

Marcadores moleculares (MM)

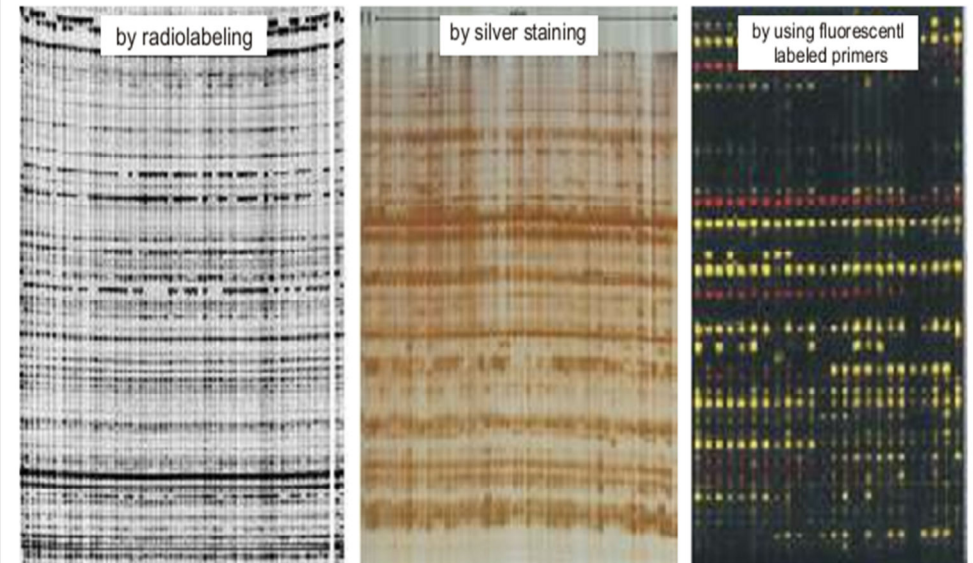
Un **marcador genético** o **marcador molecular** es un segmento de ADN situado en un lugar específico del genoma (locus) y cuya herencia genética se puede rastrear. Puede ser un "gen" o cualquier otra porción de genoma con o sin función conocida.

Se usan para "marcar" el comportamiento de un gen o la herencia de una característica particular, como puede ser la presencia o ausencia de la característica aportada por el MM.

También pueden usarse para el mapeo genético, como el primer paso para encontrar la posición e identidad de un gen desconocido (suponiendo que se conocen regiones contiguas que se heredan juntas).

Pueden ser Nucleares, Mitocondrial, Ribosomal, o cualquier elemento extra cromosomal presente en la especie de estudio.

Different detection systems for molecular markers:



Clases de marcadores

- Morfológicos (estudios fenotípicos)
- Bioquímicos (Expresión Génica)
- Moleculares
(Secuencias de ADN/ARN)

Table 1: FIRST GENERATION DNA MARKERS

Year	Acronym	Nomenclature	Reference
1974	RFLP	Restriction Fragment Length Polymorphism	Grodzicker et al. (1974)
1985	VNTR	Variable Number Tandem Repeats	Jeffreys et al. (1985)
1986	ASO	Allele Specific Oligonucleotides	Saiki et al. (1986)
1988	AS-PCR	Allele Specific Polymerase Chain Reaction	Landegren et al. (1988)
1988	OP	Oligonucleotide Polymorphism	Beckmann (1988)
1989	SSCP	Single Stranded Conformational Polymorphism	Orita et al. (1989)
1989	STS	Sequence Tagged Site	Olsen et al. (1989)

Table 2: SECOND GENERATION DNA MARKERS

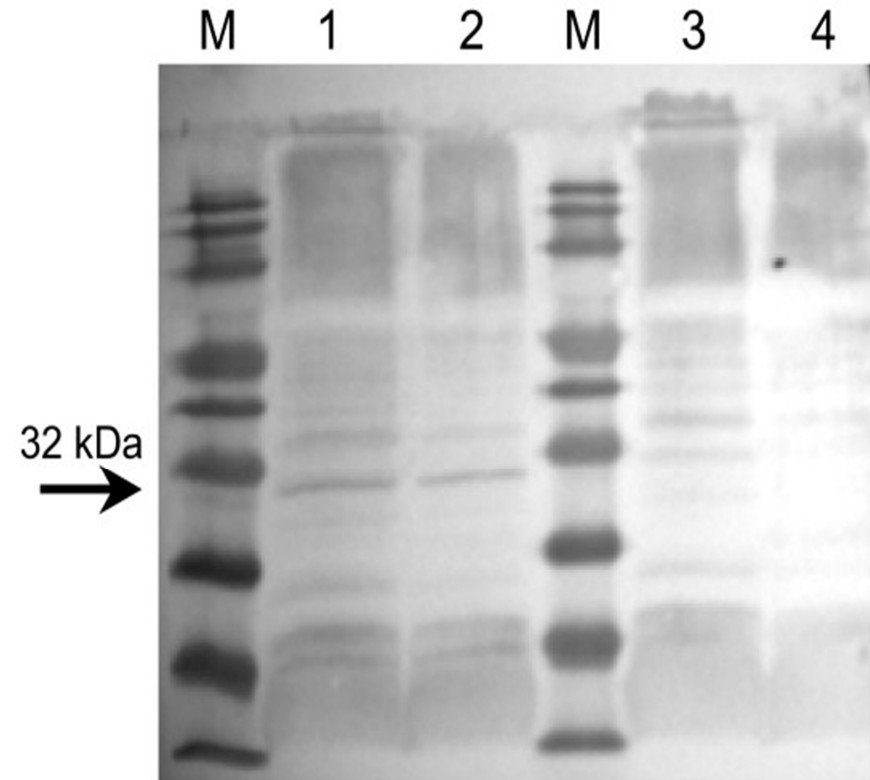
Year	Acronym	Nomenclature	Reference
1990	RAPD	Randomly Amplified Polymorphic DNA	Williams et al. (1990)
1990	AP-PCR	Arbitrarily Primed Polymerase Chain Reaction	Welsh and McClelland (1990)
1990	STMS	Sequence Tagged Micro Satellite Sites	Beckmann and Soller (1990)
1991	RLGS	Restriction Landmark Genome Scanning	Hatada et al. (1991)
1992	CAPS	Cleaved Amplified Polymorphic Sequence	Akopyanz et al. (1992)
1992	DOP-PCR	Degenerate Oligonucleotide Primer - PCR	Telenius (1992)
1992	SSR	Simple Sequence Repeats	Akkaya et al. (1992)
1993	MAAP	Multiple Arbitrary Amplicon Profiling	Caetano-Anollés et al. (1993)
1993	SCAR	Sequence Characterized Amplified Region	Paran and Michelmore (1993)

Table 3: NEW GENERATION DNA MARKERS

Year	Acronym	Nomenclature	Reference
1994	ISSR	Inter Simple Sequence Repeats	Zietkiewicz et al (1994)
1994	SAMPL	Selective Amplification Of Micro Satellite Polymorphic Loci	Morgante and Vogel (1994)
1994	SNP	Single Nucleotide Polymorphisms	Jordan and Humphries (1994)
1995	AFLP (SRFA)	Amplified Fragment Length Polymorphism (selective Restriction Fragment Amplification)	Vos et al. (1995)
1995	ASAP	Allele Specific Associated Primers	Gu et al. (1995)
1996	CFLP	Cleavage Fragment Length Polymorphism	Brow (1996)
1996	ISTR	Inverse Sequence-tagged Repeats	Rhode (1996)
1997	DAMD-PCR	Directed Amplification Of Mini Satellite DNA-PCR	Bebel et al. 1997
1997	S-SAP	Sequence-specific Amplified Polymorphism	Vaugh et al. (1997)
1998	RBIP	Retrotransposon Based Insertional Polymorphism	Flavel et al. (1998)
1999	IRAP	Inter-retrotransposon Amplified Polymorphism	Kalendar et al. (1999)
1999	REMAP	Retrotransposon-Microsatellite Amplified Polymorphism	Kalendar et al. (1999)
1999	MSAP	Methylation Sensitive Amplification Polymorphism	---
2000	MITE	Miniature Inverted-repeat Transposable Element	Casa et al. (2000)
2000	TE-AFLP	Three Endonuclease AFLP	van der Wurff et al. (2000)
2001	IMP	Inter-MITE Polymorphisms	Chang et al. (2001)
2001	SRAP	Sequence-related Amplified Polymorphism	Li and Quiros (2001)

Tipos

- La principal razón para usar caracteres moleculares es que son universales. Por ejemplo, las regiones del genoma nuclear de genes ribosomales son muy útiles para reconstrucciones filogenéticas y donde los caracteres morfológicos presentan limitación en resolución.
- Los “mejores” MM son los que distinguen múltiples alelos por locus (p.e., altamente polimórficos) y codominante (cada alelo se puede observar)



Los más "famosos" MM

- Isoenzimas
- PCR & SEQ
- SSR (simple seq. repeat, microsatellites)
- SNP (Single Nucleotide Polymorphisms)
- RFLP (Restriction Fragment Length Polymorphism)
- RAPDs

