GBS-RAD-sequencing:

Genómica de poblaciones en organismos no-modelo

Pablo Saenz Agudelo

Cómo caracterizar la diversidad genómica en organismos no modelo?

Cómo caracterizar la diversidad genómica en organismos no modelo?

Secuenciar genoma referencia

US$muchos miles!

Cómo caracterizar la diversidad genómica en organismos no modelo?

Secuenciar genoma referencia → US$ muchos miles!

X 100 Re-secuencia → US$100-500 por muestra

Que es RAD-sequencing?

X 100

Secuenciar

1-10% de del genoma

Aprox
US$6,000

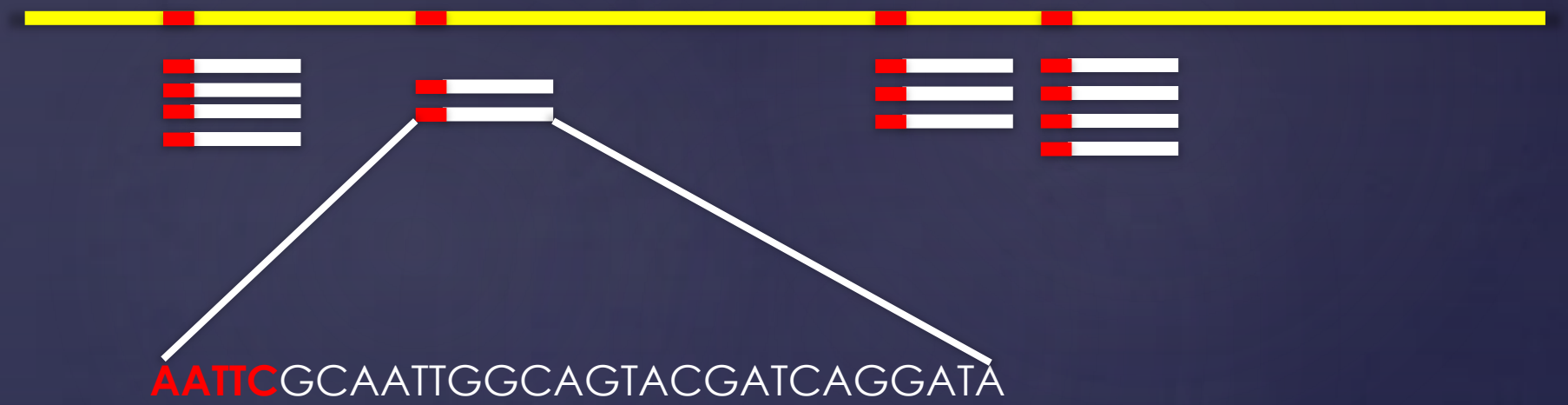
RAD-seq es un método para secuenciar una pequeña fraccíon del genóma de varios individuos al mismo tiempo usando NGS.

Como funciona?

Segmento de un genoma

Como funciona?

Segmento de un genoma



Variaciones

- RAD-seq (Baird *et al*. 2008)
- ddRAD-seq (Peterson *et al*. 2012)
- 2bRAD-seq (Wang et al 2012)
- ezRADs (Toonen et al. 2014)

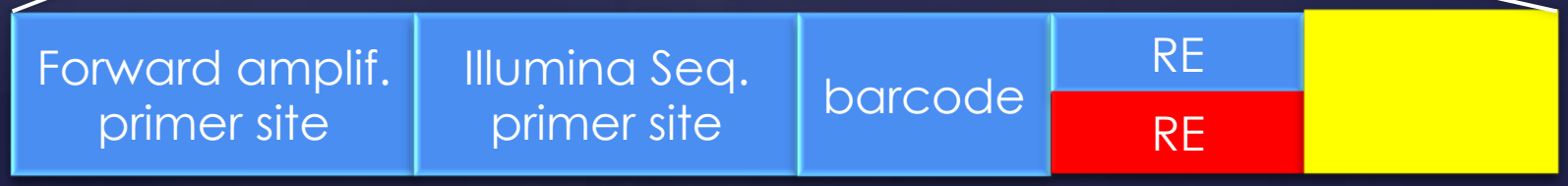
RAD-seq

Digestion con una enzima de restricción

RAD-seq

Digestion con una enzima de restricción

Adición de adaptadores especiales “P1”



RAD-seq

Se combinan las muestras, se rompen los fragmentos aleatoriamente "shear", se selecciona por tamaño.

Adición de un segundo adaptador "P2"

RAD-seq

Se combinan las muestras, se rompen los fragmentos aleatoriamente "shear", se selecciona por tamaño.

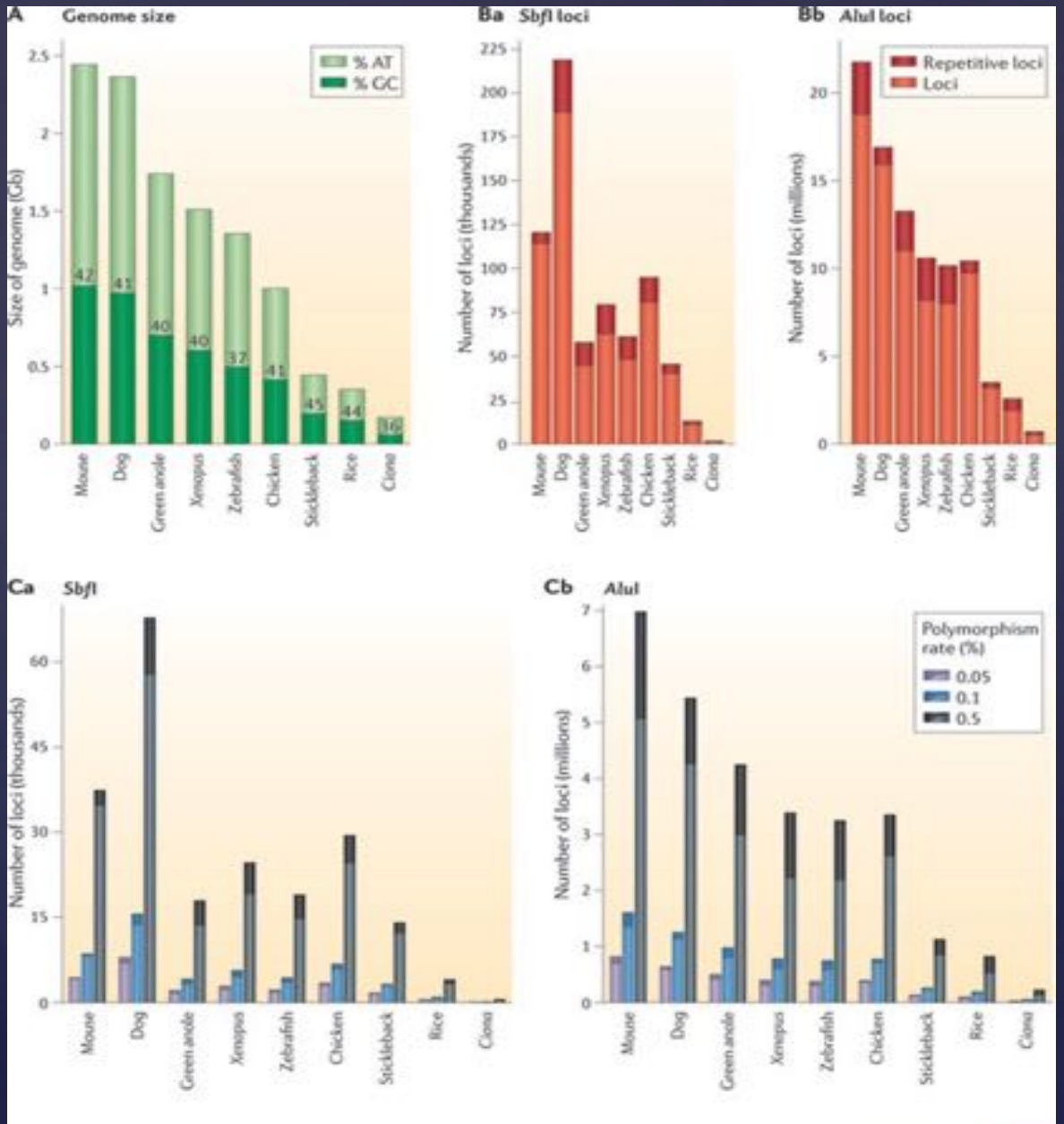
Adición de un segundo adaptador "P2"

