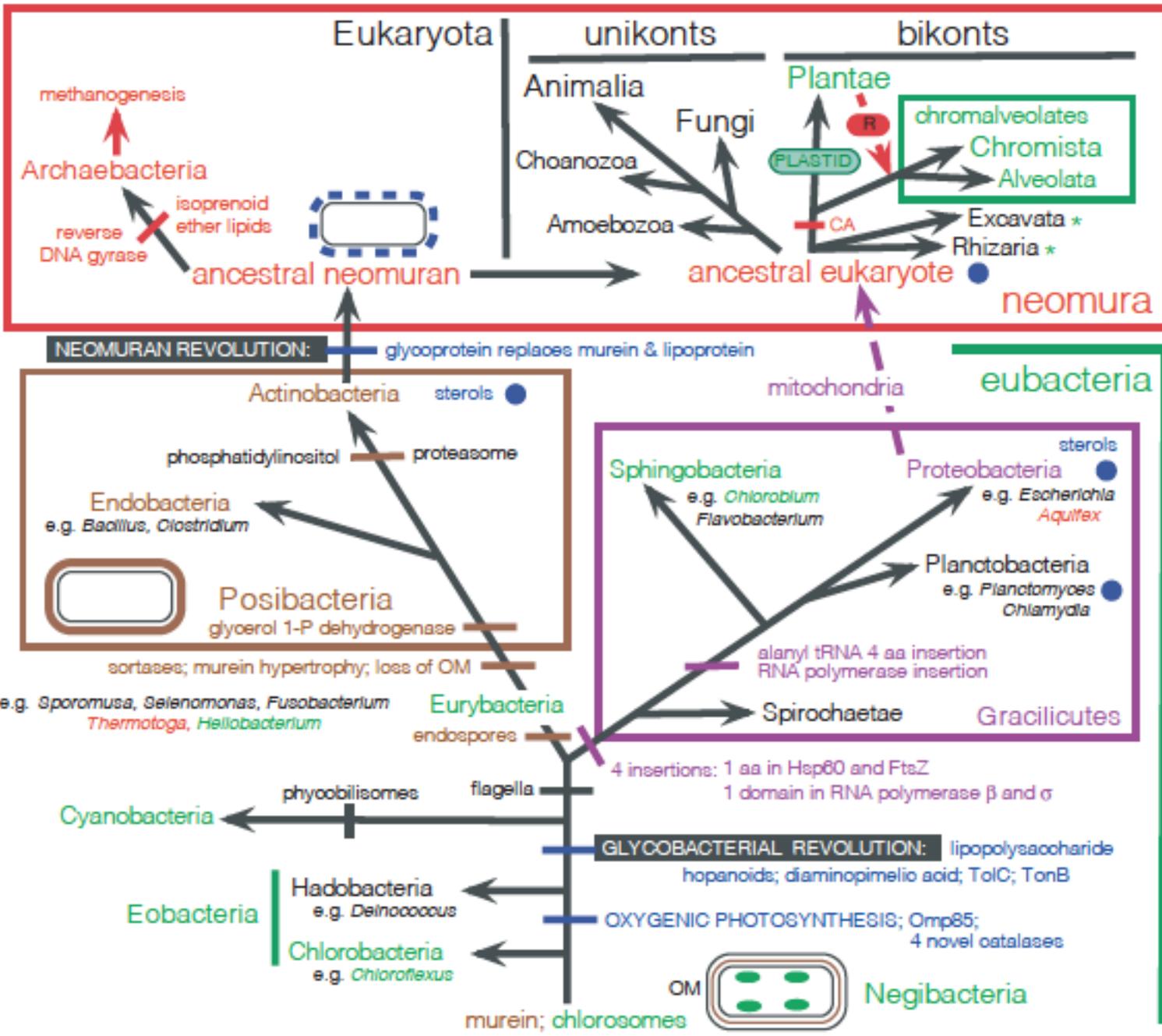
Propuesta evolutiva de Pla2g2

(Koluradov, I. et al., 2019)



Qué modelo explica mejor los datos de la familia de genes de la fosfolipasa A2?

- 1.-Neofuncionalidad
- 2.-Duplicación
- 3.-Subfuncionalidad
- 4.-Genes bifuncionales
- 5.- Opción 2 y luego 1
- 6.- Opción 1 y luego 2
- 7.-Opciones 1 al 4
- 8.-Opción 3 y luego 2?
- 9.-Modelo Pluralista?



Árbol de la vida

(Cavalier-Smith, T. 2006)

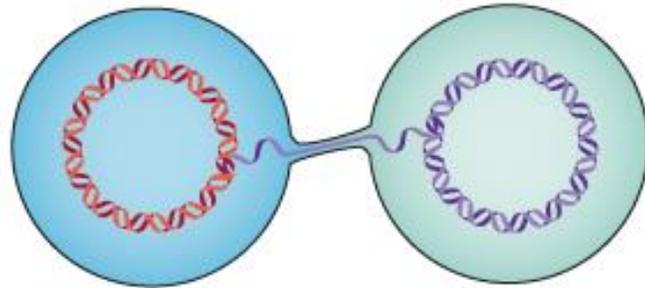
Evidencia de Transferencia Horizontal de ADN

- Evolución de bacterias por mecanismos de movimiento de ADN
- Secuencias mitocondriales al núcleo en plantas
- Movimiento de elementos transponibles entre especies de animales
- Genes derivados de THG con funciones varias en plantas
- Genes derivados de THG funcionales en vías metabólicas de eucariotas

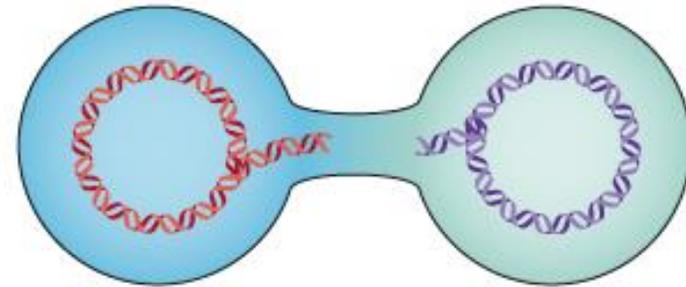
Transferencia horizontal de genes (THG)

(Soucy, S.M. et al. 2015)