Table 1 | Comparison of different molecular markers

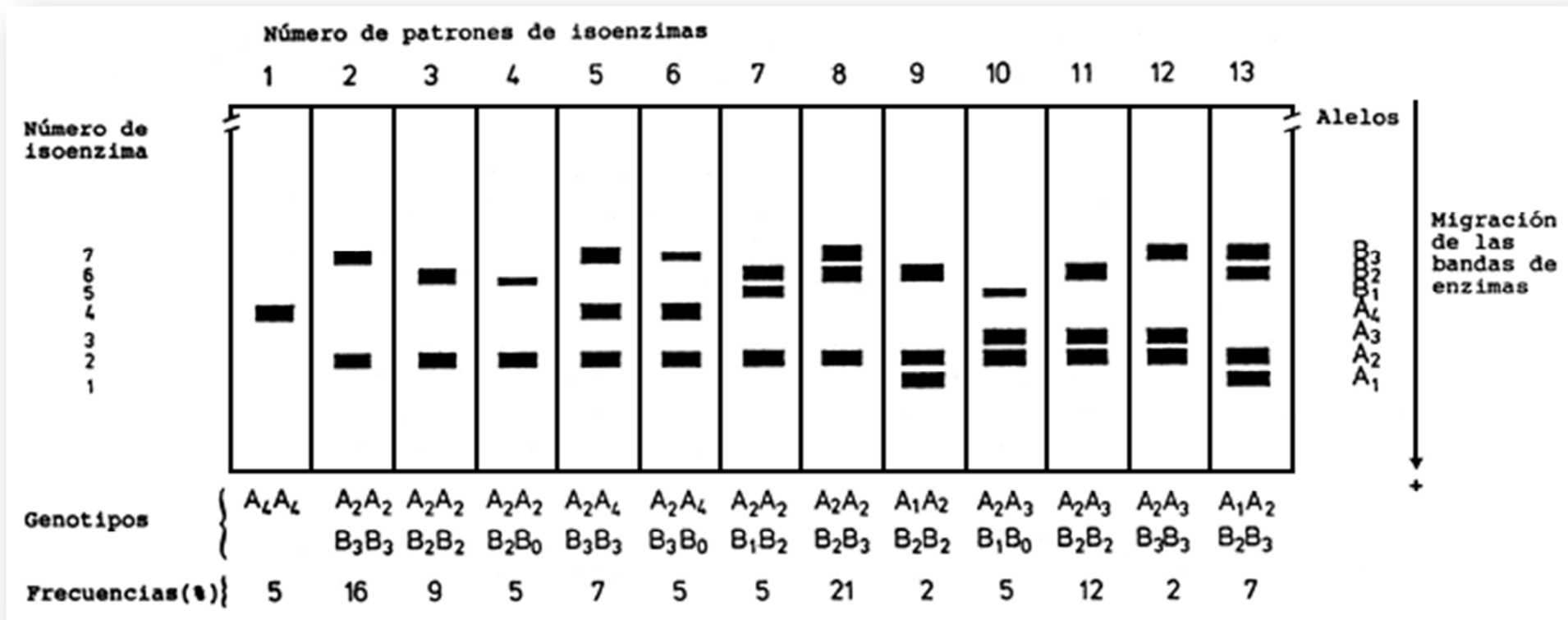
Marker	Advantages	Disadvantages
SNPs	<ul style="list-style-type: none"> • Low mutation rate • High abundance • Easy to type • New analytical approaches are being developed at present • Cross-study comparisons are easy; data repositories already exist 	<ul style="list-style-type: none"> • Substantial rate heterogeneity among sites • Expensive to isolate • Ascertainment bias • Low information content of a single SNP
Microsatellites	<ul style="list-style-type: none"> • Highly informative (large number of alleles, high heterozygosity) • Low ascertainment bias • Easy to isolate 	<ul style="list-style-type: none"> • High mutation rate • Complex mutation behaviour • Not abundant enough • Difficult to automate • Cross-study comparisons require special preparation
Allozymes	<ul style="list-style-type: none"> • Cheap • Universal protocols 	<ul style="list-style-type: none"> • Requirement for fresh or frozen material • Some loci show protein instability • Limited number of available markers • Potentially direct target of selection
RAPDs and derivatives	<ul style="list-style-type: none"> • Cheap • Produces a large number of bands, which can then be further characterized individually (for example, converted into single locus markers) 	<ul style="list-style-type: none"> • Low reproducibility • Mainly dominant • Difficult to analyse • Difficult to automate • Cross-study comparisons are difficult
DNA sequencing	<ul style="list-style-type: none"> • Highest level of resolution possible • Not biased • Cross-study comparisons are easy; data repositories already exist 	<ul style="list-style-type: none"> • Still significantly more expensive than the other techniques

RAPD, randomly amplified polymorphic DNA; SNP, single nucleotide polymorphism.

Isoenzimas o aloenzimas son variantes de la misma enzima

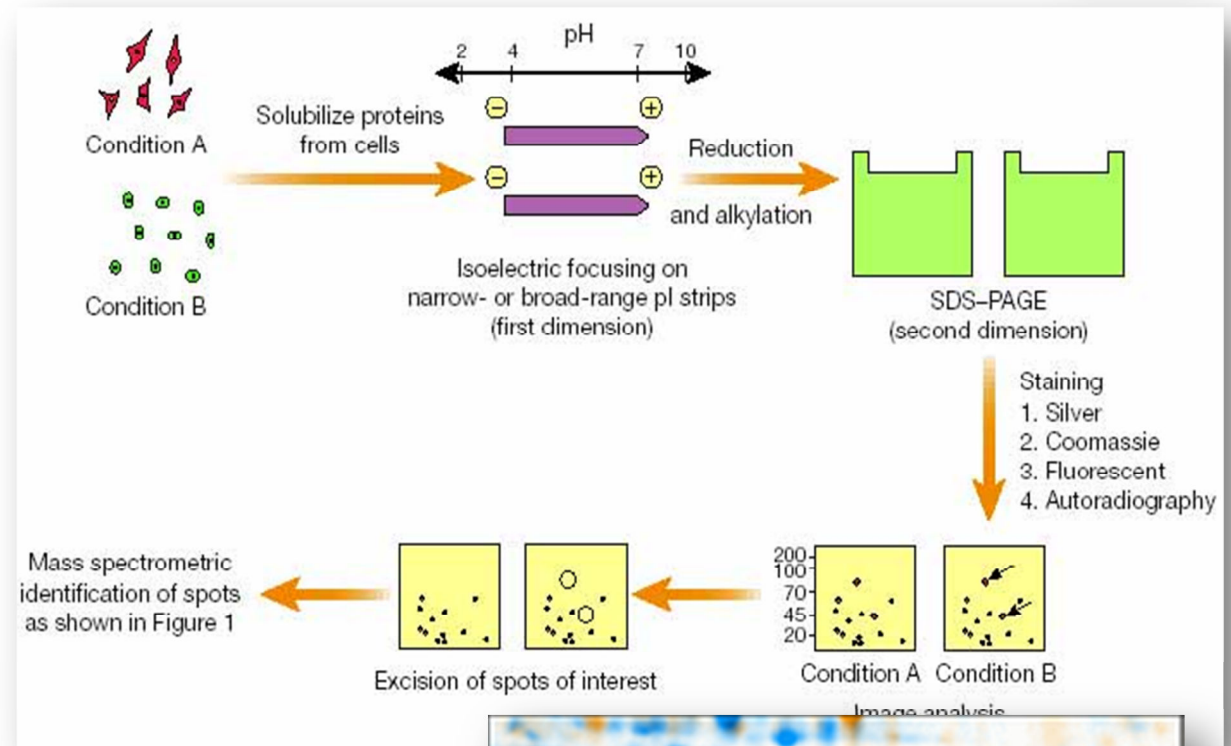
Son enzimas que difieren en la secuencia de aminoácidos, pero que catalizan la misma reacción química. Estas enzimas suelen mostrar diferentes parámetros cinéticos (*i.e.* diferentes valores de K_M), o propiedades de metabólicas diferentes.

En genética de poblaciones es esencial un estudio de las causas y los efectos de la variación genética dentro y entre las poblaciones

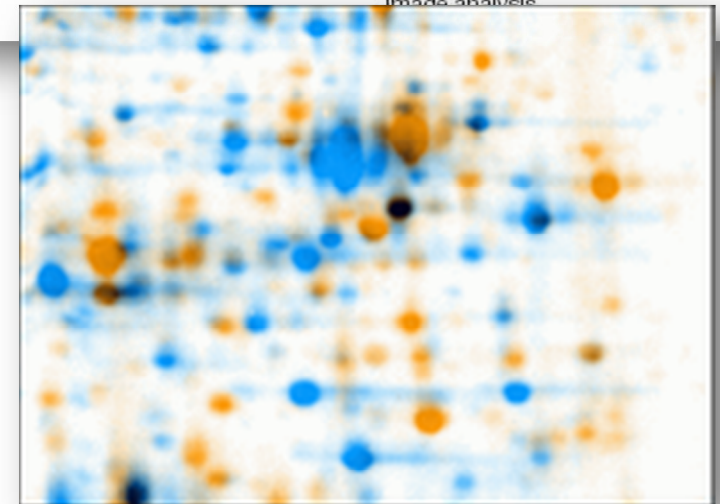


Electroforesis bidimensional en geles de poliacrilamida (2-D PAGE)

2D PAGE permite la separación de cientos o miles de proteínas ("spots") en un único gel, mostrando un patrón característico dependiendo de sus características fisicoquímicas.

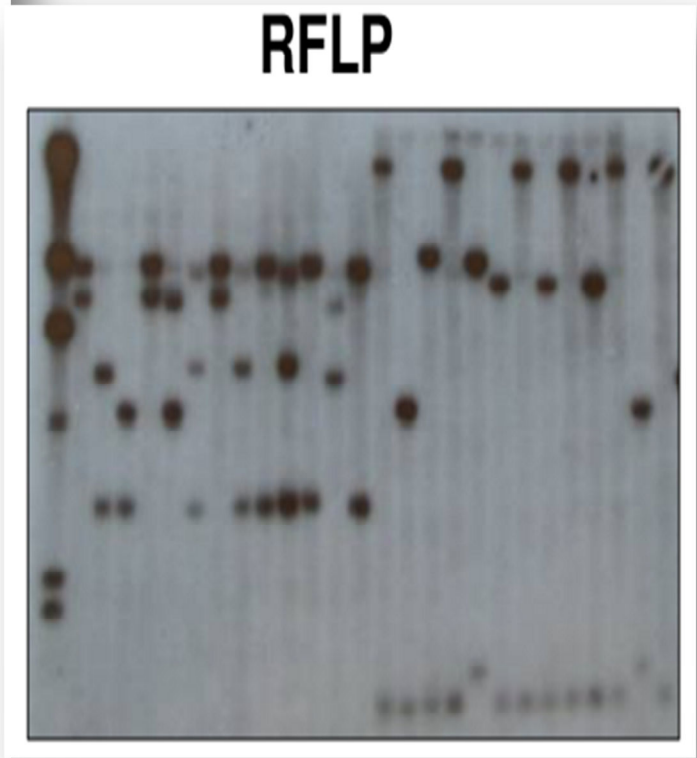
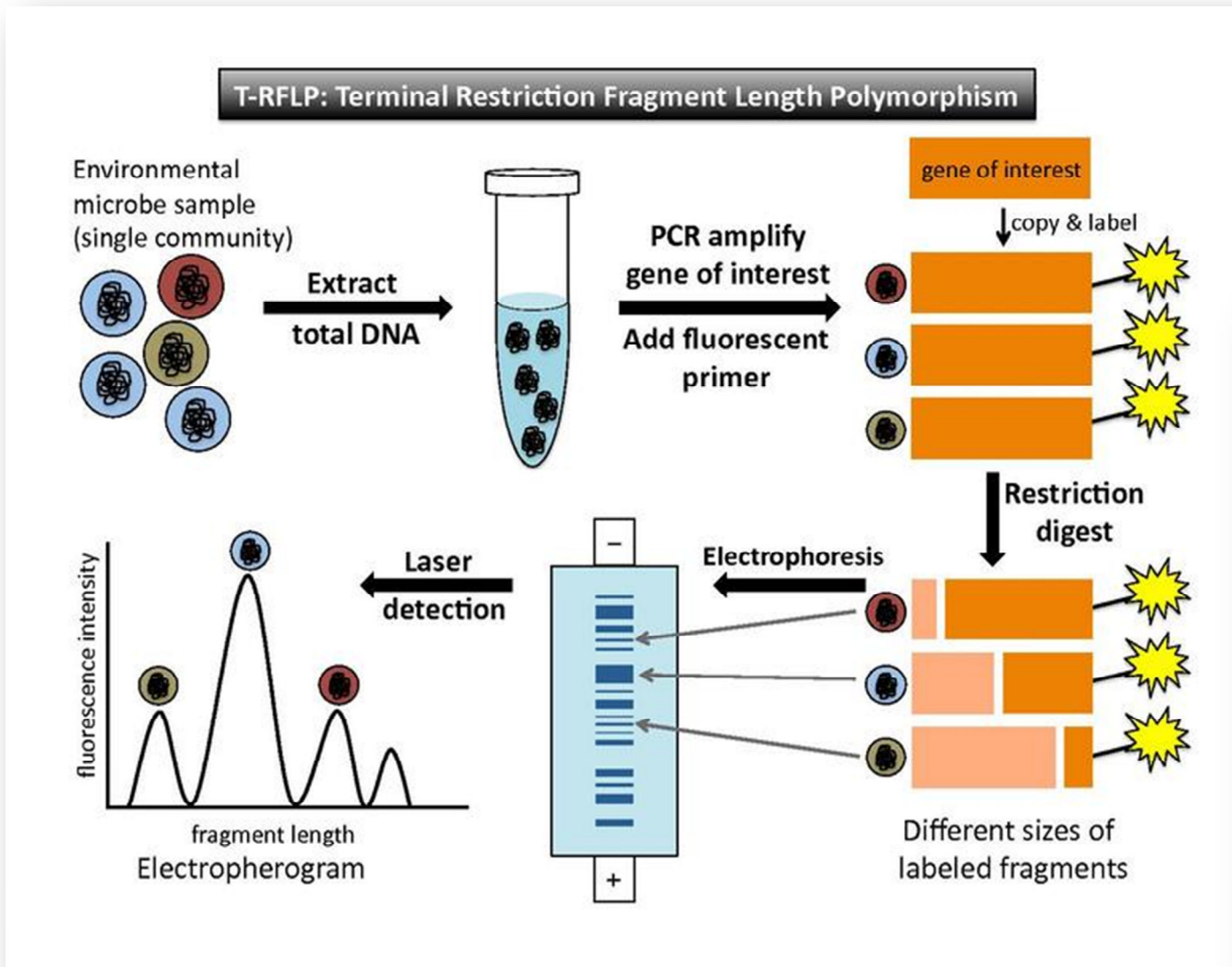


