

# **CONCEPTOS Y TÉCNICAS DE BIOTECNOLOGÍA I**

Departamento de Fisiología, Biología Molecular y Celular  
DFBMC, FCEN-UBA

## **Guía de Trabajos Prácticos**

<http://www.fbmc.fcen.uba.ar/materias/biotec1>

## **TRABAJO PRÁCTICO N°1**

### **Producción de amilasas fúngicas**

#### **Objetivos**

1. Aislar cepas fúngicas capaces de producir amilasas.
2. Producir extractos enzimáticos con actividad amilasa.
3. Inmovilizar los extractos enzimáticos conservando su actividad.
4. Determinar la actividad hidrolítica del almidón en los extractos de cada cepa.

#### **Introducción**

La molécula de almidón puede ser degradada por un conjunto de enzimas que incluyen:  $\alpha$ -amilasa,  $\beta$ -amilasa, glucoamilasa y enzimas desramificadoras. Cada una de dichas enzimas posee una actividad característica, generando diferentes productos.

Las amilasas, tanto de origen fúngico como bacteriano, son producidas a gran escala para su utilización en diversos procesos industriales:

- Producción de jarabes de maíz de alta maltosa y alta fructosa
- Producción de espesantes para la industria alimenticia
- Producción de etanol (alimentación y combustible)
- Acondicionamiento de la masa en la producción de pan
- Aditivo en detergentes
- Remoción de almidón (desencolado) en la industria textil
- Suplemento en alimentación animal

#### **PARTE I: *Screening* y aislamiento de hongos productores de enzimas amilolíticas**

##### **Día 1**

Se utilizarán como fuente de hongos amilolíticos, semillas secas de cereales (maíz, trigo) y legumbres (lentejas, arvejas, porotos, etc). Para inducir la germinación de esporas, se colocarán entre 4 y 5 semillas en placas de petri conteniendo YPD agar (1% m/v Extracto de levadura, 2% m/v peptona, 2% m/v glucosa, 2% m/v agar) suplementado con 5 % NaCl (para evitar el desarrollo de Mucorales) y 100mg/L de tetraciclina o cloranfenicol (antibióticos de amplio espectro). Incubar a 28°C durante 7 días.

**NOTA: Observar y anotar las morfologías de crecimiento (aspecto, color, tamaño). Hay diferencias entre las semillas utilizadas? Hubo crecimiento bacteriano? Por qué?**

##### **Día 2: *Screening* de cepas productoras de amilasas y producción de esporas**

Las colonias de hongos (preferentemente las verdes sobre las negras o blancas) se repican con un escarbadietes o ansa estéril cada una por separado y en paralelo a:

- 1) Cajas de Petri con agar YPD
- 2) Cajas de Petri con agar almidón (1% m/v almidón, 2% m/v agar).
- 3) Tubos en pico de flauta con agar salvado (5% m/v salvado triturado, 2% m/v agar)  
Incubar todo a 28°C durante 4-7 días.

**NOTA:** Para que se hacen los repiques en YPD, en agar almidón y salvado? Es correcto hacerlos todos el mismo día? Por qué?

Día 3a: Screening de cepas productoras de amilasas (2ª parte)

Las cepas productoras de enzimas hidrolíticas del almidón se detectan observando el halo de degradación del almidón. Aunque el halo transparente puede a veces observarse directamente, para facilitar la detección, se revelará la presencia de almidón en el medio por el agregado de solución de  $I_2/KI$  (0.026 %  $I_2$  + 0.26 % KI, reactivo de Lugol) que se acompleja con el polisacárido dando coloración azul.

**NOTA:** Observar y medir los halos de degradación de almidón y el crecimiento de cada cepa de hongo (comparar el crecimiento en agar almidón y en YPD).

**Parte II: Producción de amilasas e inmovilización**

Día 3b: Producción de enzimas por fermentación en sustrato sólido

Las esporas de los picos de flauta correspondientes a cepas con actividad amilolítica, se resuspenden en 5 ml de solución acuosa de Sarkosyl 0,05 % (v/v). Se inocula 1 ml de la suspensión de esporas en Erlenmeyers de 250 ml conteniendo 5 g de salvado y 5 ml de agua destilada. Se incuba a 28°C durante 5 días.

**NOTA:** Qué es y qué función cumple el sarkosyl? Por qué la producción de enzimas se realiza utilizando un sustrato sólido y no un medio líquido?

Día 4a: Extracción de las enzimas

Las enzimas extracelulares producidas en la fermentación del salvado por el hongo se extraen con 50 ml de NaCl 1 % (m/v). Agitar durante 30 min a temperatura ambiente. Filtrar por papel de filtro y conservar el eluido (EE).

Día 4b: Inmovilización enzimática

Para la inmovilización de las enzimas se utilizará una matriz de alginato de sodio. El alginato es un polisacárido soluble en agua producido por algas que polimeriza en presencia de iones divalentes como el  $Ca^{2+}$ , generando una matriz en la que quedan atrapadas las enzimas.

- 1) Alicuotar 5 ml del extracto enzimático en un tubo Falcon o vaso de precipitados
- 2) Agregar 5ml de alginato de sodio 2% (1 % final) y homogenizar.
- 3) Utilizando una pipeta automática de 1 ml gotear 3 ml de la solución de alginato-enzima en un vaso de precipitado conteniendo 20-30 ml de 200mM  $CaCl_2$ , contando las gotas de forma de poder estimar el número equivalente a 1 ml. Deben producirse esferas de 3-5 mm de diámetro que polimerizan al entrar en contacto con la solución de  $CaCl_2$ .
- 4) Estabilizar las esferas incubando 1 h en la solución de  $CaCl_2$  a temperatura ambiente y lavar con buffer acetato 0,1 M, pH5.

**Parte III: Cuantificación de la actividad enzimática**

La actividad amilolítica se determinará mediante una ensayo colorimétrico en el que se cuantificará la presencia del almidón remanente de la reacción de degradación, utilizando el reactivo de Lugol

- 1) Preparar una placa de 96 wells con 100 µl de Solución Reveladora R, por pocillo.

Solución Reveladora (R): 0,006 % I<sub>2</sub> + 0,06 % KI en HCl 0,02 M

- 2) Preparar en tubos eppendorf, 1ml de cada punto de la curva patrón de almidón utilizando la Solución Sustrato (S) como stock y el buffer acetato como diluyente, de acuerdo a lo indicado en la tabla más abajo.

Solución sustrato (S): 0,1 % almidón (1mg/ml), 0,15M NaCl, 0,1M acetato de sodio pH 5

Buffer Acetato de sodio pH 5: 0,15M NaCl, 0,1M acetato de sodio pH 5

Almidón	%	0,1	0,08	0,06	0,04	0,02	0
	mg/ml	1	0,8	0,6	0,4	0,2	0
Preparación de la curva patrón	Sn sustrato (ml)	1	0,8	0,6	0,4	0,2	0
	Buffer (ml)	0	0,2	0,4	0,6	0,8	1
A 655nm	1						
	2						
	Promedio						

- 3) Pre-incubar un tubo Falcon con 20 ml de Solución sustrato (10 mg de almidón) en un baño a 28C (disponer de 2 tubos falcon por cada extracto enzimático a medir: uno para el soluble y el otro para el inmovilizado) y un tubo con 20 ml de buffer acetato SIN almidón

Tratar de hacer los siguientes pasos en paralelo:

- 4) Agregar 50 µl del extracto enzimático soluble a uno de los falcon conteniendo la Solución sustrato, mezclar y tomar el tiempo. Tomar 100 µl a los 2, 5, 10 y 20 min y agregarlos por duplicado a los pocillos correspondientes de la placa de 96 wells de acuerdo a la grilla.
- 5) Agregar 50 µl del extracto enzimático al falcon con 20ml de buffer acetato SIN almidón y tomar 100 µl para medir OD. Este será el blanco para la enzima.
- 6) Agregar los 20 ml de Solución sustrato a los 3ml de enzima inmovilizada y tomar 100µl por duplicado a los 5, 10, 20 y 30 min para medir A
- 7) Cargar la placa con las reacciones de enzima + sustrato y la curva patrón de almidón y medir absorbancia entre 580 y 655 nm (preferentemente 640nm), dependiendo de la disponibilidad del lector de microplacas.

	Tiempo	Extractos Solubles					Extractos Inmovilizados					Curva patrón	
		0'	2'	5'	10'	20'	5'	10'	20'	30'			
A	Extracto 1											0	0
B	Duplicados											0,2	0,2
C	Extracto 2											0,4	0,4
D	Duplicados											0,6	0,6
E	Extracto 3											0,8	0,8
F	Duplicados											1	1
G	Extracto 4												
H	duplicados												
		1	2	3	4	5	6	7	8	9	10	11	12

El t<sub>0</sub> es el punto 0,5 de la curva patrón: máxima cantidad de almidón.  
El blanco del inmovilizado es equivalente al punto 0 de la curva patrón

**Cálculo de la actividad enzimática:**

- 1) Trazar la curva patrón y determinar la constante que vincula la absorbancia con la concentración de almidón
- 2) Calcular los mg de almidón hidrolizados con los datos medidos de A para cada tiempo utilizando la constante calculada en el punto anterior.
- 3) Determinar:
  - i) UA/ml
  - ii) UA totales en el extracto EE
  - iii) Rendimiento obtenido en el proceso de inmovilización (%)

**1UA (Unidad amilolítica): la cantidad de enzima capaz de degradar 1mg de almidón en 1 minuto en las condiciones de reacción (28<sup>0</sup>C, pH5).**

5) Ubicar la publicación pionera en identificar la asociación de este gen con metástasis en cáncer de colon (Nat Med. 2009). Explorar en el vínculo “Similar articles” para NM\_182762. Resultados: el buscador devuelve los artículos que fueron encontrados usando un algoritmo que busca referencias que poseen palabras similares en el campo título, resumen y Medical SubjectHeadings (MeSH).

6) Volver al acceso NM\_182762 y explorar el vínculo “Protein”, en “Related information”. Resultado: Obtenemos la secuencia aminoacídica (NP\_877439) correspondiente al acceso NM\_182762.

7) Volver al acceso anterior y explorar el vínculo “Run blast”. Resultado: Obtenemos una vista gráfica de los alineamientos más significativos a la secuencia de interés.