RAD-seq

Amplificación selectiva por PCR de fragmentos RAD

Longitud de secuencias en Illumina

Efecto de la enzima de restricción



Davey et al 2011

dd-RAD-seq

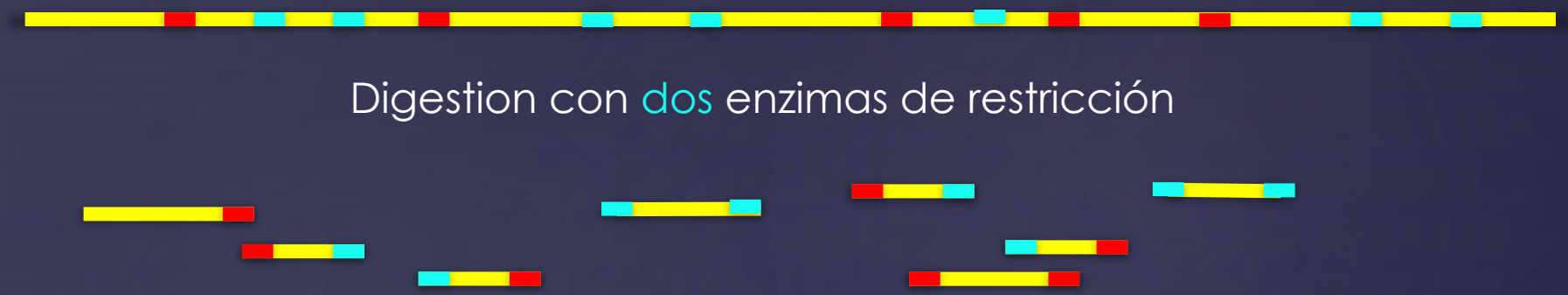
Digestion con dos enzimas de restricción

dd-RAD-seq

Digestion con dos enzimas de restricción

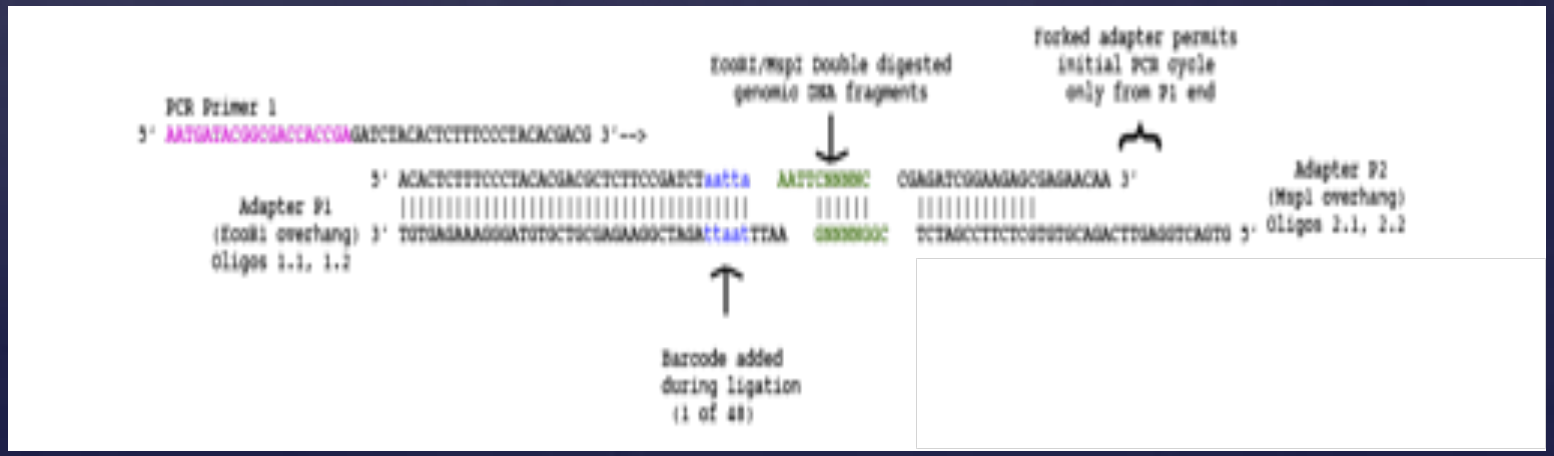
Adición de adaptadores especiales “P1 y P2”

dd-RAD-seq



Digestion con dos enzimas de restricción

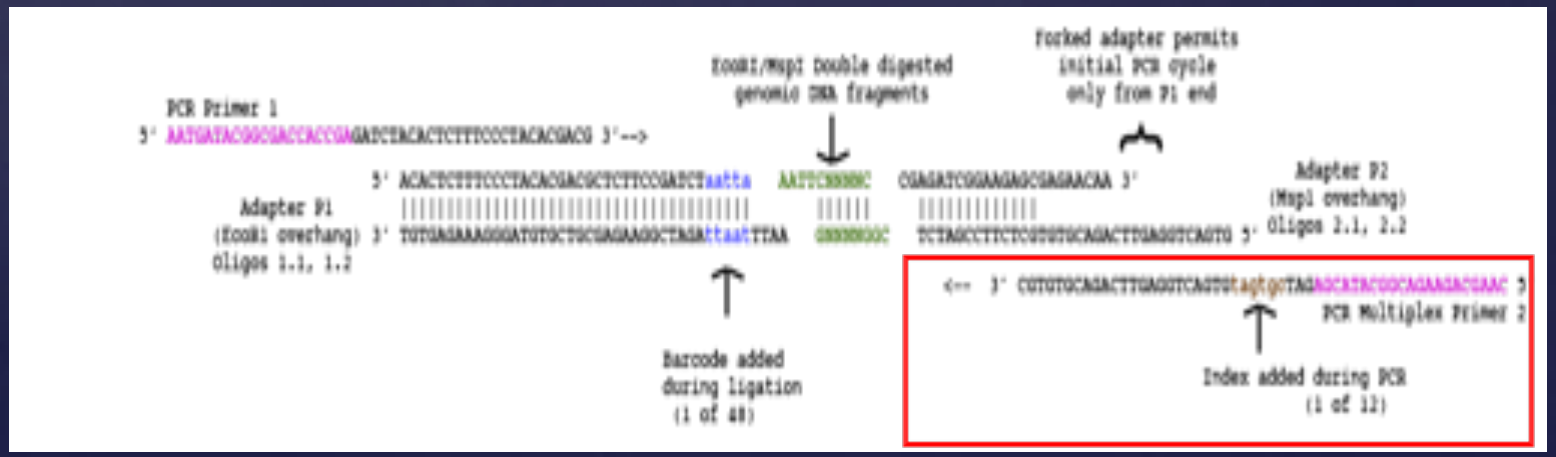
Adición de adaptadores especiales “P1 y P2”



dd-RAD-seq

Se combinan las muestras con diferentes P1, se selecciona por tamaño.

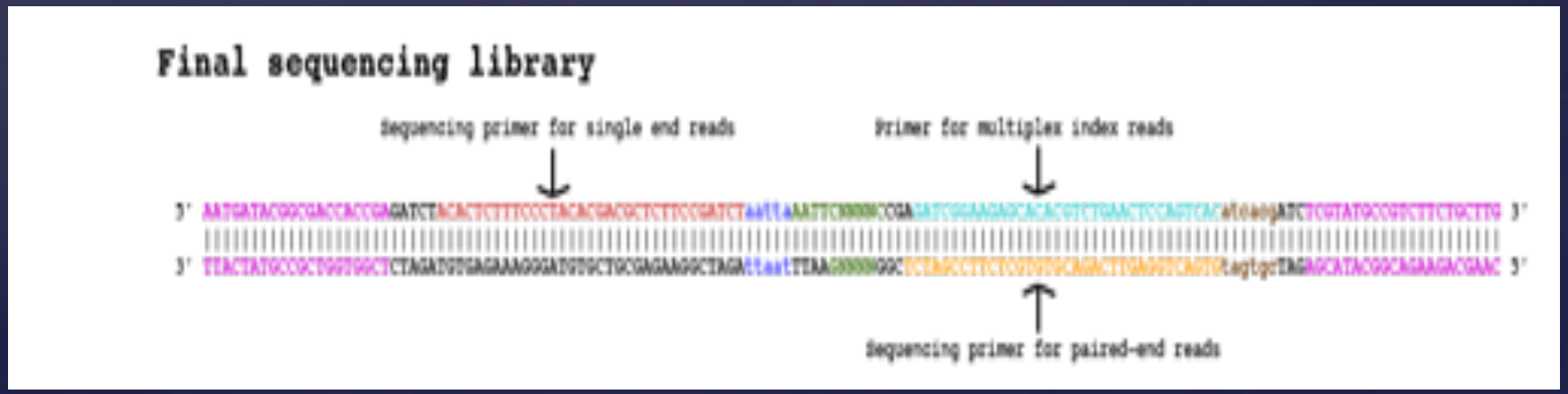
Amplificación por PCR con uno de los primers que incluye otro codigo de barras



