

細菌遺傳學、細菌基因體學 (11-16-2010, 11-17-2010)

Chapter 14

Bacterial Genetics: 14.4 Creating genetic variability 11-16-2010 http://life.nctu.edu.tw/~hlpeng/ 彭慧玲 分機:56916 (hlpeng@mail.nctu.edu.tw)

課程綱要與評分標準

11/16, 11/17	Bacterial genetics and microbial genomics
11/23, 11/24	Microbial taxonomy, Archaea
11/30	Fungi
12/7	Exam I- 20%
12/8	Bacteria (I)
12/14, 12/15	Bacteria (II)
12/21, 12/22	Viruses (I)
12/28, 12/29	Viruses (II), viroids, virusoids, and prions
1/4, 1/5	Microbial diseases and their control
1/11	Exam II- 20%

<u>出席率 10.5% (21 h)</u>

NDM-1 (New Delhi Metallo-beta-lactamase)

 In December 2009, a carbapenem-resistant *K. pneumoniae* infection acquired in New Delhi, India

 Resistant to multiple antibiotics

 Also reported in *E. coli* and *Acinetobacter baumannii* The Lancet infectious disease 2010



NDM-1 Kp01

超級細菌

Horizontal (Lateral) Gene Transfer (HGT) in bacteria

- Haploid geneome
 Exogenote (Donor)
 Endogenote (Recipient)
- Merozygote
 - Transconjugants
 - Transformants
 - Transductants



Figure 14.11

Recombination at the molecular level

- Homologous recombination (Reciprocal recombination)
 - Double-Strand Break Model
 - results from <u>DNA strand</u> <u>breakage</u> and <u>reunion</u>, leading to crossing-over



Figure 14.12

Nonreciprocal Homologous Recombination

 incorporation of single strand of DNA into chromosome, forming a stretch of heteroduplex DNA

proposed to occur during bacterial transformation Association of homologous segments

Strand separation and pairing

Endonuclease nick at the arrow on donor strand

Endonuclease nicks host strand

Gaps in strand filled and ligated



Figure 14.13

Site-specific recombination

- insertion of viral genome into host chromosomes
 - insertion of nonhomologous DNA
 - Only a small region of homology is required
 - Transposition carried out by transposable elements
 - Simple transposition
 - Replicative transposition

Transposable Elements

A ESTA MARTERA ESTA A DA ALCERTA DA CARTA MARTINA DA CARTA MARTINA DA ARTINA DA CARTA DA CARTA DA CARTA MARTERA

- transposition
 - the movement of pieces of DNA around the genome
- transposable elements (transposons)
 - segments of DNA that carry genes for transposition
- widespread in bacteria, eukaryotes and archaea

Types of transposable elements

IS element 2 IR + transposase gene Transposons 2 IS elements additional genes **Replicative transposons** Resolvase gene



Simple Transposition- cut and paste

- Generation of direct repeats in host DNA flanking a transposon

Figure 13.20



Replicative Transposition

Figure 14.16

insertion generates direct repeats of flanking host DNA
usually transposon replicated, remaining in original site, while duplicate inserts at another site
Tn3



Effects of transposition

mutation in coding region arrest of translation or transcription interruption of promoter regions... interruption of ribosomal binding sites... activation of genes provide strong promoter sequences generation of new plasmids Multi-drug resistance plasmids

R Plasmid results from Tn3 Transposition



Plasmids

small, autonomously replicating DNA molecules that can exist independently or, as <u>episomes</u>, integrate reversibly into the host chromosome

conjugative plasmids such as the F plasmid can transfer copies of themselves to other bacteria during conjugation

R plasmid, virulence plasmid, metabolic plasmid...

