



UNIVERSIDAD DE SONORA

POSGRADO EN BIOCIENCIAS

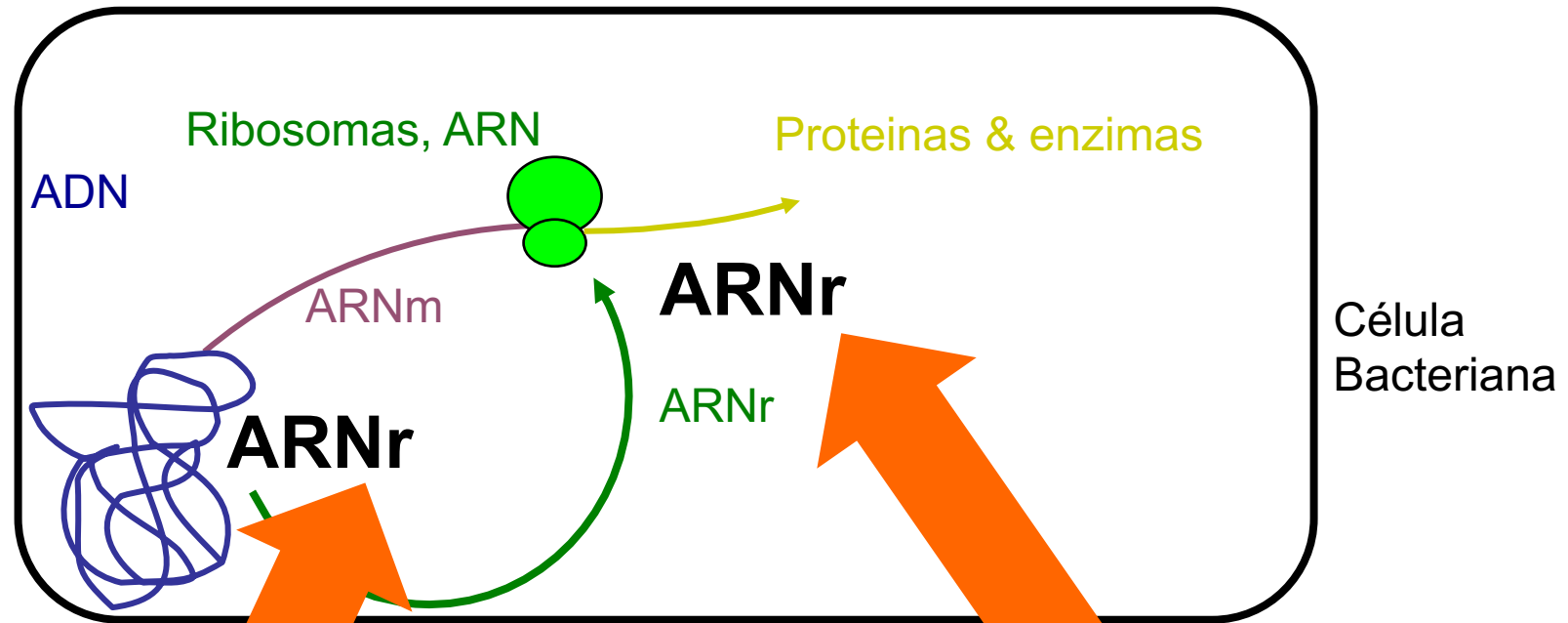
MICROBIOLOGÍA MOLECULAR (2442)

UNIDAD 9. MICROARREGLOS o Microarray

Elaboró: Dra. Kadiya del C. Calderón Alvarado

Todos los organismos necesitan sintetizar proteínas

Están activas, realizando alguna función?



Célula
Bacteriana

Metabólicamente
activas

ARN

Sólo las células activas
producirán proteínas → ARN

ADN

Células muertas o
vivas

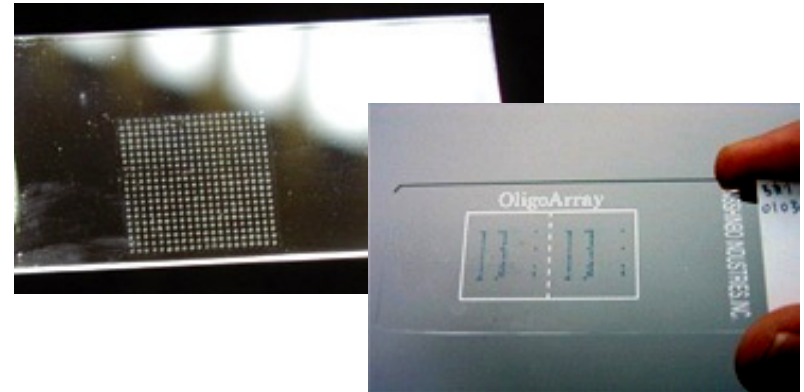
Abundancia

Chips y microchips (microarrays) de ADN

CHIP de ADN: conjunto **ordenado** (*array*) de hebras sencillas de ADN, unidas a una superficie, de manera que un grupo de genes de interés, o incluso el genoma completo de un organismo, o el metagenoma de un entorno, está representado:

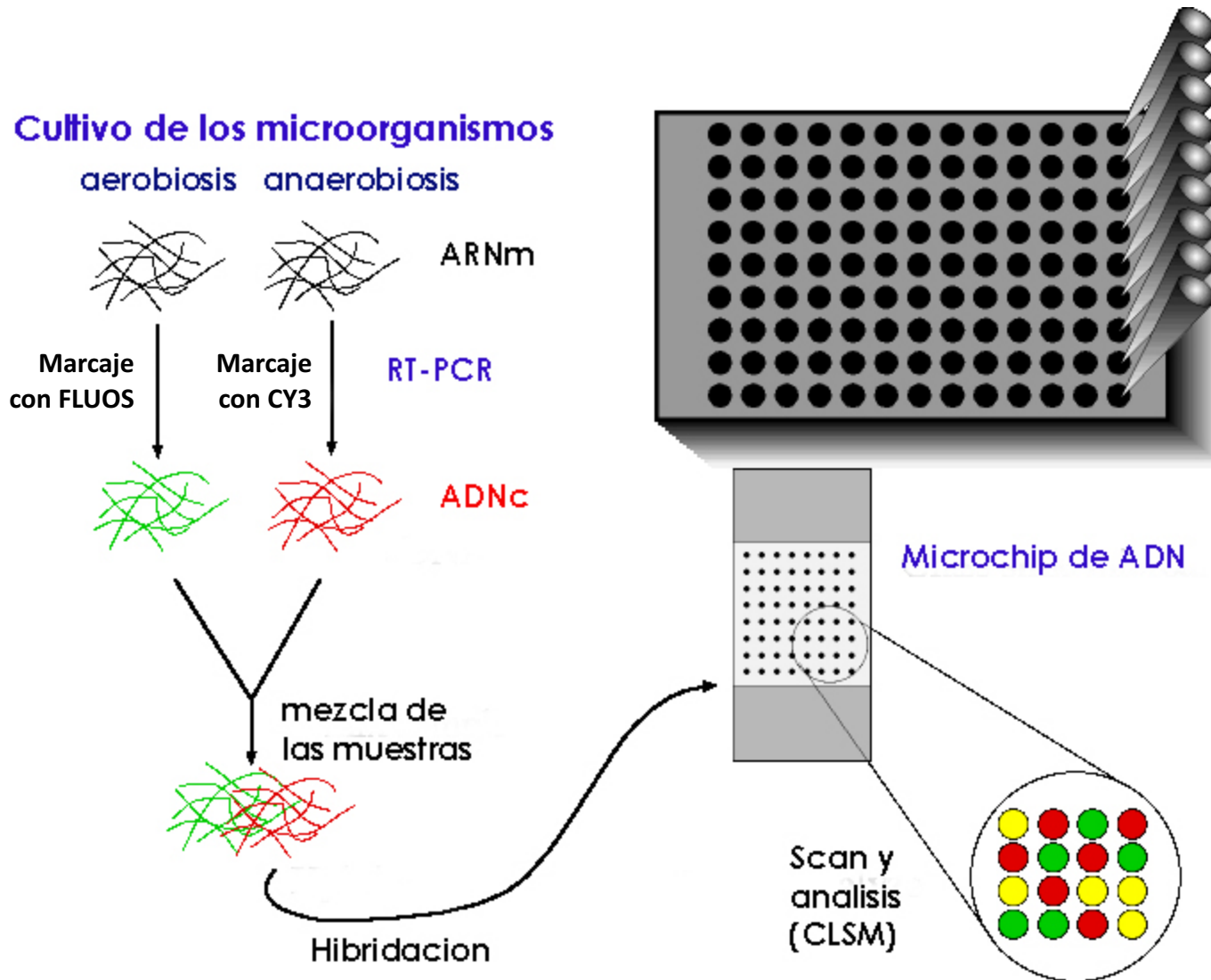


Miles de pequeños puntos conteniendo secuencias de ADN conocidas



- Técnica → emparejar cada hebra de ADN con su complementaria, si ésta se halla presente en una determinada muestra.
- Cuando dos hebras complementarias se encuentran, se unen entre ellas (**hibridación**).
- Si una de las hebras está **marcada** mediante un **colorante fluorescente**, es muy sencillo detectar aquellas hebras de ADN que han hibridado entre sí.
- Como cada hebra del chip es conocida, se puede **identificar** qué moléculas de ADN estaban presentes en la muestra.

Ejemplo básico: monitorización de la expresión génica de un microorganismo (levadura) bajo distintas condiciones de cultivo (aerobiosis y anaerobiosis).



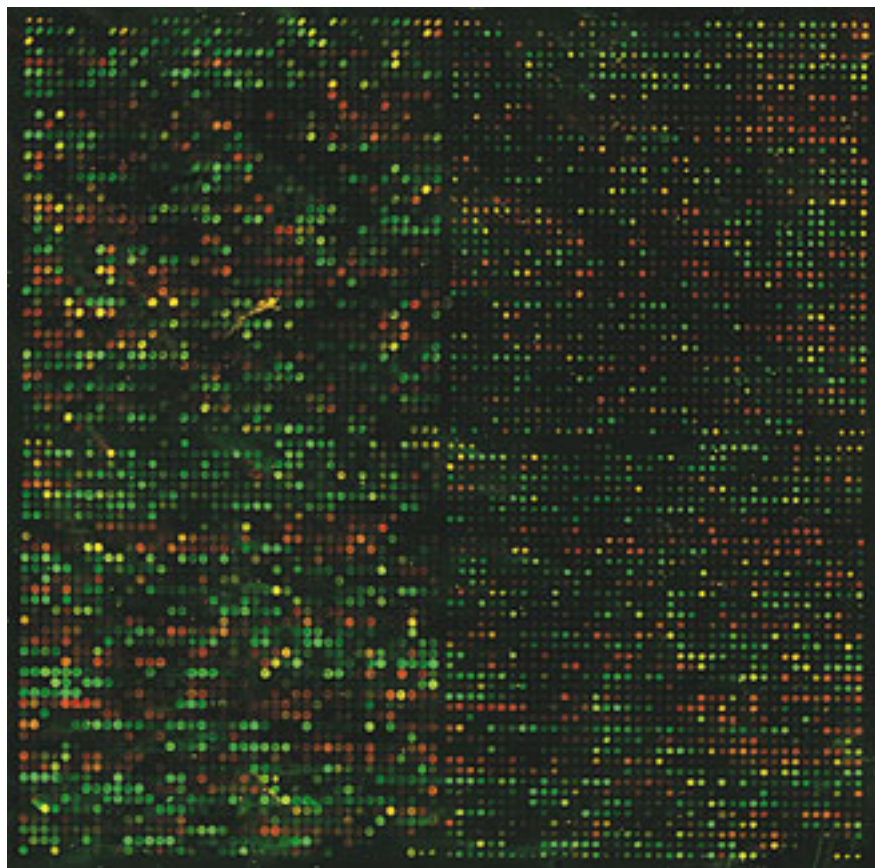
Clasificación de microarrays según:	Características	Nombre
Material inmovilizado	• Oligonucleótidos. • cDNAs. • Proteínas. • Tejidos.	• Gene Chips. • cDNA Array. • Protein Chip. • Tissue Chip.
Diseño del array o surtido	Personalizado	Arrayers
	Industrial	Cassettes
Fabricación	Densidad de integración	Alta: microarrays
		Baja: macroarrays
Impresión	Generación de la sonda	"in situ"
		"depositadas"
Soporte / Tipo de unión	No poroso / Covalente	Soporte rígido (Cristal y Plástico)
	Poroso / No covalente	Membranas
Aplicación	• Secuenciación por hibridación. • Detección de cambios en la expresión génica. • Cuantificación de la expresión génica.	• Chips de secuenciación. • Chips de hibridación comparativa. • Chips de expresión.

Tabla 1. Criterios de clasificación en Biochips (adaptado de <http://infobiochip.isciii.es>).