Se hace en dos etapas, primera separa los componentes de la muestra según un criterio (carga, tamaño o pI), segunda según un parámetro distinto del anterior. Al combinar los modos de separación se consigue el máximo de resolución por electroforéticas.



Programas para Análisis de geles 2D

RFLP (Restriction Fragment Length Polymorphism)

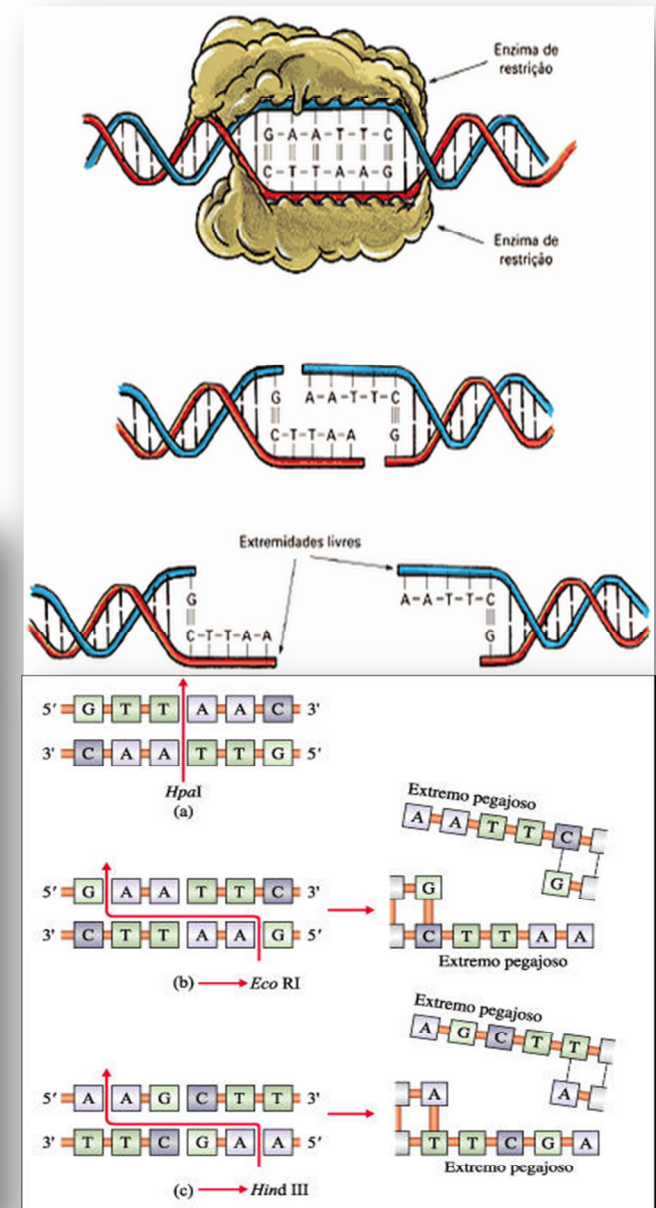
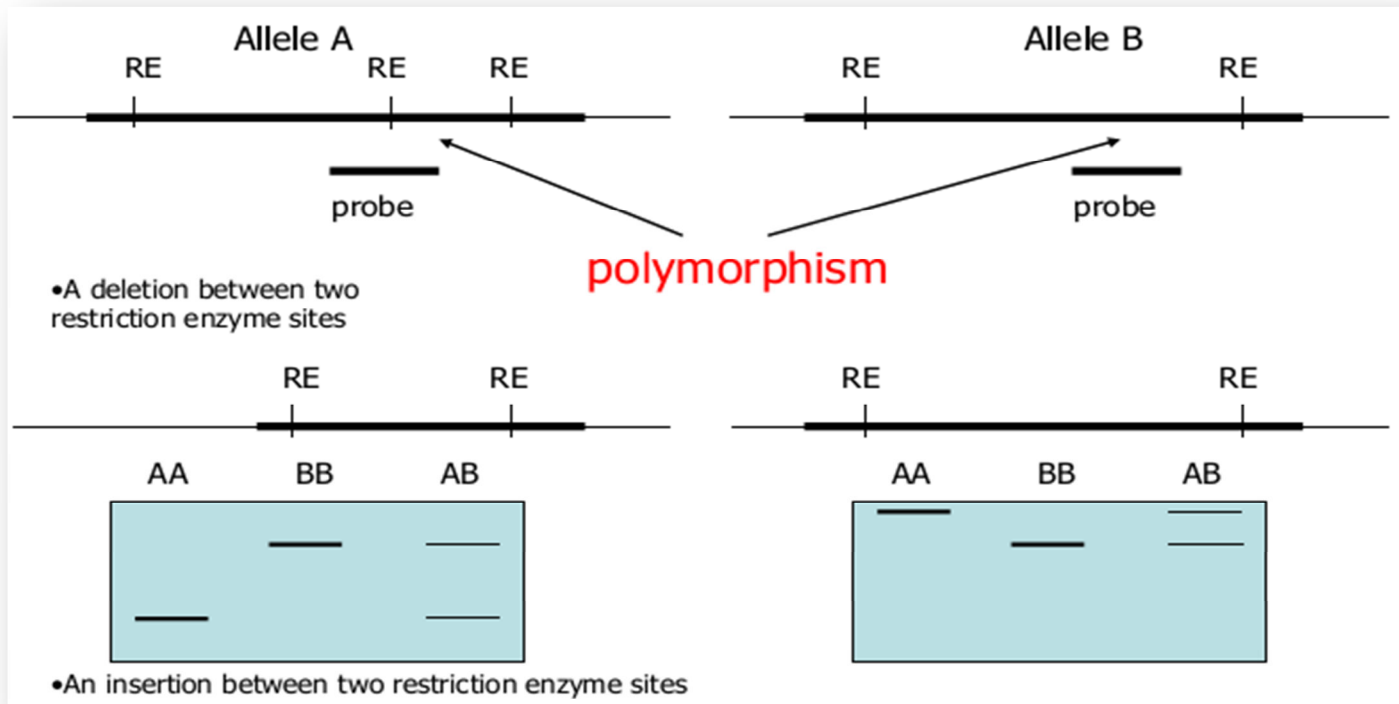


Enzimas de restricción (endonucleasas): puede reconocer una secuencia específica de nucleótidos en el ADN y cortarlo.

Lista de algunas enzimas de restricción

Programas para generar mapas de restricción

Restriction Map / RestrictionMapper / NEBcutter2



PCR & SEQ (Polymerase Chain Reaction)