Bacterial conjugation

The transfer of genes between bacteria that depends on direct cell to cell contact

> mediated by the sex pilus encoded by F plasmid



Fig. 14.19



F plasmid integration



17



 transfer of DNA between cells discovered by Lederberg and Tatum (1946)

The U-tube experiment

- cell-to-cell contact is necessary for the DNA transfer, which demonstrated by Bernard Davis (1950)



F-mediated conjugation

F⁺→ F⁻

- A polar gene transfer
 replicated by rollingcircle and the duplicate is transferred
- recipients usually become F⁺
- donor remains F⁺



The type IV secretion system

Tra proteinsTraA pilin

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



Fig. 14.22

Rolling-circle replication

 A 1S tail is generated → 2S by synthesis of a complimentary strand Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



Hfr cells

Integration of F to chromosome to generate Hfr cell
 *h*igh *f*requency of recombinants



(a) Insertion of F factor into chromosome

$HFr \rightarrow F^{-}$ conjugation

a complete copy of the F factor is usually not transferred

10.7



23

Fig. 14.24 (b)

F' cells

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

An error excision caused the A gene of Hfr cell is picked up by the F factor (red)



Fig. 14.25(a)

F' conjugation

F' x F- Mating

- formed by incorrect excision from chromosome
- some of the F factor is left behind in the host chromosome
- some host genes have been removed along with some of the F factor
 - these genes can be transferred to a second host cell by conjugation



DNA transformation

 uptake of DNA fragments and incorporation into recipient
 competent cell
 uptake of a plasmid



Transformation in Streptococcus pneumoniae

Aarc+ Receptor Concernance of the second la c **DNA fragment** binds to a cell surface receptor. The other states of the state An extracellular endonuclease outs the DNA into Sauce" smaller fragments. free co One strand is degraded and a single strand is transported into Uptake the cell. system laic. lac: The DNA strand aligns itself with a homologous region on the bacterial chromosome. Address ** lesc: The DNA strand is incorporated into the bacterial chromosome via homologous recombination. Heteroduplex The heteroduplex. DNA is repaired in a way that changes /ac* strand to create a /ac* gene. Report T Transformed Fig 14.27 cell Photos and the Photos

DNA Uptake Systems

Natural competence

G(-) Neisseria sp.

G(+) Bacillus sp.

Fig 14.28



Transduction- transfer of genes by phages



29

Generalized transduction

- any part of bacterial genome can be transferred
- occurs during lytic cycle
- during viral assembly, fragments of host DNA mistakenly packaged into phage head





Specialized Transduction

- carried out only by temperate phages that have established
 lysogeny
- only specific portion of bacterial genome is transferred
- occurs when prophage is incorrectly excised

Fig 14.31



Transduction for phage Lamda

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



Transmission of drug resistance

- The origin of drug resistance
 - R plasmid
 - Tn5, Tn21, Tn551
 - Intergon
 - several resistance genes carried together as gene cassettes



Integron + ((

A gene cassette typically consists of little more than a single promoter-less gene and a recombination site

- *intl* encodes an integrase of tryosine recombinase family
- attC and attl
- At least 130 different cassettes that carry known or predicted antibiotic resistance genes, along with many cassettes of unknown function

FEMS Microbiol Rev 33 (2009) 757–784

attC

Three characterized antibiotic resistanceconferring regions from *K. pneumoniae* 05-506

A UTI (urinary tract infection) isolate of KPC strain 05-506, from a Swedish patient of Indian origin traveled to New Delhi, India, which carries a new subgroup of metallolactamase

NUMPER OF STREET, S



NDM-1 producing E. coli

 Emergence of NDM-1-producing multidrug resistant Escherichia coli in Australia


Chapter 16

Microbial Genomics-1



REV ANNER RED AND THE BOULD BY BURNESS OF MINES HER AN ENDER MENNES AND AN AND AN AN AN AND AND AN AN AN AN AN

study of <u>molecular organization</u> of genomes, their <u>information content</u>, and <u>gene products</u> they encode

Determining DNA Sequences

 Sanger Method (1975)
 uses dideoxynucleoside triphosphates (ddNTP) as chain terminator
 automated systems
 use dideoxynucleotides labeled with fluorescent dyes



Chain termination method

- Four reactions
 - 1S DNA template
 - primer
 - DNA polymerase
 - 4 dNTP
 - one ddNTP
- random insertion of ddNTP generates different lengths of DNA fragments
- fragments separated electrophoretically
- sequence read

Fig. 16.2

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display

Isolated unknown DNA fragment

5' C A C T T A G C C G A T C C 3' G T G A A T C G G C T A G G Original DNA to be sequenced

ONA is denatured to produce single template strand.