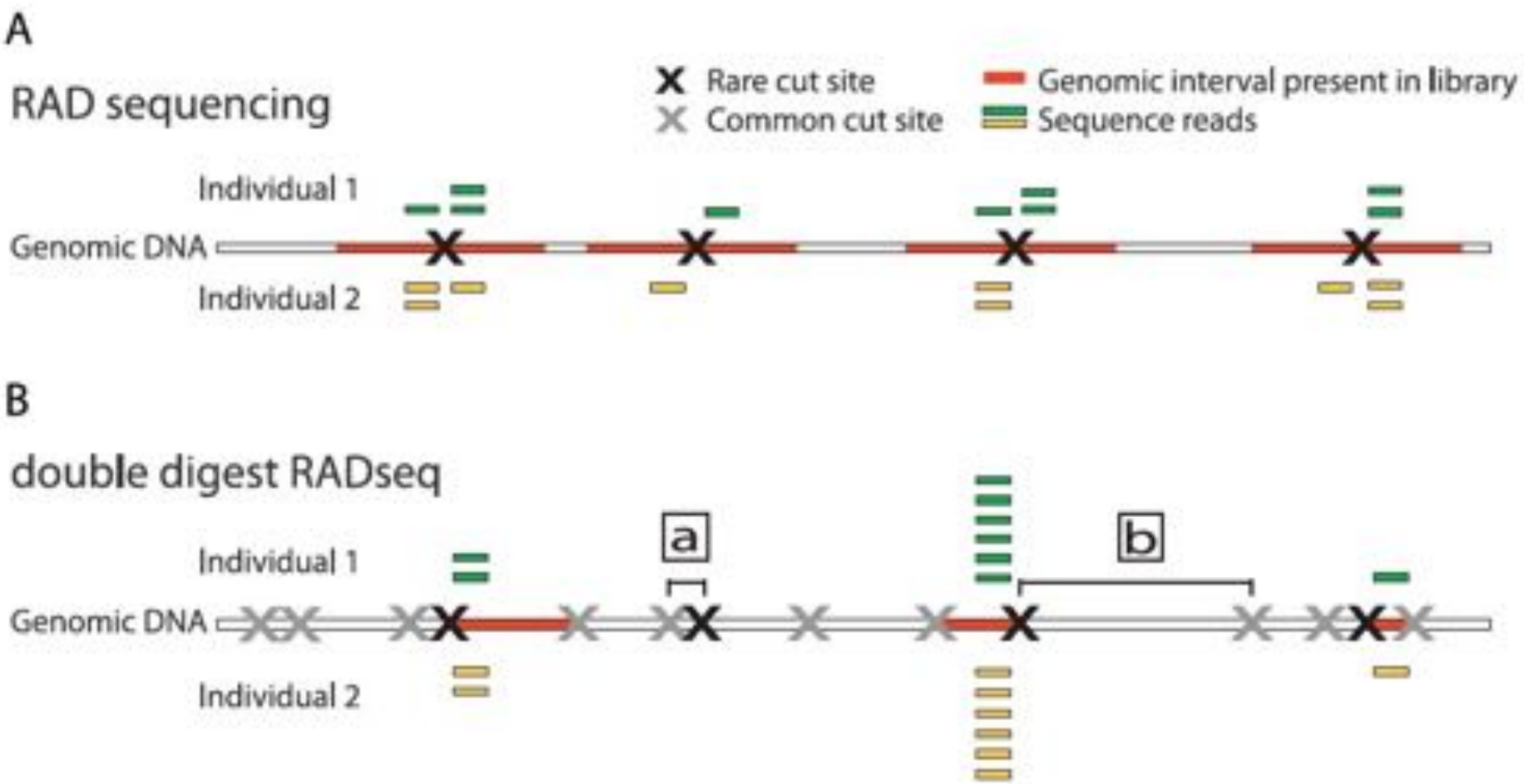
dd-RAD-seq

Se combinan las muestras con diferentes P1, se selecciona por tamaño.

Combinación de pares de P1 y P2 ➔ multiplicación del número posible de muestras a combinar.

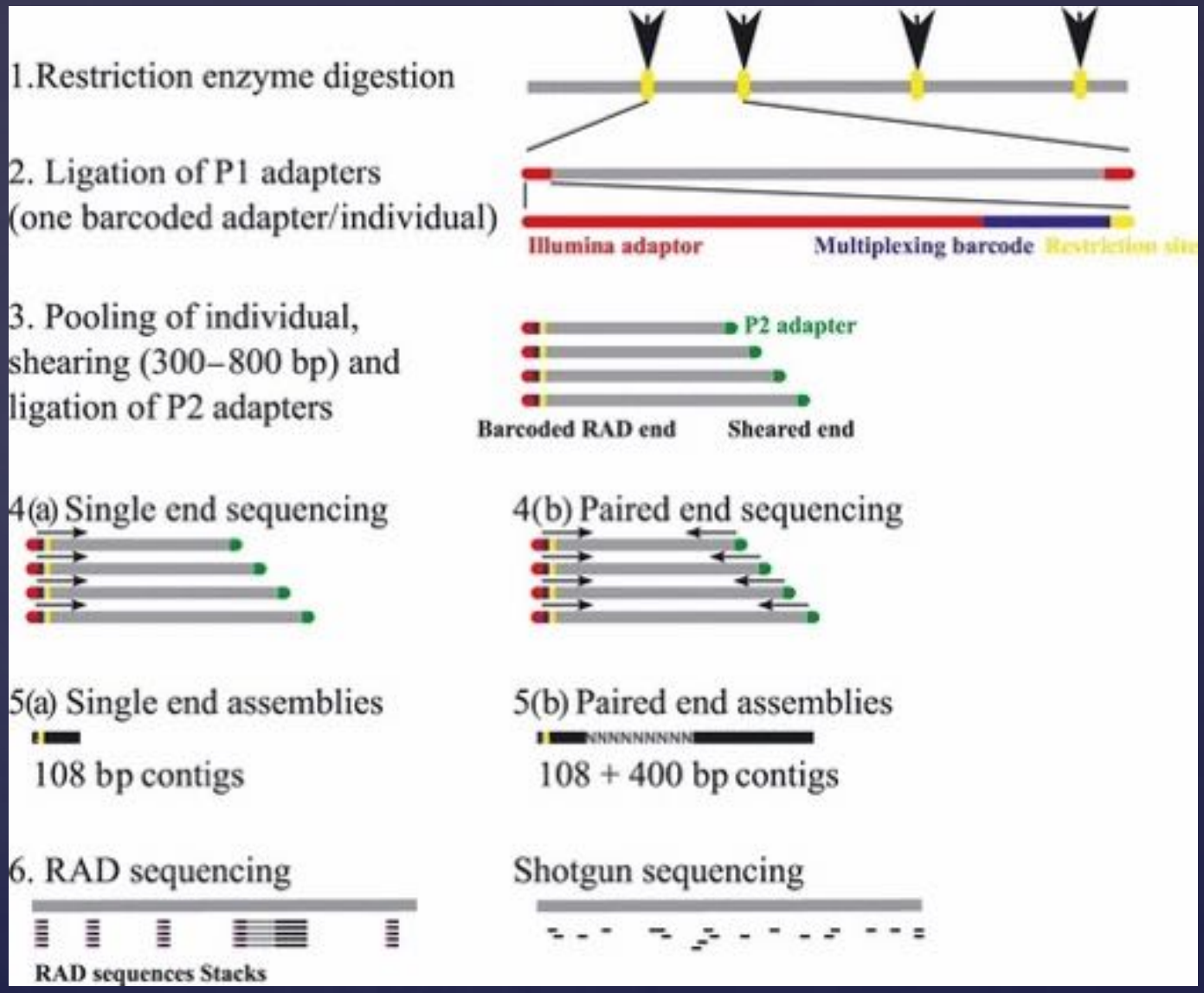


RADs vs ddRADs



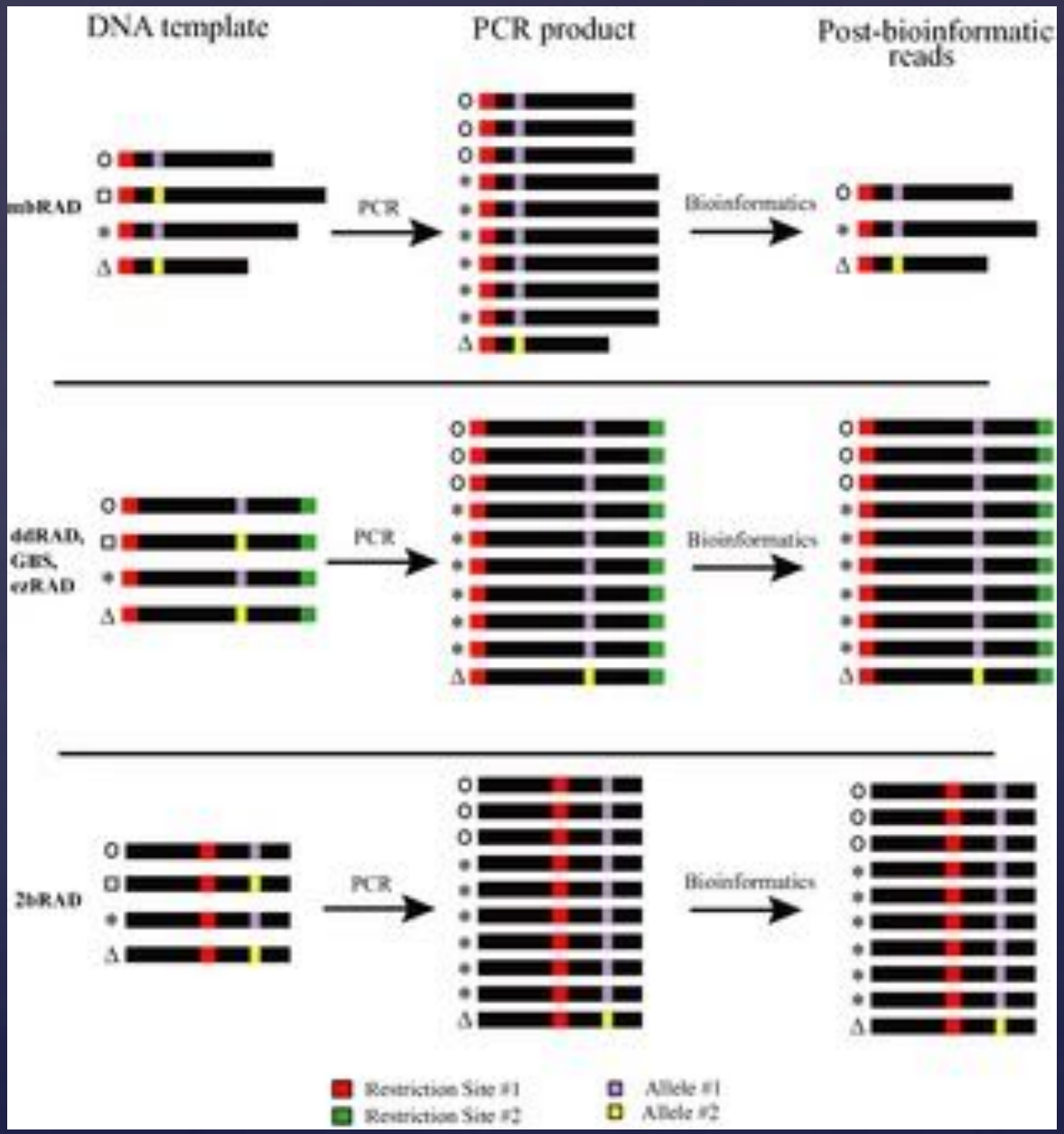
Peterson et al 2012

RADs vs ddRADs



Rowe al 2012

RADs vs ddRADs



Andrews et al 2014

Algunas observaciones sobre el uso de RADseq

MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2017) doi: 10.1111/1755-0998.12662