### **Punto 3: Identificar dominios conservados**

8) Volver al acceso correspondiente a la proteína codificada por NM\_182762 y explorar el vínculo “Identify Conserved Domains”. Resultado: obtenemos una lista de dominios conservados identificados en la secuencia de interés.

### **Punto 4: Identificar mutaciones puntuales.**

9) Volver al acceso correspondiente a NM\_182762 y explorar el vínculo SNP. Resultado: Obtenemos una lista de las variaciones alélicas depositadas por laboratorios independientes a la base de datos dbSNP.

### **Punto 5: Visualizar el contexto genómico del gen y obtener la secuencia completa de la región.**

10) Explorar el vínculo “Show in genome viewer” para el link de la proteína NP\_877439, codificada por NM\_182762. Una vez visualizada la región en la cual está ubicado el gen de interés, componer el mapa adecuado para obtener la información necesaria para llevar adelante estudios mas profundos sobre el mismo. Resultado: Este recurso nos permite acceder a las referencias que usa GenBank para estructural la región cromosómica de interés, la cual puede ser usada para seleccionar marcadores, obtener los clones que representan la región y otros recursos que facilitan el estudio detallado de la zona de interés.

## **Segunda parte. Edición y análisis de secuencias de ADN**

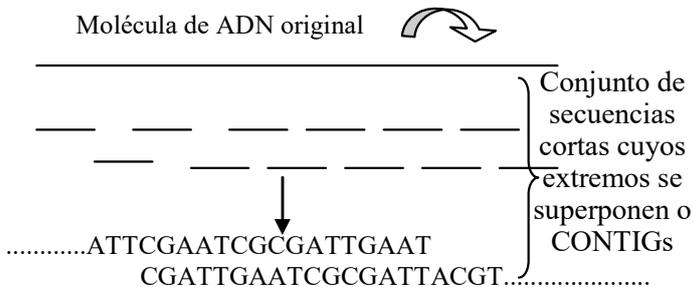
*V. Confalonieri & V. Lia y modificado por R. Heinz*

### **Objetivos**

- 1) Entrenarse en la interpretación de cromatogramas de secuencias provenientes de reacciones de secuenciación automática.
- 2) Entrenarse en el uso de programas de edición de secuencias y en la construcción de CONTIGs de secuencias.
- 3) Realizar una búsqueda de homología de secuencias en una base de datos mundial.

### **Introducción**

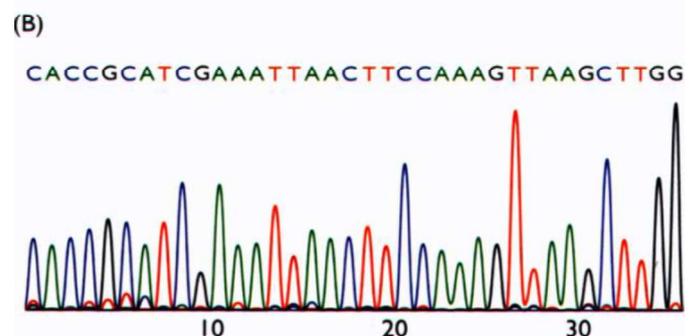
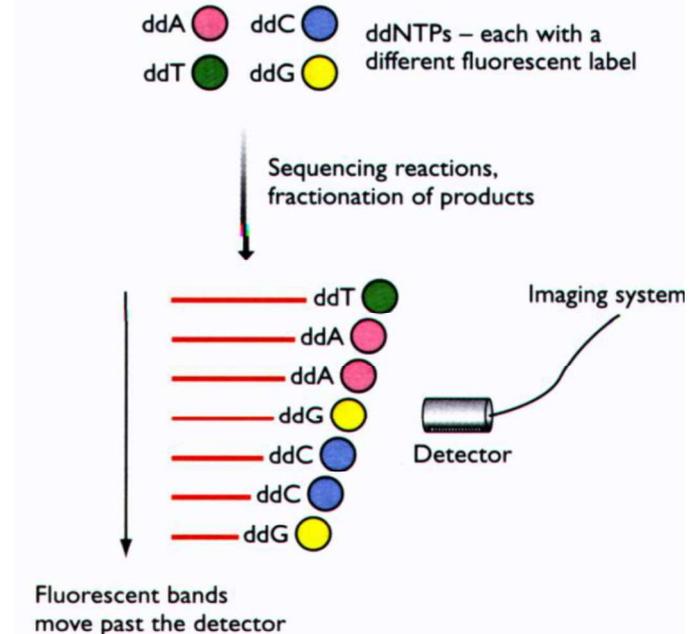
El objetivo último de un Proyecto Genoma es obtener la secuencia de ADN completa de un organismo. En el marco de este tipo de proyectos, la tecnología automatizada se ha convertido en una de las herramientas de elección ya que permite la recolección de datos a gran escala y en poco tiempo. La secuenciación automática del ADN es un ejemplo de aplicación de este tipo tecnologías, la cual está reemplazando cada vez más a la forma manual debido a su costo relativamente bajo, rapidez y mayor precisión en la obtención de los datos.



**Figura 2.**

La secuenciación automática se basa en el método de terminación de la cadena por incorporación de nucleótidos dideoxi (ddNTPs), dando como resultado la síntesis de productos de reacción que tienen el mismo extremo 5' terminal, pero diferentes extremos 3' (ver figura 1).

A principios de los 80, Lee Hood y colaboradores fueron los pioneros en el desarrollo de esta metodología. Ellos aplicaron por primera vez un sistema de detección fluorescente en paralelo a una electroforesis en gel, lo que permitió precisamente la automatización del método. Este instrumento comenzó a ser comercializado por Applied Biosystems Incorporated (ABI), y se convirtió en el caballito de batalla en los comienzos del Proyecto Genoma Humano. El sistema ABI utiliza los cuatro dinucleótidos marcados con fluorocromos que emiten a longitudes de onda distintas. De esta manera, puede llevarse a cabo una única reacción de secuenciación incluyendo las cuatro bases marcadas en un mismo tubo, y resolver los productos de distinta longitud en



una sola calle del gel. El sistema ABI permite, además, la obtención instantánea de los datos. Esto se logra a través de la excitación de la marca fluorescente activada por un lector láser que se encuentra en el punto justo de la corrida electroforética, lo cual maximiza la resolución de los fragmentos. Usando este método, se pueden obtener fragmentos confiables de secuencias de algo más de 500 pares de bases (pb) en menos de cuatro horas. Más recientemente se desarrollaron secuenciadores capilares que permiten la corrida simultánea de 96 muestras que en solo dos horas arrojan secuencias de hasta 1000 pb.

### Secuenciación Automática

Los productos de reacción que se siembran en el gel de secuenciación, provienen de un tipo particular de PCR (cycle sequencing) que implica, a diferencia de la PCR común, la amplificación desde un solo iniciador (one-side PCR). Además, utiliza una Taq polimerasa modificada, y ciclos de temperatura que generan una gran cantidad de productos de cadena terminales a partir de una pequeña cantidad de templado. En la actualidad existen otras técnicas más sofisticadas que pueden implementarse para secuenciar el ADN (p. ej.: pirosecuenciación, basada en detección quimioluminiscente de nucleótidos incorporados o la utilización de CHIPS de secuenciación, descrita en Brown, 1999 o Griffiths *et al*, 2000). Aún con el desarrollo de estas nuevas técnicas de secuenciación, la obtención de la secuencia completa de una molécula larga de ADN debe reconstruirse a partir de secuencias más cortas. Un conjunto de secuencias cortas que se superponen en sus extremos se denomina un CONTIG de secuencias (ver figura 2). Es en este punto (y en muchos otros) en donde interviene la informática como herramienta indispensable para el desarrollo de Proyectos Genómicos. Como explicamos anteriormente existen programas que permiten buscar superposiciones en secuencias y de este modo reconstruir la secuencia original (Alineamientos: locales y globales). Estos últimos se basan en la comparación de son los que se utilizan para la búsqueda de extremos coincidentes en la construcción de CONTIGs, y también en estudios de reconstrucción filogenética. Ambos programas, el CLUSTAL y el BLAST serán utilizados para el desarrollo de este trabajo práctico.

### Acceso a la información sobre Genética Molecular vía Internet

En la década de 1990 se vivió un gran desarrollo en todo lo relacionado con la informática y la comunicación de la información en general. La genética molecular es una de las tantas disciplinas que se ha visto favorecida por este vertiginoso avance. En un principio, el objetivo de los científicos de esta especialidad era el de almacenar y diseminar la información que se fuera obteniendo sobre datos de secuencias de ADN. Esta tarea se llevaba a cabo a través de redes informáticas subvencionadas por gobiernos, y que conectaban distintos centros académicos. Actualmente, en cambio, se utiliza INTERNET operando a través de una poderosa interfase gráfica, la World Wide Web (WWW). Esta red permite la apertura de archivos como objetos multimedia conectados por medio de direcciones de hipertexto. Diversos programas informáticos, llamados buscadores de red (web browsers) permiten a los usuarios usar la interfase con la WWW, lo que, como es bien sabido, se logra a través de computadoras personales conectadas a INTERNET. Es decir que actualmente, cualquier persona, pertenezca o no a un centro académico, puede acceder a una base de datos de secuencias de ADN. Existen muchas vías para acceder, pero uno de los más convenientes es a través del GenBank. Se muestra una serie de direcciones Web de utilidad para los genetistas moleculares.

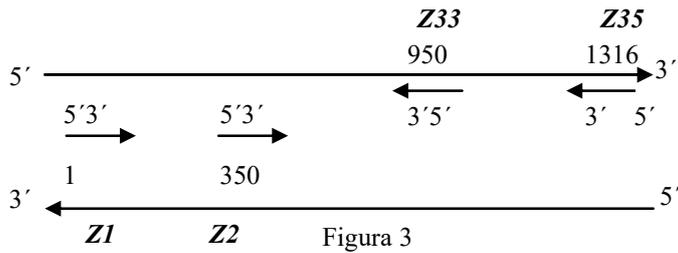
### Desarrollo del Trabajo Práctico

Los alumnos deberán editar una secuencia incógnita de aprox. 1300 pares de bases, utilizando para esto el programa BioEdit Sequence Alignment Editor ([www.mbio.ncsu.edu/RnaseP/info/programa/BIOEDIT/bioedit.html](http://www.mbio.ncsu.edu/RnaseP/info/programa/BIOEDIT/bioedit.html)). Este programa (de utilización gratuita) permite abrir gráficos de cromatogramas de secuencias, editarlos, armar CONTIGs (usando para esto subprogramas como el CLUSTAL y el CAP) y hacer búsquedas en INTERNET de secuencias homólogas mediante BLAST.