3' GTGAATCGGCTAGG 5'

3 Labeled specific primer molecule hybridizes to the DNA strand.

Primer

ONA polymerase and regular nucleotide mixture (dATP, dCTP, dGTP, and dTTP) are added; ddG, ddA, ddC, and ddT are placed in separate reaction tubes with the regular nucleotides. The dd nucleotides are labeled with some type of tracer, which allows them to be visualized.



5 Newly replicated strands are terminated at the point of addition of a dd nucleotide.



Schematic view of how all possible positions on the fragment are occupied by a labeled nucleotide





provides the correct DNA sequence.

2

Automated Sanger DNA sequencing



Post-Sanger DNA sequencingpyrosequencing

- Each bead coated with PCRamplified chromosomal fragments (300-500 bp)
- pyrophosphatase sequencing
 - $(DNA)_n + dNTP \rightarrow (DNA)_{n+1} + PPi by$ DNA polymerase
 - PPi + APS (adenosine phophosulfate) → ATP + SO₄²⁻ by <u>ATP sulfurylase</u>
 - ATP+luciferin+O2→ AMP +ppi +oxylucerin +CO2+ light by <u>luciferase</u>



A Pyrogram

Fig. 16.5



Roche 454 Genome Sequencing

Nature Rev Microbiology (2008) May



SOLEXA sequencing

- A glass slide attached with the PCR amplification of immobilized DNA fragnents
- Sanger-type sequencing
 - 4 fluorophore terminators



SOLiD sequencing

- Supported oligo ligation detection
- The DNA fragments are lengthen by
 - the addition of a short piece of DNA (the anchor primer)
 - eight-base oligonucleotides (by every possible combination of A-T-C-G) and each has a fluorescently labeled A, T, C, or G in the 4th and 5th positions
 - Ligase
 - Exciting with laser light



Whole-genome shotgun sequencing

- Library construction to provide the templates
 - genomic clones
- random sequencing
- fragments alignment and gap closure
 - Contig- a set of larger, contiguous nucleotide sequences
- Editing (proofread)
 47
 1000 kb/ 24 h



1995, J. Craig Venter and Hamilton Smith (TIGR: the institute of genomic research)

Single-Cell genomic sequencing

- Multiple strand displacement (NDA)
 - DNA from single cell (down to 10⁻¹⁵ g)
 - random primers (hexamers)
 - phage phi29 DNA polymerasedNTPs
- Many new strands (12 kb~100 kb) are synthesized → cloning → sequencing



Fig. 16.7

Bioinformatics

Combine Biology, mathematics, computer science, and Statistics to generate data on genome content, structure, and arrangement

- also provides data on protein structure and function
- DNA sequence data stored in large databases
 International Nucleic Acid Sequence Data Library (GenBank)
 analysis of genome data using computers
 in silico analysis

Genome annotation

NEW YORK WARRANT CONTRACTOR STOLEN AND AND ADDRESS AND ADDRESS AND ADDRESS AND ADDRESS AND ADDRESS AND ADDRESS

- To determine location of genes and nature of genes or presumed genes
 identifies open reading frame (ORF) in genome
 - a reading frame > 100 codons that is not interrupted by a stop codon
 - there is an apparent ribosomal binding site at the 5' end and terminator sequences at the 3' end
- To assign tentative function of gene

Finding potential protein coding genes - CDS (coding sequences)