Using a butterflyfish genome as a general tool for RAD-Seq studies in specialized reef fish

JOSEPH D. DIBATTISTA,*† PABLO SAENZ-AGUDELO,*‡ MAREK J. PIATEK,§¶ XIN WANG,* MANUEL ARANDA* and MICHAEL L. BERUMEN*
*Red Sea Research Center, Division of Biological and Environmental Science and Engineering, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia, †Department of Environment and Agriculture, Curtin University, PO Box U1987, Perth WA 6845, Australia, ‡Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Valdivia 5090000, Chile, §Computational Bioscience Research Center, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia, ¶Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2016) doi: 10.1111/1755-0998.12588

Draft genome of an iconic Red Sea reef fish, the blacktail butterflyfish (*Chaetodon austriacus*): current status and its characteristics

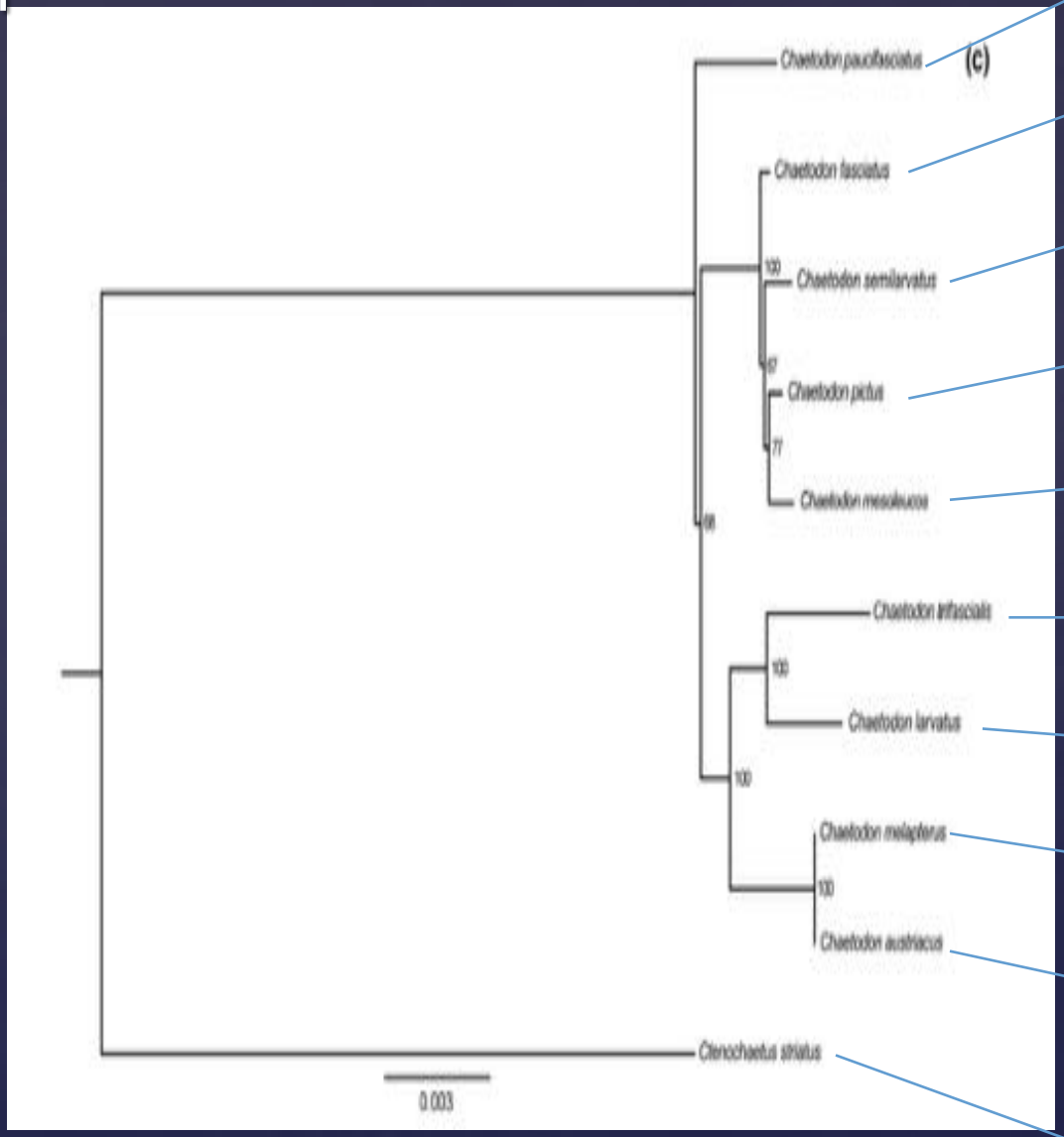
JOSEPH D. DIBATTISTA,*† XIN WANG,* PABLO SAENZ-AGUDELO,*‡ MAREK J. PIATEK,§ MANUEL ARANDA* and MICHAEL L. BERUMEN*
*Division of Biological and Environmental Sciences and Engineering, Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia, †Department of Environment and Agriculture, Curtin University, PO Box U1987, Perth, WA 6845, Australia, ‡Instituto de Ciencias Ambientales y Evolutivas, Universidad Austral de Chile, Valdivia 5090000, Chile, §Computational Bioscience Research Center, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia

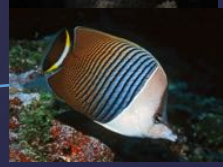
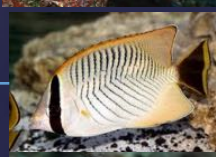
Algunas observaciones sobre el uso de RADseq

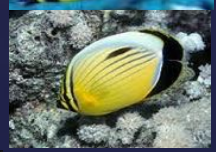
Utilidad de un genoma de referencia para análisis RADseq en otras especies?

Distribución de SNPs (zonas codificantes vs no codificantes).

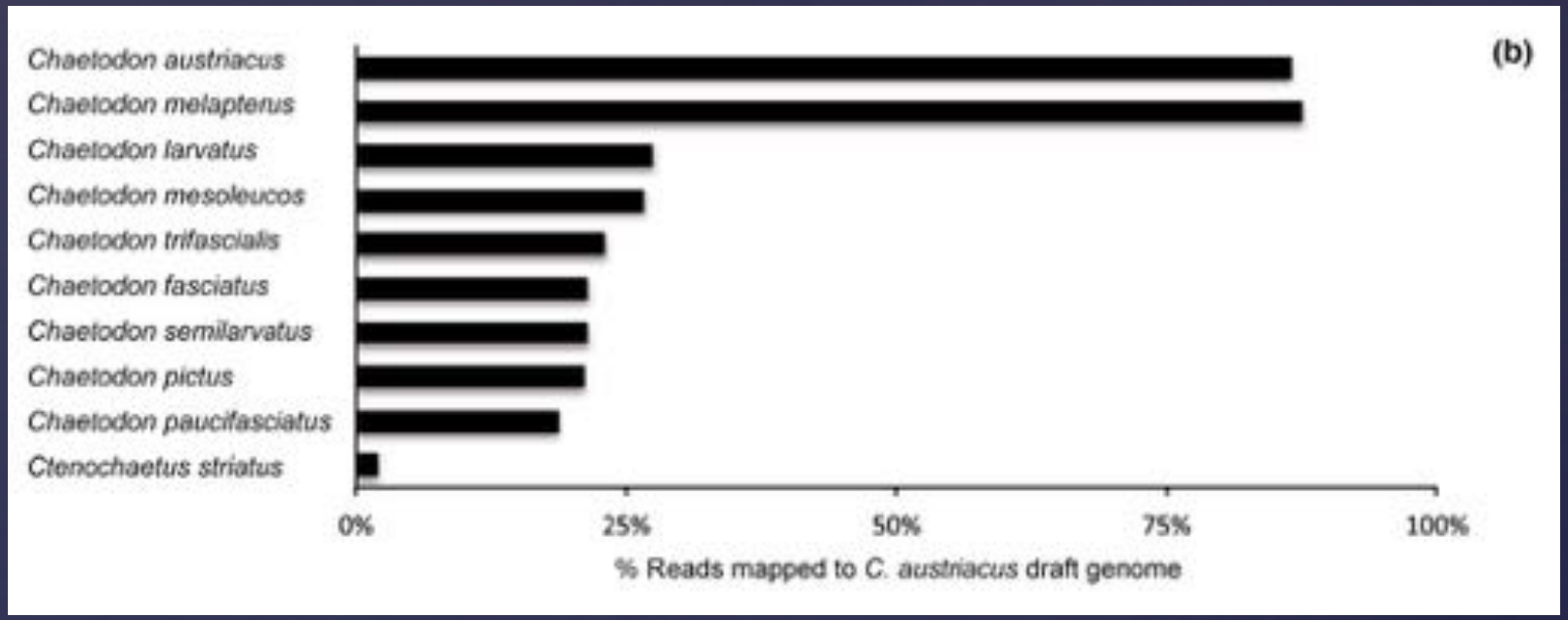
Algunas observaciones sobre el uso de RADseq