Dirección de Internet (http:/)	Descripción de la página
<a href="http://www.nih.gov/">www.nih.gov/</a>	National Institute of Health
<a href="http://www.ncbi.nlm.nih.gov/">www.ncbi.nlm.nih.gov/</a>	National Center for Biotechnology Information
<a href="http://www.ncbi.nlm.nih.gov/BLAST/">www.ncbi.nlm.nih.gov/BLAST/</a>	Sitio para iniciar una búsqueda BLAST
<a href="http://www.ncbi.nlm.nih.gov/ncicgap/">www.ncbi.nlm.nih.gov/ncicgap/</a>	Proyecto Anatomía del Genoma del Cancer.
<a href="http://gdbwww.gdb.org/">gdbwww.gdb.org/</a>	La base de datos del Genoma
<a href="http://www.genome.ad.jp/">www.genome.ad.jp/</a>	Japan GenomeNet
<a href="http://www.tigr.org/">www.tigr.org/</a>	El instituto para la investigación del Genoma Humano
<a href="http://www-genome.wi.mit.edu/">www-genome.wi.mit.edu/</a>	Instituto Tecnología de Investigación del Genoma, Massachusetts.
<a href="http://www.bio.net/">www.bio.net/</a>	BIOSCI bio-net newsgroup.
<a href="http://www.elsevier.com/locate/tto">www.elsevier.com/locate/tto</a>	Elsevier Publishing technical Tips.



La secuencia incógnita se construirá a partir de 4 secuencias más cortas que fueron amplificadas a partir de cebadores externos (Z1 y Z35) y cebadores internos (Z2 y Z33), como se indica en la siguiente figura.



Los cromatogramas correspondientes a los cuatro productos de amplificación se encuentran en 4 archivos diferentes cuyos nombres y ubicación en la computadora serán indicados por el docente.

Para poder visualizarlos y analizarlos se seguirán los siguientes pasos:

- 1) Ingresar al programa BioEdit
- 2) Abrir la carpeta que contiene los cuatro archivos con los cromatogramas (seleccionando FILE, OPEN de la barra de herramientas), y dentro de ella, abrir uno de los archivos de secuencias. Se abrirá una ventana que contiene una secuencia y su cromatograma correspondiente (bajo el título ABI chromatogram) y otra ventana que contiene la secuencia escrita solamente o "sólo texto" (bajo el título DNA sequence), como lo muestra la figura 4. En ambos casos, dichas secuencias se encuentran numeradas cada 10 pares de bases.

- 3) Observar los cromatogramas en cuanto a los colores usados para cada base, forma de los picos, picos superpuestos y posiciones con base indeterminada (N).

Identificar las regiones del cromatograma que permitan definir la secuencia con mayor certeza. Esta elección se basará en la presencia de picos bien definidos, baja frecuencia de superposición de picos en una misma posición que den lugar a bases no definidas (N), baja frecuencia de bases repetidas en tandem muchas veces.

- 4) Registrar las posiciones iniciales y finales de la porción utilizable.
- 5) Dentro de cada región elegida (porción utilizable), ubicar las bases ambiguas (N) y resolverlas en base al cromatograma. Para esto, se debe primero transformar al cromatograma en editable, seleccionando VIEW, EDITABLE SEQUENCE.

- 6) Guardar los cambios. Para esto seleccionar la opción EXPORT as FASTA del menú FILE de la barra de herramientas y adjudicarle un nombre de archivo.

- 7) Repetir este procedimiento para cada secuencia.

- 8) Eliminar las regiones no utilizables de los cromatogramas de la siguiente manera: a) abrir los archivos recién creados; b) ubicarse sobre el nombre de la secuencia que aparece a la izquierda de la ventana (Ej.: PT/Z1); c) seleccionar la opción EDIT SEQUENCE del menú SEQUENCE de la barra de herramientas, lo que permitirá que se abra una segunda ventana; e) pintar con el botón derecho del "mouse" las bases a borrar. Eliminar con DELETE, primero las últimas de la secuencia y después las primeras; f) oprimir "APPLY and CLOSE". Las modificaciones se habrán incorporado a la ventana de secuencias "sólo texto"; g) guardar los cambios con FILE, SAVE; h) repetir esto con cada secuencia.

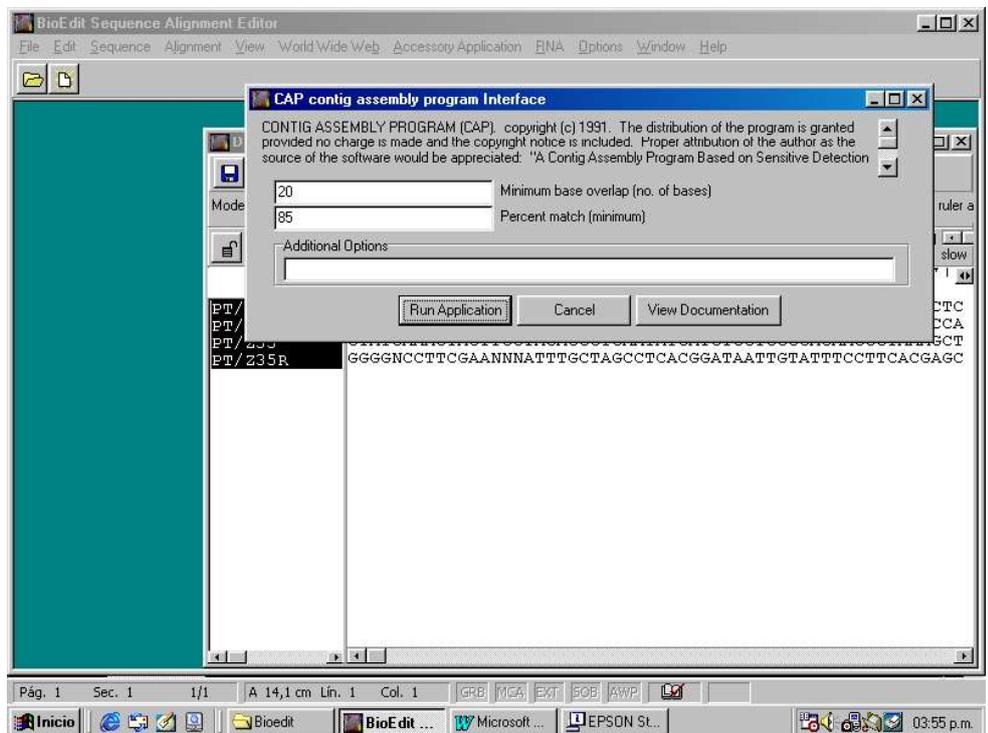
#### Armado de los CONTIGs de secuencias

- 1) Agrupar todas las secuencias en un mismo archivo. Para esto, ubicarse en cada ventana, pintar el nombre de la secuencia ubicado a la izquierda, seleccionar la opción COPY SEQUENCE, del menú EDIT y pegarlo en el archivo de destino con la opción PASTE SEQUENCE, del mismo menú EDIT.

- 2) Guardar el nuevo archivo (FILE, SAVE).

- 3) Para poder armar los CONTIGs, las cuatro regiones secuenciadas deben corresponderse a la misma hebra y deben estar escritas en la misma dirección (ej 5' 3'). Dada las características de los productos de amplificación de los cuales provienen estas secuencias (figura 3), habrá dos hebras que deberán ser modificadas para que cumplan con ambos requisitos. Elegir las secuencias a modificar (pintar nombre de la izquierda) y seleccionar de la barra de herramientas SEQUENCE, NUCLEIC ACID, REVERSE COMPLEMENT. Luego guardar con FILE, SAVE.

- 4) Seleccionar las cuatro secuencias (pintando los nombres en la columna izquierda) y seleccionar ACCESORY APPLICATION y luego CAP (Contig Assembly Program). Aparecerá una ventana como lo muestra la figura 5.



Elegir una superposición mínima de 10 bases y 85% - 90% de homología. Seleccionar RUN APPLICATION. Cerrar la ventana de DOS. Modificar los parámetros de tal manera de aumentar la rigurosidad y porcentaje de superposición mínima. ¿Pudieron armarse los CONTIGs?

5) Analizar la secuencia en las regiones de superposición. ¿Existen discrepancias en estas regiones o hay un 100% de coincidencia? ¿Qué criterio utilizaría para resolverlas?

6) Modificar el CONTIG (que aparecerá como una secuencia más por debajo de las otras en la ventana “solo texto” bajo el nombre de contig) en caso de encontrar errores. Para esto seleccionar la opción EDIT SEQUENCE.

#### Búsqueda de marcos abiertos de lectura (ORF)

El programa buscará posibles codones de iniciación (AUG) y codones STOP (UAG, UAA, UGA), de manera de determinar posibles marcos de lectura abiertos. Para ello seleccionar el CONTIG en la columna de la izquierda de la ventana “solo texto”, y elegir la opción SEQUENCE, NUCLEIC ACID, SORTED SIX-FRAME TRANSLATION. ¿Por qué encontró más de un marco de lectura posible?

#### Identificación de la secuencia incógnita: búsqueda de homología en INTERNET

Se procederá a identificar la secuencia con que trabajamos mediante un alineamiento local, utilizando el programa BLAST. Para ello seleccionar el CONTIG en la columna de la izquierda de la ventana “solo texto” y seleccionar COPY SEQUENCE.

Abrir ventana de Internet y copiar la dirección de NCBI: <http://www.ncbi.nlm.nih.gov/>

Elegir la opción BLAST. Realizar una búsqueda en blastn y otra en blastx, pegando en cada caso la secuencia del CONTIG obtenido.

¿Corresponderá esta secuencia a una región codificante? ¿pudo homologarse a algún gen de algún organismo? ¿a cuál? ¿existe homología con un único organismo? ¿Por qué?. Compare los valores de significancia o “scores” obtenidos con ambas búsquedas.

#### Bibliografía

Brown, T.A. 1999. *Genomes*. John Wiley & Son Inc., Publication. New York.

Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C., Gelbart, W.M. 2000. *An Introduction to genetic analysis*. W.H. Freeman Publ., New York.

Higgins, D.G. y Sharp, P.M. 1988. CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* 73: 237-244.

Higgins, D.G. y Sharp, P.M. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *CABIOS* 5: 151-153.

Hillis, D.M., Moritz, C., y Mable, B.K. 1996. *Molecular Systematics*. Sinauer Associates, Inc. Sunderland, USA.

Miesfeld, R. L. 1999. *Applied Molecular Genetics*. John Wiley & Son Inc., Publication. New York.

### Tercera parte.

Ana Distéfano, Paula Fernández, Norma Paniego

#### **Tutorial BLAST**

##### **Objetivo**

Familiarizarse con el uso de herramientas bioinformáticas y consulta de bases de datos disponibles a través de internet, empleadas frecuentemente en el estudio molecular de genomas virales.

##### **Desarrollo**

Se analizarán cuatro clones obtenidos a partir de una genoteca del MRCV, estudiando el nivel de identidad con otras secuencias depositadas en bases de datos públicas y ubicando la posición de los clones en el genoma del virus.

1. Identificar la carpeta Práctica bioinformática en el Escritorio de la computadora. Abrirla y localizar los archivos Clon A.txt, clon B.txt, clon C.txt, clon D.txt, allí se encuentran la secuencia de cuatro clones correspondiente a la genoteca del MRCV. Abrir dichos archivos con el block de notas o Word.

2. Buscar en la base de datos Genbank utilizando el programa BLAST la identidad de los cuatro clones. A la secuencia enviada, la denominaremos secuencia query. Para ello acceder a través de la página del NCBI, Seleccionar Nucleotideblast (blastn).

3. Copiar la secuencia de cada archivo y pegarla en la ventana del BLAST “enter query sequence”. En program selection elegir “somewhat similar sequence (Blastn)”. Enviar búsqueda (clic en “blast”).

4. Interpretar, analizar y discutir el resultado de las búsquedas con los docentes.

##### >ClonA

```
CCCGTTTCGGCAAAGTCACAGCGAAAGAAAAATTTCTTGATATTGAAAAGAATATGTTACAAGTATGGCAGAATCAAGCTCTTAAGTAAAGCGAATACGC
CTCCAAATTAATACTATTGAAAGAAAGTCTATGGATGGGATAATTGCCAATGCCGTTAGAGATTAGAAAAACATTATTTGGAGAAATACGGTTAACCACAGA
AATAATTGAATCATAATTAAGACATTGCTTTAGTTCTCAATAATGCTAGCGATTAAATTTAAAACTATTCGAAAACGACGGTGGTATTGAAAAATCTT
GCATCATTCCAATTCAGTAATTGAAAAAGACATTTTGGCGTACTACATCCGATTG
```

##### >ClonB

```
TTGAAAAAGACATTTTGGCGTACTACATCCGATTGAAATCTTTGATTTAATGCTTTTGGTATCGATGAAGTTAACAATTTAGCTATAAAATCCTAAAAATCC
GTTACCTGCCAAATTCATTGTTTCAGAAATTTAGAGATACCATAGATAACTTAAAAAGATGAAAGTTCATATGTTAAAGCTAAATCATTGTCAAAGAATAT
GGTATAGTTTGAATTGAATGAAAAGACAATAGTTAATTGGCGTAGATAGAATGTTAATAAGTTAAGTATTCAGCATAGGTTTGGAACTTTTGAAGTCA
AATGTTTGAAGTTTCTTATGAGGCAAATTTACAGAGGATCAGTCATATGTTAATTAACCAGAATGTGTCTCGAAGAAAAGTATTCAGAAGCGTGGTG
TTGATACTGTTTGTGGCAAATTTACGTGAAATTTACATCATAATGGTTTCCTTTCCATTTTATCATCATACATAGTTCAAATTTGGAGTCTAAGATA
GTACTCGTATTTCATCCACTGATAA
```

##### >ClonC

```
ATAATCTCAATAAATATTTGGATGGAAATCGTGTATTGATGAGTTAAATAGTATTGCTGTGCTACGTTCCCGCAAATTTGAAGTTTCATTTGGACTCATGT
ATGTCAATCATAATCGATCGAAAATTGATCAATCCTTGACAGTAAGAACTGAATCATTGGAAGAGCGATTATCTTTGGATTCAATATTCTTGAAACATGTA
```

## Genómica Aplicada 2018

### TP n°3- Metagenómica: Búsqueda de enzimas glicosilhidrolasas

#### Objetivo:

Recuperar enzimas glicosil hidrolasas novedosas de microorganismos no cultivables a partir de datos metagenómicos.

#### Introducción:

El agotamiento gradual de las reservas de combustibles fósiles, junto con la creciente preocupación por el deterioro del medio ambiente, han dado lugar al gran interés actual por la producción de combustibles a partir de fuentes renovables. El mayor limitante en la obtención de energía a partir de biomasa es la degradación de los compuestos lignocelulósicos para dar azúcares fácilmente fermentables. La necesidad de utilizar grandes cantidades de enzimas, debido a su baja actividad catalítica, hacen que la hidrólisis sea costosa y los combustibles obtenidos poco competitivos en el mercado. Los microorganismos, dada su enorme diversidad fisiológica y funcional, son una potencial fuente de nuevas enzimas capaces de catalizar los pasos limitantes de la producción de químicos a partir de biomasa no comestible. Para el descubrimiento de nuevas enzimas pueden aplicarse distintas estrategias. La más antigua consiste en el aislamiento de nuevas especies con la capacidad buscada codificada en su genoma. Este método, si bien ha permitido avanzar en el desarrollo de nuevos procesos, suele ser tedioso y desfavorable desde el punto de vista de económico. Posteriormente surge la idea de analizar los genes de una comunidad microbiana como un conjunto, su metagenoma, para encontrar aquellos genes que codifiquen las enzimas buscadas.

Si la comunidad es adecuadamente seleccionada, su ADN estará enriquecido genes codificantes de la función deseada. El ADN metagenómico puede fragmentarse, insertarse en vectores de expresión y transformarse en células hospedadoras apropiadas con el fin de construir una biblioteca de expresión y realizar una selección basada en función. Para que una enzima sea detectada exitosamente es necesario que el inserto contenga el gen o el operon completo y que este se exprese y modifique post-transduccionalmente en forma correcta en el hospedador.

Por otra parte, el ADN metagenómico puede secuenciarse directamente para encontrar los genes que codifican las nuevas enzimas ya sea predicción ab initio o por homología con enzimas ya secuenciadas, realizando una selección basada en secuencia. Este método elimina las limitaciones de la selección por función pero no permite encontrar genes completamente diferentes, carentes de homología con enzimas conocidas.

Los tres métodos tienen sus ventajas y limitaciones por lo que siguen siendo utilizados en la actualidad. En este trabajo práctico nos concentraremos en la selección basada en secuencia de glicosil hidrolasas (GHsas), a partir de un metagenoma obtenido de los microorganismos adheridos a fibras de pasto `fastq_to_fasta -i SRR094429_trimm.fastq -o SRR094429_trimm.fasta -v -Q 33` varilla incubado en el rumen de una vaca. Las GHsas son las enzimas capaces de hidrolizar el enlace glicosídico que mantiene unidas las unidades de glucosa tanto en la celulosa como en el almidón.

Los procedimientos experimentales realizados fueron, en resumen, los siguientes:

### **1- Secuenciación**

Se colocaron en el rumen de una vaca (mediante una fístula) bolsitas red de nylon rellenas de pasto molido y se incubaron 72 hs. Luego se retiraron las bolsas y se lavaron profusamente para obtener solamente los microorganismos adheridos firmemente al pasto a partir de los cuales se extrajo el ADN metagenómico. Con este se construyeron 4 bibliotecas, dos de “genoma completo” con fragmentos de 200 y 300 bps secuenciadas con Illumina HighSeq 2000 con 2 x 125bp. Y dos bibliotecas “jumping” con insertos de 3 y 5 kb secuenciadas en Illumina GAII por 2 x 75bp.

Los datos a utilizar corresponden al artículo “Metagenomic Discovery of Biomass-Degrading Genes and Genomes from Cow” Hess /Rubins 2011 y pueden descargarse del NCBI SRA023560 con acceso libre. Vamos a trabajar con una de las corridas de 2x125 cuyo número de acceso es: [SRR094429](#). Los datos se descargan en formato fastq, debemos submuestrearlos en R para poder trabajar con un volumen manejable. Estos pasos ya fueron realizados por que son muy demandantes de capacidad de procesamiento y memoria.

```
#!/fastq-dump SRR094429.sra
```

```
#R
```

```
#library(ShortRead)
```

```
#SS<-FastqSampler("SRR094429.fastq", n=1e6, readerBlockSize=1e8, verbose=TRUE,
```

```
#ordered=FALSE)
```

```
#srrss<-yield(SS)
```

```
#writeFastq(srrss,"SRR094429SS.107.fastq", compress=FALSE)
```

### **2- Controlar la calidad de las secuencias y eliminar adaptadores y secuencias de baja calidad**

```
cd '/home/eva/Desktop/Shared_Folder/GenómicaAplicada2018/TP_bioprospeccion'
```

```
perl /home/eva/Desktop/Programas/NGSQCToolkit_v2.3.3/QC/IlluQC.pl -se SRR094429SS.106.fastq N  
A
```

```
java -jar /home/eva/Desktop/Programas/Trimmomatic/trimmomatic.jar SE -threads 2
```

SRR094429SS.106.fastq SRR094429SS\_trimm.fastq LEADING:3 TRAILING:15  
SLIDINGWINDOW:4:15 MINLEN:30 ILLUMINACLIP:TruSeq3-SE.fa:2:30:10

Obtendremos un archivo llamado SRR094429\_trimm.fastq

### **3- Separaremos las secuencias correspondientes a ADNr16S**

Las lecturas correspondientes al gen del ARNr 16S deben eliminarse porque no aportan a la recuperación de genes enzimáticos y complican el ensamblado al estar presentes en más de un operón y contener regiones conservadas entre organismos. Para esto transformamos fastq a fasta con una rutina del fastx\_toolkit, Q 33 refiere a la codificación de la calidad de secuencia.