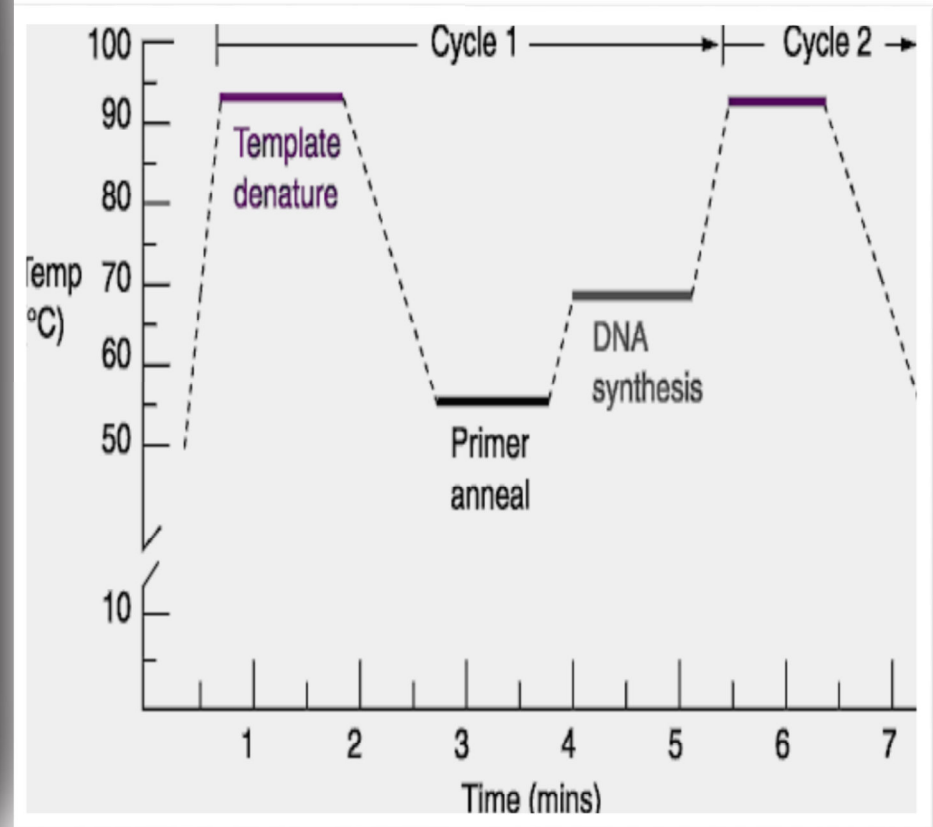
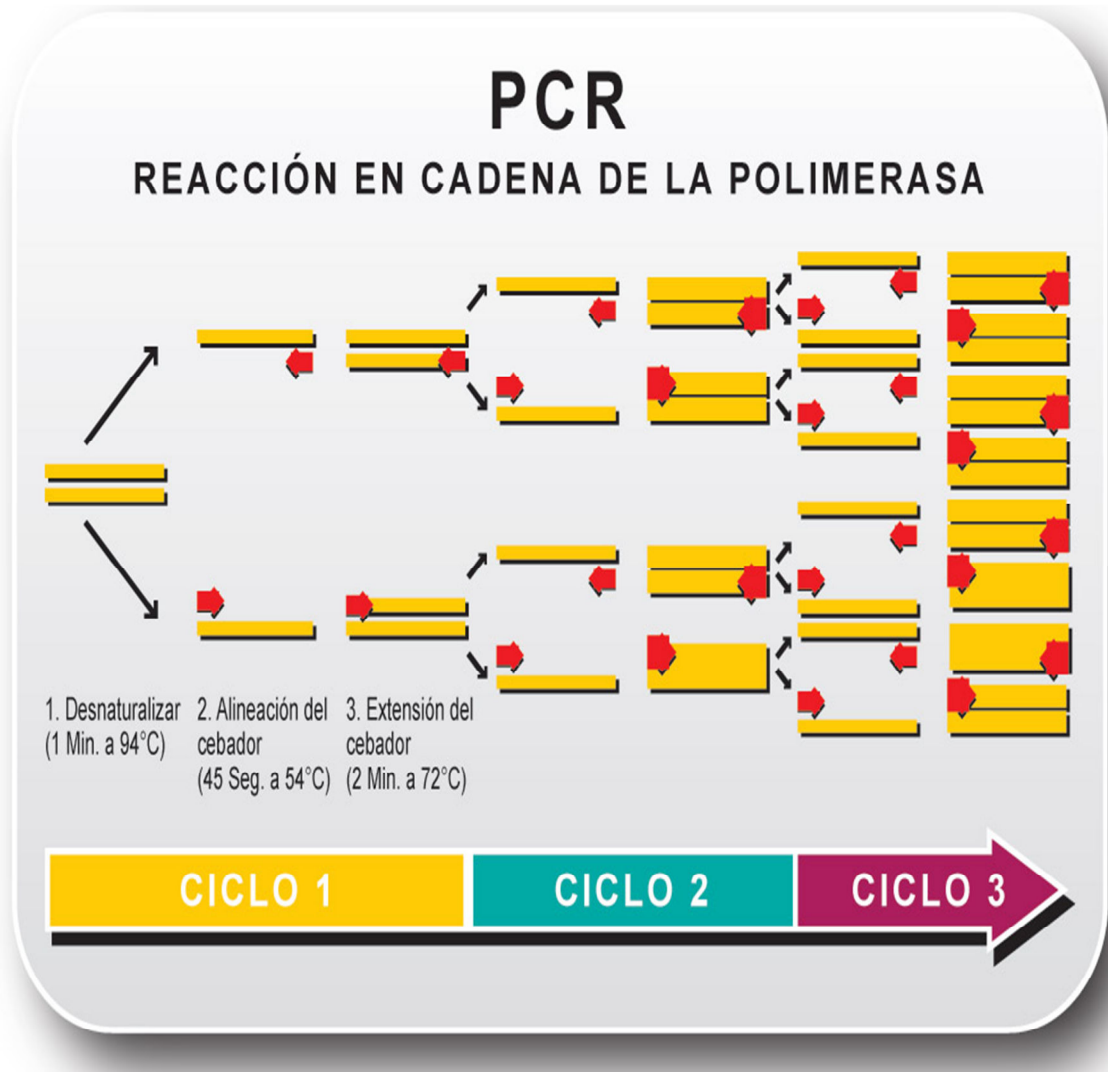


Tabla 1. Condiciones estándar para un PCR
(cálculos para 50 µl en cada tubo)

	Concentración inicial*1	Concentración final en la reacción	Cantidad para un tubo	Cantidad para 10 tubos
dNTPs	10mM (todos) *2	200µM varía de acuerdo al magnesio	1 µl	10 µl
Magnesio	25mM	1.5mM puede probarse de 1 a 4 mM	3 µl	
Oligo forward	10µM	1µM puede probarse de 0.1 a 1 µM	5 µl	
Oligo reverse	10µM	1µM puede probarse de 0.1 a 1 µM	5 µl	
Enzima	5U/µl	1U	0.2 µl	
Buffer	10x	1x	5 µl	
Agua	--	--	29.8 µl	
ADN (agregar después de dividir el mix)	0.1mg/ml	0.1mg genómico (máximo 500 ng; de bacteria de 1 a 10 ng y si es plásmido de 0.1 a 1 ng)	11 µl	

Condiciones estándar de la PCR

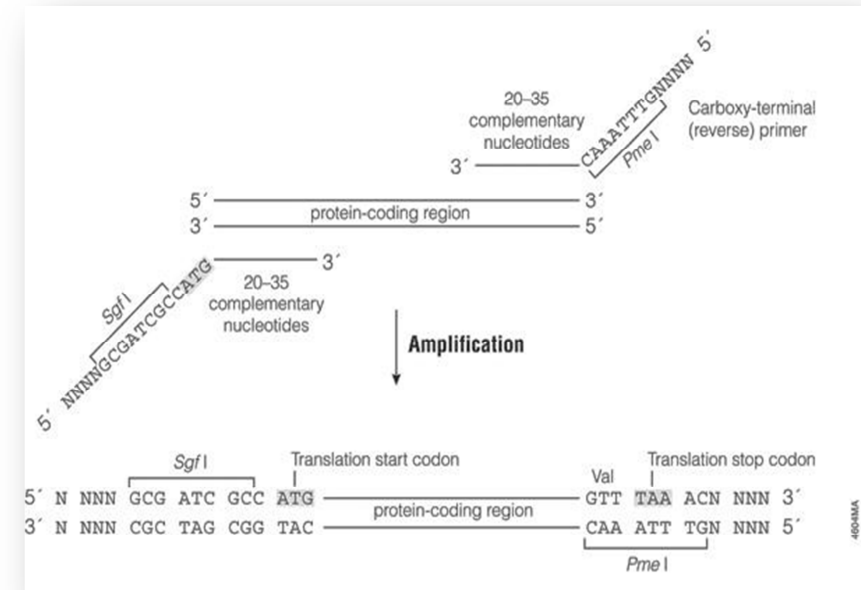
Cycles	Temperature	Time
Denaturation/Activation		
1	95°C	4 min ^a
Amplification Program		
30 - 40 ^{b,c}	Denaturation	95°C 30 s
	Annealing	45 to 65°C ^d 30 s
	Elongation	72°C 45 s-3 min ^e
Final Elongation		
1	72°C	7 min

- ^a) This step serves two purposes. It activates the FastStart Taq DNA Polymerase as well as denaturing the template. For specific applications or to increase yields, you may vary the activation time from 2 min (for sensitive templates) to 10 min (for complex mixtures, e.g., for multiplexing up to 14 bands with 28 primers).
- ^b) Thirty (total) cycles is usually enough to produce an adequate amount of product if the starting template contains >10⁴ copies of the target. For low concentrations of target or complex (e.g., multiplexing) target mixtures, increase the total number of cycles to 40.
- ^c) Gradually increasing elongation time ensures a higher yield of amplification products. To increase the yield of amplification products, modify the amplification program so the elongation time increases gradually (e.g., by 5 seconds/cycle) after 15 cycles of amplification. (For details, see the FastStart Taq DNA Polymerase package insert.)
- ^d) Optimal annealing temperature depends on the melting temperature of the primers and the system used.
- ^e) Elongation time depends on fragment length. Use 1 min/kb for fragments up to 3 kb.

Manual de la PCR de Roche

Los Oligo (Primers)

- La longitud de 18 a 24 bases (en casos especiales de mayor tamaño). Se pueden agregar secuencias adicionales en el extremo 5' no tomarse en la T_m.
- Evitar los nucleótidos degenerados en el extremo 3' (la cola 3' puede ir de 7 a 9 nucleótidos).
- Por lo general se diseñan con un contenido de G/C entre 40 y 60 % (Dependerá de la secuencia blanco y su contenido de G/C). Recomendado esto para la cola 3'.
- LA T_m debe estar entre 55 – 65°C La diferencia entre la T_m de los primers no debería exceder los 5 °C.
- Se recomienda que los primers deban iniciar y terminar con 1 o 2 bases púricas, principalmente en la cola 3'.
- Se deben descartar la formación de estructuras secundarias internas entre los pares de primers (Dímeros de primers y Horquillas >3 pb).



$$T_m \text{ (Básica)} = (x_A + x_T) * 2 + (x_G + x_C) * 4$$

$$T_m \text{ (Salt-Adjusted)} = 81.5 + 16.6x(\log_{10}\{[Na^+] + [K^+]\}) + 0.41x(\%G+C) - 675/n$$

Programas

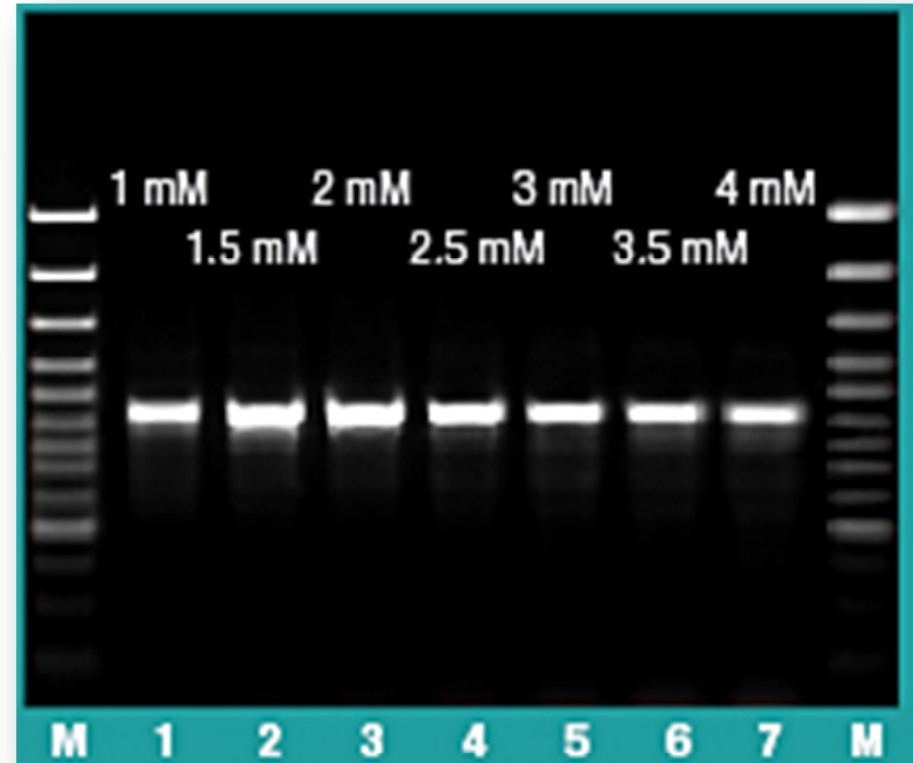
[Primer-BLAST](#) / [Primer3](#) / [BioMath](#) / [SciTools](#)

MgCl₂

La ADN_{Pol} requiere de Mg⁺⁺ para actuar como un cofactor durante la reacción de la PCR, sin el Mg⁺⁺ la ADN polimerasa es inactiva

Muchos de los componentes de la PCR unen Magnesio, la matriz de ADN, agentes quelantes, como el EDTA o Citrato, dNTP's, proteínas y residuos de los procesos de aislamiento del ADN, esto hace que la [Mg⁺⁺] descienda durante la PCR.