Modern Drug Discovery - September/October 1999

Modern Drug Discovery, **1999**, 2(5) 30-31, 33-34, 36. (Figure 1)

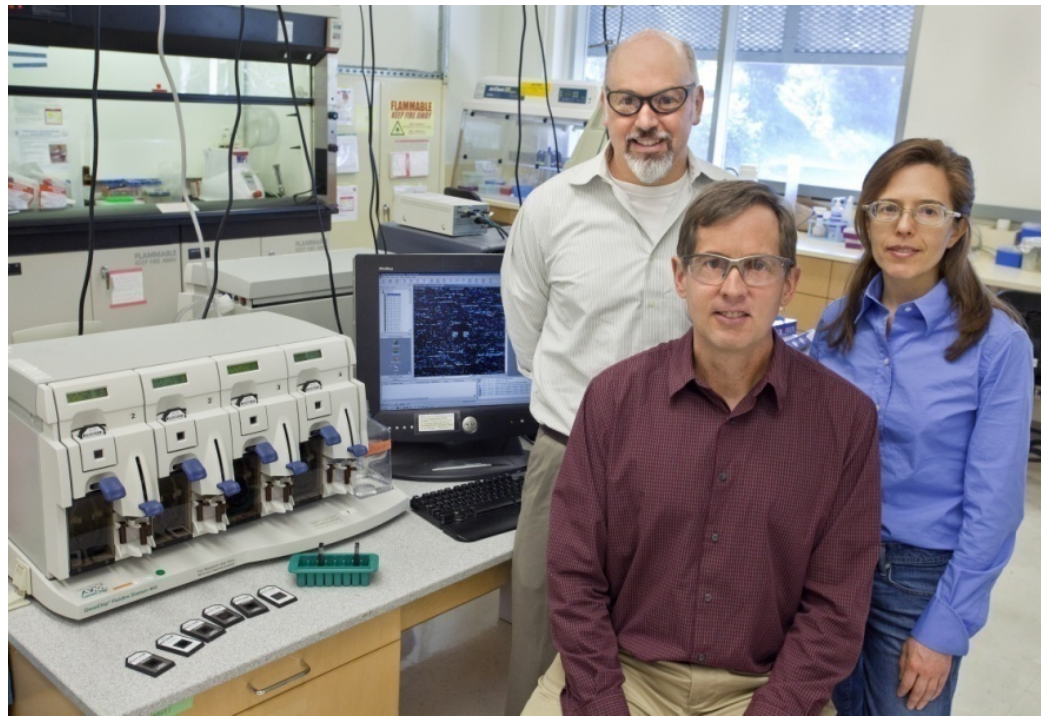
© 1999 American Chemical Society.



Microchip de ADN. Este chip contiene más de 6000 genes—virtualmente todos los genes de la levadura de panadería *Saccharomyces cerevisiae*. La actividad genética en las células es medida mediante la extracción de su ARNm y la posterior conversión a ADNc combinada con el marcaje con un colorante fluorescente. La mezcla de los ADNc es añadida al microchip, y cada ADNc hibrida con la región que contiene la secuencias complementaria. En el ejemplo, las regiones rojas corresponden a genes expresados cuando las células crecen en presencia de glucosa, las verdes representan los genes cuya expresión se reprime en presencia de glucosa, y las amarillas (verde+rojo) representan los genes que se expresan por igual en ambas condiciones.

Tipos de chips útiles en estudios de ecología y biorremediación:

- **Arrays de oligonucleótidos filogenéticos:** contienen secuencias signatura (=identificativas) del ARNr de **grupos específicos** de organismos.
- **Arrays de genoma de comunidades:** contienen secuencias signatura muy específicas de **genes de especies cultivadas**.
- **Arrays de genes funcionales:** contienen dominios conservados de **genes involucrados en rutas metabólicas específicas** de los ciclos de los elementos (C, N, S y metales)



Ejemplos de chips y microchips de ADN

Bacillus-PhyloChip. Identifica cepas de 5 especies de *Bacillus* (*Bacillus anthracis*, *Bacillus cereus*, *Bacillus mycoides*, *Bacillus medusa* and *Bacillus subtilis*).

Burkholderia-Phylochip. Detecta y diferencia miembros del genero *Burkholderia*.

Compost Community-Microarray. Caracteriza las comunidades bacterianas en el compost

ECC-PhyloChip. Detecta y diferencia 19 especies de *Enterococcus*

Freshwater Sediment-Microarray. Caracteriza las comunidades bacterianas de los sedimentos de agua dulce.

Human Intestinal Bacteria Chip. Detecta 20 especies bacterianas del intestino humano (*Bacteroides thetaiotaomicron*, *B. vulgatus*, *B. fragilis*, *B. distasonis*, *Clostridium clostridiiforme*, *C. leptum*, *Fusobacterium prausnitzii*, *Peptostreptococcus productus*, *Ruminococcus obeum*, *R. bromii*, *R. callidus*, *R. albus*, *Bifidobacterium longum*, *B. adolescentis*, *B. infantis*, *Eubacterium biforme*, *E. aerofaciens*, *Lactobacillus acidophilus*, *Escherichia coli*, *Enterococcus faecium*)

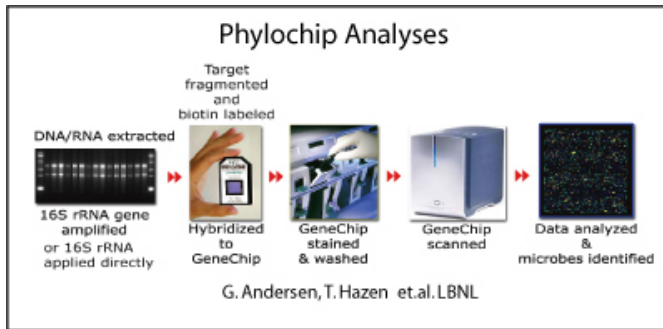
Nitrifier-Microarray. Detecta bacterias nitrificantes

RHC-PhyloChip. Identifica todos los miembros conocidos (cultivables o no cultivados) del Orden *Rhodocyclales* (RHCs).

SRP-PhyloChip. Identifica todos los linajes conocidos de procariotas reductores de sulfato.

Ejemplo de la aplicación de los microchips de ADN en estudios de tratamiento de aguas residuales

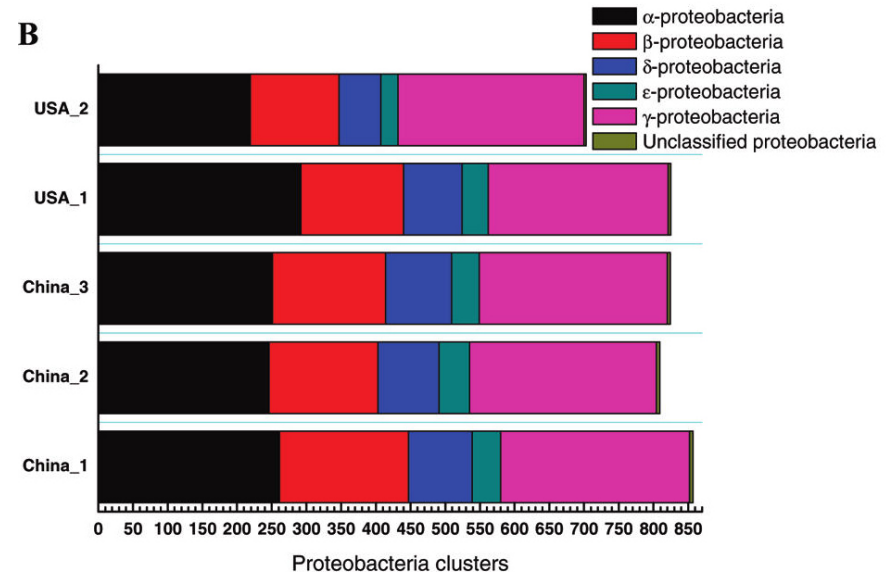
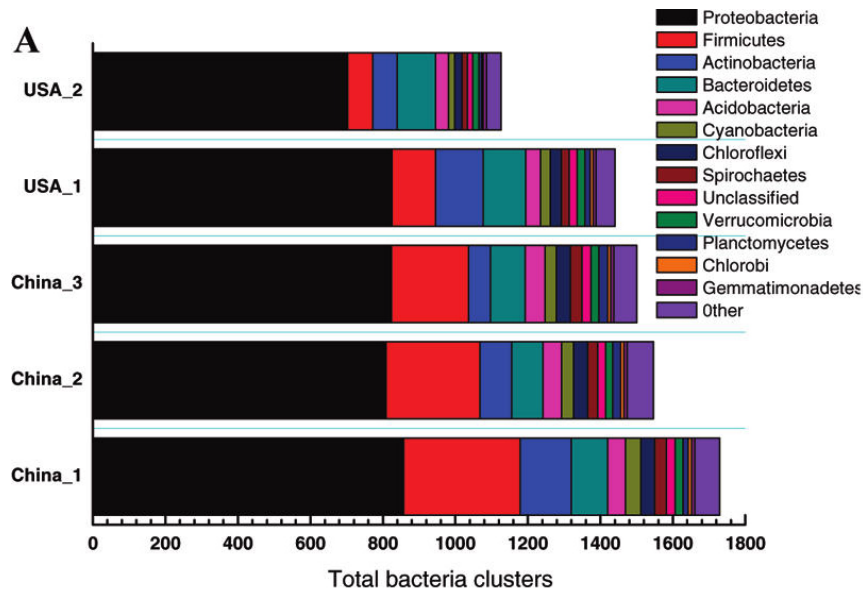
Composición de comunidades bacterianas en tratamientos de agua residual

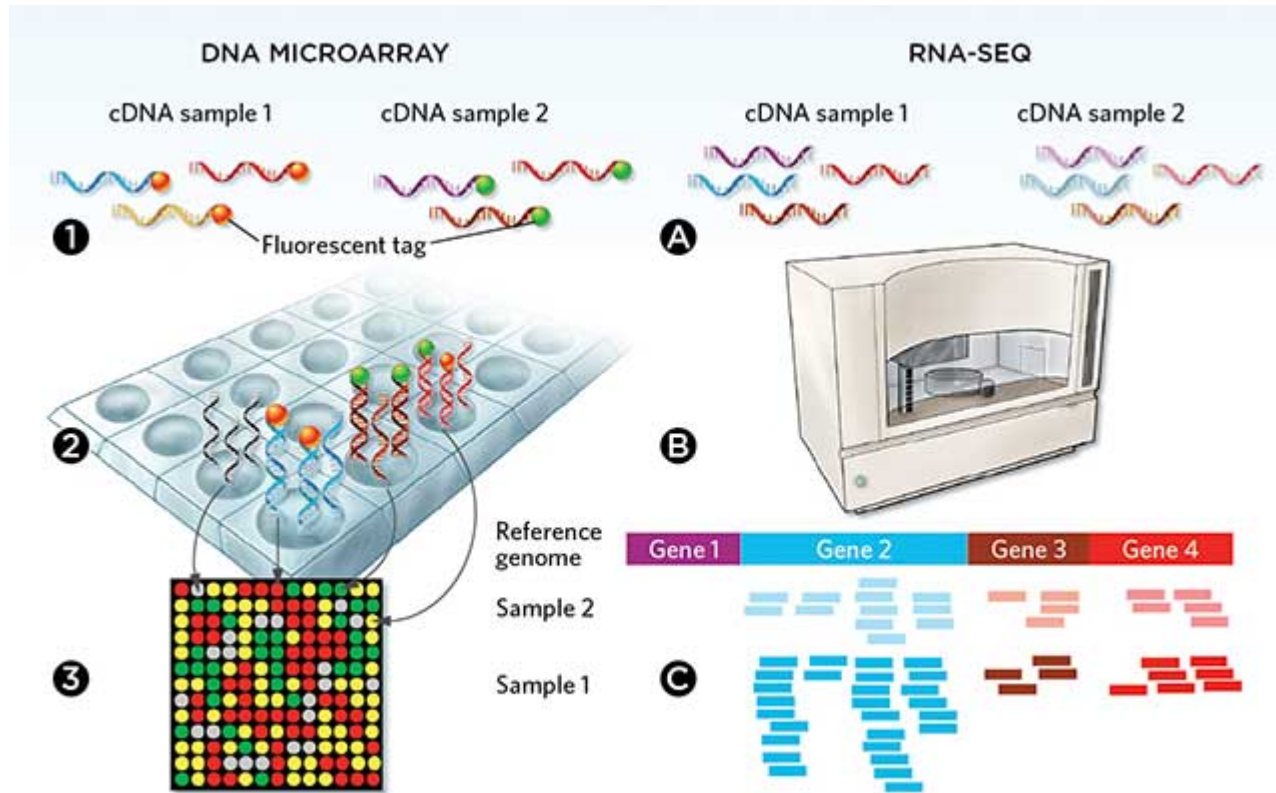


Microarray de alta densidad

Cada PhyloChip incluye 506944 sondas, 297851 tienen como diana la secuencia del ARNr 16S, divididos en 8935 grupos de secuencias con divergencia <3%. El resto de sondas son controles y estándares.