ERREY AND ET ALS MADE THE ERRE OF THE TOP THE RECTORNAL COMPANY AND IN THE RECTOR OF THE RE



Alignment of the conserved regions

									C	Con	sen	ISU	s h	elic	al r	egi	on									
							4														L					
							. ()																		
	Escherichia coli	Ρ	F	RI	FΙ	E	E	Е	к	к	G	F	Ŀ		-	K	R	L	F	G	G	270	D			
	S. enterica serovar	Ρ	F	RI	FΙ	E	Е	Е	κ	κ	G	F	Ŀ		-	K	R	L	F	G	G	27(D			
Crom negative	Typhimurium	Ρ	F	RI	F۱	/ E	Е	Е	к	κ	G	F	Ŀ		-	K	R	L	F	G	G	27(C			
Gram-negative	Yersinia pestis	Е	F	RI	Fι	. Т	E	Α	K	Κ	G	1	F.		-	K	R	L	F	G	G	276	6			
Bacteria	Vibrio cholerae	Ρ	н	RI	Fι	_ D	V	Q	Q K K G F L Q R L F G G R E 271 E K K S F F K R L F G G 271 E K K G F F S K L F G G 269 T K K G F W S R L F G G 245																	
	Pseudomonas aeruginosa	Е	М	RI	FΙ	. E	Α	Ε	K	Κ	S	F	F -		-	K	R	L	F	G	G	27	1			
	Neisseria gonorrhoeae	Ρ	М	RI	F 1	ΓТ	۷	Е	к	κ	G	F	F -		-	S	K	L	F	G	G	269	Э			
	Xylella fastidiosa	Г	Α	11	EE	ΕТ	Q	т	ĸ	κ	G	F	w -		-	S	R	L	F	G	G	24	5			
Gram-positive	Staphylococcus aureus	Ρ	L	M	S I	E	Т	κ	ĸ	А	G	FI	F /	A F	R L	K	Q	L	F	S	G	Κ	266	3		
Bacteria	Listeria monocytogenes	Ρ	F	EI	K١	ΥE	т	Q	-	т	G	F I	1	A A	A I	K	K	1	F	s	κ	26	3			
	Clostridium acetobutylicum	Ρ	L	Q	VL	. E	E	Q	Ν	K	G	M	M	AK	()	K	S	F	F	G	٧	R	S	268		
Hyperthermophilic	F Bacillus subtilis	Ρ	L	KI	R 1	Y G	-	Ε	ĸ	к	G	L	Ŀ		-	S	R	L	L	G	G	262	2			
Bacteria	Aquifex aeolicus	L	Е	NI	DF	= V	т	٧	S	к	G	LI	1 1	DT	L	K	D	F	F	s	к	L	ĸ	RG	27	1
	☐ Thermotoga maritima	Е	D	E		K I	1	R	K	Е	S	F	1	DK	(]	K	R	L	F	R	М	Y	263	3		
Archaea -	Methanocaldococcus jannaschii	Ρ	A	E	V	K E	K	κ	ĸ	Е	G	A	L	AK	(N	1 L	R	1	F	R	R	R	263	3		
	Archaeoglobus fulgidus	т	Ρ	PI	EF	PE	S	Ρ	v	K	R	1	F.		-	K	Α	L	F	G	G	Κ	R	264		
0	☐ Pvrococcus furiosus	L	D	NI	RH	K R	R	G	v	L	G	F			-	L	R	F	F	G	V	E	295	5		
Spirochetes -	Borrelia burgdorferi	T	E	1	A	ΕT	G	G	Ĺ	S	G	F	ί.		-	R	R	i.	F	G	R	E	W	E 3	04	
	Treponema pallidum	Ŷ	Ē	VI	N	E	T	G	N	ĸ	G	L	LI	KF	R V	0	Q	F	L	т	G	S	E	EN	V	286
Chloroplasts -	Mesostiama viride	P	s	P	s i	D S	A	P	S	R	G	w	F /	AA	À I	R	R	L	W	S	27	4	_			
	Nephroselmis olivacea		-						~			D	-							-						
												Put	ativ	e n	iem	ibra	ne									
												tarç	jetii	ng s	seq	uen	ce									