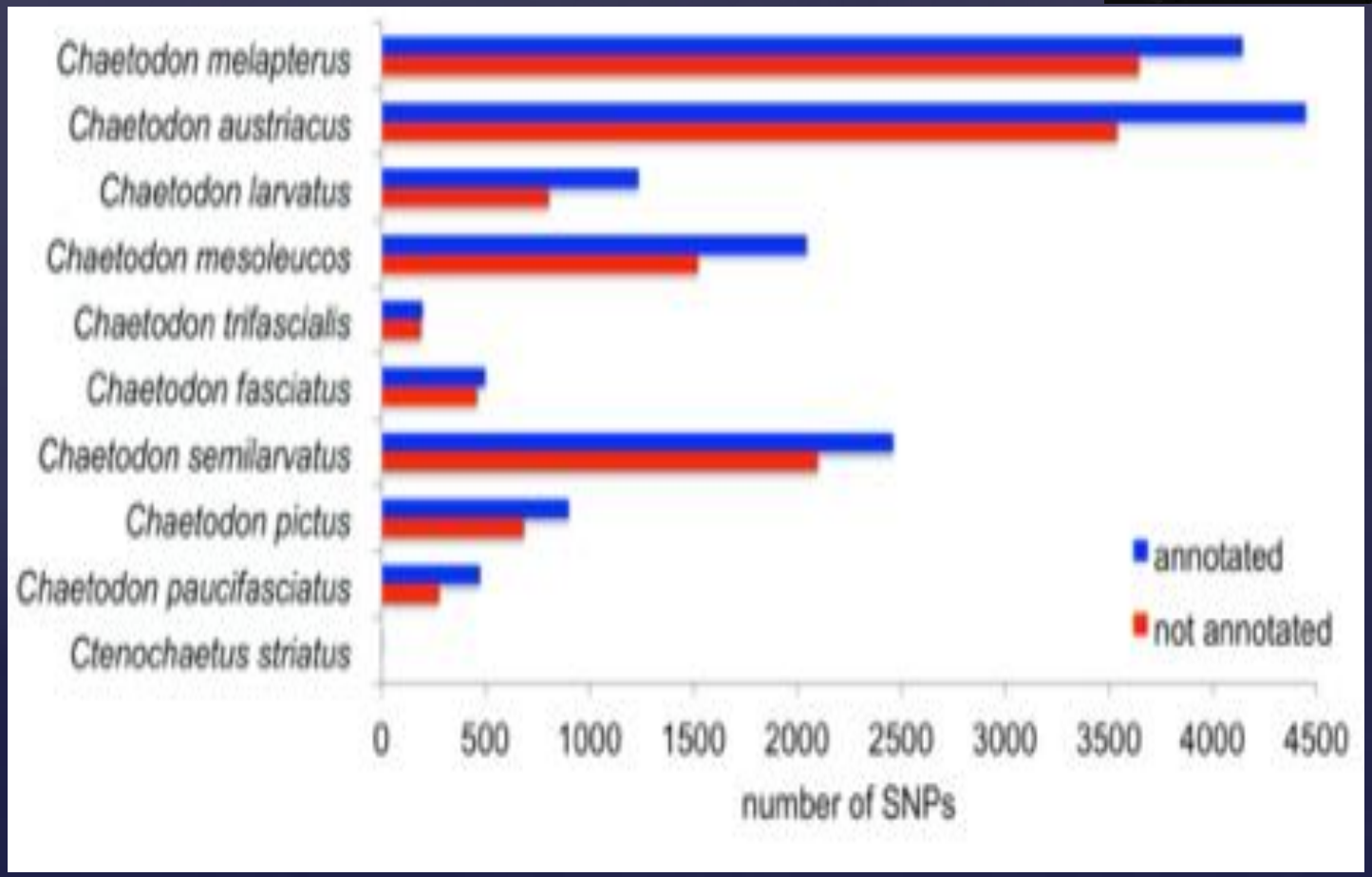


Algunas observaciones sobre el uso de RADseq

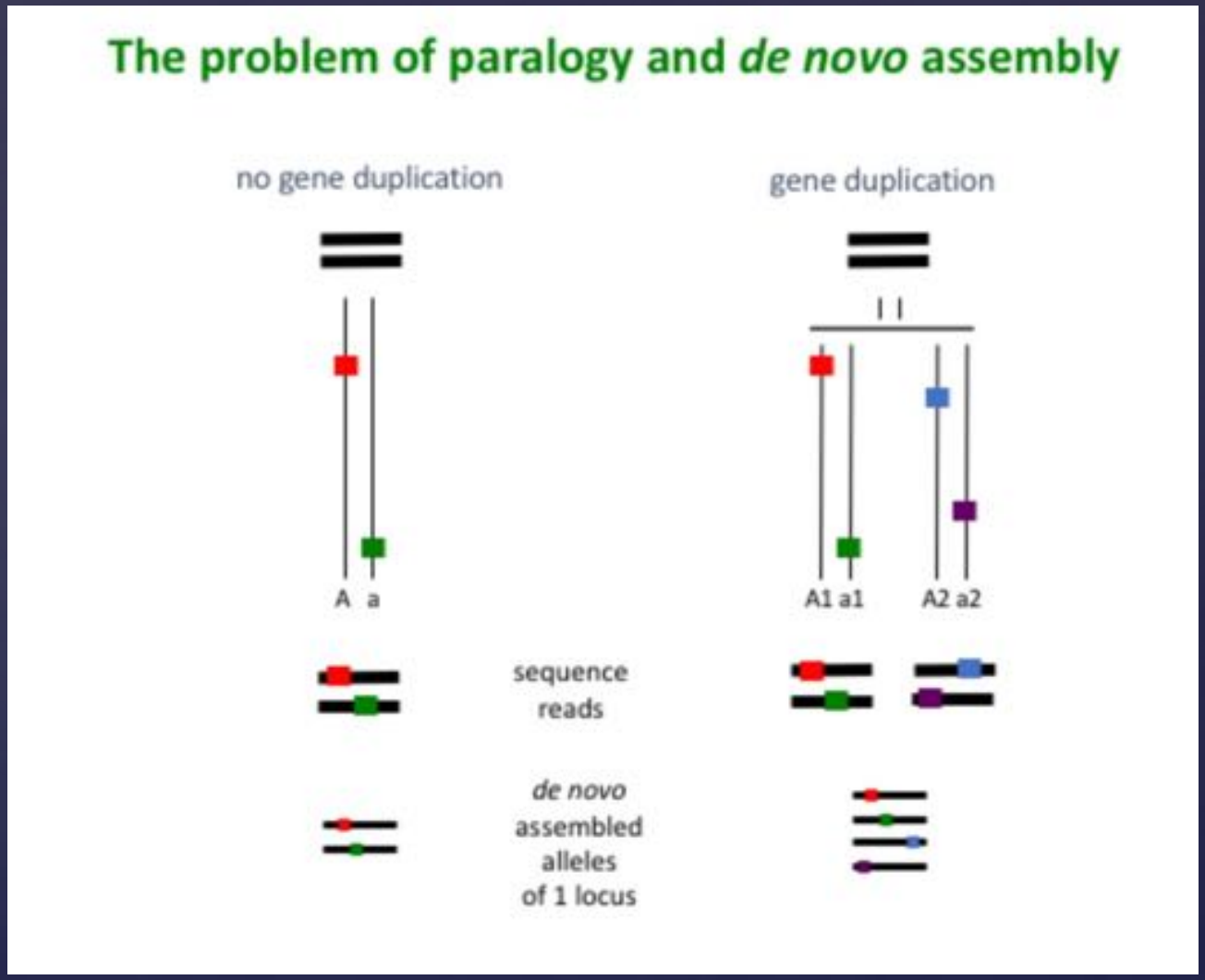


dd-RADseq (*SphI* and *MluCI)*

Algunas observaciones sobre el uso de RADseq



Algunas observaciones sobre el uso de RADseq



Source: https://github.com/u-genoma/BioinfinvRepro/blob/master/Unidad8/GBS-RAD/slides/Slide12.png

Varias opciones para análisis de datos

	Stacks Catchen et al 2013	ipyRAD Eaton et al 2014	Tassel Bradbury et al 2007
Better for	Population genetics non-model spp.	Phylogenies non-model spp.	Population genetics GBS of crops
Type of variation	SNPs, indels	SNPs, indels	SNPs, indels
Method	*de novo* reference genome	*de novo*, reference, reference addition and reference subtraction	*de novo* (not supported anymore) reference genome
Wet lab Protocols	RAD-seq, ddRAD, DaRT, 2bRAD and GBS	RAD, ddRAD, GBS, NextRAD, RApture	GBS
Pair-end?	Yes	Yes	No in de novo
output	**VCF, plink**, Stacks custom, structure, **GenPop**, phylip, PHASE, Beagle	**VCF**, ipyrad custom, **phylip**, geno, structure, **nexus**,	VCF, plink, h5
Comments	Estimates Fst, π, He	Jupyter notebooks	Graphical interface useful to "visualize" SNPs