Luego utilizo ParallelMeta para seleccionar las lecturas de 16S, con la base de datos -b B de bacteria.

```
ExtractRNA_plus -i SRR094429_trimm.fasta -o SRR094429_trimm_RNA
```

Entonces elimino las secuencias de ARN

```
cd SRR094429SS_trimm_RNA
```

```
cat meta.rna| awk -F "|" '/^>/ { print $NF }'>rna.names
```

```
mv rna.names ~/Desktop/GenomicaAplicada2018/TP_bioprospeccion
```

```
cd ..
```

```
grep -v -f rna.names SRR094429SS_trimm.fastq> SRR094429SS_trimm_noRNA.fastq
```

```
mv SRR094429SS_trimm_noRNA.fastq ~/Desktop (velvet no puede generarlos archivos en la sf_)
```

```
cd ~/Desktop
```

### **4- Ensamblado de las secuencias del metagenoma**

El primer paso consiste en indicarle al ensamblador cuales son los archivos de secuencia que queremos procesar, que contienen y construir la tabla de k-meros. Esto se hace con velvet, en este caso elegimos utilizar distintos largos de k-mer para encontrar luego el óptimo, probaremos de 27 a 35 dando saltos de dos en dos. Podríamos contar con datos adicionales de bibliotecas con distinto tamaño de inserto y lecturas apareadas, eso también debería indicarlo en este paso.

```
velveth velvet_output 27,35,2 -short -fastq SRR094429SS_trimm_noRNA.fastq
```

Luego velvetg se encarga de generar los gráficos de de Bruijn que posteriormente serán manipulados.

Como se trata de un metagenoma seteamos la cobertura esperada en automático, ya que depende de la abundancia de cada genoma en particular. Usamos el k-mer de 29 porque es que utilizaron en el estudio original, pero podemos comparar lo que ocurre con los otros largos de palabra que elegimos en velvet. En general un k-mer más largo es más específico pero a su vez tiene menos cobertura, es una solución de compromiso que depende del caso.

```
cd velvet_output_29
```

```
cp meta-velvetg.Graph2-stats.txt /media/sf_Shared_Folder/GenómicaAplicada2018/TP_bioprospeccion
```

En R vamos a chequear las coberturas de los contigs ensamblados con metavelvet usando el [script\\_coverage.R](#).

Luego vamos a descomponer el metagenoma y tomar un genoma, en realidad un conjunto de contigs, de cobertura 9.

```
meta-velvetg velvet_output_29 -exp_covs 49
```

## **5-Predicción genes**

A partir de acá buscar archivos velvet\_output originales

En el siguiente paso tenemos que encontrar los marcos abiertos de lectura. Lo haremos usando MetaGeneMark

```
cd /media/sf_Shared_Folder/GenómicaAplicada2018/TP_bioprospeccion/archivos/velvet_output_29
```

```
gmhmmp -a -d -k -m ~/Desktop/Programas/MetaGeneMark_linux_64/mgm/MetaGeneMark_v1.mod
```

```
metavelvetg.contigs.fa -o metagenemark_output
```

```
less metagenemark_output
```

## **6- Anotación de genes**

Vamos a anotar los genes que predijo metagenemark por método ab initio en CAZy que es una base de datos de enzimas modificadoras de carbohidratos (recuerden que buscamos glicosilhidrolasas).

<http://mothra.ornl.gov/cgi-bin/cat/cat.cgi?tab=ORTHOLOGS>

Al finalizar nos enviarán un correo con los resultados.

Mientras tanto probaremos una estrategia alternativa para la anotación de genes con el programa ParallelMeta.

```
cp metavelvet.contigs.fa /Result/Contigs/contigs.fa
```

```
parallel-meta-func -S 1 -i meta-velvetg.contigs.fa
```

## **7-Informe**

Construir una tabla con las proteínas predichas, la proteína de mayor homología hallada en NCBI y su porcentaje de similitud.



# **AGROBIOTECNOLOGÍA**

**CURSO 2016**

**TRABAJOS PRÁCTICOS**

DEPARTAMENTO DE FISIOLOGÍA,  
BIOLOGÍA MOLECULAR y CELULAR  
FACULTAD DE CIENCIAS EXACTAS Y NATURALES  
UNIVERSIDAD DE BUENOS AIRES

# **MÓDULO IV**

## **MARCADORES MOLECULARES**

## TP6. MARCADORES MICROSATÉLITES Y AFLP

### INTRODUCCIÓN

#### **Microsatélites**

Los microsatélites son regiones del genoma que contienen arreglos de secuencias simples en *tandem* de mono, di, tri, tetra o pentanucleótidos que se repiten entre 10 y 100 veces. Los marcadores microsatélites, también denominados *Simple Sequence Length Polymorphism* (SSLP) o *Simple Sequence Repeats* (SSR), se encuentran distribuidos por todo el genoma de la mayoría de las especies eucarióticas y son detectados mediante la técnica de PCR. El polimorfismo (ocurrencia de diferentes alelos en la población para el mismo locus) está dado por diferencias en el número de repeticiones. Se supone que estas diferencias son causadas por errores en la replicación del DNA debido a que las repeticiones en *tandem* “confunden” a la DNA polimerasa celular durante la replicación, que agrega nucleótidos extra. En el caso de que se presenten cambios en un gran número de repeticiones, se supone que la variación proviene del cruzamiento desigual entre cromosomas homólogos. Además de los SSR nucleares, se han encontrado secuencias repetidas en cloroplastos (cpSSR), que se diferencian de las primeras en que suelen repetir una sola base.

Los microsatélites se detectan mediante amplificación por PCR usando *primers* específicos que hibridan en la región que flanquea al *tandem* de repeticiones (ver Figura 2). Los diferentes tipos alélicos se observan mediante electroforesis en geles de poliacrilamida en condiciones desnaturalizantes. También es posible analizar los patrones de microsatélites mediante el empleo secuenciadores automáticos, para lo cual uno de los *primers* debe estar marcado con un compuesto fluorescente.

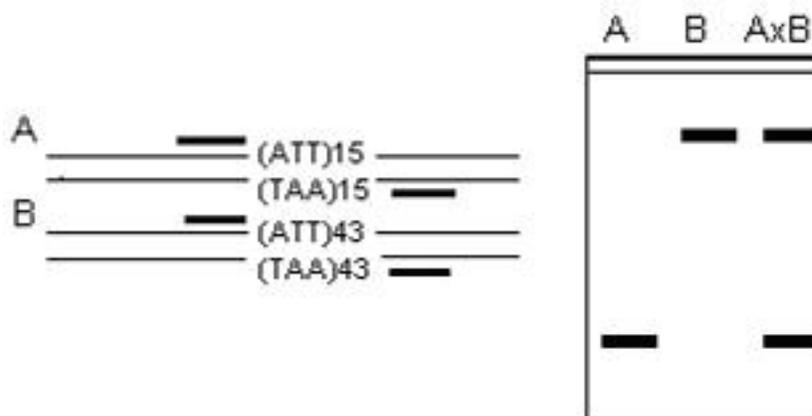


Figura 2: Esquema de microsatélites. El tamaño de los productos de amplificación depende del número de repeticiones de la secuencia base, y es ese el principio de los polimorfismos detectados.

Los microsatélites son herramientas muy útiles como marcadores genotípicos debido a que son relativamente abundantes, multialélicos (hipervariables con respecto al número

de repeticiones), heredados en forma estable, y codominantes (se distinguen los alelos de un locus), permitiendo distinguir materiales estrechamente relacionados. Desde el punto de vista metodológico, son fáciles de detectar mediante PCR, requiriendo poca cantidad de DNA para el análisis y poco equipamiento.

Quizás una de las pocas desventajas que tiene esta técnica es que se requiere del desarrollo previo de los *primers* que flanquean las secuencias repetidas, lo cual es costoso y lleva cierto tiempo. Otras desventajas son: la existencia de alelos nulos, el bandeo generado por los errores de la Taq polimerasa y la dificultad para determinar la tasa y el modelo de mutación de estas secuencias. Estos marcadores se emplean en: mapeo de genes, mejoramiento asistido por marcadores, identificación de cultivares, cuantificación de la variabilidad genética, estimación de la relación entre individuos (parentesco), estudios de genética de poblaciones, conservación de germoplasma, etc.

### **Polimorfismos de Longitud de Fragmentos Amplificados (AFLP: *Amplified Fragment Length Polymorphisms*)**

La técnica de AFLP es otra metodología que permite analizar la variabilidad a nivel de DNA, permitiendo obtener un gran número de marcadores moleculares distribuidos en el genoma. El ensayo de AFLP combina la especificidad, resolución y poder de muestreo de la digestión con enzimas de restricción con la velocidad y practicidad de detección de los polimorfismos vía PCR. De esta manera, esta técnica explora simultáneamente el polimorfismo debido a presencia o ausencia de un sitio de restricción, tal como ocurre con los RFLP, y la ocurrencia o no de amplificación a partir de secuencias arbitrarias, como en el caso de los RAPDs. Desde su desarrollo y divulgación, esta técnica ha sido utilizada en forma creciente con fines de fingerprinting, mapeo genético localizado y construcción de mapas genéticos, principalmente en especies de plantas cultivadas que presentan una baja tasa de polimorfismo a nivel de DNA.

El análisis mediante AFLP consiste esencialmente de cuatro etapas, representados esquemáticamente en las Figura 3 y 4

1) En la primera etapa el DNA genómico total a analizar es digerido con dos enzimas de restricción, una enzima de corte raro y otra de corte frecuente. Una enzima de corte raro reconoce de 6 a 8 pares de bases (por ejemplo, *EcoRI*, *Not I*). Por otro lado, una enzima de corte frecuente reconoce 4 pares de bases (por ejemplo, *Mse I*, *Rsa I*). De esta forma, se generarán tres clases de fragmentos que diferirán en cuanto a sus extremos según la enzima que haya cortado.

2) En la segunda etapa, se ligan adaptadores específicos a los extremos generados por la digestión del paso anterior. Estos adaptadores tienen secuencias complementarias a dichos extremos. Por ejemplo, adaptadores *EcoRI* específicos se ligan a los extremos resultantes luego del corte de la enzima *EcoRI*, en tanto que adaptadores *Mse I* específicos se ligan a los extremos producidos por el corte con *Mse I*. Estos adaptadores poseen de 20 a 30 pares de bases. Las secuencias de estas bases son diferentes para cada adaptador y no regeneran el sitio de corte de la enzima. Con este paso, un gran número de fragmentos podría ser amplificado mediante PCR dirigida con *primers* específicos diseñados a partir de la secuencia de los adaptadores.