Chapter 16

<u>**16. Microbial Genomics- 2**</u> 11-17-2010

Functional genomics

Determination of how genome works from alignment of gene sequences paralogs – genes arose from gene duplication orthologs – genes very similar and are predicted to have same function involves analysis of translated amino acid sequence of presumed genes to understand protein structure and function motif, a short pattern of amino acids, may represent a functional unit within the protein, such as the active site

Physical map of the Haemophilus influenza genome

1st annotated genome Color-code genes 1/3 of unknown function (white regions) 65 regulatory genes (dark blue) > 700 complete bacterial and archaeal sequences now available



Reprinted with permission from Fleishchman, R.E. et al. 1995. Whole-Genome Random Sequencing and Assembly of Haemophilus Infl uenzae Rd., Science 269.496– 512, fr g. 1, page 507. © 1995 by the AAAS. Photo by the Institute for Genomic Research

Treponema pallidum-syphilis 梅毒



respiratory electron transport

56

First synthetic genome

BEREV ANDER VESS AND DER BUTTEN VERSCHWINDER DER STATISMEN DER STATISMEN UND DER VERSCHWINZEN UND DER STATISMEN

Mycoplasma mycoides JCVI-syn1.0



- 1.08 million bp synthesized chromosome of a modified *M. mycoides* genome at the J. C. Venter Institute (San Diego, CA, and Rockville, MD, USA)
- ~ 1 kb→ 10 kb in *E. coli* → 100 kb in yeast →assembly of 10 X ~100 kb fragments in yeast→ mycoplasma
- The new cells have expected phenotypic properties and are capable of <u>continuous</u> <u>self-replication</u> (>30 divisions) and <u>making</u> <u>new set of proteins</u>.



DNA microarray analysis (Gene Chips)

NET ANDER FOR ET MAN DER BETTE BETTER VON DIE VERSten Andere Bereichen B

- Determine gene expression at a specific time
 - spotted arrays
 - prepared by robotic application of <u>DNA probe</u>
 - each DNA probe represents a single gene or ORF
 - PCR product, cDNA or oligonucleotide
 - oligonucleotide probes from eukaryotes
 - expressed sequence tags (EST)

photolithography

Construction of DNA chip-photolithography



International and the second second

Monitoring changes of gene expression





Hierarchical cluster analysis of gene

expression

- Analysis of gene expression of *D*. *radiodurans* following exposure to γ-radiation
 - Each group of genes has been scored for relatedness
 - A tree has been generated with indication of <u>correlation</u> <u>coefficient</u> (r value)
 - Genes <u>induced</u> (red) or repressed (green) after radiation