El conjunto de sondas representa los 121 ordenes del dominio Bacteria





www.the-scientist.com/lab-tools/an-array-of-options-35381

Methods for RNA extraction

Applied and Environmental
Microbiology

Rapid Method for Coextraction of DNA and RNA from Natural Environments for Analysis of Ribosomal DNA- and rRNA-Based Microbial Community Composition

Robert I. Griffiths, Andrew S. Whiteley, Anthony G. O'Donnell and Mark J. Bailey
Appl. Environ. Microbiol. 2000, 66(12):5488. DOI:
10.1128/AEM.66.12.5488-5491.2000.

ARN + ADN

Digestión de ADN

ARN

http://tools.lifetechnologies.com/content/sfs/manuals/cms_055740.pdf

TURBO DNA-free™ Kit

TURBO™ DNase Treatment and Removal Reagents

Catalog Number AM1907

Publication Number 1907M Revision G

Product description

Ambion® TURBO DNA-free™ DNase Treatment and Removal Reagents are designed to remove contaminating DNA from RNA preparations, and to subsequently remove the DNase and divalent cations from the sample. The included TURBO™ DNase (patent pending) is an engineered version of wild type DNase I with 350% greater catalytic efficiency. TURBO™ DNase has a markedly higher affinity for DNA than conventional DNase I, and is thus more effective in removing trace quantities of DNA contamination. In addition, TURBO™ DNase maintains up to 50X greater activity than DNase I in solutions containing physiological salt concentrations. The TURBO™ DNase provided in the kit is overexpressed in an animal-free system, and is then extensively purified in a bovine-free process and tested. It is guaranteed to lack any contaminating RNase activity. The kit also includes an optimized DNase reaction buffer that contains a small molecule enhancer that extends the activity of the TURBO™ DNase enzyme by 100-fold or more. Using TURBO DNA-free™, contaminating DNA is digested to levels below the limit of detection by routine PCR (Figure 1). The DNase is then removed rapidly and easily using a novel method which does not require phenol/chloroform extraction, alcohol precipitation, heating, or the addition of EDTA (see Table 1 on page 2). TURBO DNA-free™ treated RNA is suitable for endpoint or real-time RT-PCR, microarray analysis, RPAs, Northern, and all other RNA analysis methods.

In addition to removing the DNase enzyme, DNase Inactivation Reagent also removes divalent cations, such as magnesium and calcium, which can catalyze RNA degradation when RNA is heated with the sample (Figure 2 on page 3).

ARN

Retrotranscripcion

ADNc

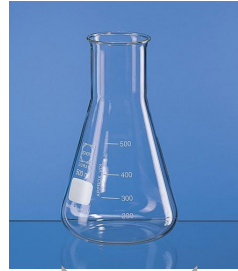
<http://www.lifetechnologies.com/be/en/home/life-science/pcr/reverse-transcription/reverse-transcriptase-enzymes/superscript-iii-reverse-transcriptase.html>

Primers aleatorios
Primers específicos
polyT (Eucariontes)

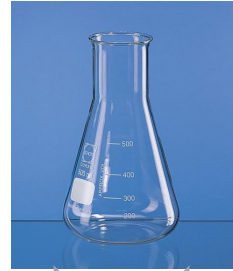
Microarrays para determinar expresión

Usar ADNc como molde en vez de ADN

Condición A



Condición B



Extracción ARN

Digestión ADN

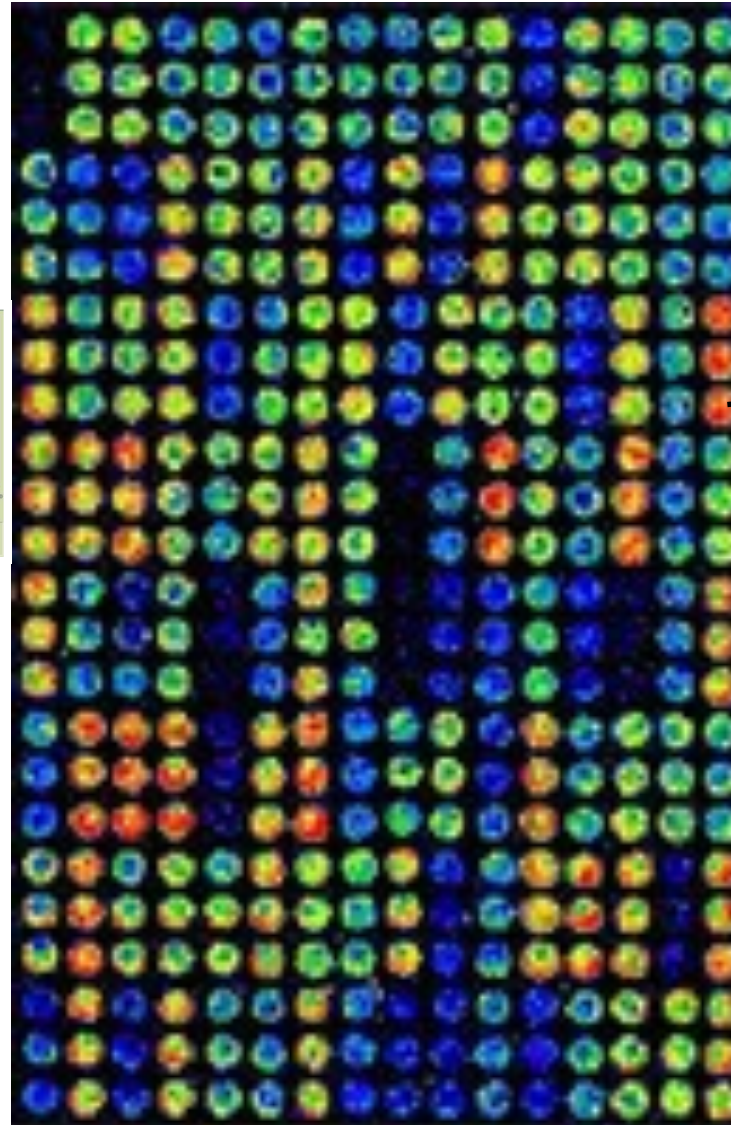
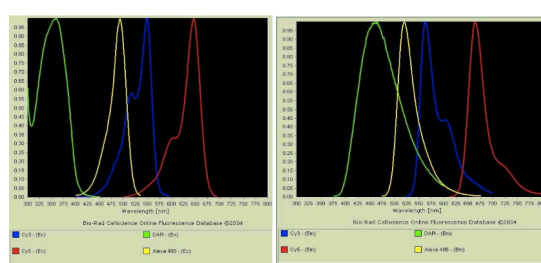
Síntesis de ADNc

Marcaje de ADNc

Hibridación de ADNc

Microarrays aplicados para la expresión de genes

Fluoróforos



Condición A
verde Cy3 575 nm

Condición B
Rojo Cy5 675 nm

SONDA

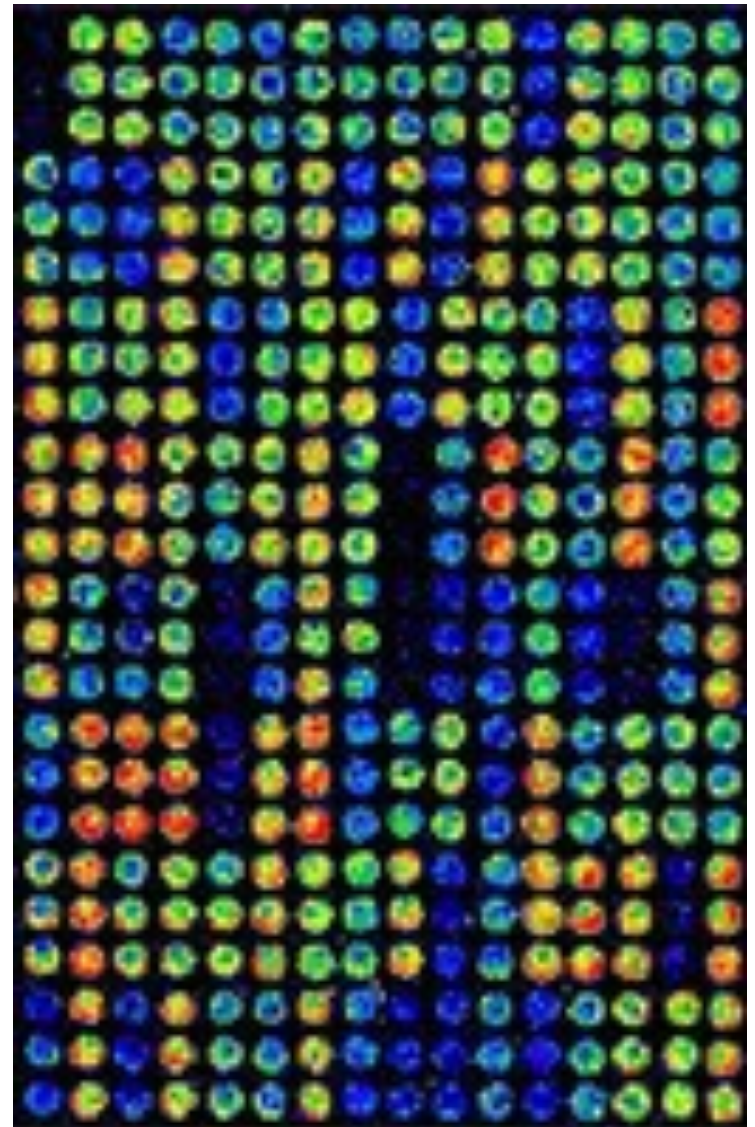
Resultados:

1 **no hay señal:** el gen no está activo.

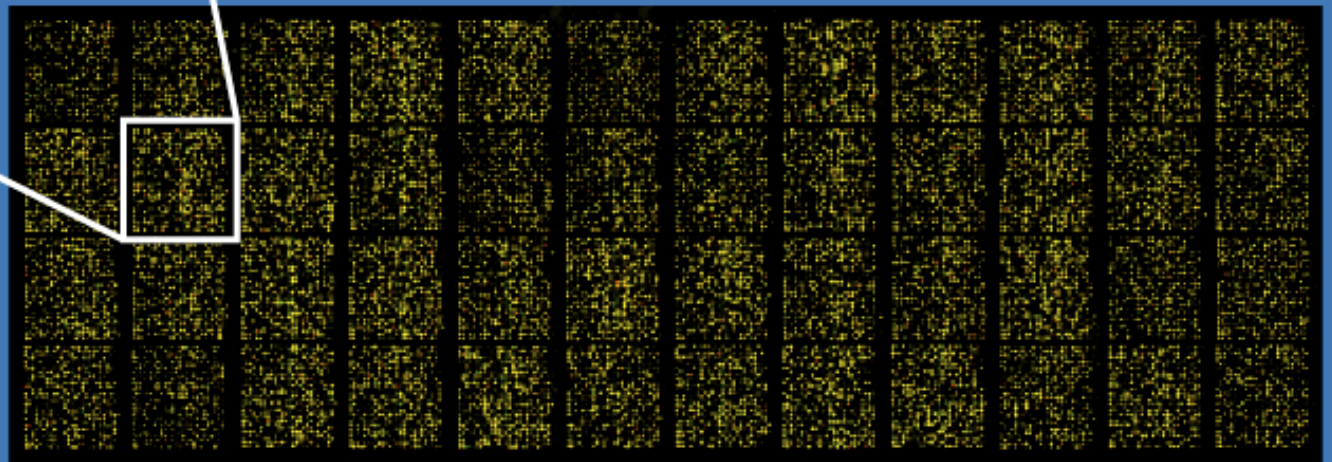
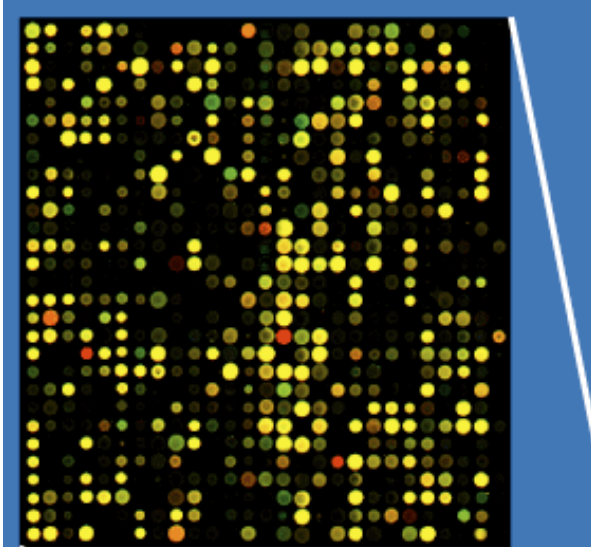
2 CY3 (verde): solo el tratamiento 1 resulta en la expresión del gen.