Mientras más magnesio reacción de PCR será más "rápido". Lo que puede ser malo, ya que la polimerasa trabajará muy rápidamente lo que puede generar errores de lectura. Al contrario de bajas [Mg⁺⁺] harán que no toda la Taq se active y trabaje muy lento lo que afectará el rendimiento final de la amplificación. Por ello es indispensable optimizar la [Mg⁺⁺] para cada nuevo ensayo (1 - 9 mM) final.



Parameter

Optimal Concentration or Condition

Template DNA

10 pg - 500 ng complex DNA (*e.g.*, human genomic DNA)
10 pg - 100 ng cDNA or plasmid

FastStart Taq DNA
Polymerase

between 2.0 and 5.0 U per 50 μ l reaction

Primers

between 200 and 500 nM (final concentration of each primer)

dNTPs

between 100 and 500 μ M

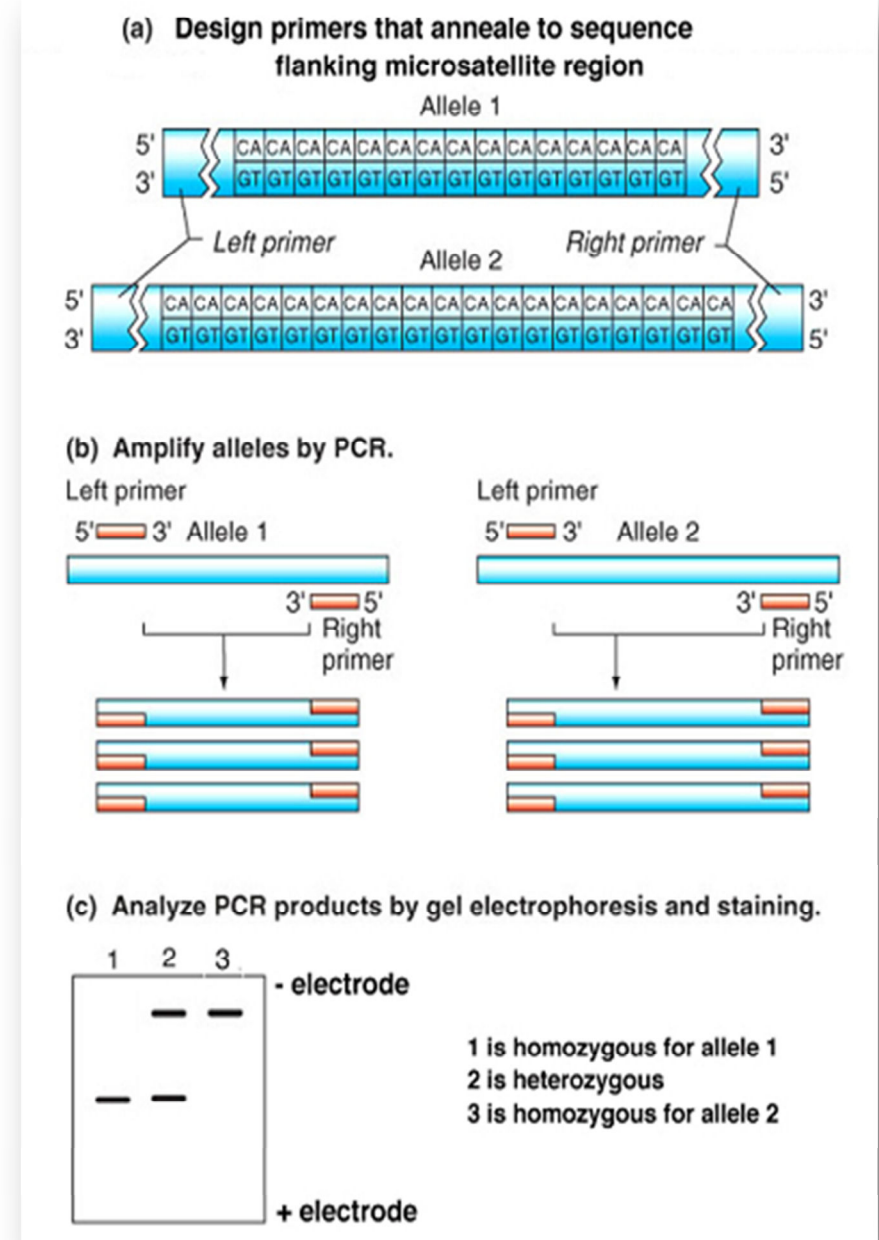
MgCl₂

between 1 and 4 mM^a

- a) For optimization, use the PCR reaction buffer without MgCl₂ (vial 3, supplied with enzyme) and titrate the Mg²⁺ concentration in 0.5 mM steps. Use the 25 mM MgCl₂ solution (vial 4, supplied with enzyme) to produce the following concentrations:

Microsatélites (SSR -simple sequence repeat)

- Secuencias simples de ADN altamente hipervariabilidad abundantes y distribuido en el genoma de eucariotas (algunos procariontas). Se desconoce su función, aunque se cree que intervienen en el empaquetado y la condensación del ADN.
- Consisten en segmentos cortos de ADN con motivos repetidos de 1 a 6 pb. Están precedidos por secuencias conservadas (útiles para la PCR)
- El polimorfismo está determinado por número de repeticiones en un determinado locus,

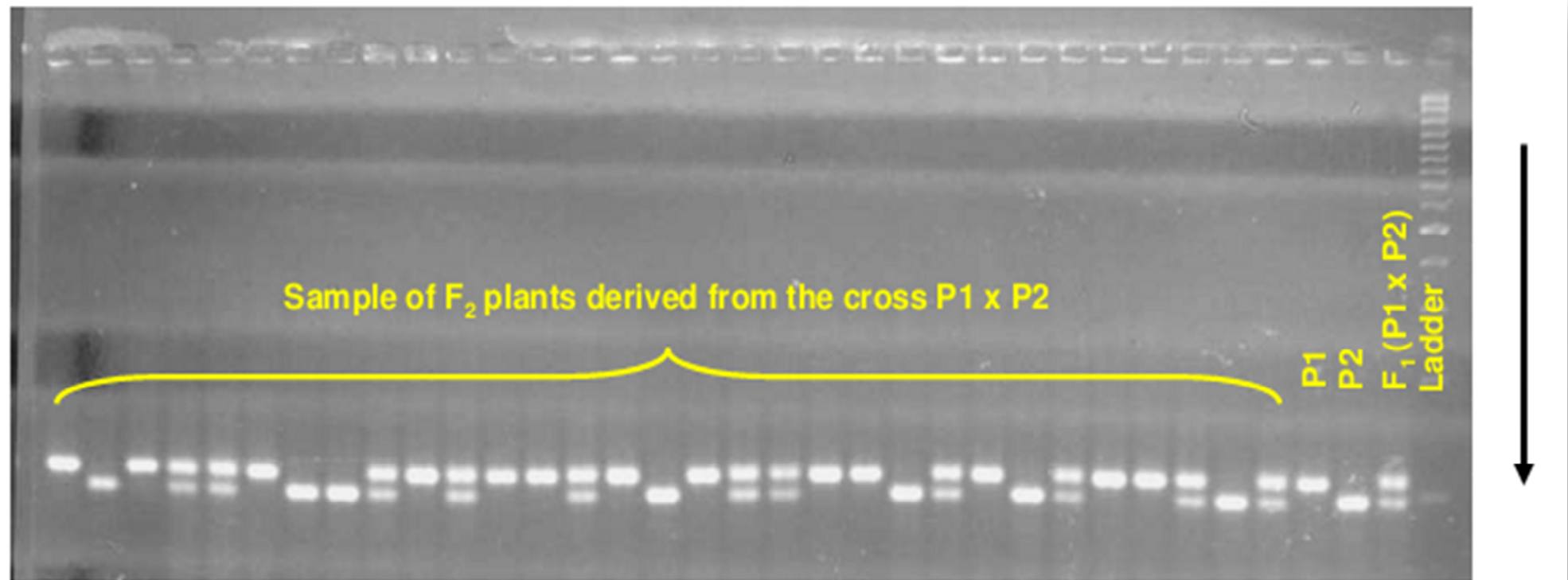


siendo el nivel de variación la longitud de los fragmentos. Se estima la tasa de mutación entre 10^{-2} - 10^{-5} por generación.

- Sus ventajas están en que son multialélicos, polimórficos, herencia mendeliana codominante, interpretación sencilla de los resultados, alta reproducibilidad y se requieren pequeñas cantidades de ADN.
- Útiles para estudios poblacionales, conservados entre poblaciones cercanas o especies muy emparentadas.
- Aplicados en el mapeo de ligamiento, relaciones genéticas entre poblaciones, pruebas de paternidad, entre otros.
- Son abundantes en peces, insectos himenópteros y mamíferos. Menos en aves, plantas y lepidópteros

TIPOS de Microsatélites (SSR)

- **Perfectos:** CACACACACACACACACA
- **Compuestos:** CACACACACACACACAGAGAGAGAGAGAGA
- **Interrumpidos:** CACACACACACACACACATTCACACACACACACACACA



High concentration (3%) gel electrophoresis.

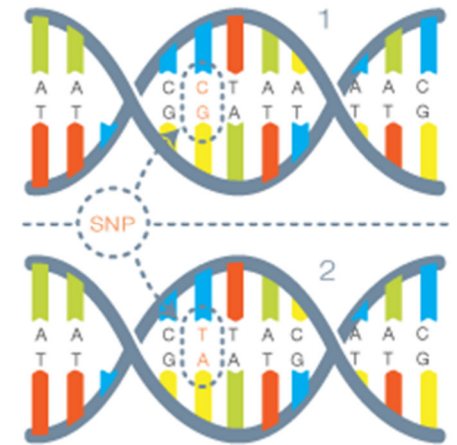
Marker umc1001. Maize.