ight © The McGraw-Hill Companies, Inc. Permission required for reproduction or

23	Time (h)	Gen	e#, putative function	Rati
.83	0.10v0194	A. recA-li	ike activation pattern	(fold)
		- DR0911	DNA-directed ma polymerase beta subunit, rpoC	1.99 (±1.37)
h	And and a second se	- Di 220	Tellurium resistance protein TerB	3.13 (±1.49)
		 DR2221 	Tellurium resistance protein TerE	5.24 (±2.94)
		- DRB0069	Storman	3.18 (±1.39)
	and the second s	= DR0261	Extracellular nuclease with Hibronectin III domains	4.37 (±1.21)
	the second se	- DBA0344	LEXA repressor HTH+protease lexA	1.80 (+1.08)
		- DR0099	SsDNA-binding protein, ssb	3.01 (±1.20)
F14r 🗖	and the second se	- DR2129	Ribosomal component L17, rp/Q	5.92 (±2.09)
		- DR2128	RNA polymerase alpha subunit, /poA	4.03 (±2.80)
	and the second se	- DR0324	Probable glutamate formiminotransferase	3.30 (±1.47)
	The second se	- DR2337	Uncharacterized protein	7.41 (±5.71)
	THE R. LEWIS CO., LANSING MICH.	- DR1835	PprA protein, involved in DNA damage resistance Paotein export membrane protein	3.52 (±1.94)
	A DECISION OF THE OWNER.	= DB1771	INRA ARC tamix ATPase unrA-1	3.52 (+1.15)
		- DRA0345	Predicted esterase	10.05(±4.39)
		- DR0422	Trans-aconitate methylase	18.85 (±7.46)
	and the second se	- DR1143	Uncharacterized protein	8.85 (±4.26)
		- DR0003	Uncharacterized protein	14.03 (±5.53)
	and the second division of the second divisio	- DR1776	Nudix family pyrophosphatase	4.70 (±2.83)
	and the second se	- DR2340	RecA, recA	7.98 (±3.86)
	and the second se	= DR1610	Teichola and biomethadia protein, march	4.13 (±1.67)
	and the second se	- DR0696	V-type ATPase synthese subunit K	7.19 (+2.14)
1110		- DR0421	Uncharacterized protein	4.94 (±2.30)
	and the second se	- DR1775	Superfamily Thelicase, uvrD	3.30 (±1.69
1114	and the second se	- DR1561	UDP-N-acetylglucosamine 2-epimerase, wecB	6.00 (±1.40)
	Statement of the local division of the local	- DR2285	MutY, A/G-specific adenine glycosylase, mutY	2.36 (±0.40)
	And Personal Property lies in the left	- DR2356	Nudix family hydrolase	3.35 (±0.45)
	Statement of the local division of the local	- DR2275	Excinuclease ABC subunit 8, wr8	4.93 (±1.81)
		- DR0206	Uncharacterized protein	5.45 (±2.65)
	DOLLAR STREET	- DR0204	Uncharacterized membrane protein	6.01 (±1.35)
	The second se	- DB0000	Exchanged and suburit G, MFG	3.78 (±0.42)
		- DR0205	ABC transporter ATPase	4.10 (+2.45)
		- DR1357	ABC transporter, permease subunit	6.79 (±2.56)
		- DR2482	Predicted transcription regulator	5.75 (±2.92)
		- DR2483	MorAnuolease	5.43 (±1.22)
		- DRA0008	Conserved membrane protein	6.60 (±2.00)
	and the second se	- DRA0234	Uncharacterized protein	12.76 (±5.27)
	and the second second	= DR31359	Record ransporter, perpaismic subunit	24.83 (±11.13)
		- DR1356	ABC transporter, ATP-binding protein	9.85 (+5.98)
_ _ _		- DR80136	Putative DEAH ATP-dependent helicase, hepA	5.22 (±0.45)
I JH-		- DR1548	Bacillus ykwD ortholog, PRP1 supertamily protein	5.62 (±2.35)
Աե		- DR0207	ComEA related protein, secreted	15.47 (±8.31)
		- DRA0249	Metalloproteinase, leishmanolysin-like	6.47 (±4.43)
	The second value of the se	- DR0665	Uncharacterized protein	11.66 (±5.74)
	The second se	 DR0596 DR0912 	Resolvasome RuvABC, subunit B, ruvB DNA-directed ma polymerase beta subunit, rpoB	3.22 (±1.31) 3.19 (±0.80)
71	E	B. Growth	n-related activation pattern	
-		- DR1172	Lea76/LEa29-like desiccation resistance protein	2.66 (±0.60)
		- DR0461	Bacillus yacB ortholog	2.58 (±0.81
		 DR1595 	6-phosphogluconate dehydrogenase, gnd	2.30 (±0.52)
-C	and the second se	 DRA0043 	TDP-rhamnose synthetase	5.08 (±2.12)
		- DRA0042	Glucose-1-phosphate thymidylyltransferase, r/bA	3.70 (±1.19)
		- DRA0031	Glucose-1-phosphate thymidylytransferase	2.48 (±1.64)
		 DRA0065 DR2263 	Chromosomai protein HU HupA, AupA Ractariofamilio, Inco chalation contain	7.71 (±2.07)
		- DRA0275	Soluble cytochrome C	4.80 (+1.22)
		- DR1279	Superoxide dismutase (Mn)	3.91 (±1.43)
77		C. Repres	ssed pattern	
		 DR1126 	RecJ like DHH superfamily Phosphohydrolase	0.33 (±0.12)
		 DR1337 	Transaldolase, tal	0.25 (±0.05)
1 <u>1</u>		 DR0728 	Fructokinase, cscK	0.37 (±0.13)
		 DR0977 DR1747 	Phosphoenolpyruvate carboxykinase, pokA	0.48 (±0.22)
		= 0H1742	Catalase, CATX, ketA	0.42 (±0.12)
		- DR1146	GSP26 general stress like protein	0.25 (+0.06)
LL LL-		- DR0493	Formamidopyrimidine-DNA glycosidase, mutW	0.46 (±0.09)
		- DR0674	Argininosuccinate synthase, ASSY, argG	0.35 (±0.15)
		DR2620	Cytochrome oxidase subunit I, COX1, caaA	0.45 (±0.25)
62	1 - 5			
0.2)		
		1		
		PTCIS		