3 CY5 (rojo): solo el tratamiento 2 resulta in expresión de gen.

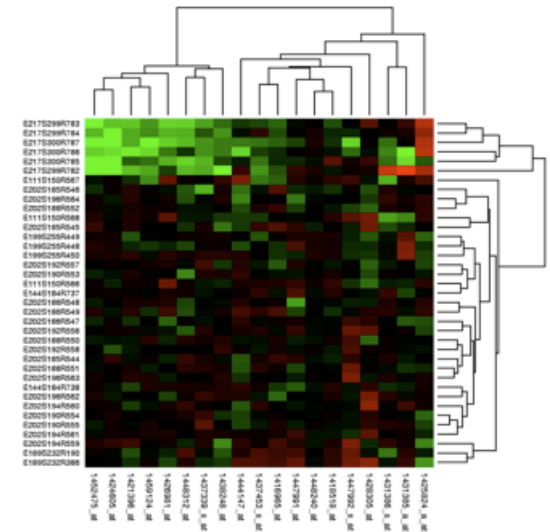
4 CY3 y CY5 (verde + rojo= amarillo): ambos tratamiento tienen la expresión del gen



¿Cuántos genes se pueden imprimir?



Aplicaciones en microbiología

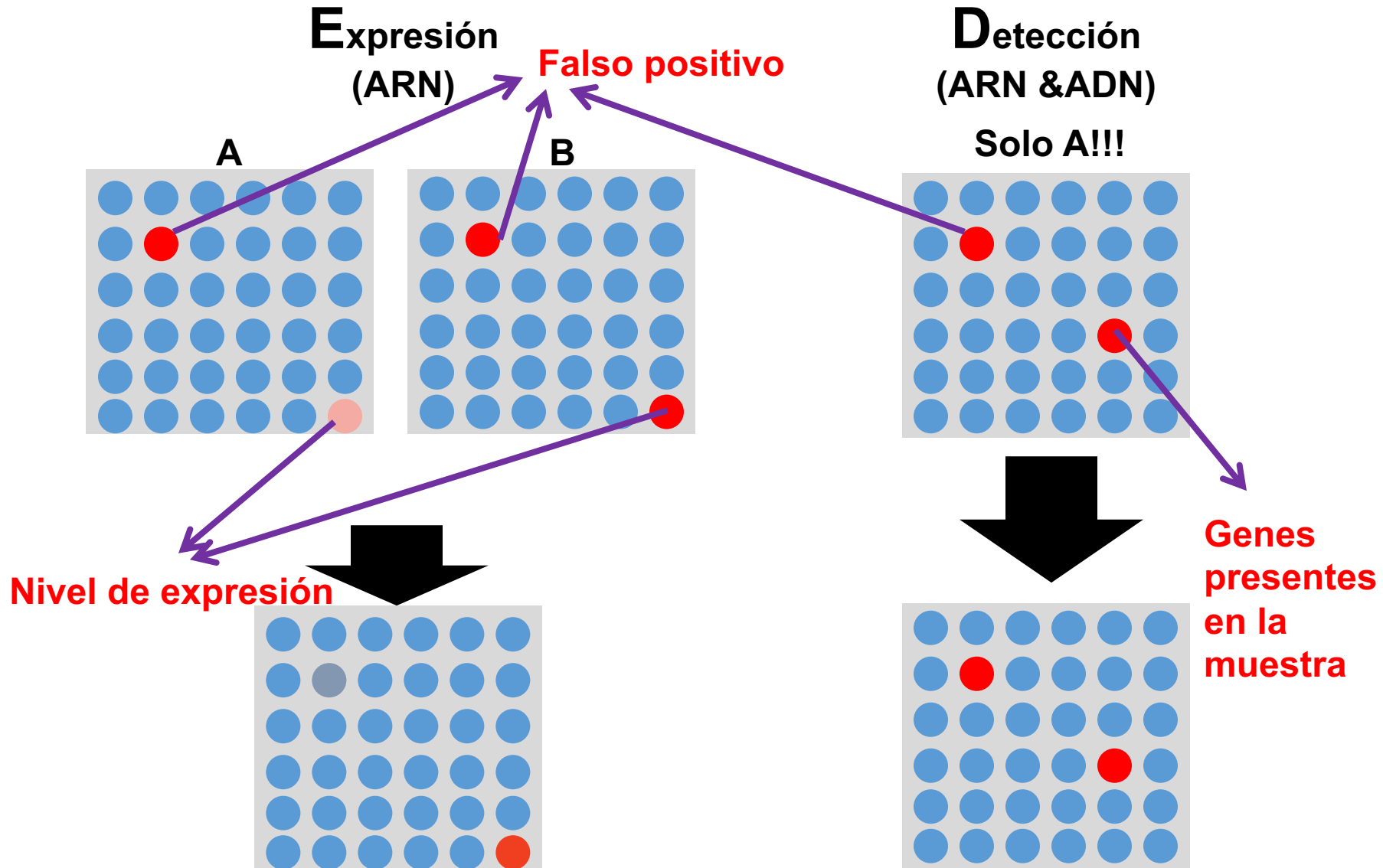


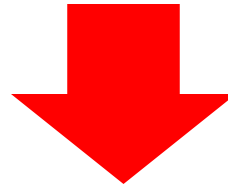
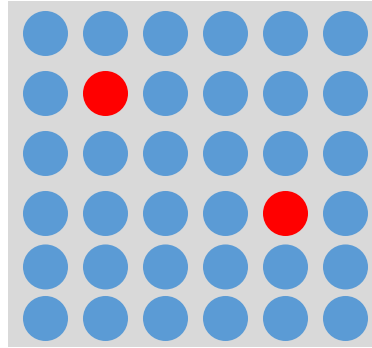
Expresión en micro-arrays

- Genoma entero de solo una especie
- Monitoreo de cambios en la expression ARNm:
- Sin señal: el gen no se está expresando
 - CY3 (verde):
 - CY5 (red):
 - CY3 & CY5 (amarillo):

¿Sólo es válido para cultivos puros?

Microarrays: expresión vs detección





**Por lo tanto, los microarreglos se usan para hacer
barridos o *screenings* de genes**

¿Qué es un screening?

“Screening”

Caso A

Estudio sobre búsqueda de petróleo
en el océano



El screening debe dar casi el 100 % de las localizaciones donde puede encontrarse, no importa si no hay yacimiento viable. Al menos la información recavada es cierta y no se pierde dinero.



Para muestras ambientales es
una excelente herramienta

Caso B

Casos de hepatitis en pacientes



El screening nos indicará todos los pacientes infectados con Hepatitis.

Cómo se calcula el resultado de los microarrays?

		Matríz de confusión			
		actual value			
		<i>p</i>	<i>n</i>	total	
prediction outcome	<i>p'</i>	True Positive	False Positive	P'	
	<i>n'</i>	False Negative	True Negative	N'	
total		P	N		

**Terminology and derivations
from a confusion matrix**

true positive (TP)

eqv. with hit

true negative (TN)

eqv. with correct rejection

false positive (FP)

eqv. with false alarm, Type I error

false negative (FN)

eqv. with miss, Type II error

sensitivity or true positive rate (TPR)

eqv. with hit rate, recall

$$TPR = TP/P = TP/(TP + FN)$$

false positive rate (FPR)

eqv. with fall-out

$$FPR = FP/N = FP/(FP + TN)$$

accuracy (ACC)

$$ACC = (TP + TN)/(P + N)$$

specificity (SPC) or True Negative Rate

$$SPC = TN/N = TN/(FP + TN) = 1 - FPR$$

positive predictive value (PPV)

eqv. with precision

$$PPV = TP/(TP + FP)$$

negative predictive value (NPV)

$$NPV = TN/(TN + FN)$$

false discovery rate (FDR)

$$FDR = FP/(FP + TP)$$

Matthews correlation coefficient (MCC)

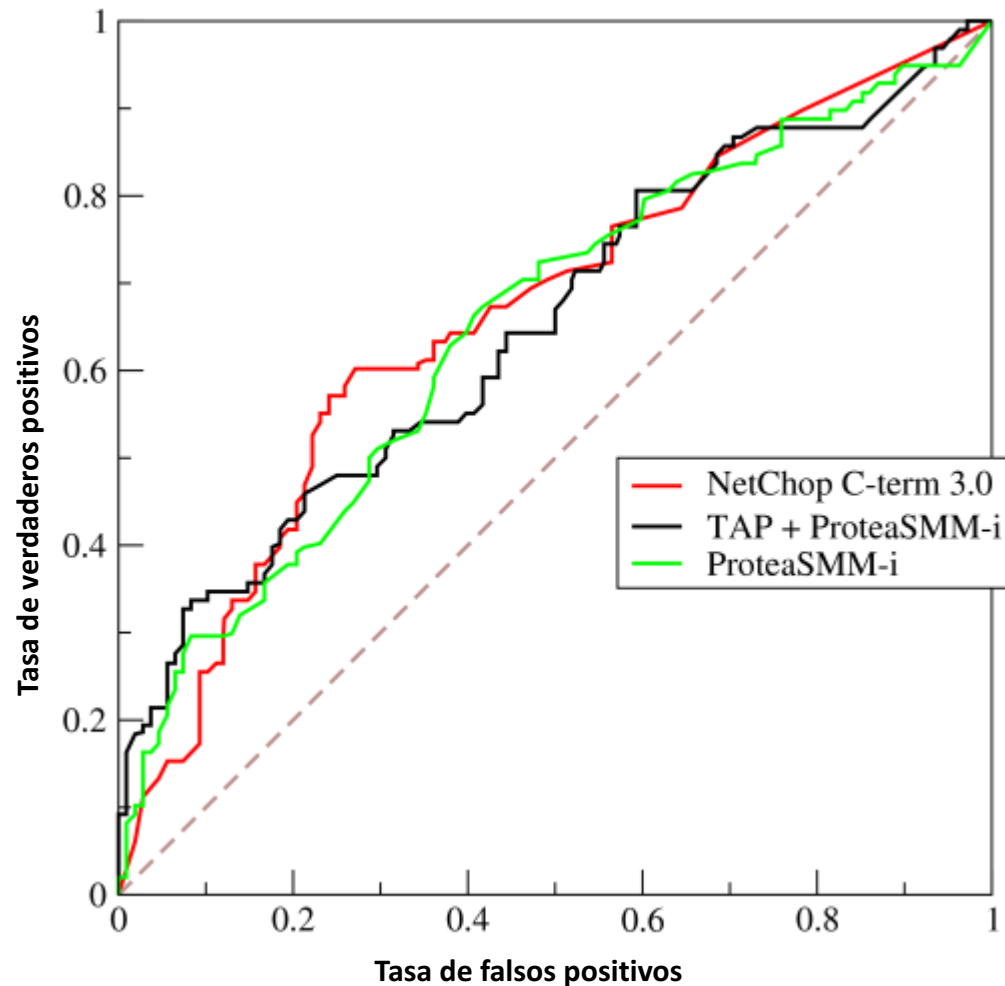
$$MCC = (TP * TN - FP * FN) / \sqrt{P * N * P' * N'}$$

F1 score

$$F1 = 2TP/(P + P') = 2TP/(2TP + FP + FN)$$

Source: Fawcett (2006).

Transformación a Características Operacionales Recibidas (ROC) o curva ROC



¿Cuál es relación entre la curva ROC y la matriz de confusión en **microarray?**



Ejemplo:

Microarray de *Burkholderia xenovorans* LB400

La cepa está completamente secuenciada.

Conocemos el “**actual value**”

(qué es verdadero y qué es falso)

Definiciones:

Verdaderos positivos: Señales que son ciertas (el blanco es el genoma) y **aparecen** como positivas en el total de datos del microarray.

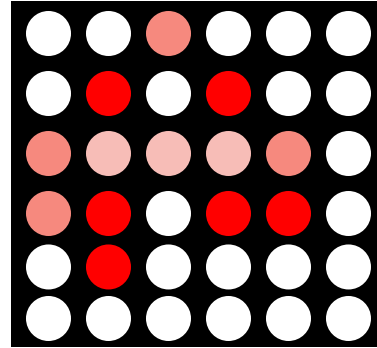
Falsos negativos: Señales que son verdaderas (el blanco es el genoma) **pero NO aparecen como positivas** en el total de datos del microarray.

Falsos positivos: Señales que **no son verdaderas** (el blanco **NO está** presente en el genome) pero **aparecen** como positivo en el total de datos del microarray.

Verdaderos negativos: Señales que **no son verdaderas** (el blanco **NO está** presente en el genome) pero **NO aparecen** como positivo en el total de datos del microarray.

		actual value		
		<i>p</i>	<i>n</i>	total
prediction outcome	<i>p'</i>	True Positive	False Positive	P'
	<i>n'</i>	False Negative	True Negative	N'
total		P	N	

¿Cómo extrae la información de un microarray?