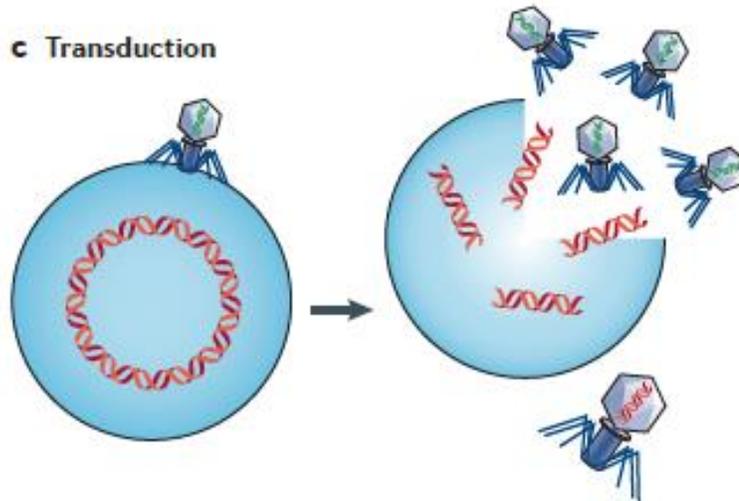
a Conjugation



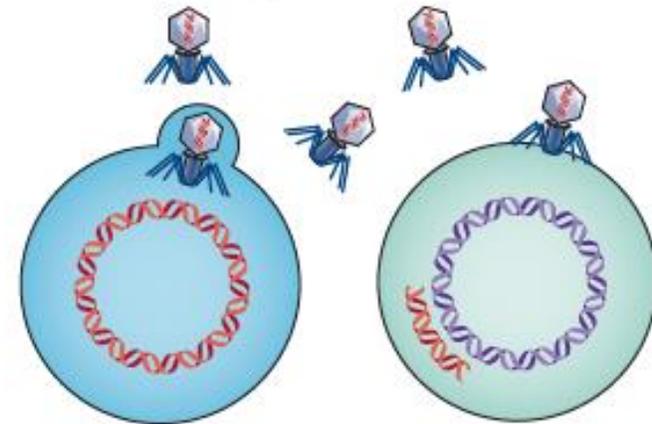
b Cell fusion



c Transduction

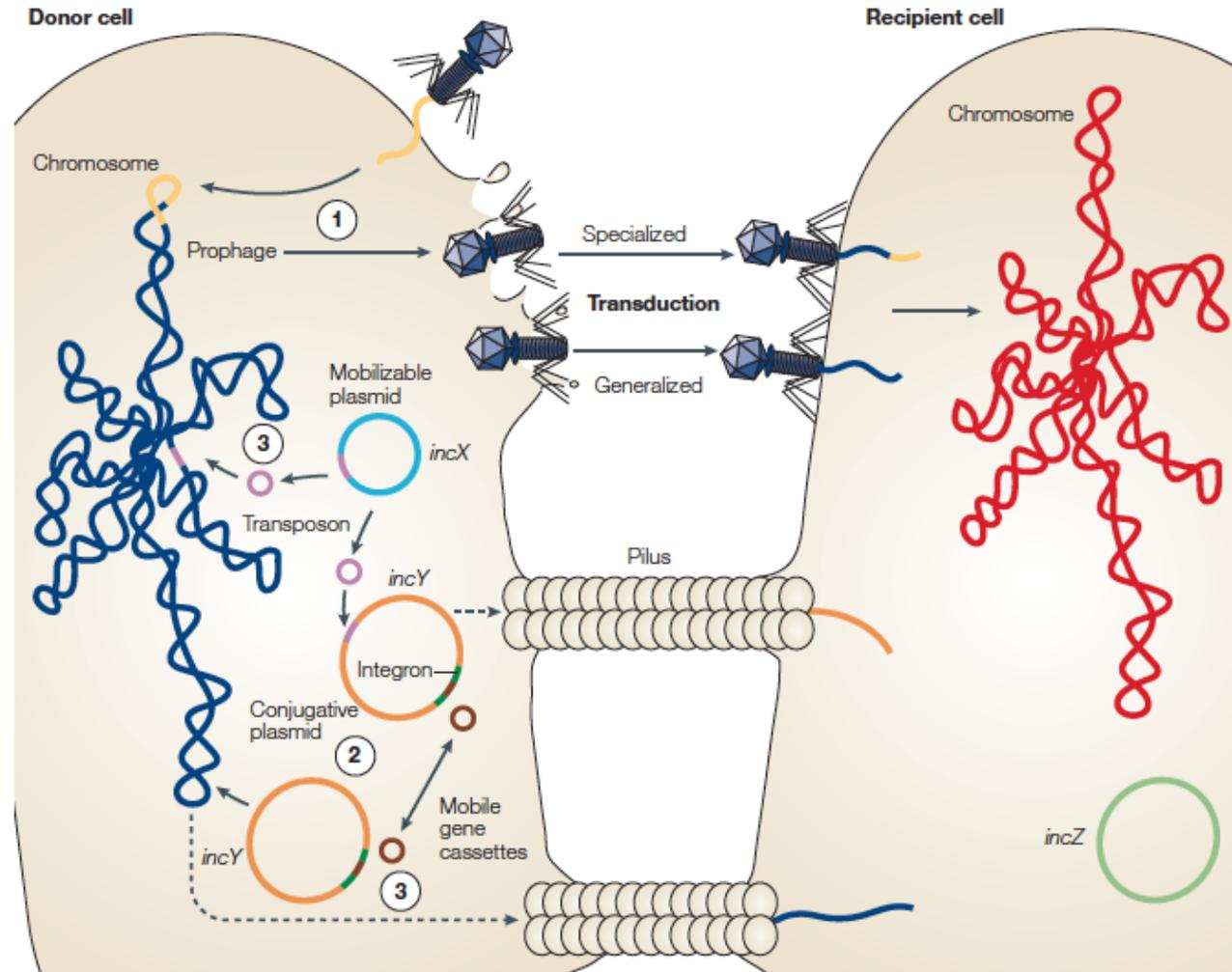


d Gene transfer agents