Proteomics

- The study of the proteome (the entire collection of proteins)
 - often analyzed by two-dimensional gel electrophoresis
 - Functional proteomics to determine what is actually happening in cell proteome
 - Structural proteomics to resolve 3 dimensional structure of the proteins
 - protein modeling (on the basis of the assumption that proteins folds into a limited number of shapes)

Two-dimensional gel electrophoresis

Fig 16.14

63

Load a mixture of proteins onto an isoelectric focusing tube gel. pH 4.0 Proteins migrate until they reach the pH where their net charge is 0. At this point, a single band could contain two or more different proteins. pH 10.0 Lay the tube gel onto an SDS-PAGE gel and separate proteins according to their molecular mass. SDS-gel pH 4.0 pH 10.0 日日日 200 kDa 10 kDa (a) The technique of two-dimensional gel electrophoresis

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display



(b) An autoradiograph of a two-dimensional gel. Each protein is a discrete spot.

(b): © Tyne/Simon Fraser/Photo Researchers, Inc.

Tandem Mass Spectrometry

- unknown spot from 2-D gel is cut and cleaved
- fragments are analyzed by mass spectrometer
- mass of fragments is plotted
- protein tentatively identified from probable amino acid composition



Fig 16.15

ChIP-Chip analysis

- probing DNA-protein interaction
 - EMSA (electrophoretic mobility shift assay) or called gel mobility shift assay
 - ChIP- chromatin immunoprecipitation
 - search for protein binding DNA

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display



Microbial cells are grown under the conditions of interest.

- A cross-linking agent such as formaldehyde is added so that the protein is stably attached to the DNA.
- The cells are broken and the DNA is fragmented by sound waves, a process known as sonication.
- Antibodies are added and bind to the target protein. This increases the weight of the protein-DNA complex, so it can be selectively precipitated, a process known as immunoprecipitation.
- Cross-linking is reversed by heating, which also denatures the DNA.

The DNA is fluorescently labeled and hybridized to a microarray for identification.