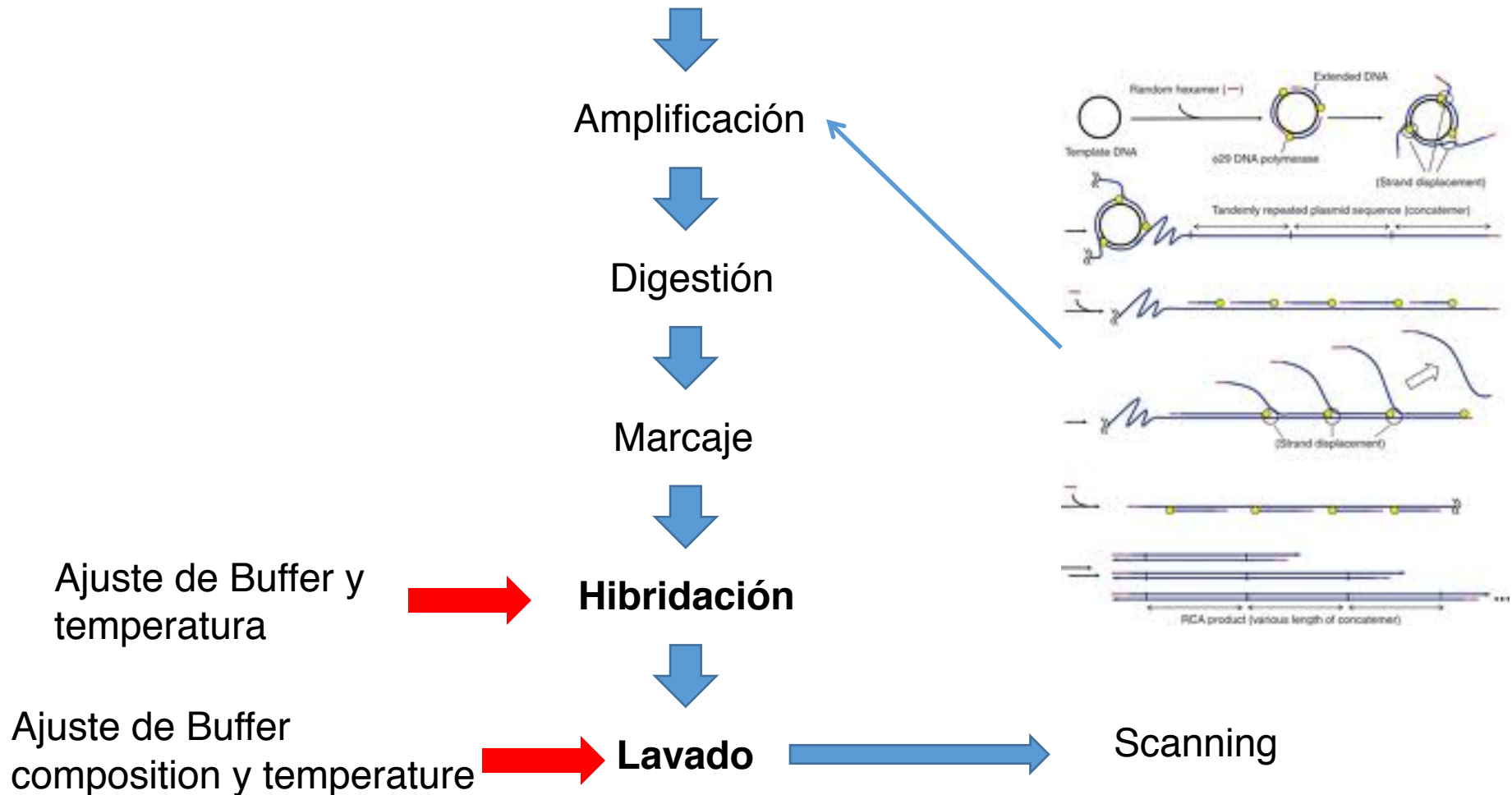
Sólo de una manera: Ordenando por intensidad

pero.... ¿Hasta donde fijar el corte o *cut off*?

y.... ¿Cómo eliminar los falsos positivos del set de datos?

Para eliminar los falsos positivos en el set de datos: Diseño experimental:

Muestra de ADN (ambiental o de cultivo puro)



Qué necesita ser optimizado?

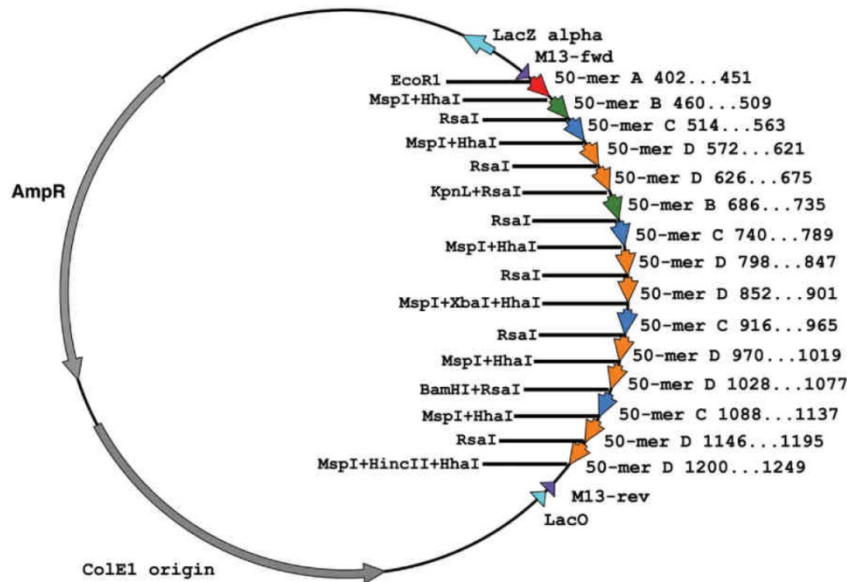
-Para eliminar los falsos positivos.....

Author	Journal	Página	Temperatura de hibridación
Kimes et al.	Env Microbiol 2010	552	35°C
Hinsa-Leasure et al.	FEMS Microbiol Ecol 2010	106	39°C
Liang et al.	ISME J 2011	405	42°C
Xu et al.	ISME J 2010	1062	45°C
Van Nostrand et al.	Appl. Env. Microbiol. 2011	3861	50°C

-Para establecer el nivel de corte: Controles positivos Internos

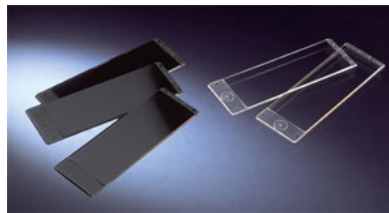
-Para establecer el nivel de corte cut off: necesario usar controles internos positivos

¿Se puede construir una curva estandar para cada fracción del microarray?



Un plásmido sintético con alto número de copias:

- A: 1 copia
- B: 2 copias
- C: 4 copias
- D: 8 copias



Impresión de 50-mers del microarray

Adición de Plasmido a la mezcla de reacción

Muestra de ADN

Amplificación

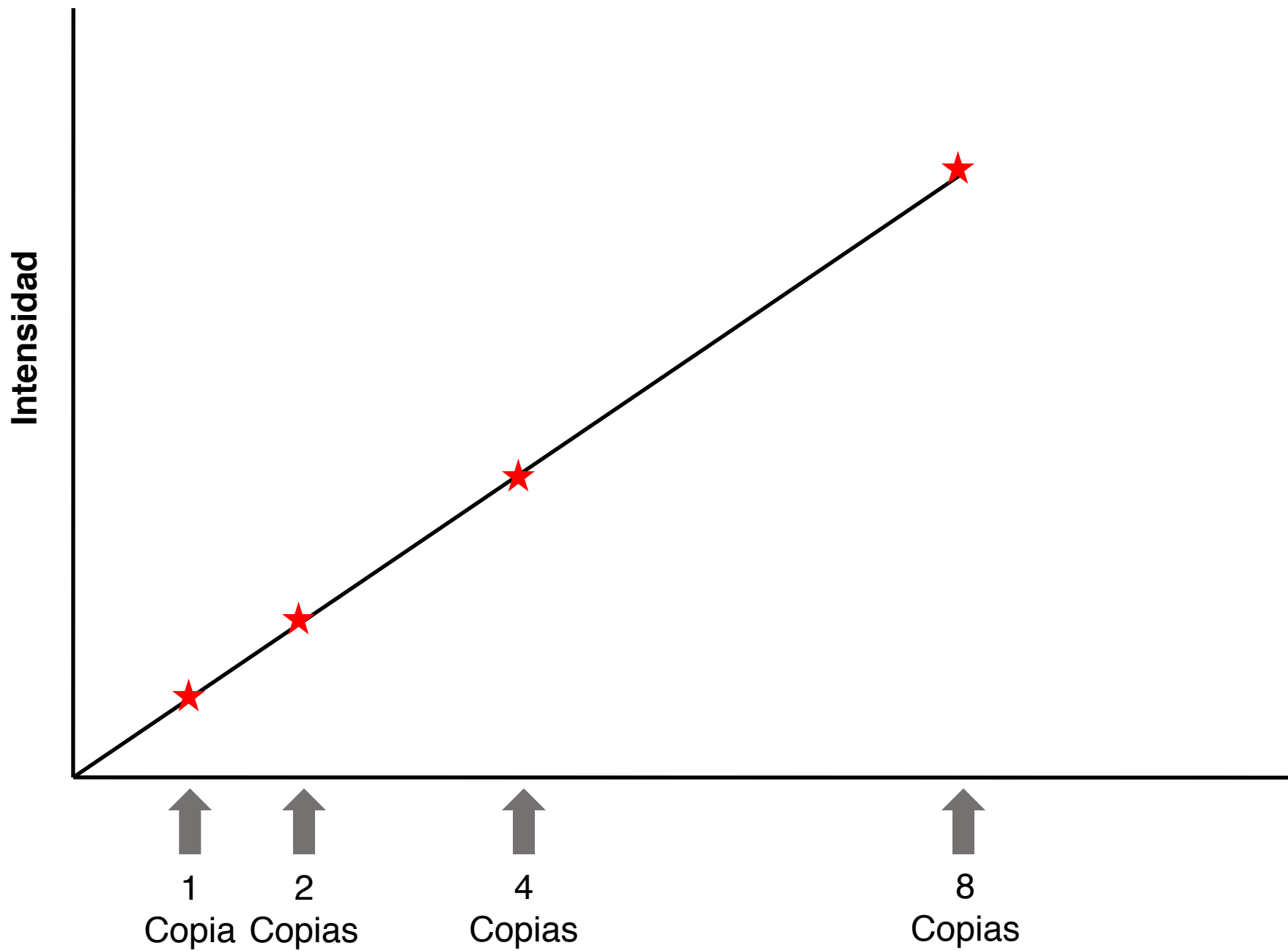
Digestión

Marcaje

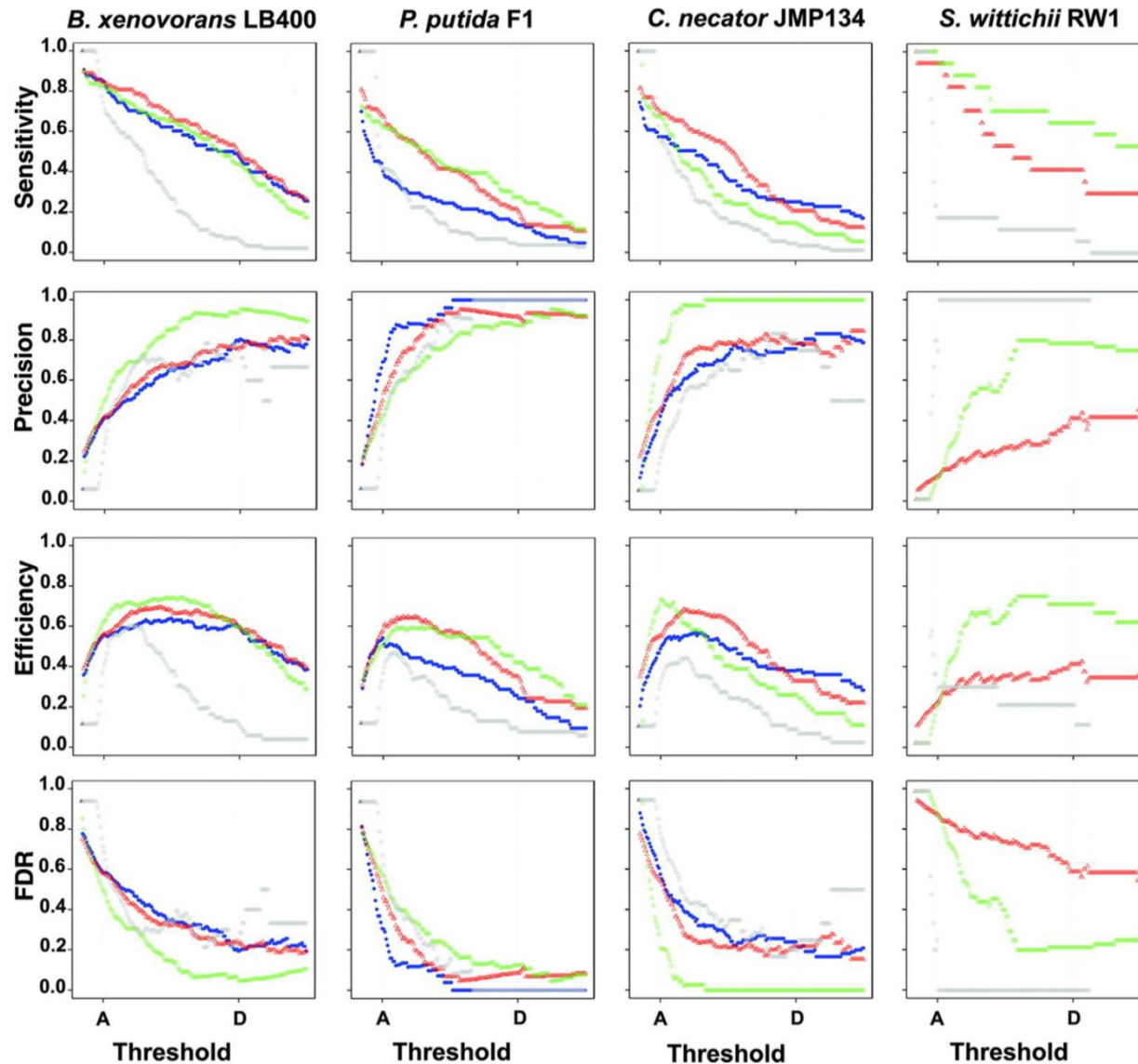
Hibridación

lavado

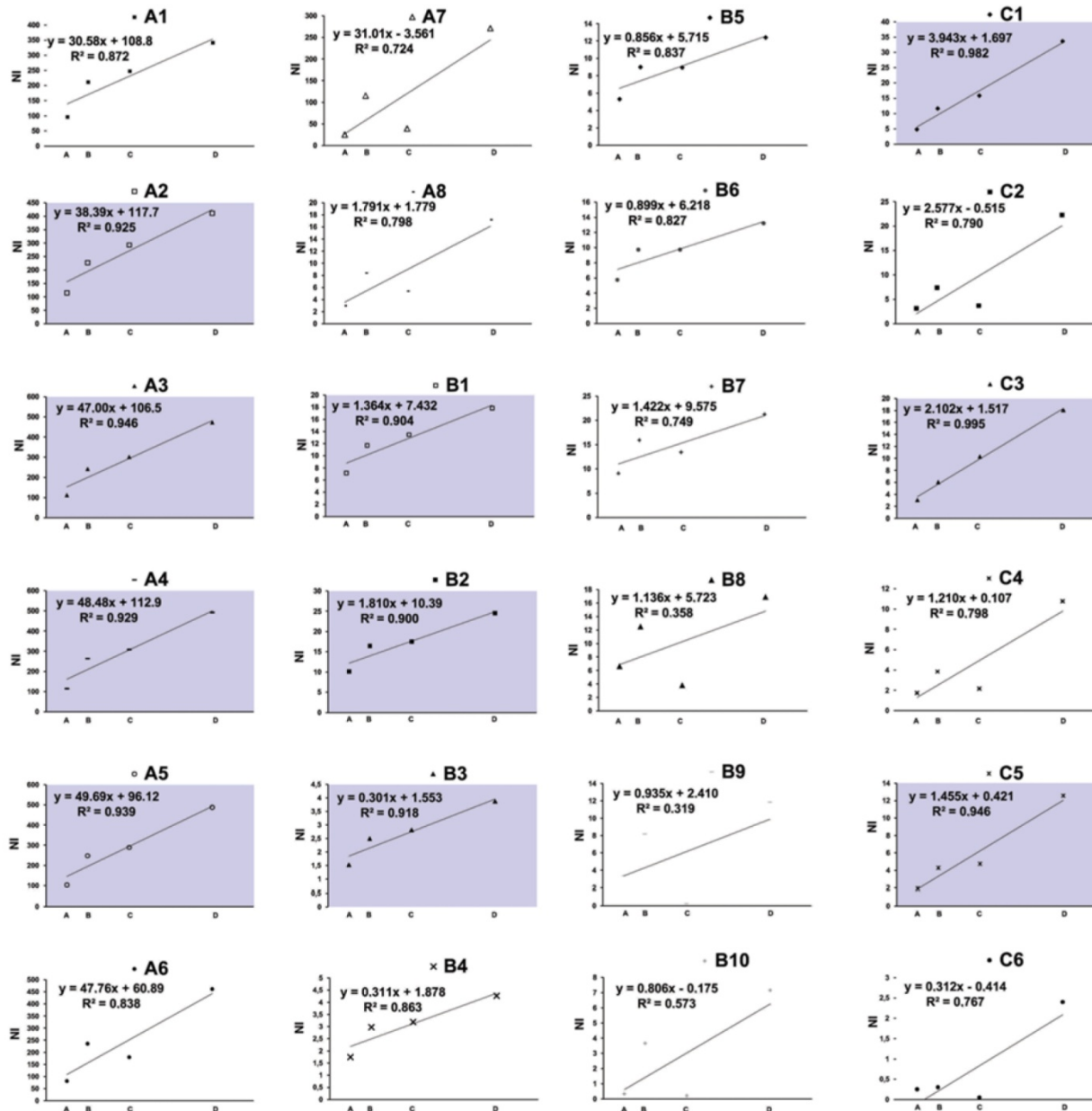
Scanning



CURVA ROC



Standard curve for each microarray



¿Cuáles son las sondas que se pueden imprimir en el microarray?



AMERICAN
SOCIETY FOR
MICROBIOLOGY

Applied and Environmental
Microbiology



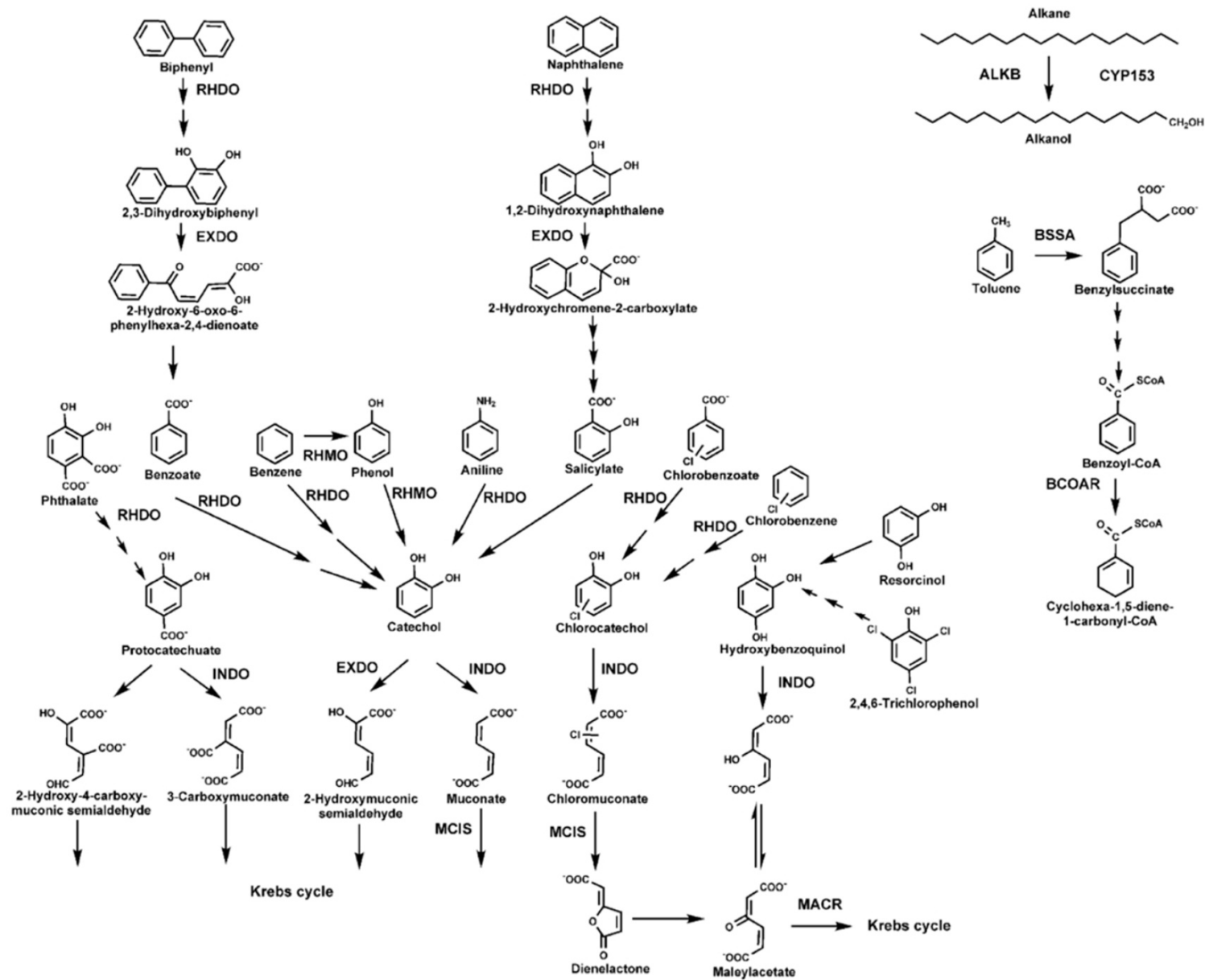
CrossMark
click for updates

Linking Microbial Community and Catabolic Gene Structures during the Adaptation of Three Contaminated Soils under Continuous Long-Term Pollutant Stress

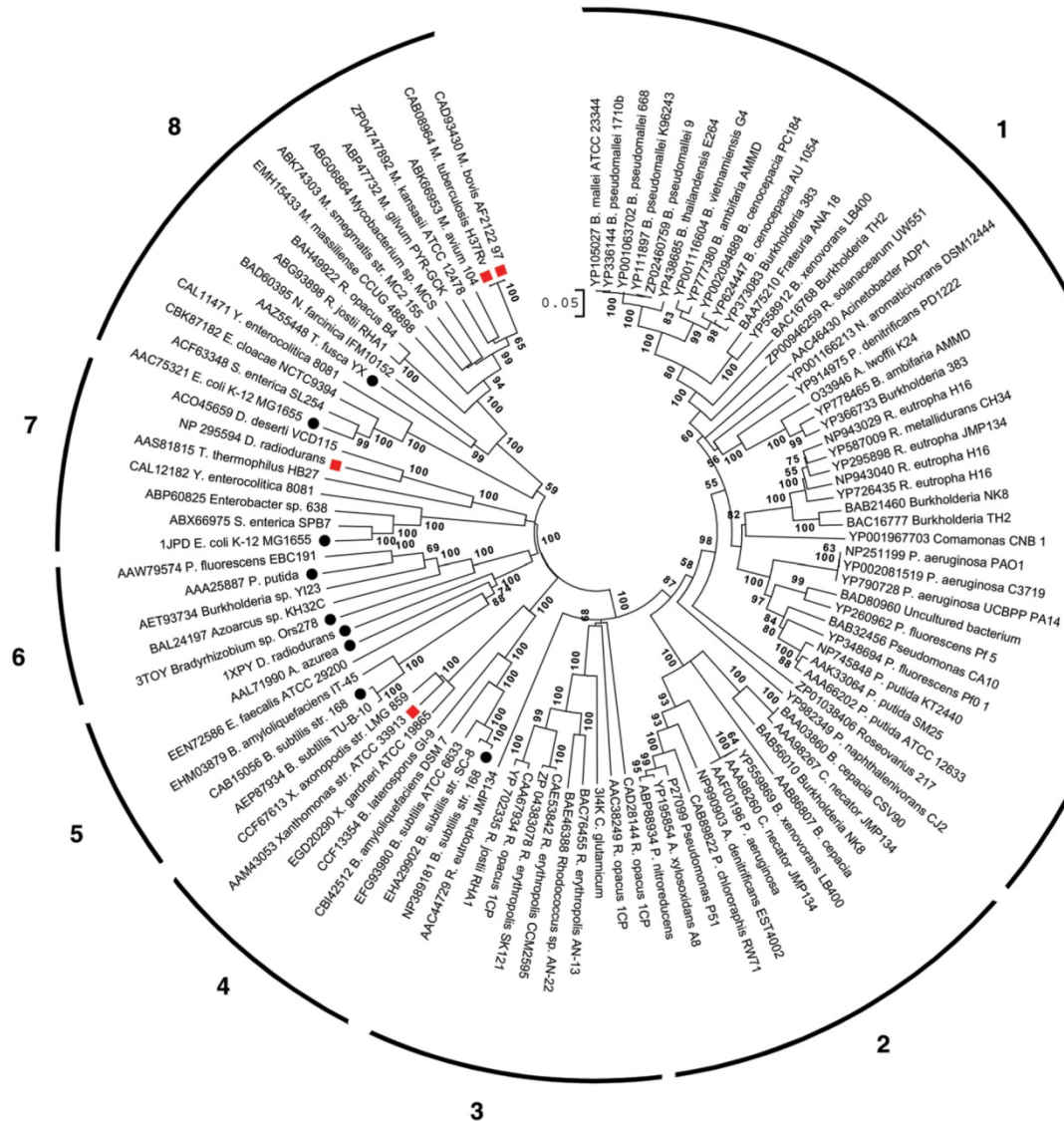
Daiana Lima-Morales,^{a*} Ruy Jáuregui,^{a*} Amelia Camarinha-Silva,^{a*} Robert Geffers,^b Dietmar H. Pieper,^a Ramiro Vilchez-Vargas^{a*}