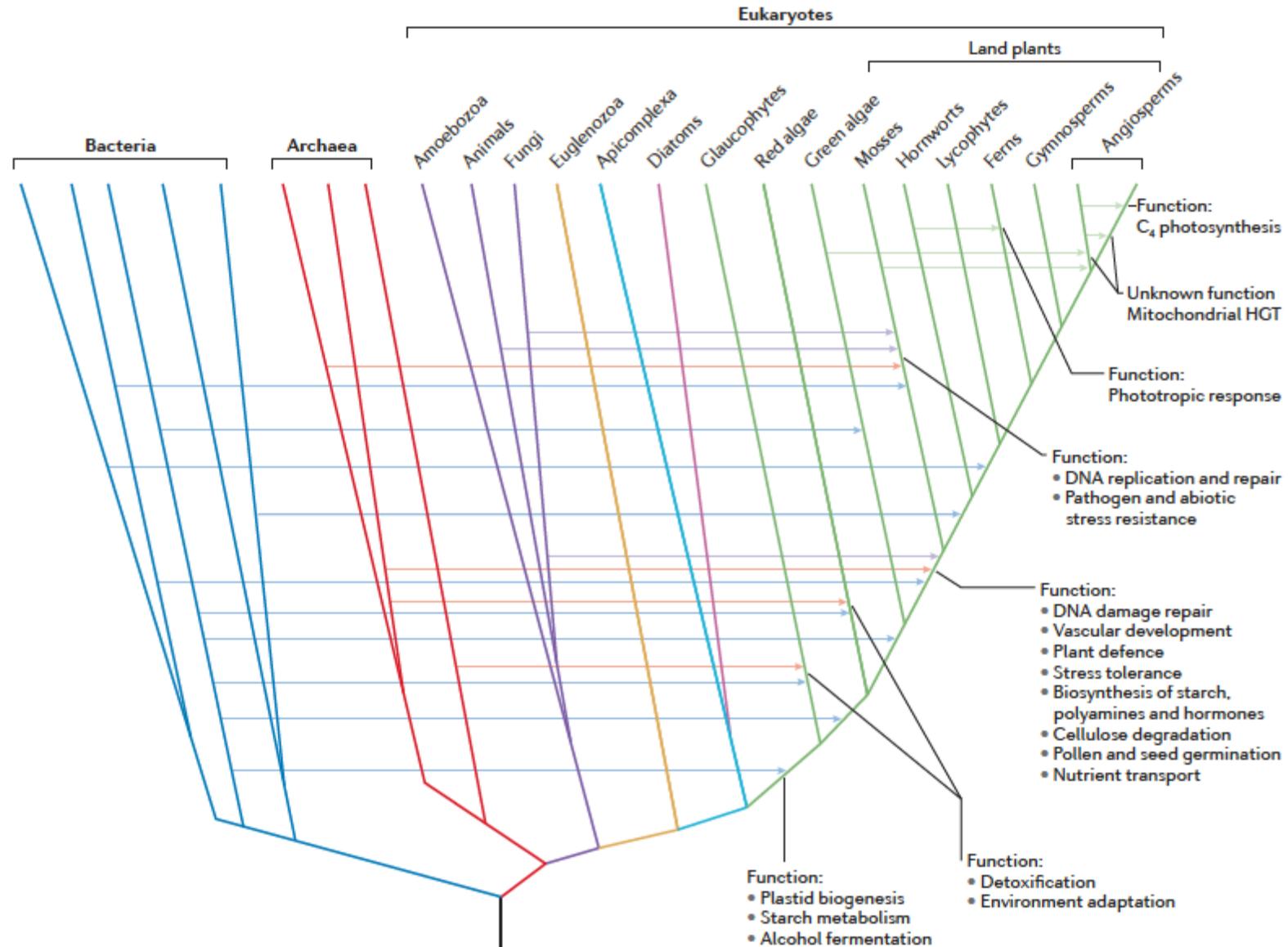


THG en bacterias está muy bien documentado: adquisición de ventajas y resistencia

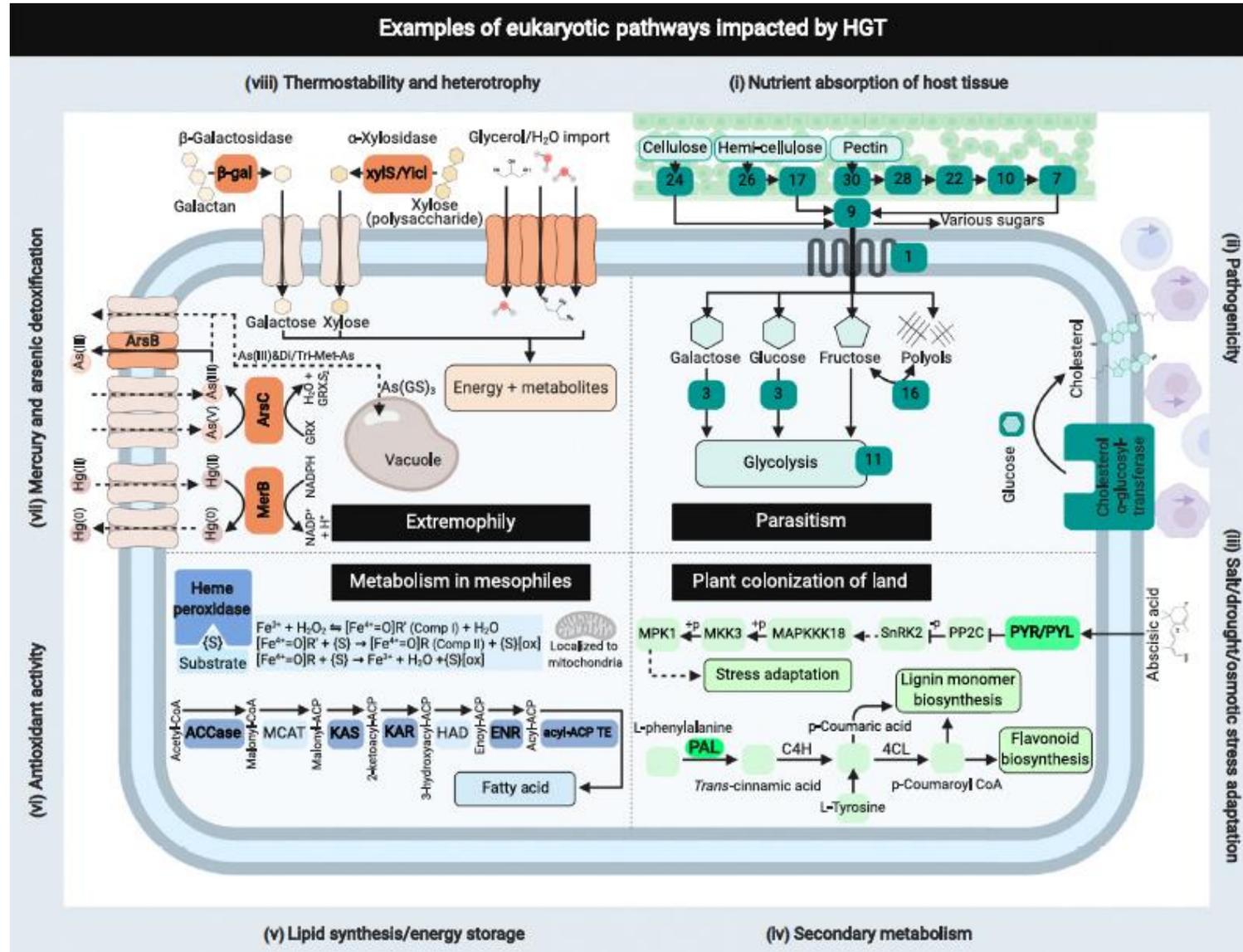
(Frost, L.S. et al. 2005)