El número de fragmentos generados sería excesivamente grande para poder ser observados individualmente aún en un gel de alta resolución. Por lo tanto, son necesarias algunas etapas de selección de una subpoblación de fragmentos, obteniendo un número tal que permita una adecuada resolución en un gel.

3) La tercera etapa consiste en la selección de los fragmentos a ser amplificados. Esta es realizada a través de una estrategia que se basa esencialmente en el concepto de los polimorfismos generados por la técnica de RAPD. Esto es, la utilización de *primers* que además de contener una secuencia específica de 20 a 25 nucleótidos complementaria a la secuencia de los adaptadores, poseen de 1 a 3 nucleótidos adicionales de secuencia arbitraria en su extremo 3'. La especificidad en esta posición es crucial ya que es justamente a partir del extremo 3' que se inicia la polimerización. De esta manera, los *primers* sólo van a hibridar con aquellos fragmentos que poseen secuencia complementaria a su secuencia arbitraria adicional. Por lo tanto, cuando se realiza la PCR y solamente una subpoblación de fragmentos es amplificada ocurre una acción selectiva. Esta acción se produce en dos etapas. En la primera, preselectiva (preamplificación), los *primers* utilizados contienen solamente un nucleótido arbitrario adicional (*primers*+1). En la segunda, de mayor intensidad de selección, los *primers* utilizados poseen otros dos nucleótidos arbitrarios, totalizando así tres nucleótidos adicionales (*primers*+3).

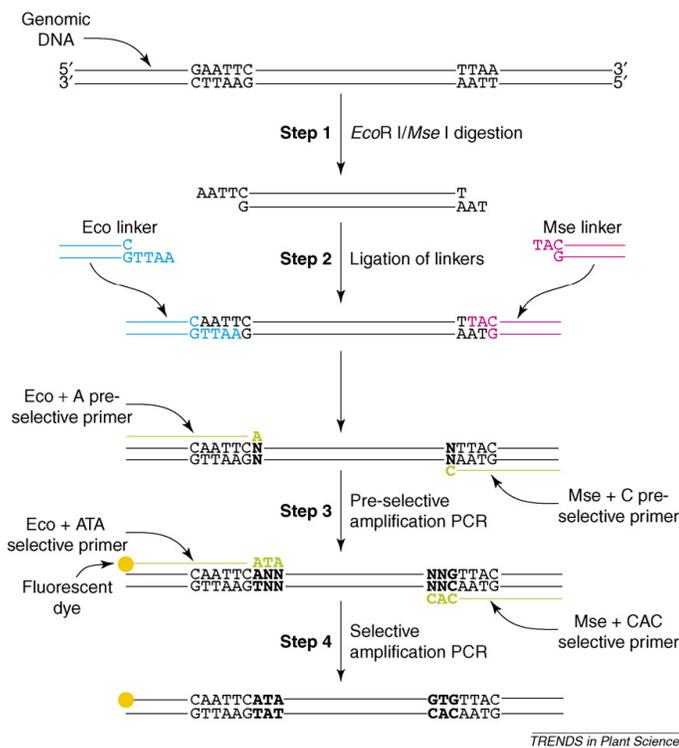
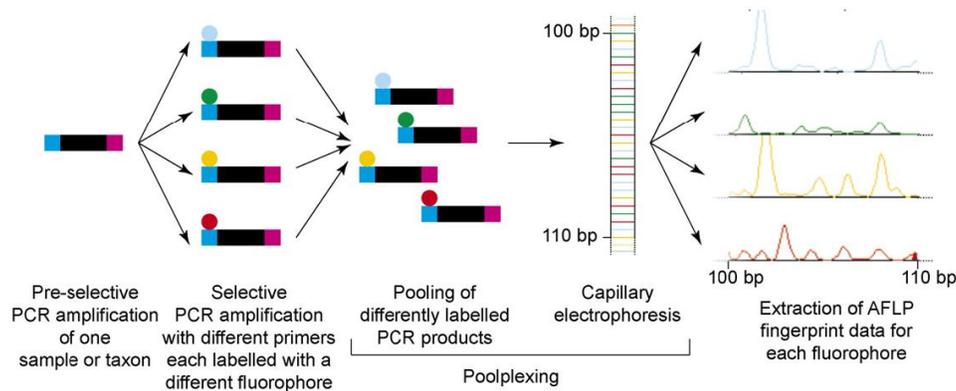


Figura 3: Representación esquemática del proceso de análisis de marcadores AFLPs

4) En esta cuarta y última etapa, la subpoblación de fragmentos amplificados es separada mediante electroforesis de alta resolución en geles de poliacrilamida y visualizada mediante tinción con plata (o mediante autoradiografía si se utiliza en la última reacción de PCR un *primer* marcado radioactivamente, o mediante secuenciador automático si se utiliza un *primer* fluorescente). La cantidad de bandas (50-100 o más) y

el perfil de las mismas depende de las enzimas usadas y de las bases al azar agregadas. Varios *loci* polimórficos pueden encontrarse en un mismo gel.

Una variante de la técnica de AFLP, desarrollada en la última década, es la marcación de los fragmentos amplificados con fluoróforos, permitiendo la visualización de los mismos mediante electroforesis capilar, utilizando secuenciadores. Esta modificación así como el desarrollo de software para su lectura permite realizar análisis en forma más procesiva. La figura 4 esquematiza el proceso de detección de fragmentos marcados con fluoróforos.



TRENDS in Plant Science

Figura 4: Representación de la metodología de AFLP fluorescente.

El polimorfismo entre fragmentos AFLP resulta de mutaciones de punto, inversiones, deleciones e inserciones que llevan a: a) pérdida o ganancia de un sitio de restricción reconocido por las enzimas utilizadas o b) alteración de la secuencia reconocida por los nucleótidos arbitrarios en los extremos 3' de los *primers* que dirigen la PCR a partir de los adaptadores.

El concepto de “dominancia” de los marcadores RAPD se aplica también a los marcadores basados en AFLP. Así, los marcadores AFLP no permiten la detección de individuos heterocigotas, o sea, no es posible distinguir si una banda en el gel es el resultado de la amplificación de uno o dos alelos. Esta técnica es sumamente usada por su versatilidad y alta reproducibilidad en estudios de *fingerprinting*, mapeo de genes simples y QTLs, estudios de diversidad genética, etc.

En el transcurso de la práctica, se realizarán reacciones de microsatélites para una población tipo RIL de girasol (*Helianthus annuus*) y se resolverán en geles de agarosa. Asimismo, se analizarán resultados de microsatélites para una población F2 de sorgo (*Sorghum bicolor*), obtenidos mediante el empleo de un secuenciador automático. Por otra parte, se observarán y discutirán muestras evaluadas mediante las técnicas de RAPD, microsatélites y AFLP provistas por el docente, resueltas en geles de poliacrilamida teñidos con nitrato de plata. A partir de los resultados obtenidos se discutirán las aplicaciones de estas metodologías y las ventajas y desventajas de su utilización como marcadores genéticos.

## **BIBLIOGRAFIA**

### ***Microsatélites***

- Akkaya, M.S., Bhagwat, A.A. and Cregan, P.B. (1992). Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics*, 132:1131-1139
- Powell, W. *et al.* (1996). Polymorphism revealed by simple sequence. *Science*, 1(7):234-238.
- Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, 17:6463-6471.
- Weber, J. and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, 44:388-396.

### ***AFLP***

- Muller, U.G. and Wolfenbarger, L.L. (1999). AFLP genotyping and fingerprinting. *Tree*, 14:389-394.
- Ridout, C and Donini, P. (1999). Use of AFLP in cereals research. *Trends in Plant Science*, 4:76-79.
- Savelkoul, P.H. *et al.* (1999). Amplified Fragment Length Polymorphism Analysis: the State of an Art. *Journal of Clinical Microbiology*, 3:3083-3091.
- Vos, P. *et al.* (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23:4407-4414.
- Heidi M. Meudt and Andrew C. Clarke (2007). Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *TRENDS in Plant Science Vol.12 No.3*

## Protocolo experimental para el análisis mediante SSR

### a) Material

Cada grupo recibirá ADN de girasol proveniente de una población de mapeo (líneas parentales y líneas recombinantes endocriadas), a partir del cual se harán reacciones de amplificación de microsatélites. Estas muestras serán sembradas en un gel de agarosa y visualizadas mediante tinción con bromuro de etidio.

### b) Procedimiento

#### **b.1.) Reacción de amplificación**

Cada grupo recibirá ADN de plantas individuales y llevará a cabo el siguiente procedimiento:

- 1) Rotular apropiadamente los tubos de PCR con marcador de alcohol, (utilizar uno para cada ADN y uno para el control negativo).
- 2) Pipetear el ADN (2  $\mu$ l) en los tubos correspondientes y 2  $\mu$ l de H<sub>2</sub>O en el tubo "control negativo".
- 3) Realizar una mezcla de reacción (*Premix*) según la siguiente tabla:

Reactivos	Stock	[cc] final	Vol ( $\mu$ l)
Primers microsat	[5mM]	0,25 mM	0,60 (0,3 de c/u)
MgCl <sub>2</sub>	[50mM]	1,5 mM	0,35
dNTPs	[10mM]	0,2 mM	0,25
Buffer 10x	(10x)	1x	1,20
ddH <sub>2</sub> O			7,20
Taq		1 U	0,50
		Vf:	10,00

Mantener el orden de la lista de reactivos para preparar la *Premix*. Una vez preparada, mezclar brevemente con *Vortex*, aplicar un *spin* de centrifuga y enfriar en hielo.

**PRECAUCIÓN:** Una vez que la *Taq* polimerasa fue agregada a la *Premix*, trabajar siempre en hielo, ya que la enzima pierde irreversiblemente actividad a temperatura ambiente.

- 4) Agregar 10  $\mu$ l de la *Premix* a cada tubo que ya tiene el ADN. Tapar bien. Dejar los tubos en hielo hasta colocarlos en el termociclador.