Source: https://github.com/u-genoma/BioinfinvRepro/blob/master/Unidad8/GBS-RAD/slides/Slide13.png

Análisis de datos con

Stacks

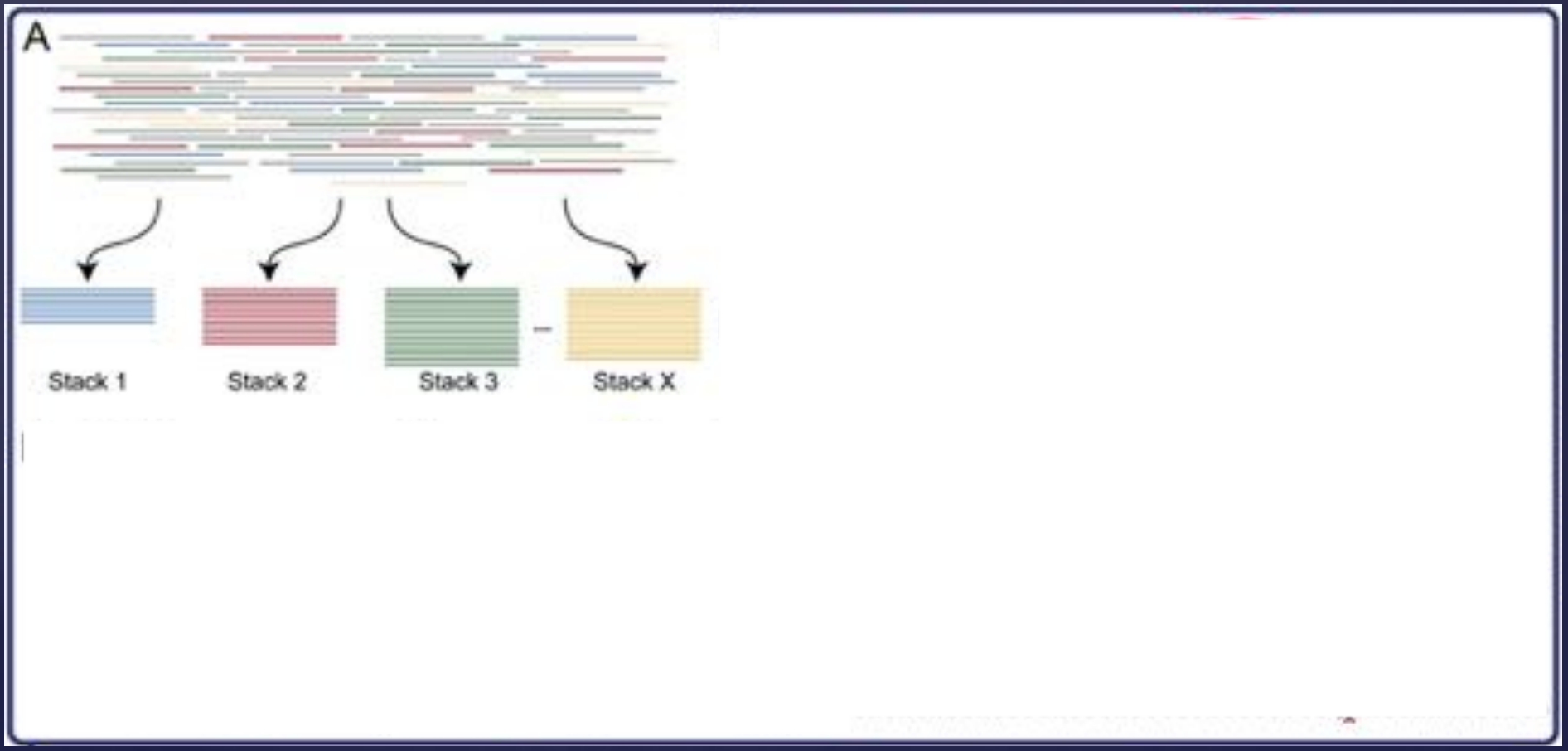
Stacks is a software pipeline for building loci from short-read sequences, such as those generated on the Illumina platform. Stacks was developed to work with restriction enzyme-based data, such as RAD-seq, for the purpose of building genetic maps and conducting population genomics and phylogeography.