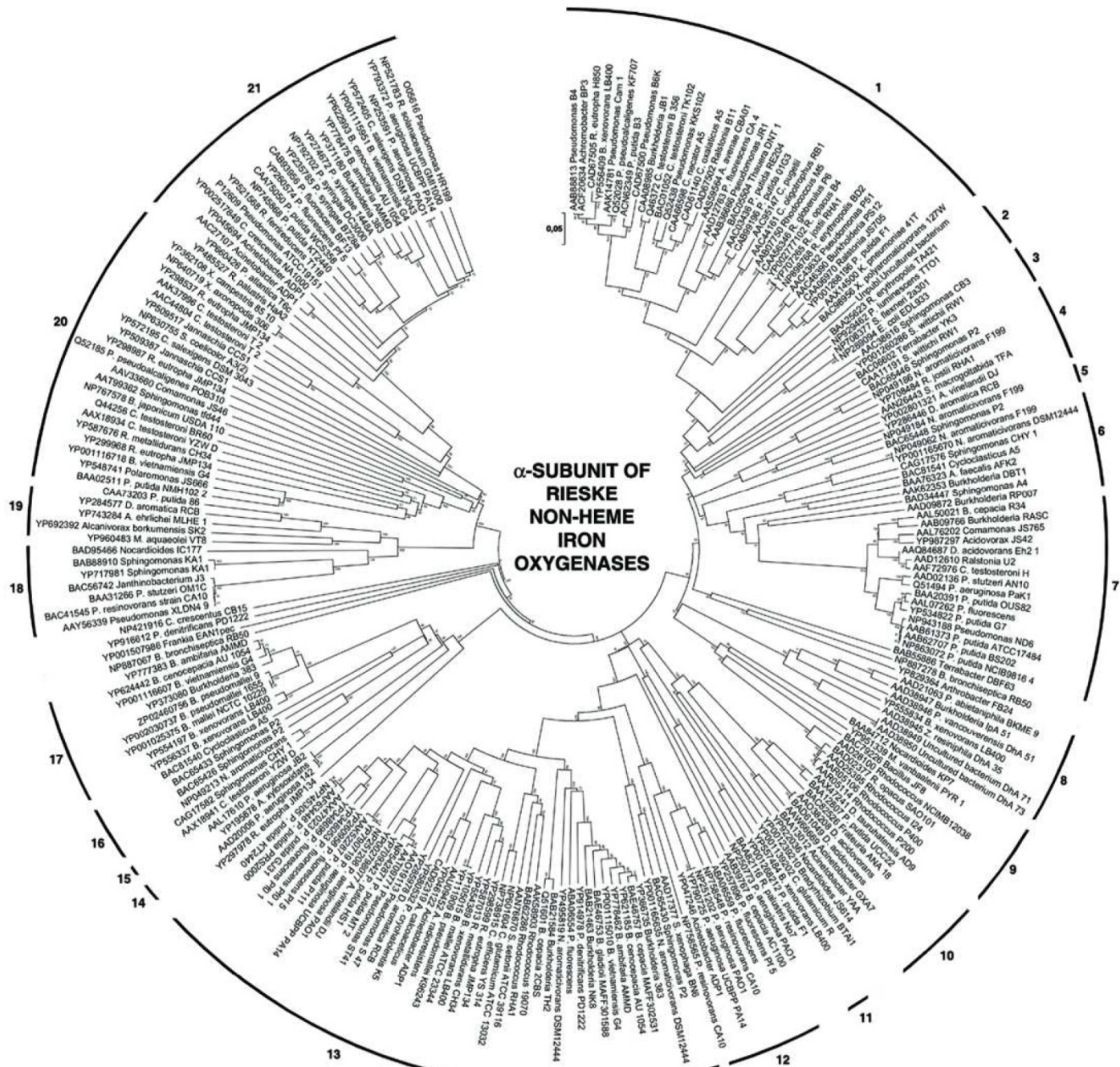
Microbial Interactions and Processes Research Group, Helmholtz Centre for Infection Research, Braunschweig, Germany^a; Genome Analytics, Helmholtz Centre for Infection Research, Braunschweig, Germany^b

Three types of contaminated soil from three geographically different areas were subjected to a constant supply of benzene or benzene/toluene/ethylbenzene/xlenes (BTEX) for a period of 3 months. Different from the soil from Brazil (BRA) and Switzerland (SUI), the Czech Republic (CZE) soil which was previously subjected to intensive *in situ* bioremediation displayed only negligible changes in community structure. BRA and SUI soil samples showed a clear succession of phylotypes. A rapid response to benzene stress was observed, whereas the response to BTEX pollution was significantly slower. After extended incubation, actinobacterial phylotypes increased in relative abundance, indicating their superior fitness to pollution stress. Commonalities but also differences in the phylotypes were observed. Catabolic gene surveys confirmed the enrichment of actinobacteria by identifying the increase of actinobacterial genes involved in the degradation of pollutants. Proteobacterial phylotypes increased in relative abundance in SUI microcosms after short-term stress with benzene, and catabolic gene surveys indicated enriched metabolic routes. Interestingly, CZE soil, despite staying constant in community structure, showed a change in the catabolic gene structure. This indicates that a highly adapted community, which had to adjust its gene pool to meet novel challenges, has been enriched.

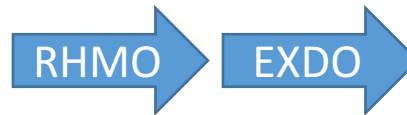


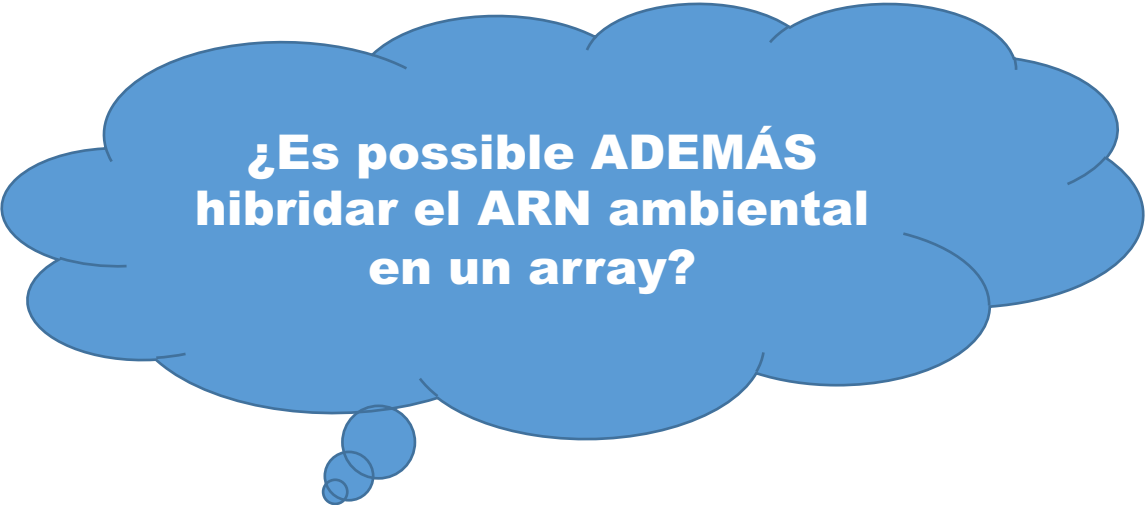
Relaciones evolutivas de miembros de la familia de las enolasas





Organización típica del genoma de los genes clave catabólicos





**¿Es possible ADEMÁS
hibridar el ARN ambiental
en un array?**

Diseño experimental de ADN

Plásmido
adicionado al
mix de reacción



Muestra de ADN



Amplificación



Digestión



Marcaje



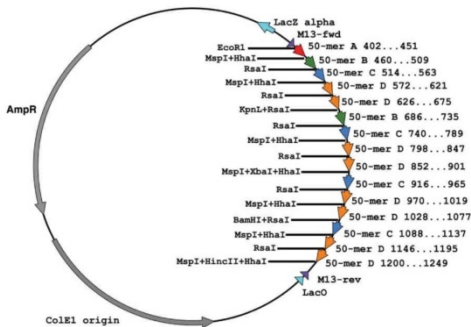
Hibridación



lavado



Scanning



Impresión
de 50-mers
del array

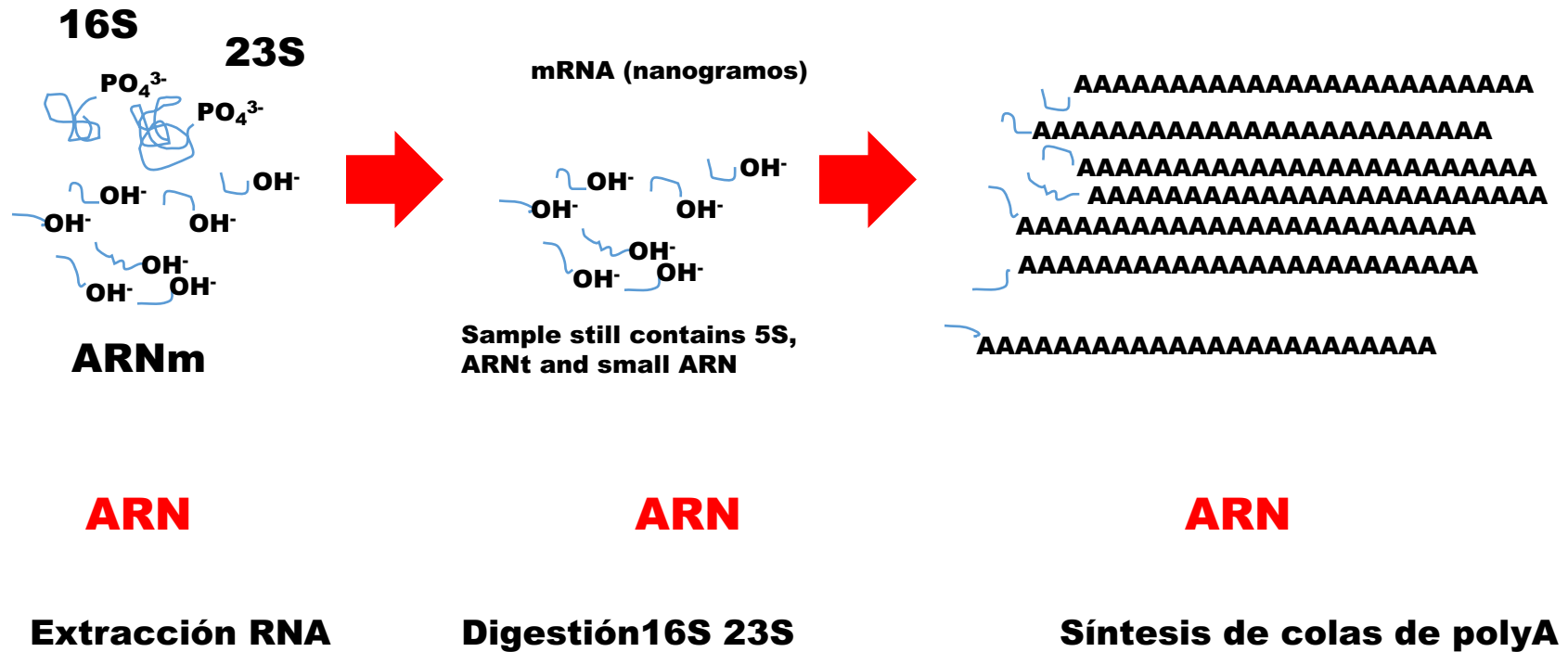
Diseño experimental de ARN

Modificación en :