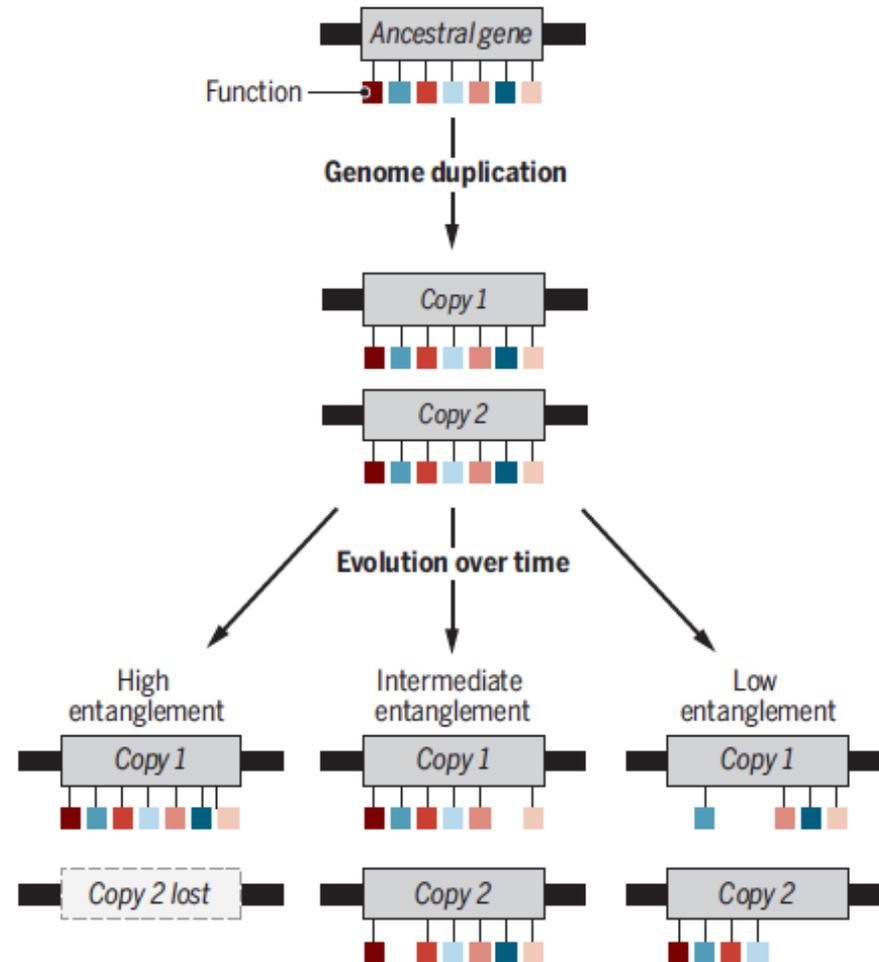
THG en plantas (Soucy, S.M. et al. 2015)



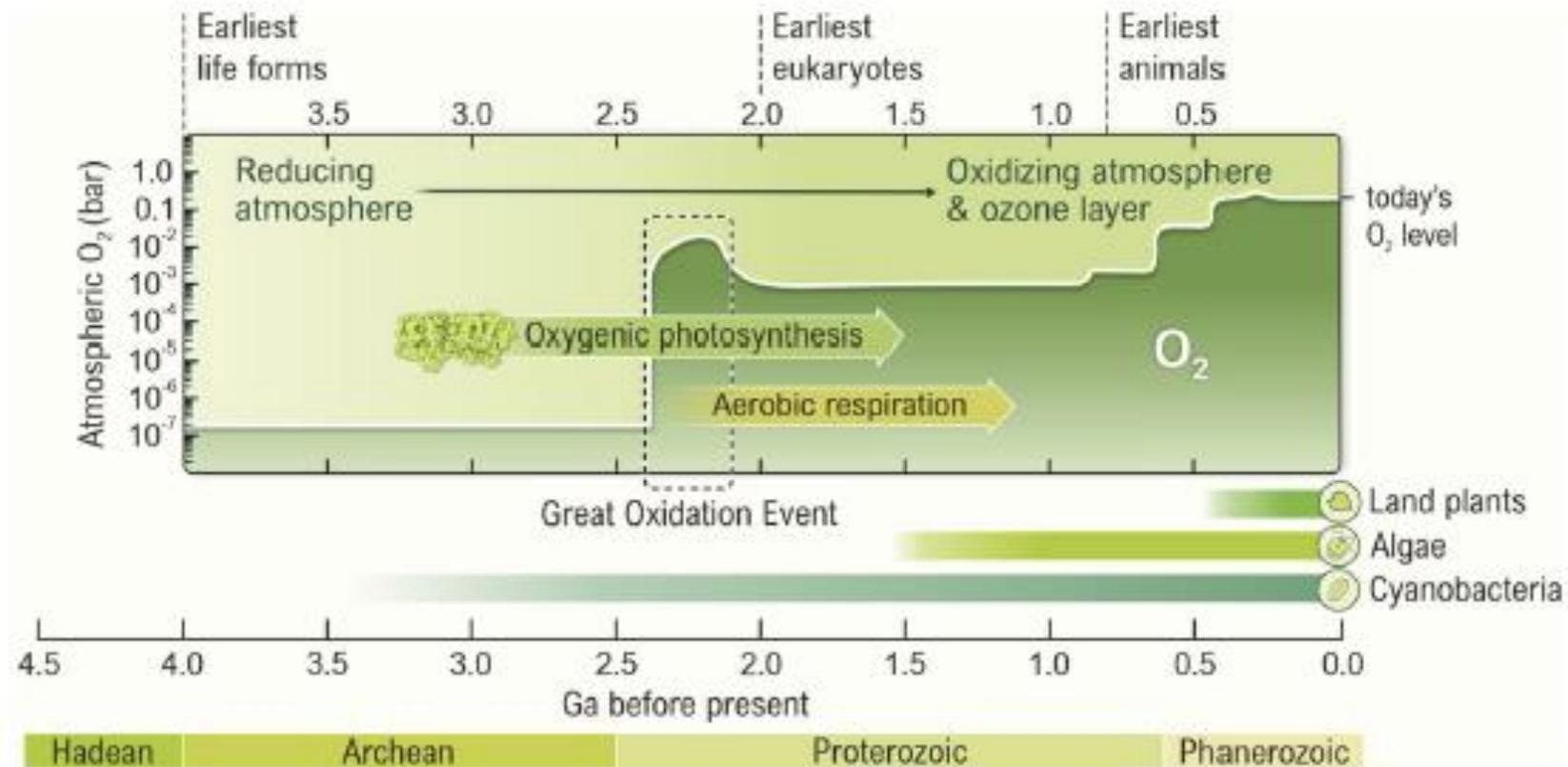
Genes funcionales derivados de THG (van Etten & Bhattacharya, 2020)



Funcionalidad de genes por duplicación de genoma (Ehrenreich, I.M. 2020)



Estudio Caso 2: Duplicación de genes y/o transferencia horizontal de genes para explicar origen de la Fotosíntesis (fotólisis del agua)