Programas

[SSR Locator](#) / [Websat](#) / [SSRfinder](#) / [Troll](#) / [IMEx-web](#)

SNP (Single Nucleotide Polymorphisms)

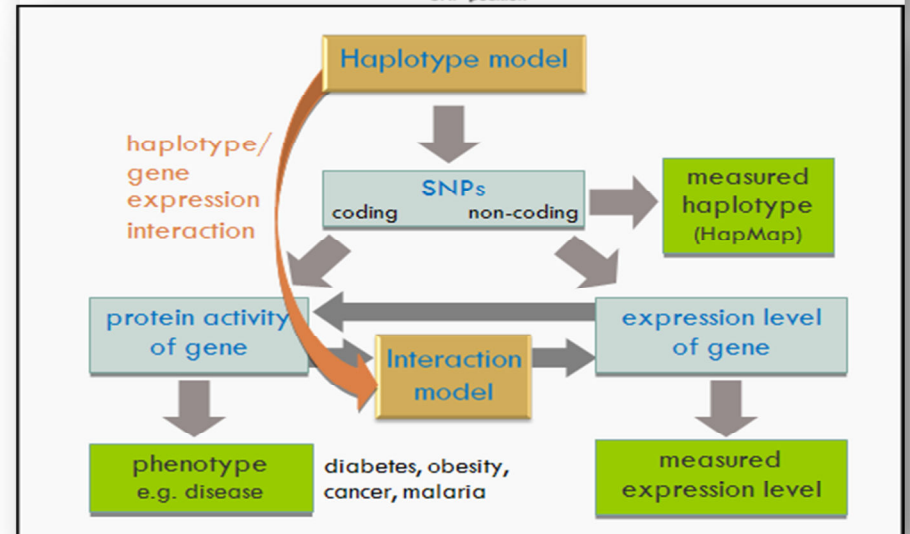
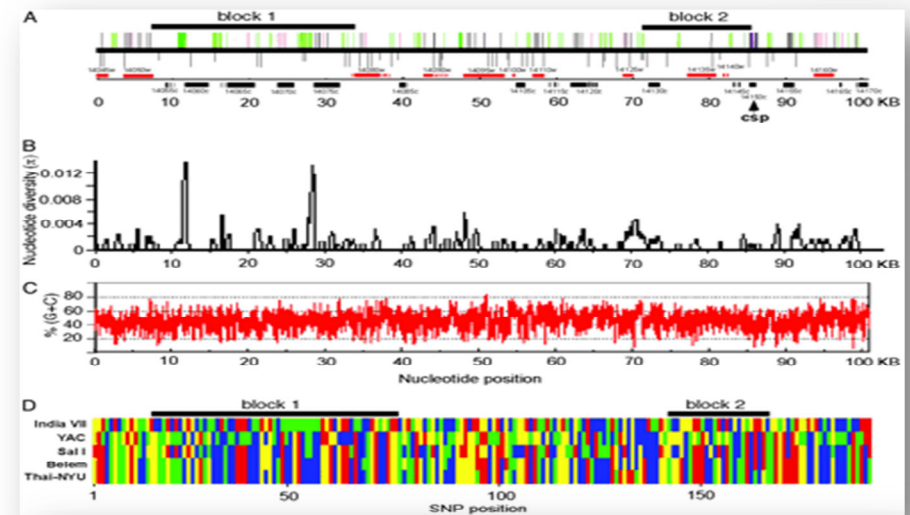
- **Polimorfismo de nucleótido simple**, variación en la secuencia de ADN que afecta a una sola base nucleotídica dentro del genoma.
- La mayoría de SNP se localiza en regiones no codificantes, y no tienen impacto en el fenotipo (SNPs sinónimos). Algunos SNPs pueden ocurrir en secuencias codificantes, y producir cambios a nivel fenotípico y posteriormente funcional (SNPs no sinónimos).
- La variación debe darse al menos en un 1% de la población para ser considerada como un SNP. Si no se llega al 1% no se considera SNP y sí una mutación puntual.
- Son muy abundantes en el genoma humano, a razón de un SNP por cada 1 300 pares de bases y representan el 90% de la variación, siendo dos tercios de estas del tipo C/T y pueden estar involucrados en la respuesta frente a enfermedades.



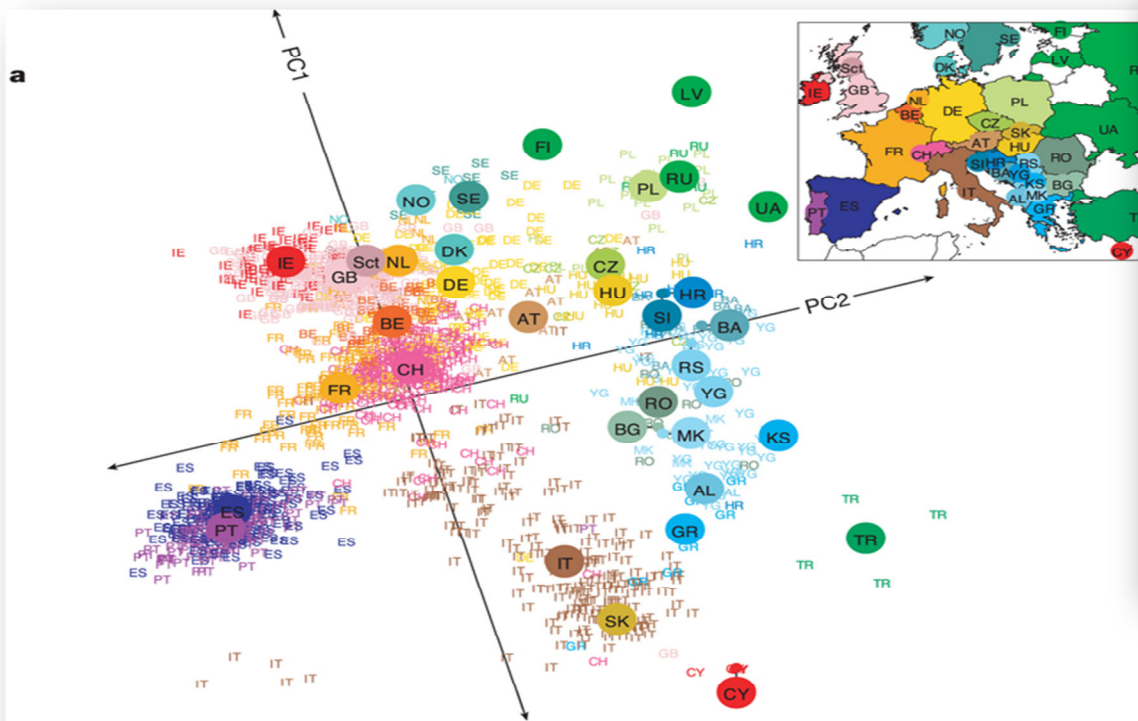
- La variación de los SNPs es muy baja, son buenos marcadores para estudios farmacológicos, evolución de poblacionales y estudios de enfermedades con herencia mendeliana como los de asociación del genoma completo GWA (Genome-wide association study).

Programas

- SNPLims
- SNPpy
- GWAS Analyzer
- SNP tools MS Excel



Bases de datos: HapMap Project / Ensembl / dbSNPs

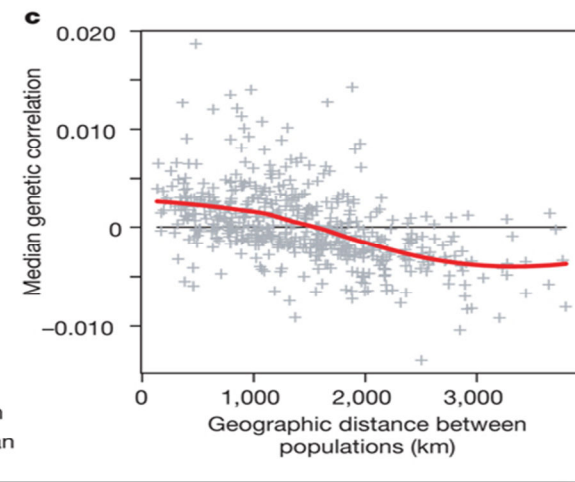
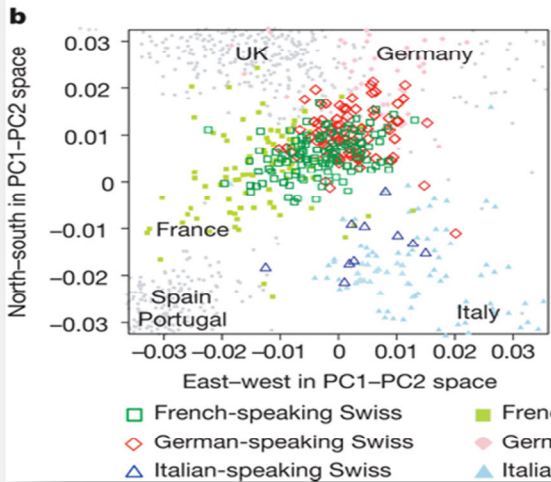


LETTERS

Genes mirror geography within Europe

John Novembre^{1,2}, Toby Johnson^{4,5,6}, Katarzyna Bryc⁷, Zoltán Kutalik^{4,6}, Adam R. Boyko⁷, Adam Auton⁷, Amit Indap⁷, Karen S. King⁸, Sven Bergmann^{4,6}, Matthew R. Nelson⁸, Matthew Stephens^{2,3} & Carlos D. Bustamante⁷

Understanding the genetic structure of human populations is of fundamental interest to medical, forensic and anthropological sciences. Advances in high-throughput genotyping technology have markedly improved our understanding of global patterns of human genetic variation and suggest the potential to use large samples to uncover variation among closely spaced populations¹⁻³. Here we characterize genetic variation in a sample of 3,000 European individuals genotyped at over half a million variable DNA sites in the human genome. Despite low average levels of genetic differentiation among Europeans, we find a close correspondence between genetic and geographic distances; indeed, a geographical map of Europe arises naturally as an efficient two-dimensional summary of genetic variation in Europeans. The resulting figure bears a notable resemblance to a geographic map of Europe (Fig. 1a). Individuals from the same geographic region cluster together and major populations are distinguishable. Geographically adjacent populations typically cluster together, and recognizable geographical features of Europe such as the Iberian peninsula, the Italian peninsula, southeastern Europe, Cyprus and Turkey are apparent. The data reveal structure even among French-, German- and Italian-speaking groups within Switzerland (Fig. 1b), and between Ireland and the United Kingdom (Fig. 1a, IE and GB). Within some countries individuals are strongly differentiated along the principal component (PC) axes, suggesting that in some cases the resolution of the genetic data may exceed that of the available geographic information.



METHODS

Sample collection and genotyping. The samples were assembled and genotyped as part of the larger POPRES project currently consisting of ~6,000 individuals from worldwide populations⁷. The subsample of European individuals used here is derived from two independent collections: the London Life Sciences Population (LOLIPOP) study²⁶, which consists mainly of European individuals sampled in London, and (2) the CoLaus study²⁷, which represents a broad set of European individuals sampled from Lausanne, Switzerland. The combined sample contains individuals with origins from across Europe (Supplementary Table 2), although origins from eastern Europe are generally less well represented (for example, Finland, Latvia, Ukraine, Slovakia and Slovenia) and some countries are not sampled at all (for example, Belarus, Estonia, Lithuania and Moldova).

Genotyping was carried out using the Affymetrix GeneChip Human Mapping 500K Array Set according to published protocol. We observe no significant differentiation in the PCA between individuals collected and/or genotyped at different times (ANOVA, $P > 0.05$). A thorough description of the collections, data processing methods and public data release is presented in ref. 7.

To prepare the sample analysed here, we used the demographic data available for each individual to create a 'geographic origin' that represents a single location from which the individual's very recent ancestry is derived. Where possible, we based the geographic origin on the observed country data for grandparents. We used a 'strict consensus' approach: if all observed grandparents originated from a single country, we used that country as the origin. If an individual's observed grandparents originated from different countries, we excluded the individual. Where grandparental data were unavailable, we used the individual's country of birth.

Nuevas técnicas

La técnica de secuenciación clásica (método Sanger) aunque es capaz de procesar con bastante eficiencia la secuenciación de pequeños fragmentos, tiene la limitante que puede ser un proceso lento y costoso para muestras grandes (por ejemplo, metagenómica).