El programa de amplificación es el siguiente:

1 ciclo	94 °C, 4 minutos (desnaturalización inicial)
10 ciclos	94 °C, 45 segundos (desnaturalización)
	64 °C, 45 segundos, descendiendo 1 °C por ciclo (hibridación con <i>touchdown</i> )
	72 °C, 1 min (extensión)
35 ciclos	94 °C, 45 segundos (desnaturalización)
	54 °C, 45 segundos (hibridación)
	72 °C, 1 minuto (extensión)
1 ciclo	72 °C, 10 minutos (extensión final)
	15 °C, infinito (conservación de las muestras)

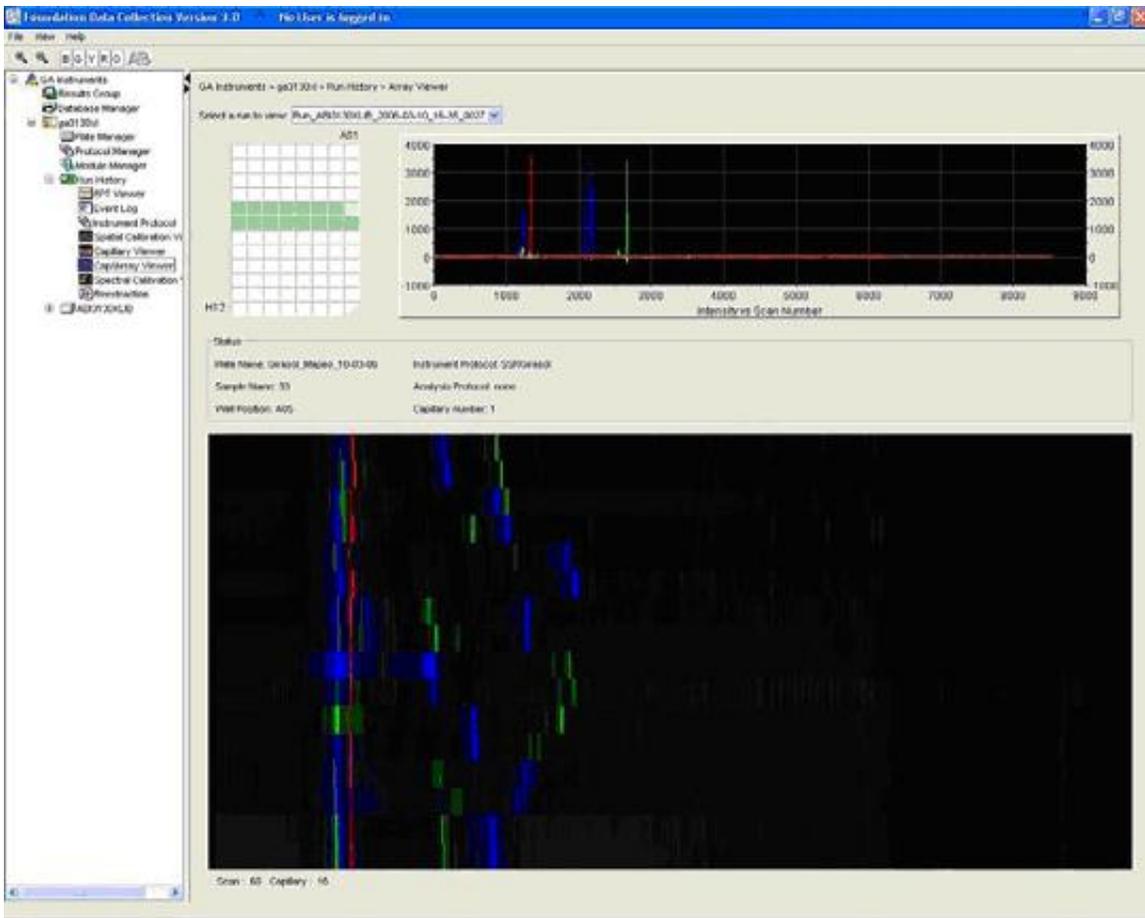
5) Una vez terminado el programa, guardar en freezer a -20°C hasta el momento de la visualización de los resultados.

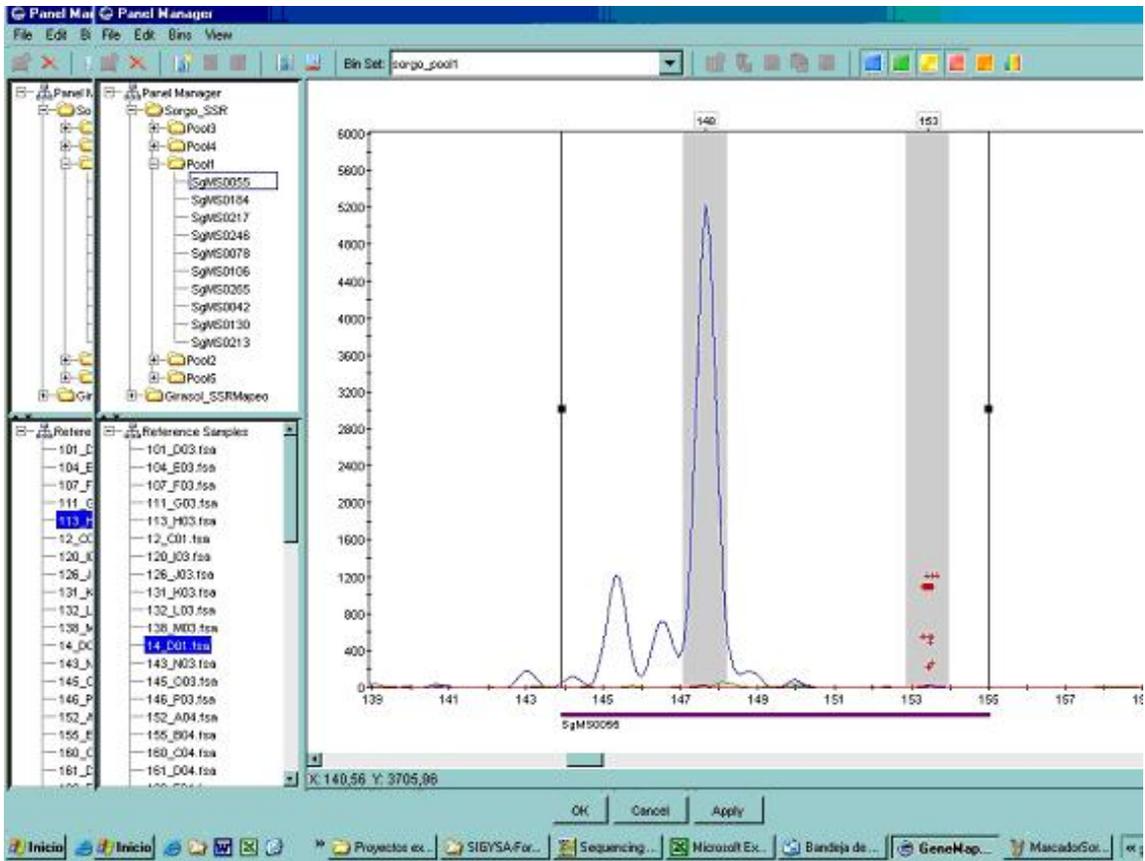
### **b.2.) Visualización de los productos de amplificación**

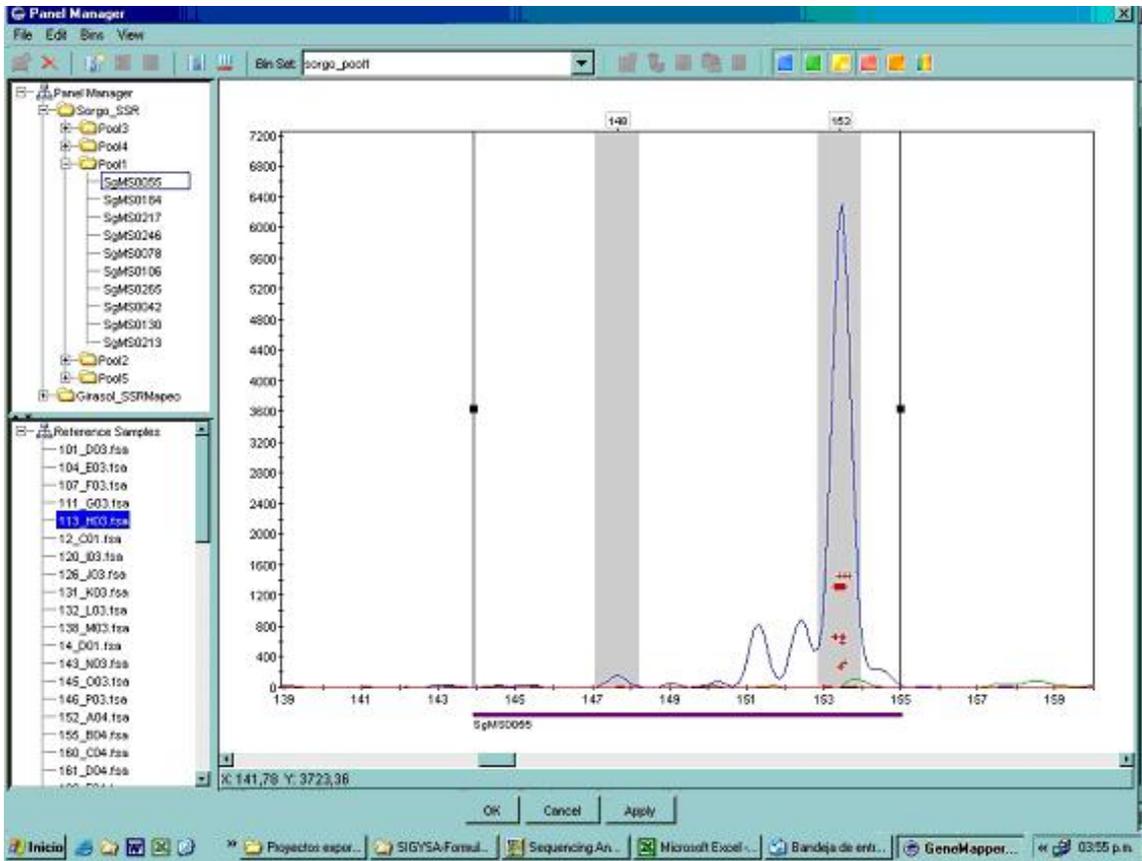
Se realizará en las mismas condiciones que para RAPD (ver b.2 en guía anterior de marcadores RAPD).