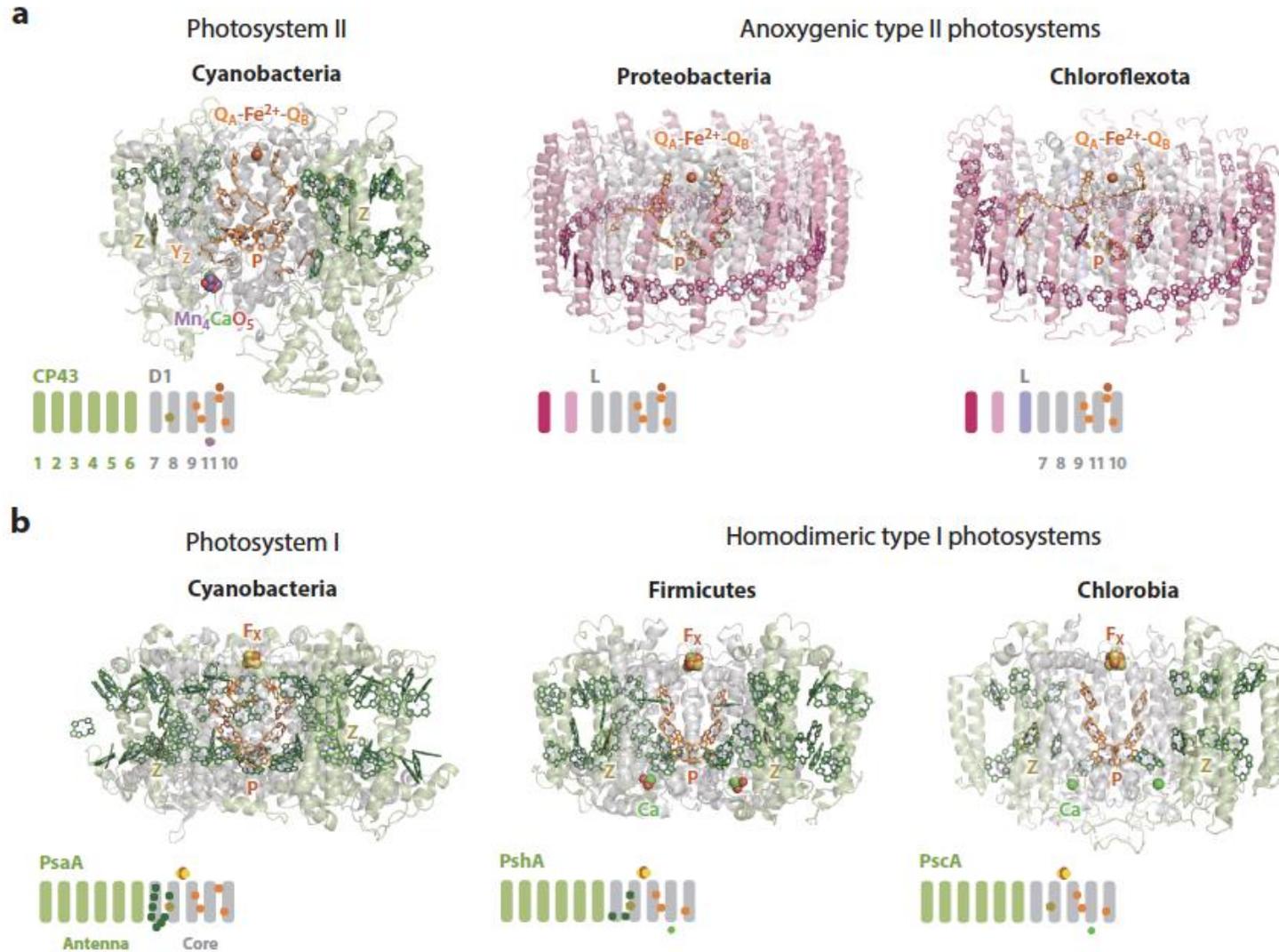


(Shevela et al. 2023)

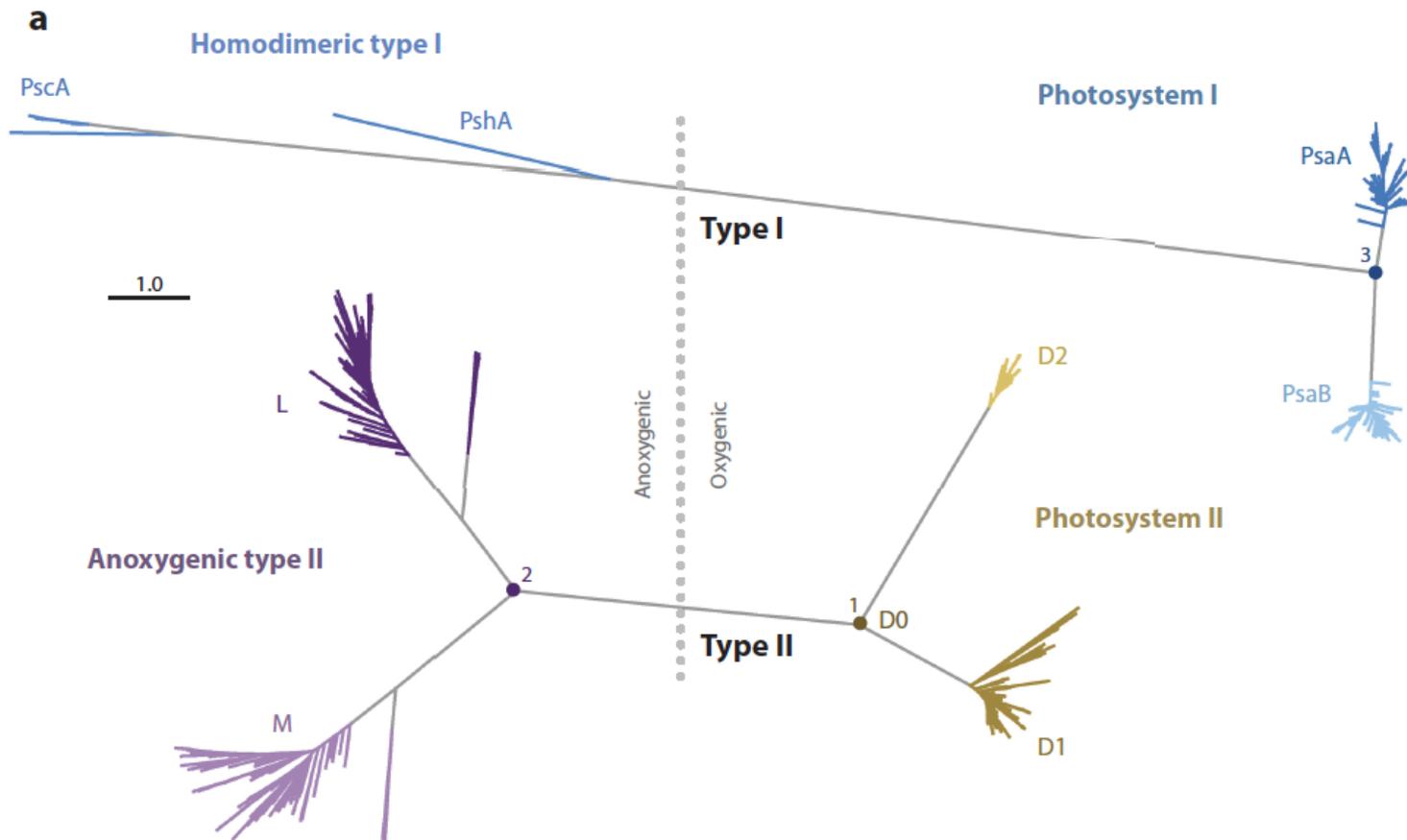
Estudio Caso 2: Duplicación de genes y/o transferencia horizontal de genes para explicar origen de la Fotosíntesis (fotólisis del agua)