En los últimos años han surgido novedosas técnicas denominadas Secuenciación de nueva generación, Next Generation Sequencing [NGS] o Massive Parallel Sequencing [MPS], que permiten a un bajo costo abarcar grandes fragmentos de ADN/día.

Rodríguez-Santiago B. et al (2012)

Tabla 1 - Comparativa entre diferentes plataformas de secuenciación

Plataforma	Tiempo carrera ^a	Reads/carrera (en millones)	Bases/read ^b	Rendimiento (Mb/carrera)
3730xl (capilares, no NGS)	2 h	0,000096	650	0.06
Ion Torrent (chip 314)	2 h	0,10	100	>10
454 GS Jr. Titanium	10 h	0,10	400	50
Starlight ^g	?	~0,01	> 1.000	?
PacBio RS	0,5 - 2 h	0,01	860-1.100	5-10
454 FLX Titanium	10 h	1	400	500
454 FLX+ ^c	18-20 h.	1	700	900
Ion Torrent (chip 316)	2 h	1	> 100	>100
Helicos ^d	N/A	800	35	28.000
Ion Torrent (chip 318)	2 h	4-8	>100	>1.000
Illumina MiSeq ^g	26 h	3,4	150+150	1.020
Illumina iScanSQ	8 días	250	100+100	50.000
Illumina GAIIx	14 días	320	150+150	96.000
SOLiD - 4	12 días	>840 ^e	50+35	71.400
Illumina HiSeq 1000	8 días	500	100+100	100.000
Illumina HiSeq 2000	8 días	1.000	100+100	200.000
SOLiD - 5500 (PI) ^g	8 días	>700 ^e	75+35	77.000
SOLiD - 5500xl (4hq) ^g	8 días	>1.410 ^e	75+35	155.100
Illumina HiSeq 2000 - v3 ^f g	10 días	≤ 3.000	100+100	≤ 600.000

h: horas; Mb: megabases.

~ Valor probablemente derivado de información no publicada indisponible en mayo de 2011.

Adaptado de Glenn, 2011⁵⁹.

^a Tiempo necesario en el instrumento para conseguir la longitud máxima de read.

^b Longitud promedio para los reads de alta calidad.

^c Actualización del instrumento FLX, verano 2011.

^d Instrumentos y reactivos no se pueden comprar ya, solo ofrecen servicios.

^e Reads alineables (número crudo de reads de alta calidad).

^f Reactivos y software TruSeq v3 anunciados, reads y rendimiento son la mitad que el HiSeq1000.

^g Información basada solamente en datos de la compañía (datos independientes todavía no disponibles).


Secuenciación de nueva generación (NGS)

A diferencia de los sistemas de secuenciación tradicionales, estas plataformas son capaces de generar paralelamente, y de forma masiva, millones de fragmentos de ADN en un único proceso de secuenciación, a muy alta velocidad y muy bajo costo.

Ideales para estudios que ambicionen procesar grandes

cantidades de ADN como son los estudios de genómica, metagenómica, exomas, genoma completo, smallRNAs, metilación, ChIP-Seq, RNA-Seq.

Platforms:



Life Tech. SOLiD™ Roche 454™ Illumina HiSeq™ Life Tech. Ion Torrent™ Illumina MiSeq™

Applications:

Amplicon Sequencing	RNA Sequencing	FDA Standard Next-Gen
Deep/Targeted Sequencing	Biomarker Detection	SNP/Variant Genotyping
Gene Panel Assays	Methylation Analysis	Microbial ID & Metagenomics

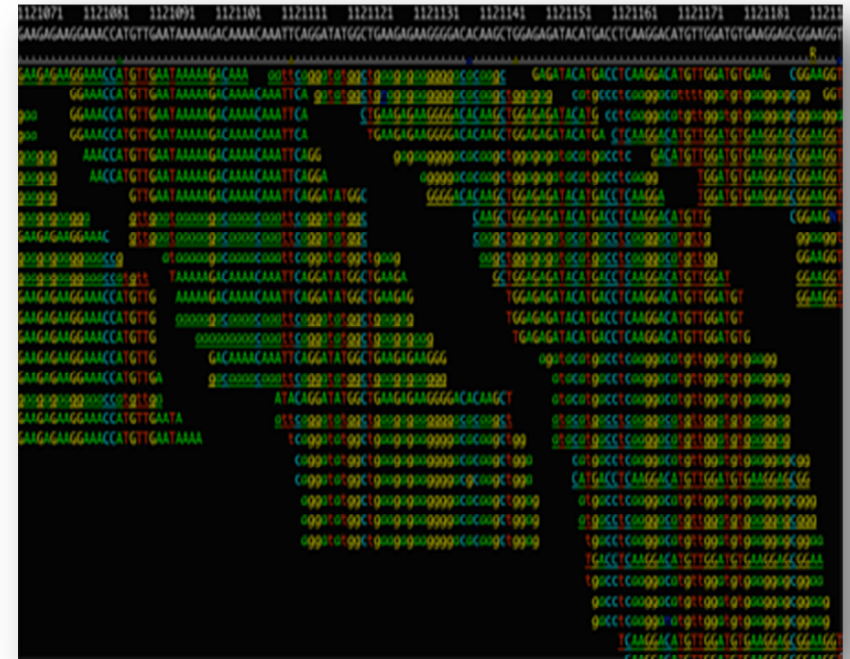
Genoma completo

La secuenciación de última generación (Next-Generation Sequencing o NGS) permite la secuenciación de genomas completos, muy rápidamente y a “bajo” costo.

Las plataformas de secuenciación masiva son capaces de producir cientos de gigabases en un único corrida, lo que permite comparar las diferencias o similitudes en genomas de interés

Metagenómica

es el estudio del conjunto de genomas de un determinado entorno (metagenoma) directamente a partir de muestras de ese ambiente, sin necesidad de aislar y cultivar esas especies.

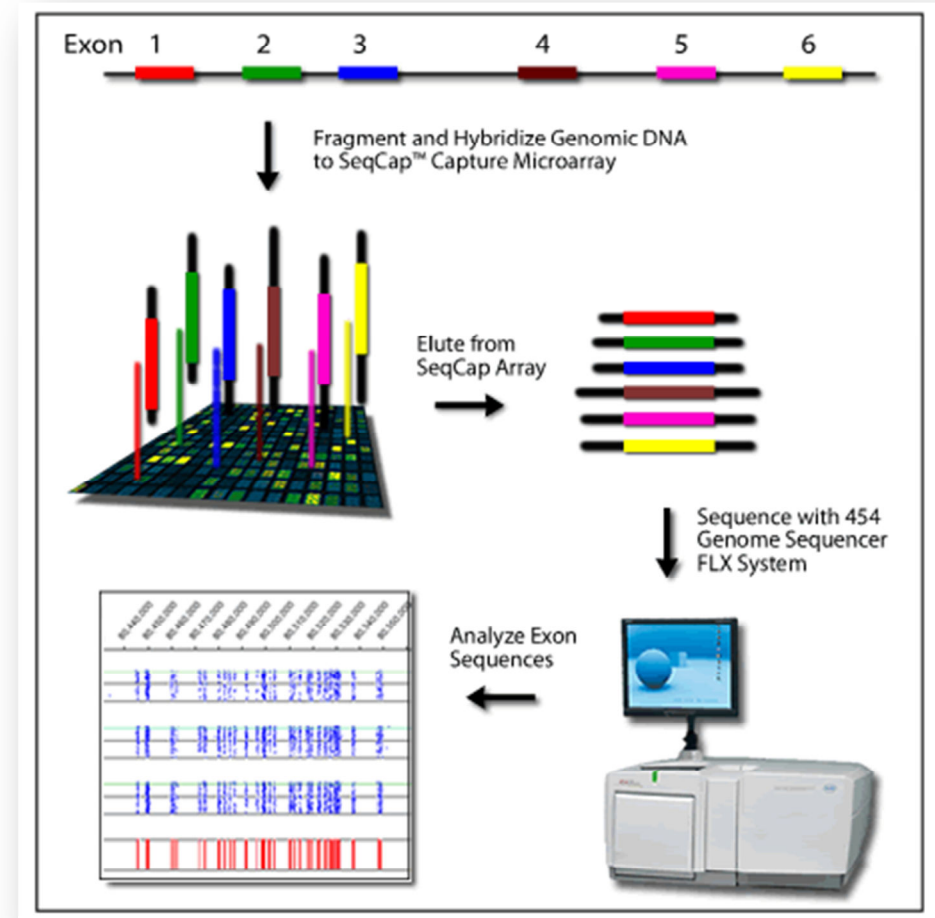


Resecuenciación dirigida

Consiste en el aislamiento, enriquecimiento y secuenciación de regiones específicas del genoma en una muestra. Esta técnica permite la detección sistemática tanto de variantes comunes como variantes raras o poco frecuentes (caracterización de grandes números de genes simultáneamente)

Exomas:

Permite la captura, el enriquecimiento y la secuenciación de regiones genómicas codificantes, lográndose la identificación de nuevos genes asociados tanto a enfermedades raras como comunes.

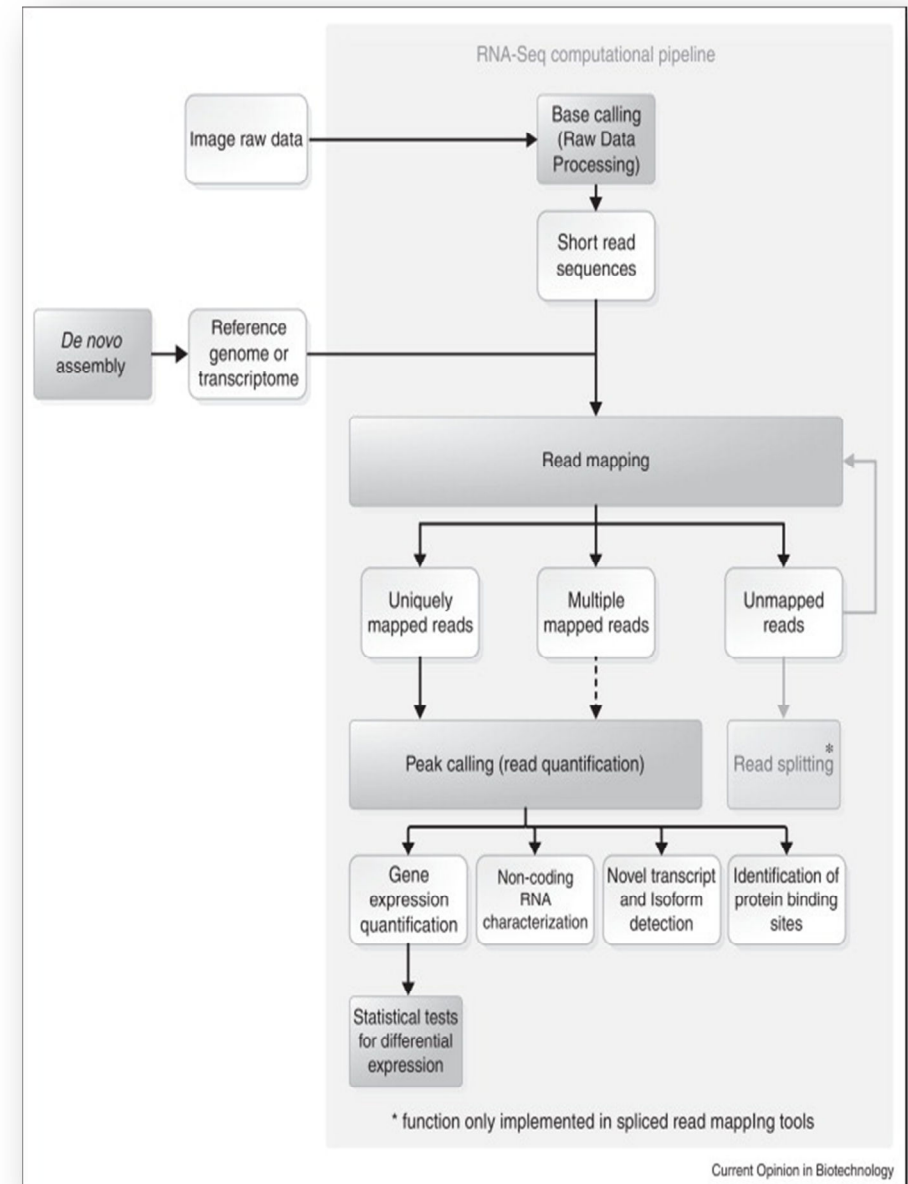


Configuradas

Resecuenciación dirigida de regiones específicas del genoma mediante el diseño de sondas configuradas.