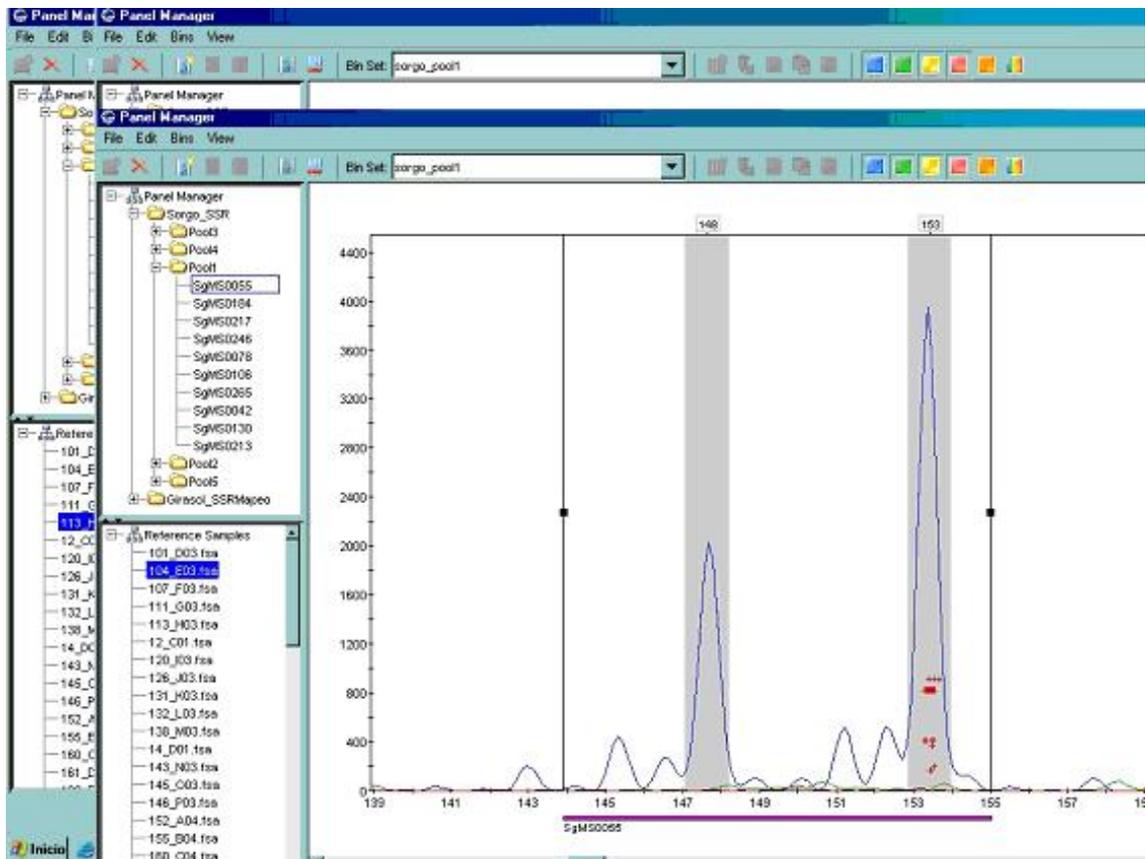
### **c) Material complementario para discutir en clase**

Resultados obtenidos mediante el empleo de un secuenciador automático para una población F2 de sorgo.









#### **d) Preguntas para responder al finalizar el TP**

- 1) Realice un cuadro comparativo para las tres metodologías discutidas en el TP: RAPDs, AFLPs y SSR, analizando las siguientes características: metodología que emplean, laboriosidad, número de loci analizados, nivel de polimorfismo, capacidad de discriminar alelos, necesidad de conocimiento previo del genoma, reproducibilidad, costo, posibilidad de automatización (y alguna otra característica que a Ud. le parezca relevante)
- 2) ¿Qué resultados obtuvo mediante el análisis de RAPDs para las variedades y especies estudiadas? ¿Qué le permiten concluir?
- 3) ¿Qué resultados obtuvo para la población analizada para un locus SSR? ¿Qué debería hacer para obtener un mapa de ligamiento de la misma?
- 4) Con respecto a los resultados obtenidos para sorgo con SSR de manera automatizada: ¿cuáles son las ventajas de esta metodología?
- 5) Enuncie al menos 3 aplicaciones de los marcadores moleculares para resolver problemas agronómicos concretos.

## TP N°4

### Ensayo para determinación de la producción potencial de biometano (bmp)

#### Objetivo:

Comparar la capacidad de producción de metano de tres sustratos alternativos.

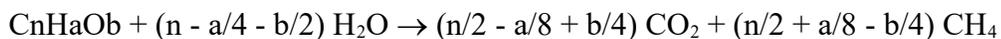
#### Introducción:

El proceso de digestión anaeróbica permite transformar materia orgánica en biogás una fuente de energía renovable. Esta transformación es mediada por varios grupos de microorganismos, que deben funcionar de forma balanceada para que tanto la degradación como la producción de biogás sean adecuadas. El tratamiento anaeróbico de residuos cumple simultáneamente el objetivo de minimizar el impacto ambiental de la disposición de los desechos y producir energía a partir de fuentes alternativas. . El proceso se realiza en forma anaeróbica y su primer paso involucra la hidrólisis de los sustratos poliméricos. A continuación se produce la fermentación de monómeros y oligómeros que implica la acidogénesis para luego producirse la acetogénesis y finalmente la metanogénesis la cual puede darse tanto por la vía acetoclástica como por la hidrogenotrófica.

Los primeros pasos son llevados a cabo por bacterias, mientras que la metanogénesis es realizada exclusivamente por arqueas. Esta red de microorganismos es interdependiente, si no se llevan a cabo hidrólisis y fermentación, no se genera biogás y si el último paso está bloqueado, la fermentación se detiene por acumulación de compuestos inhibitorios.

El biogás está compuesto fundamentalmente por metano y dióxido de carbono, la proporción entre ambos gases puede calcularse mediante la siguiente ecuación:

Rendimiento teórico de metano (McCarty 1964)



El contenido de CO<sub>2</sub>, sin embargo, depende de las condiciones del proceso. Si la fermentación predomina sobre la metanogénesis, entonces veremos que aumenta y esto será interpretado como una señal de alerta. En este trabajo práctico vamos a determinar la proporción de metano en el biogas generado a partir de distintas fuentes de desechos, comparando grasas, proteínas e hidratos de carbono.

Table 2-3 Theoretical Methane Content of Biogas

Substrate	Chemical Composition	Methane, % of Total Gas
Fat	$C_{15}H_{31}COOH$	72
Protein	$C_4H_6ON$	63
Carbohydrate	$C_6H_{12}O_6$	50

Desarrollo del trabajo práctico:

- 1- Colocar 20 mL volumen de barro anaeróbico en el fondo del la botella (digestor). Trabajar en duplicado.
- 2- Completar con 50 mL de buffer A .
- 3- Llenar la probeta con KOH y colocarla invertida en el cristallizador que contiene KOH para deteminación de metano.
- 4- Ensamblar el aparato como se muestra en la figura.
- 5- Todo el aparato deberá montarse en forma equivalente pero con H<sub>2</sub>O en el cristallizador y probeta para la determinación del total de biogás.
- 6- Agregar una cantidad 2g de gelatina o 2 g de almidón o 5 mL de crema (40%) según corresponda.
- 7- Registrar la evolución de gases durante en tiempo que dure el experimento.
- 8- Graficar producción de biogás en función del tiempo.
- 9- Graficar relación Metano/Biogás en función del tiempo.
- 10- Determinar el contenido promedio de CO<sub>2</sub>, suponiendo que el biogás se compone fundamentalmente de CH<sub>4</sub> y CO<sub>2</sub>)
- 11- Comparar con los resultados de los otros residuos y las proporciones teóricas.

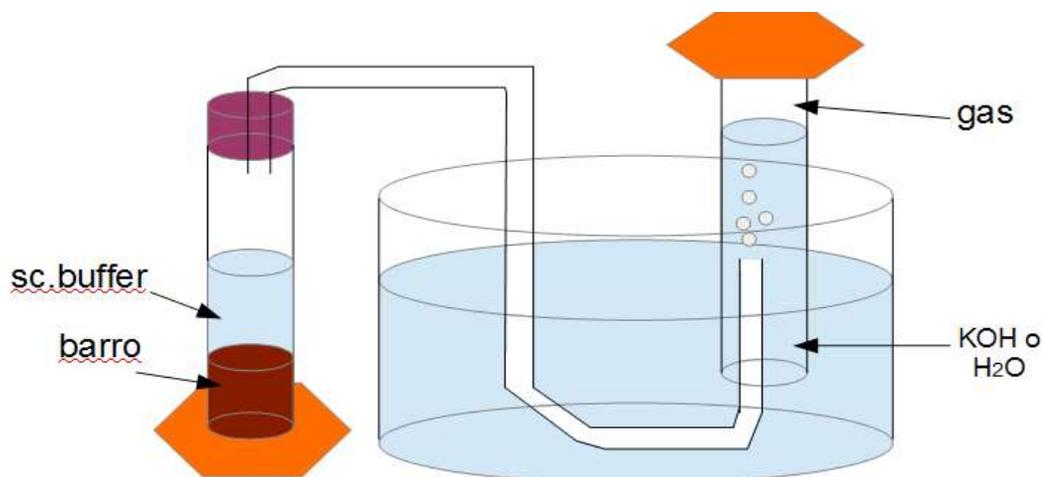


Fig. 1 -

*Armado del dispositivo para medir biogás*

Buffer A (0.1% NH<sub>4</sub>Cl, 0.04% K<sub>2</sub>HPO<sub>4</sub>, 0.01% MgCl<sub>2</sub>, pH 7.0)

### Bibliografia

Owen, W. F., Stuckey, D. C., Healy, J. B., Young, L. Y., & McCarty, P. L. (1979). Bioassay for monitoring biochemical methane potential and anaerobic toxicity. *Water research*, 13(6), 485-492.

Krich, K., Augenstein, D., Batmale, J. P., Benemann, J., Rutledge, B., & Salour, D. (2005). Biomethane from dairy waste: a sourcebook for the production and use of renewable natural gas in California. USDA Rural Development.