	Tipo II	Tipo I
Fotosíntesis oxigénica	Fotosistema II	Fotosistema I
Centro de Reacción	D1 (Mn ₄ CaO ₅)/D2	psaA, psaB
Organismos	Cianobacterias, algas y plantas	Cianobacterias, algas y plantas
Fotosíntesis anoxigénica	Anóxico II	Homodimero I
Centro de Reacción	L y M	pshA, pscA
Organismos	Proteobacteria, Chloroflexota	Chlorobia, Firmicutes

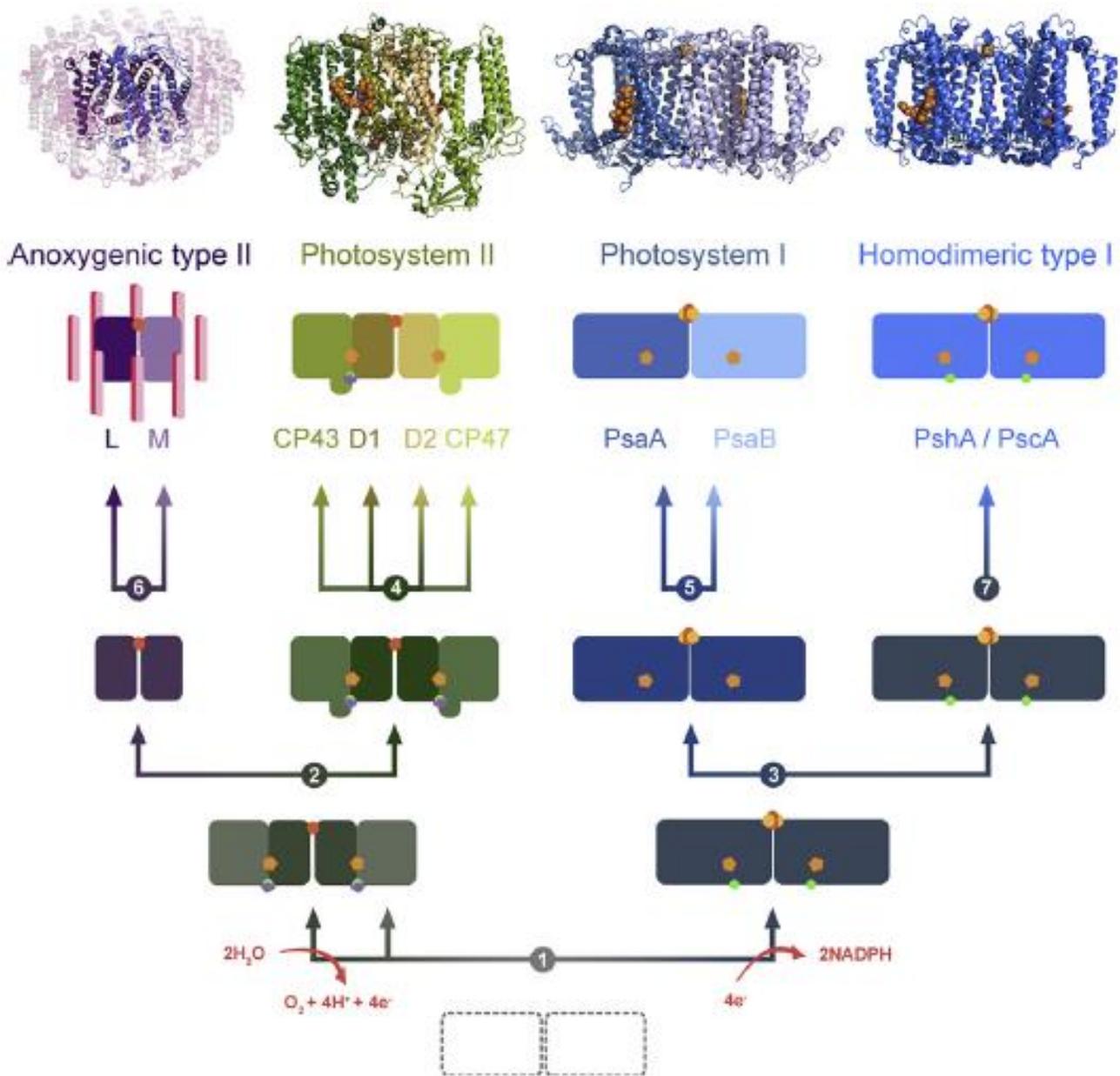
Arreglo estructural del fotosistema tipo II (a) y tipo I (b) (Oliver et al. 2023)



Filogenias de los fotosistemas tipo I y tipo II (Oliver et al. 2023)