Download Stacks

Version 1.24

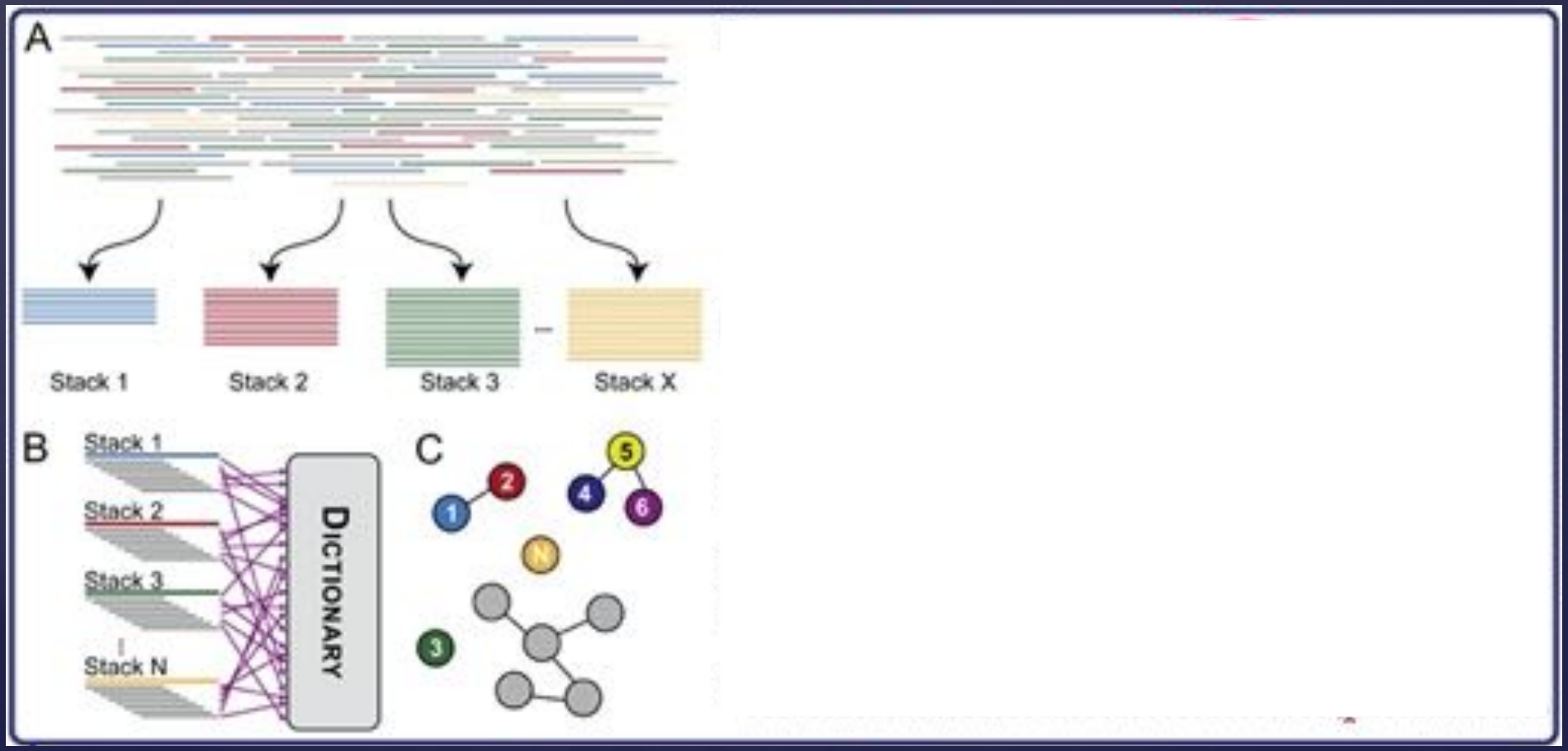
1. Demultiplexing

2. Construcción de loci



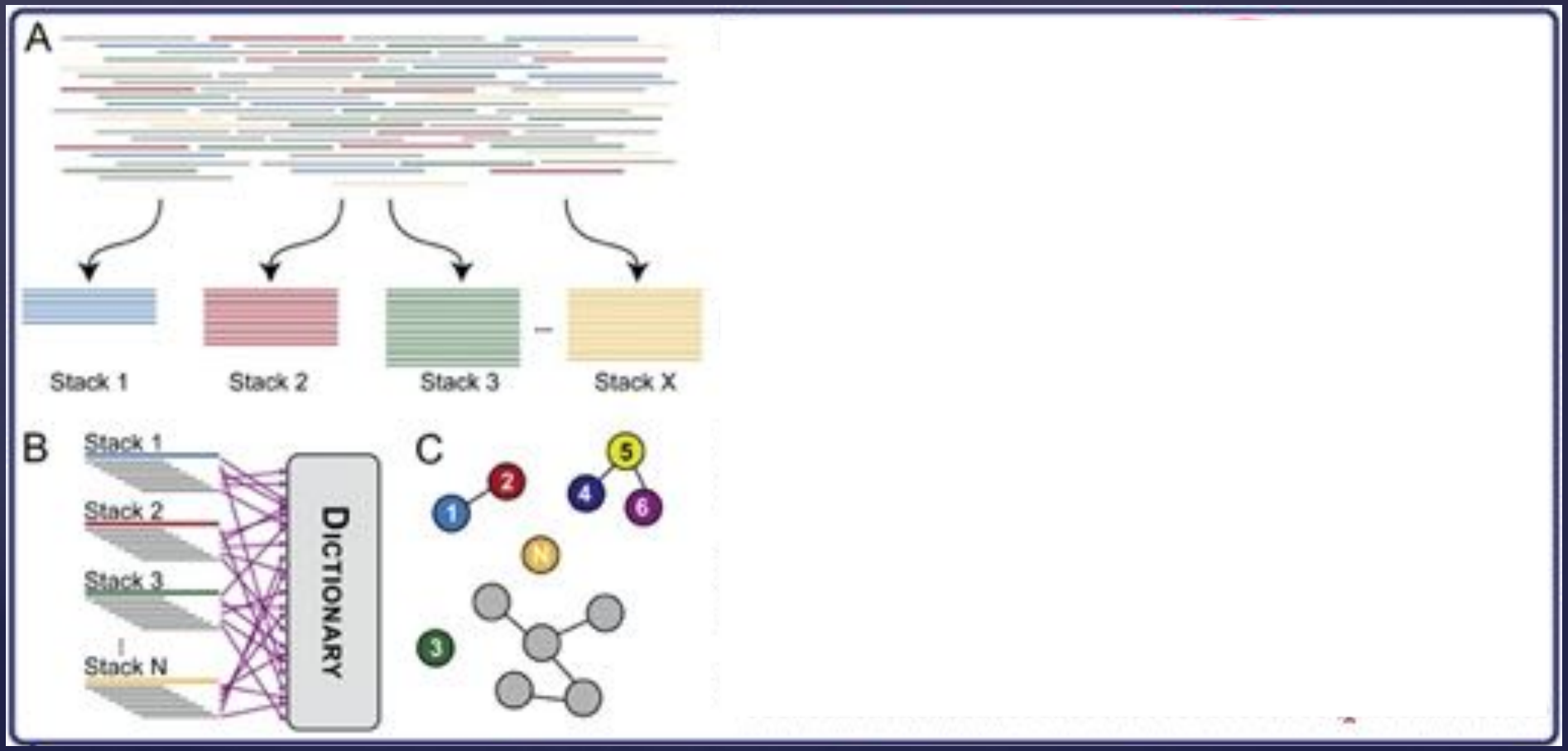
Catchen et al 2011

2. Construcción de loci



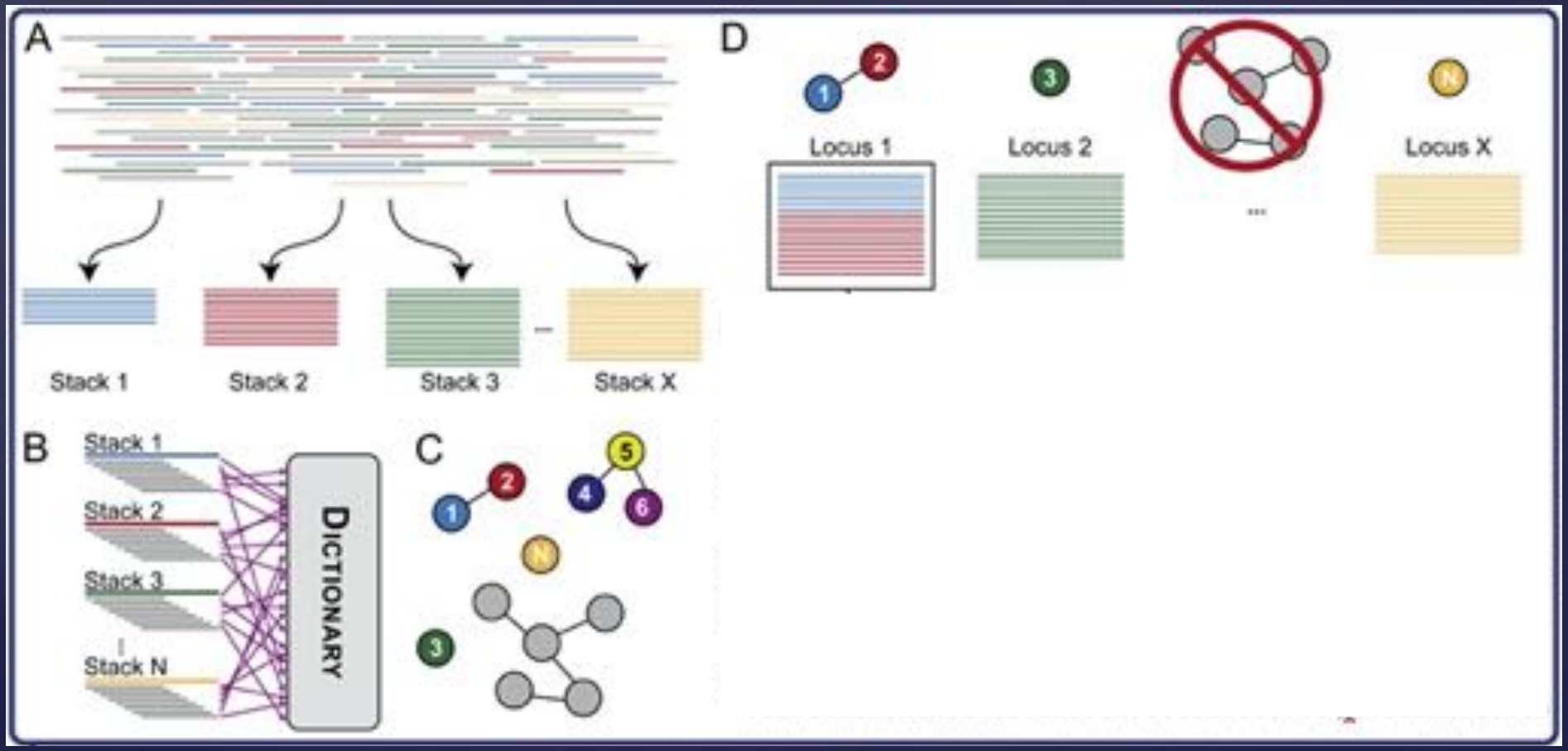
Catchen et al 2011

2. Construcción de loci



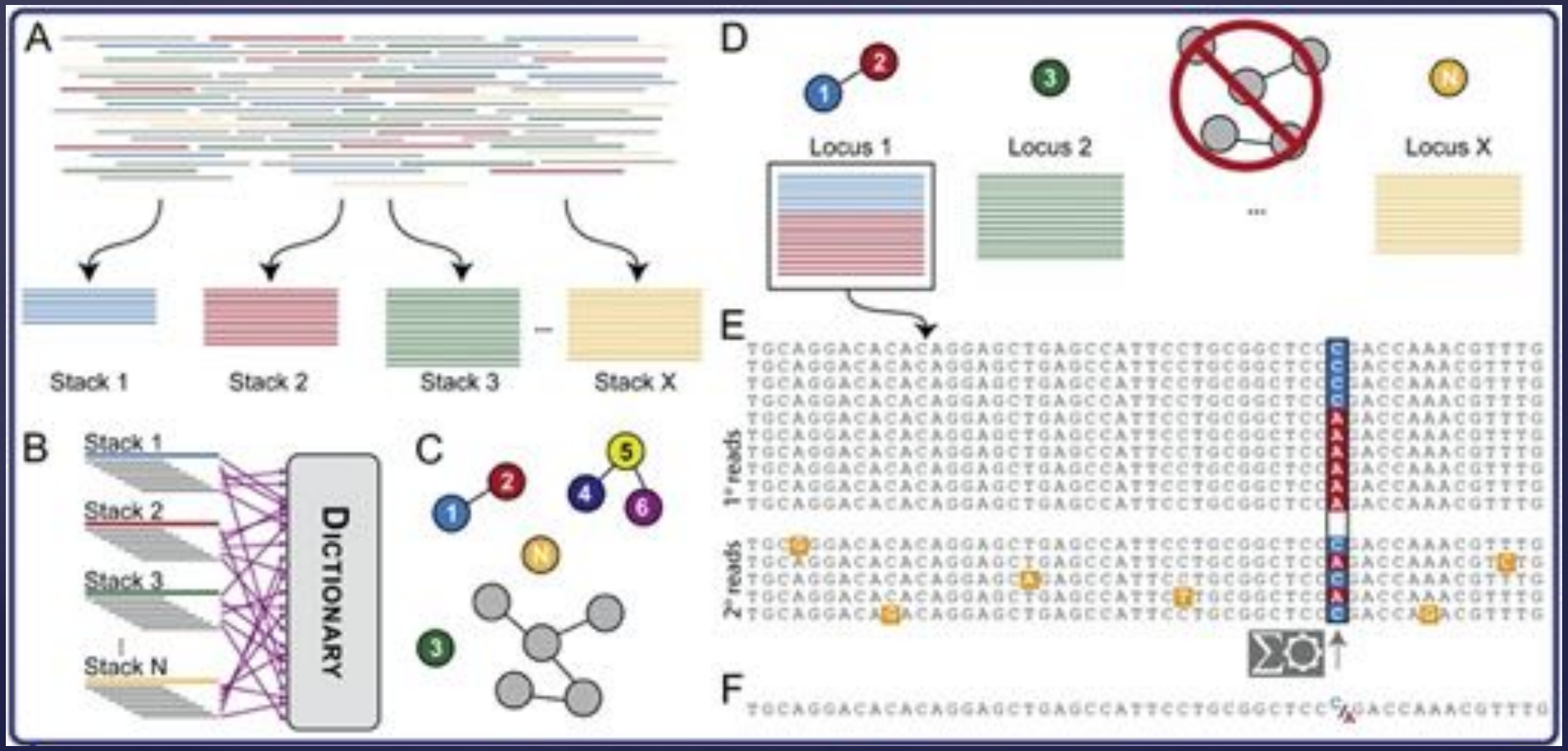
Catchen et al 2011

2. Construcción de loci



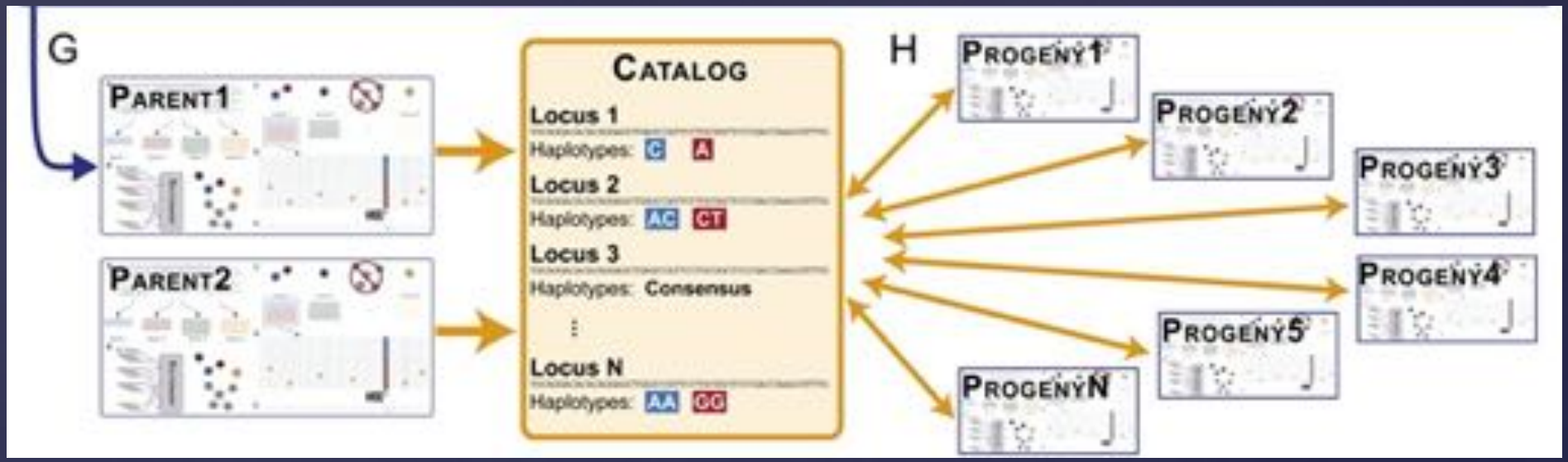
Catchen et al 2011

2. Construcción de loci



Catchen et al 2011

3. Construcción de catálogos de loci y asociar todas las muestras al catálogo

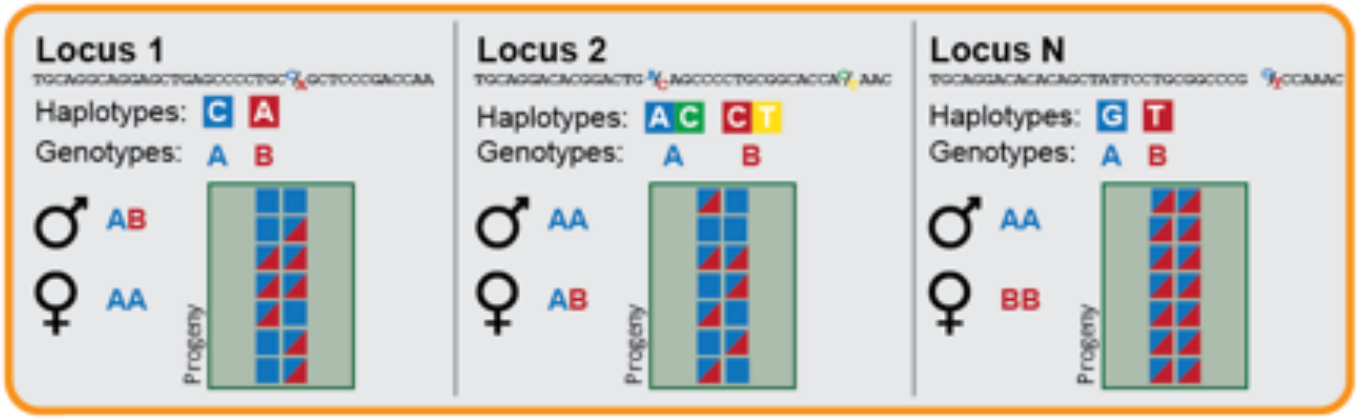


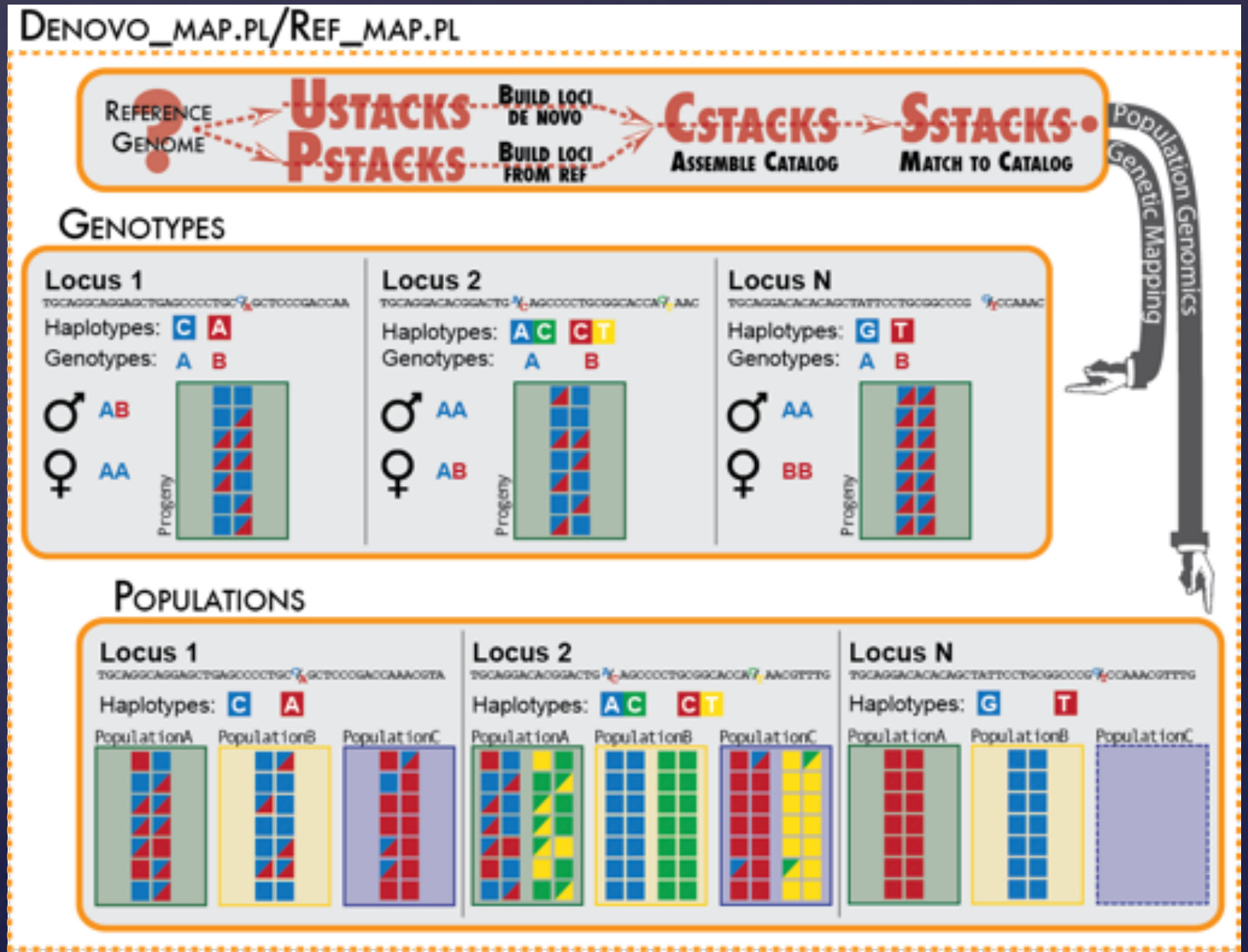
Catchen et al 2011

Denovo_map.pl/Ref_map.pl

Denovo_map.pl/Ref_map.pl

Genotypes




DENOVO_MAP.PL/REF_MAP.PL
REFERENCE GENOME
USTACKS
PSTACKS
BUILD LOCI DE NOVO
BUILD LOCI FROM REF
CSTACKS
ASSEMBLE CATALOG
SSTACKS
MATCH TO CATALOG
Population Genomics
Genetic Mapping
GENOTYPES
Locus 1
Haplotypes: C A
Genotypes: A B
AB
AA
Progeny
Locus 2
Haplotypes: A C C T
Genotypes: A B
AA
AB
Progeny
Locus N
Haplotypes: G T
Genotypes: A B
AA
BB
Progeny
POPULATIONS
Locus 1
Haplotypes: C A
PopulationA
PopulationB
PopulationC
Locus 2
Haplotypes: A C C T
PopulationA
PopulationB
PopulationC
Locus N
Haplotypes: G T
PopulationA
PopulationB
PopulationC

Preguntas?