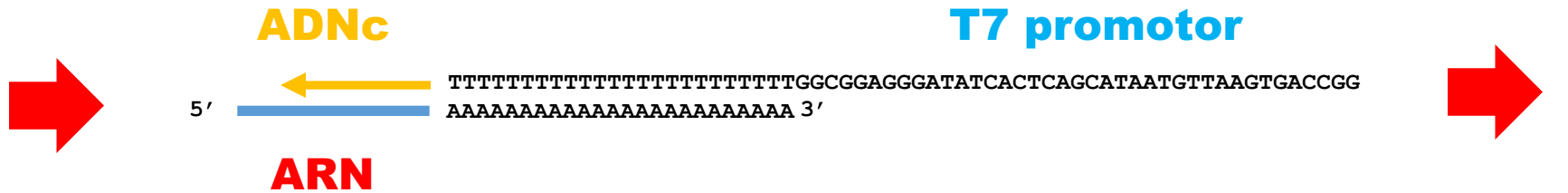
1.- El plásmido

2.- El diseño experimental

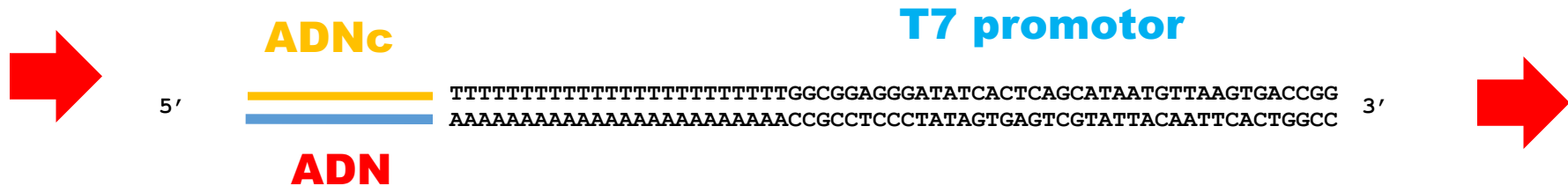
Modificación del diseño experimental



Modificación del diseño experimental

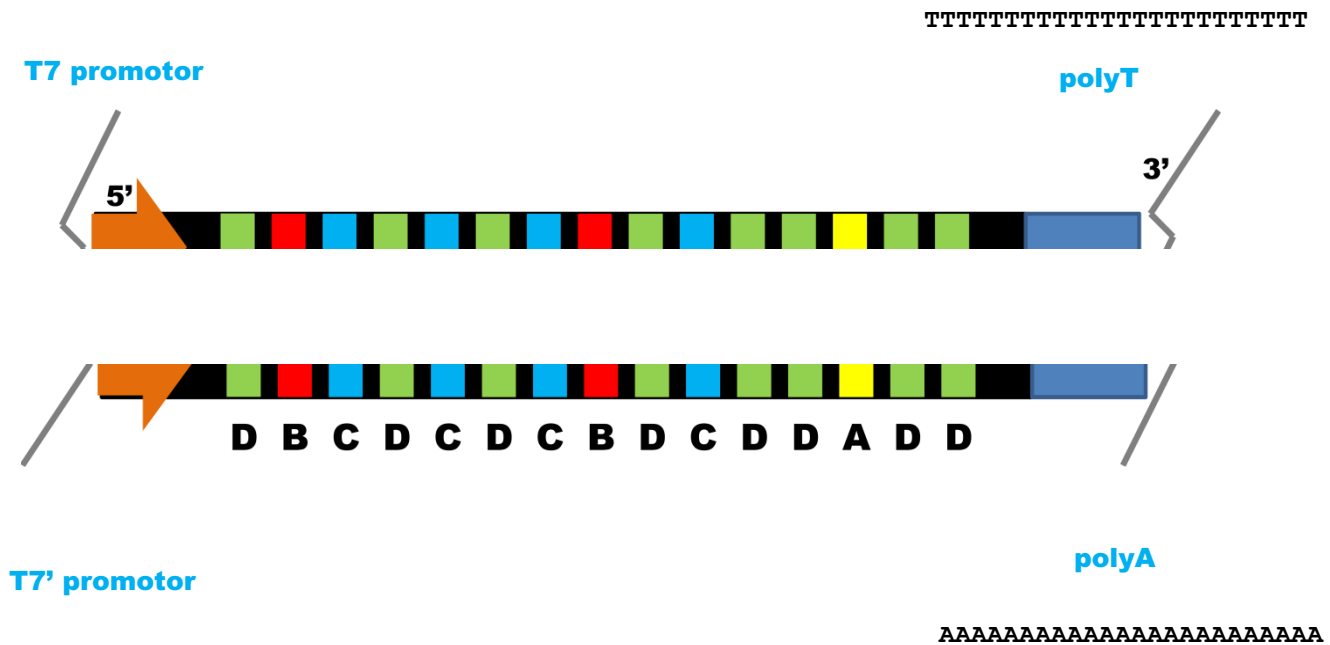


Transcriptasa Reversa

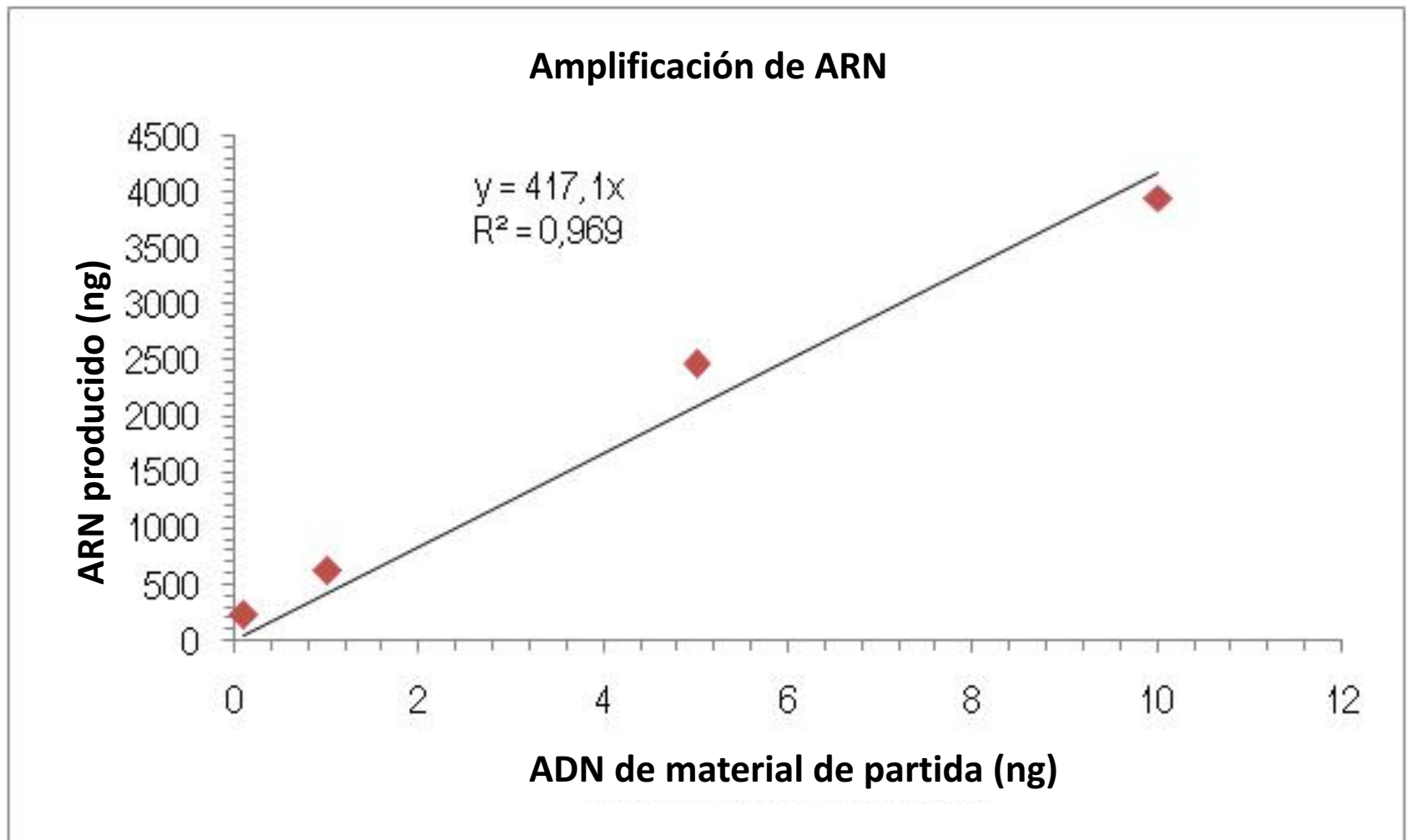


Síntesis de la segunda hebra de ADN

Modificación del plásmido (controles positivos internos)

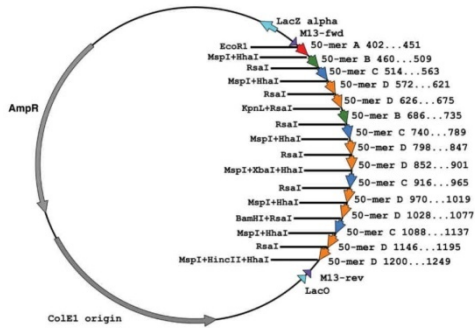


Modificación del plásmido (controles positivos internos)



Diseño experimental ADN

Plásmido
adicionado al mix de reacción



Muestra de ADN



Diseño experimental ARN

Muestra

Extracción ARN

Digestión de 16S 23S

Síntesis de colas de polyA

Transcripción reversa

Síntesis de segunda hebra de ADN

Amplificación T7 ARN

ng ADN $\longrightarrow$ μ g ARN

Transcripción reversa

- Síntesis de segunda hebra de ADN

APLICACIONES DE MICROARRAYS Y BIOCHIPS

Objetivo	Técnica	Aplicación
Caracterización del ADN	• Secuenciación por Hibridación (Sequencing by Hybridization o SBH).	• Detección de mutaciones y variaciones en el genoma.
	• Identificación por hibridación.	• Diagnóstico molecular.
Cuantificación del ADN	• Análisis de la expresión génica.	• Caracterización del perfil genético de enfermedades.
	• Descubrimiento de nuevos fármacos.	• Farmacogenómica.
		• Validación de dianas terapéuticas.
Comparación del ADN		• Caracterización de la acción molecular de compuestos activos.
	• Identificación de polimorfismos.	• Diagnóstico molecular.
	• Comparación de secuencias homólogas.	• Farmacogenómica.
		• Genómica (funcionalidad de genes).

Tabla 13. Aplicaciones de microarrays y biochips (elaboración propia).

Artículo de discusión

environmental
microbiology



Environmental Microbiology (2010) 12(12), 3089–3104

doi:10.1111/j.1462-2920.2010.02340.x

Minireview

Metabolic networks, microbial ecology and ‘omics’ technologies: towards understanding *in situ* biodegradation processes

Ramiro Vilchez-Vargas,¹ Howard Junca² and Dietmar H. Pieper^{1*}

¹Microbial Interactions and Processes Research Group, HZI – Helmholtz Centre for Infection Research, Inhoffenstraße 7, D-38124 Braunschweig, Germany.

²GeBIX Colombian Center for Genomics and Bioinformatics of Extreme Environment and Research Group Microbial Ecology: Metabolism, Genomics and Evolution of Communities of Environmental Microorganisms, CorpoGen, Carrera 5 # 66A-35, Bogotá, Colombia.

Summary

Microbial degradation is the main mechanism responsible for the recovery of contaminated sites, where a huge body of investigations is available in which most concentrate on single isolates from soils capable of mineralizing pollutants. The rapid development of molecular techniques in recent years allows immense insights into the processes *in situ*, including identification of organisms active in target sites, community member interactions and catabolic gene structures. Only a detailed understanding of the functioning and interactions within microbial communities will allow their rational manipulation for the purpose of optimizing bioremediation efforts. We will present the status of the current capabilities to assess and predict catabolic potential of environmen-

Introduction

Given the widespread contamination with aromatic and aliphatic pollutants, it is a long-held desire to treat organic and inorganic waste more efficiently and remediate polluted environments via controllable and amenable microbial activities. However, despite their promising performance in the laboratory, the application of pollutant-degrading bacteria in microcosms or near-field situations have mostly ended in disappointment (El Fantroussi and Agathos, 2005; Thompson *et al.*, 2005). Therefore, more optimal and rational use of the extremely high potential of catalytic activities in the environment has been proposed for more successful pollution treatment (Watanabe *et al.*, 2002). Presently, this potential cannot be sufficiently exploited because of the lack of knowledge on the desired catabolic activity and ecological behaviour of the microbial community (Paerl and Steppe, 2003). Pollutant degradation in contaminated environments is in many cases carried out by microbial food webs rather than single species (de Lorenzo, 2008), where key species and catabolic genes are often not identical to those that have been isolated and described in the laboratory (Jeon *et al.*, 2003; Witzig *et al.*, 2006). We now know that microbial diversity in these environments is in orders of magnitude higher than assumed from previous cultivation efforts (Leigh *et al.*, 2007). A particularly large number of novel techniques have been developed, which now allow the determination of microbial diversity and activity *in situ* at