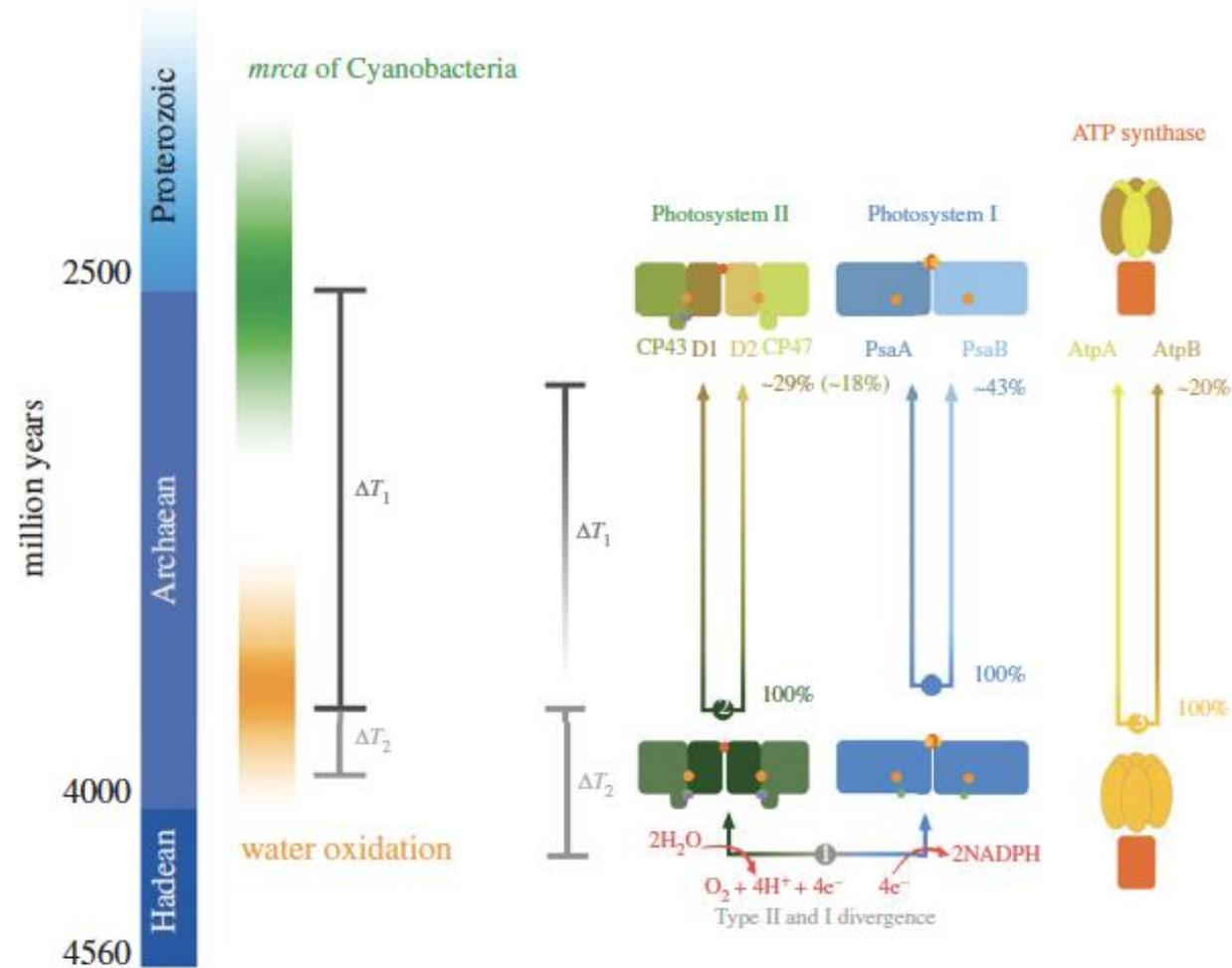


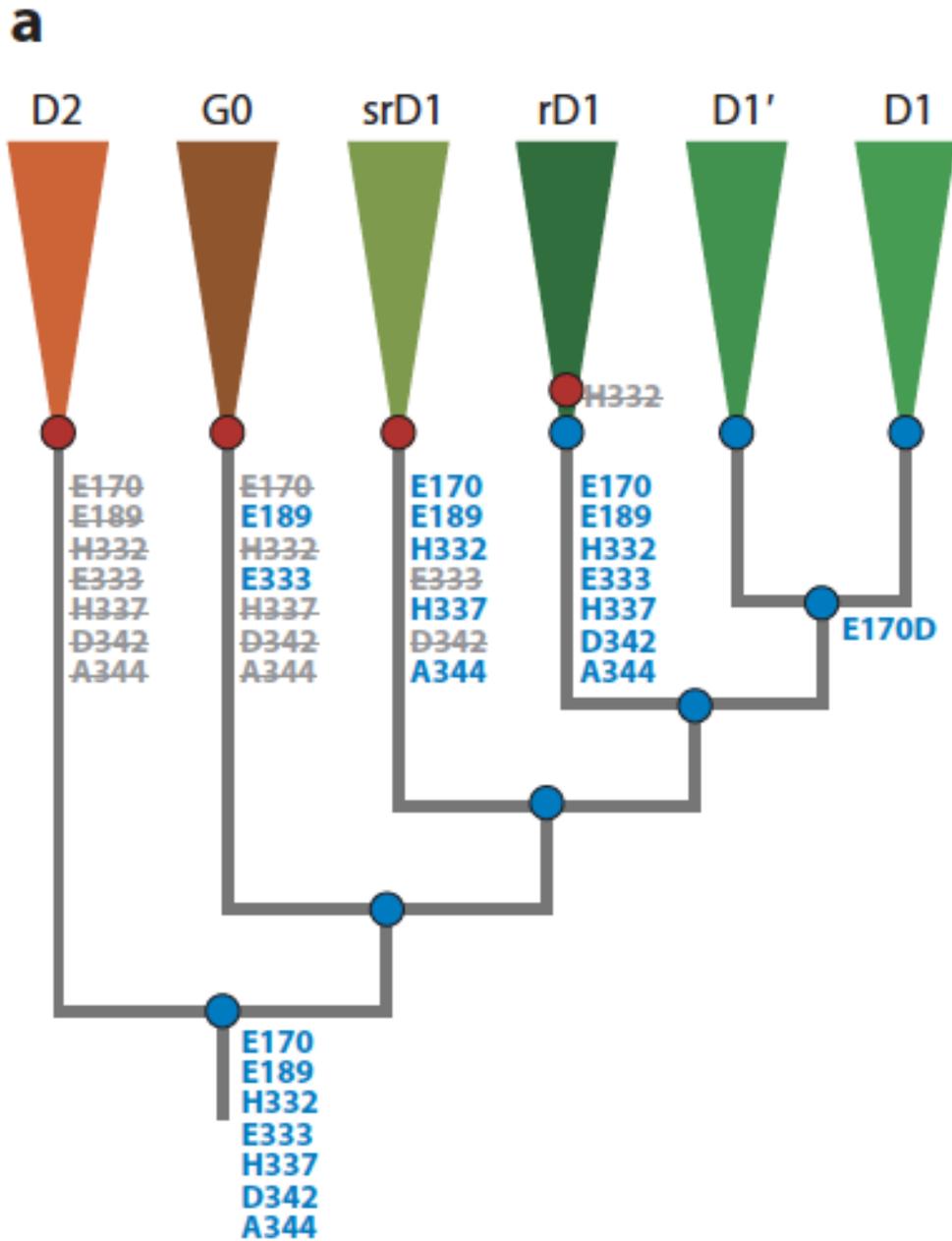
- No hay duda sobre tres eventos de duplicación de genes.
- No hay evidencia de que el PSI y PSII se originaron de fotosistemas anoxygenicos.
- La duplicaciones sugieren una historia del linaje oxigenico anterior y de divergencia a las cianobacterias



Representación estructural
de la filogenia de los
fotosistemas (Cardona & Rutherford,
2019)

Esquema de la evolución de la fotosíntesis (Cardona, T., 2019)