Transcriptómica

RNA Seq - Transcriptoma completo RNA-Seq es la secuenciación masiva de cDNA para la obtención de información global de ARNs de una muestra (mRNAs, rRNAs, tRNAs y otros RNAs no codificantes) en una única o conjunto de células, su expresión génica, tipos y sitios de splicing, fusión génica, SNVs



SAGE

Serial Analysis of Gene Expression, permite el análisis de miles de transcritos al mismo tiempo, secuenciando el extremo 5' o 3' del transcrito. Este método permite la identificación y cuantificación de transcritos.

SmallRNAs

smallRNAs, pequeños ARNs no codificantes que incluyen un gran número de moléculas con funciones y estructuras muy diversas (miRNAs, snoRNAs, piRNAs, etc).

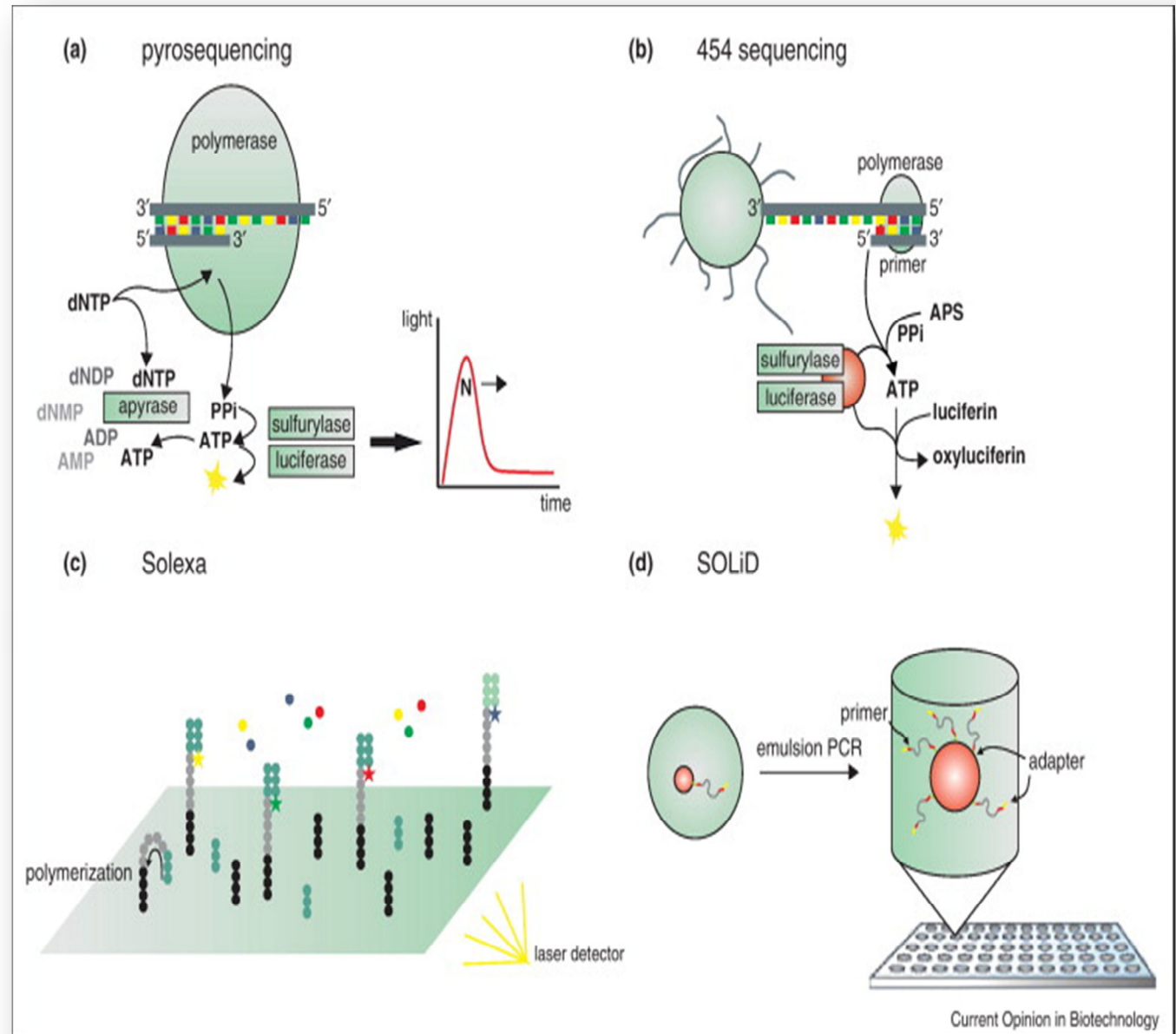
- Chip-Seq

ChIP-Seq combina el método de inmunoprecipitación de la cromatina con la NGS, de esta forma se pueden identificar zonas de interacción proteína-DNA (cistroma).

- Metilación

NGS permiten identificar patrones globales de metilación (metiloma) con una alta fiabilidad en muy poco tiempo y a bajo precio.

Comparison of Next-Generation Sequencing Systems



RNA-Seq

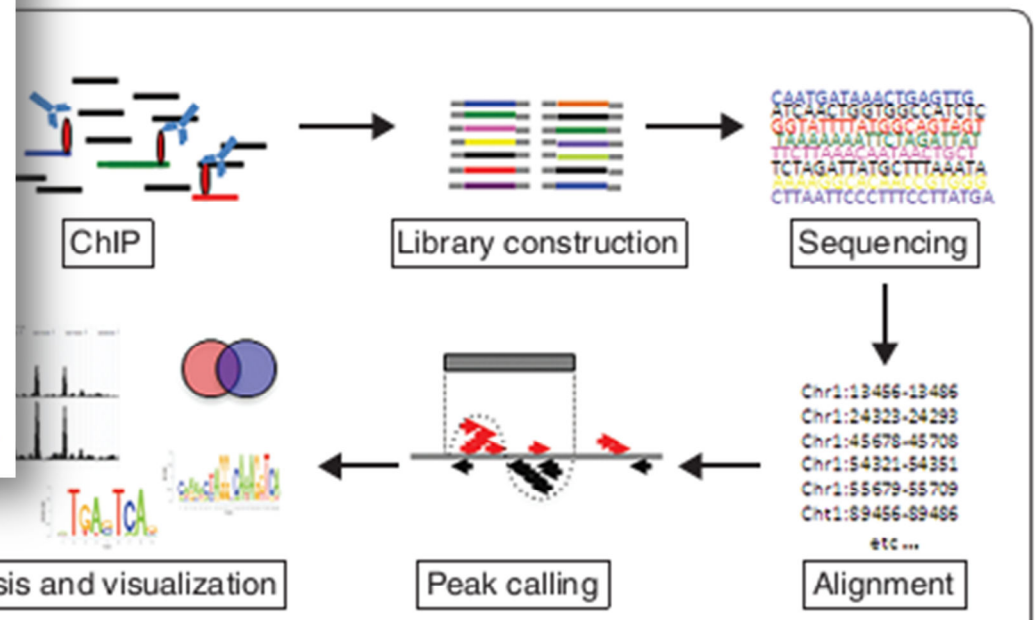
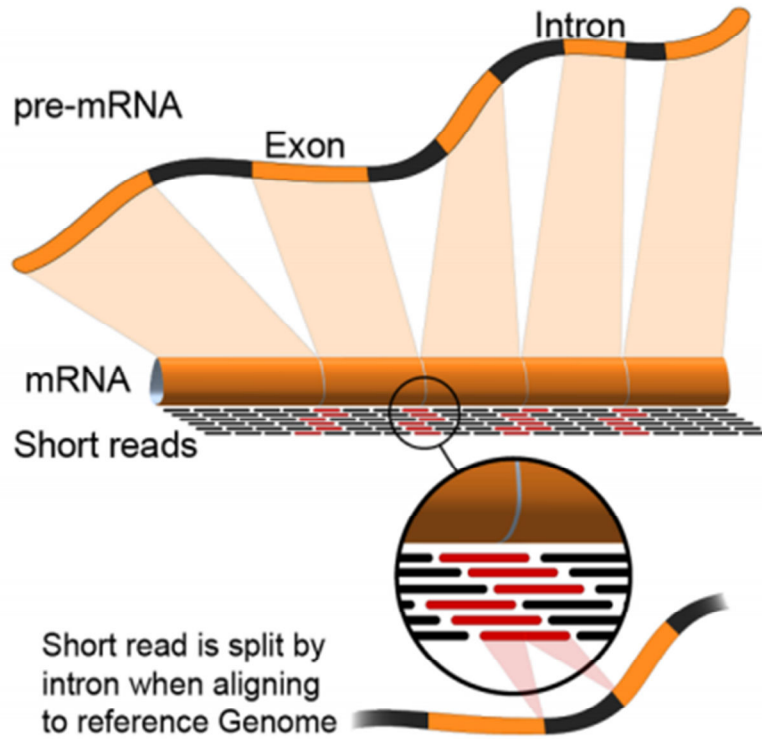


Figure 1. Flow scheme of the central steps in the ChIP-seq procedure.

RNA-Seq: a revolutionary tool for transcriptomics

Zhong Wang, Mark Gerstein and Michael Snyder 2009, Nature Genetics

Technology	Tiling microarray	RNA-Seq
Technology specifications		
Principle	Hybridization	High-throughput sequencing
Resolution	From several to 100 bp	Single base
Throughput	High	High
Reliance on genomic sequence	Yes	In some cases
Background noise	High	Low
Application		
Simultaneously map transcribed regions and gene expression	Yes	Yes
Dynamic range to quantify gene expression level	Up to a few-hundredfold	>8,000-fold
Ability to distinguish different isoforms	Limited	Yes
Ability to distinguish allelic expression	Limited	Yes
Practical issues		
Required amount of RNA	High	Low
Cost for mapping transcriptomes of large genomes	High	Relatively low