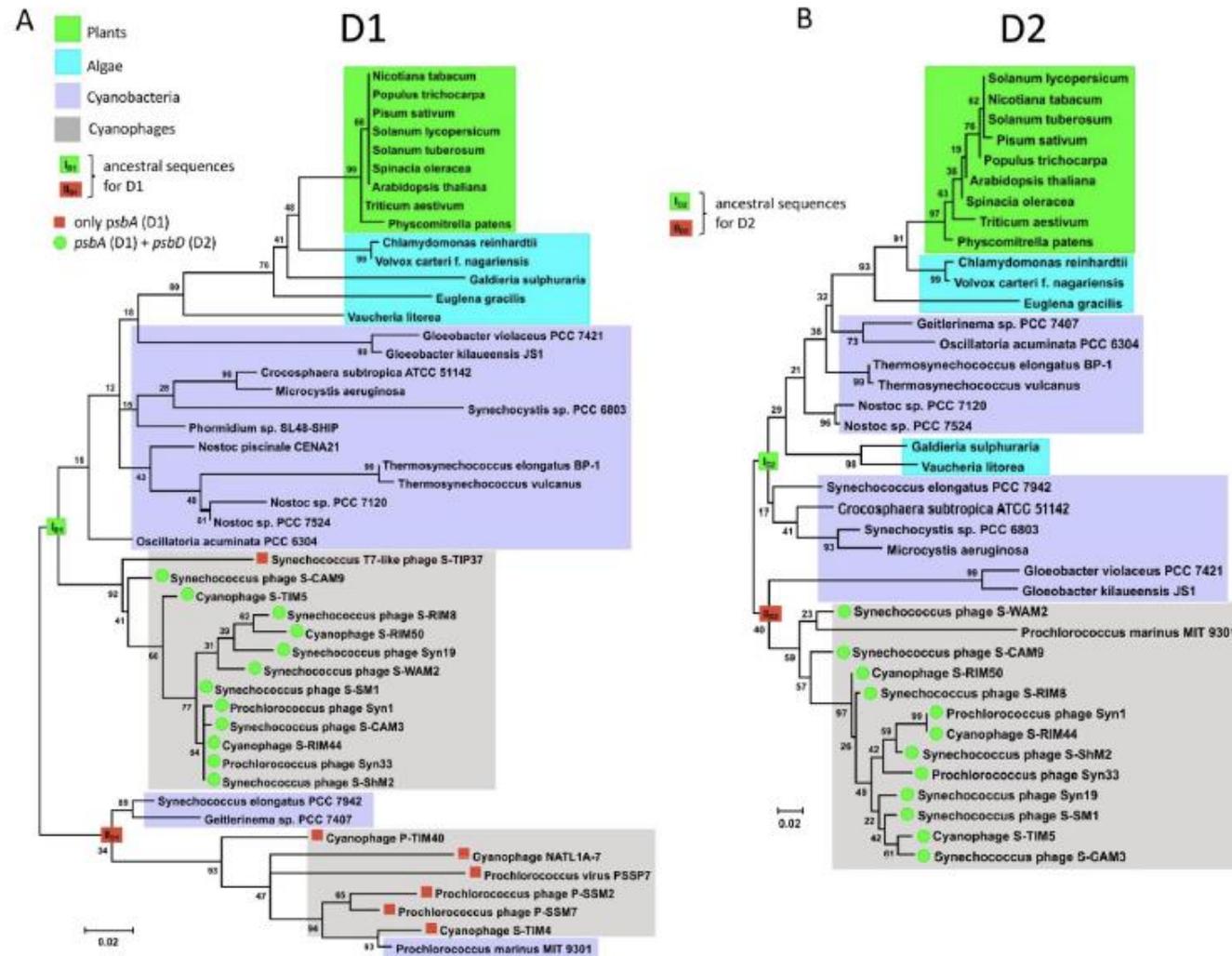


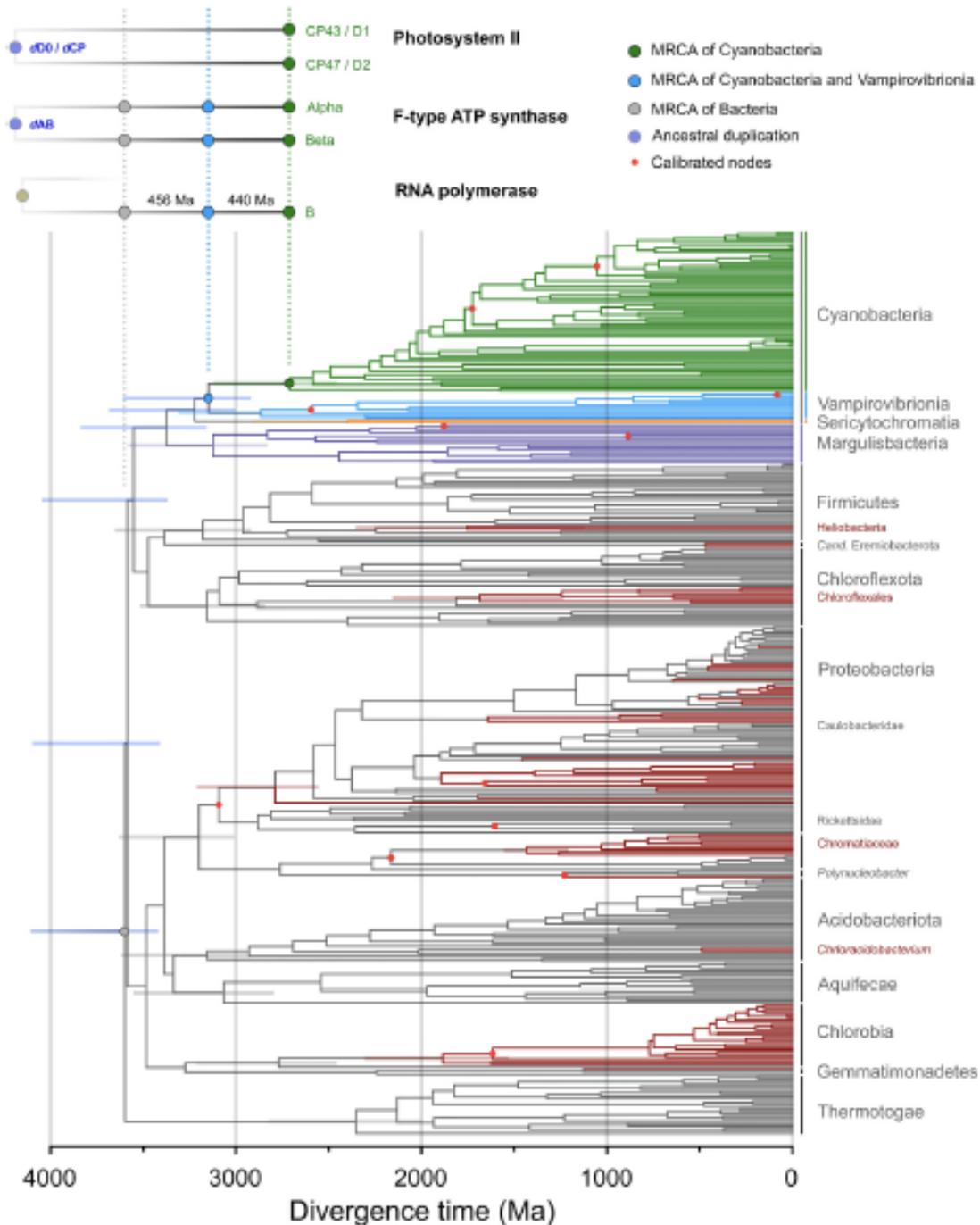
Reconstrucción de la secuencia ancestral de los ligandos Mn_4CaO_5 de D1 en la filogenia de D1 y D2 (Oliver et al., 2023)

- Residuos en azul representan secuencia ancestral
- Residuos en rojo los que se perdieron
- Circulos en azul representan capacidad de oxidar agua
- Circulos en rojo indican la pérdida de la función

Filogenia de las proteínas D1 y D2 en organismos y cianofagos

(Slesak & Slesak, 2022)





Reloj molecular de gen RpoB2 en bacterias, denotando el MRCA (ancestro común más reciente) de las Cianobacterias (Oliver et al., 2021)