Comparative Genomics-Archaea genome size



Comparative Genomics-Bacteria genome size

NUMBER NO.1	Copyright © The Me	Grav	v-Hill C	Compar	nies, Inc	. Permi	ssion re	equired	for rep	roductio	on or di	splay		
	Bacteria													
	Bradyrhizobium japonicum	-										9.11 (8401 orfs		
	Streptomyces coelicolor	_	8.67 (8319 orfs)											
	Pseudomonas aeruginosa		6.60 (5925 orfs)											
	Acarvochloris marina	_	6.50 (8557 orfs)											
	Bacteroides thetaiotaomicron		6.26 (4922 orfs)											
	Microcystis aeruginosa		5.84 (6366 orfs)											
	Escherichia coli 0157:H7		5.53 (5643 orfs)											
	Frankia sp.							5.43	(4670 c	orfs)				
	Bacillus anthracis	_	5.23 (5973 orfs)											
	Magnetospirillum magneticum	_	4.97 (4619 orfs)											
	Escherichia coli K12		4.64 (4612 orts)											
	Mycobacterium tuberculosis		4.40 (4340 01(5)											
	Clostridium difficile -		4.29 (3989 orfs)											
	Bacillus subtilis	_	4.21 (4408 orfs)											
	Caulobacter crescentus	_	4.02 (3867 orfs)											
	Geobacter sulfurreducens		3.81 (3587 orfs)											
	Mycobacterium leprae		3.27 (2749 orfs)											
	Deinococcus radiodurans		3.06 (3304 orfs)											
	Staphylococcus aureus		2.91 (2980 orfs)											
	Flavobacterium psychrophilum		2.86 (2505 orfs)											
	Synechococcus elongatus	_	2.70 (2590 orfs)											
	Sulfurovum sp.	_	2.56 (2524 orfs)											
	Thiomicrospira crunogena	2.43 (2307 orfs)												
	Chlorobium tepidum	_	2.15 (2369 orfs)											
	Lactobacillus acidophilus	_	1.99 (2035 orfs)											
	Francisella tularensis			1.8	39 (1967	7 orfs)								
	Haemophilus influenzae			1.8	3 (1774	orfs)								
	Campylobacter jejuni		1.78 (1990 orfs)											
	Prochlorococcus marinus	1.67 (2024 orfs)												
	Helicobacter pylori	-		1.60	(1628 c	orfs)								
	Aquifex aeolicus	1.55 (1663 orfs)												
	Borrelia burgdorferi	1.49 (1701 orfs)												
	Dichelobacter nodosus	-	_	1.40 (1	395 orf	s)								
	*Rickettsia conorii	_		1.27 (14	419 orfs	5)								
	*Ehrlichia chaffeensis	-	1	.18 (11	97 orfs)									
	*Chlamydia trachomatis	-	1.	04 (923	orfs)									
	*Mycoplasma genitalium	-	0.580	(564 o	rfs)									
			1	-	1	1	ż	1	-					
		0	1	2	3	4	5	6	7	8	9	10		

Genome size (Mb)

Comparative Genomics-Eukarya genome size



Core and Pan-genomes

- Core genome- set of genes found in all members of a species (or monophyletic group)
- Pan genome- combination of all different genes found in all the different strains in a given species
 - HGT (horizontal gene transfer)
 - Genomic island
 - Pathogenicity island
 - different GC contents
 - codon bias



Synteny

- Phylogenetic relationship analysis
- To examine how similar the organization of <u>orthologous genes</u> are in the compared genomes

Fig. 16.19





Comparative genomics analysis

- M. leprae- Leprosy
 - pseudogenes-degraded and nonfunctional genes
- *M. tuberculosis* (4.4 Mb) and *M. bovis* (4.3 Mb)-Tuberculosis
 - 99.5% sequence identity



Reverse Vaccinology

- Two types of reverse vaccinology
 Single species
 - genome is examined for vaccine targets
 - a pan-genomic approach



The Journal of Clinical Investigation (2009) 119: 2515-25.
Metagenomics- an environmental genomics

used to learn more about the <u>diversity</u> and <u>metabolic potential</u> of microbial communities

- shotgun sequencing DNA obtained directly from an environmental sample or series of related samples -> determine the presence and level of classes of genes
- may serve to establish hypotheses concerning interactions between community members
- being viewed as a baseline technology for understanding the ecology and evolution of microbial ecosystems
 - Phylochip

Construction and screening of genomic libraries directly from the environment





Microbiome- phylogenetic diversity of Sargasso sea microbes



Sargasso sea

 portion of the Atlantic ocean surrounds Bermuda

> The Sargasso Sea is well known for the floating mats of Sargassum weed





The human microbiome project

- Launched by the US National Institutes of Health (NIH) as one of its major roadmap initiatives (2007, earmarking ~US\$140 million)
 - to determine whether individuals share a core human microbiome
 - to understand whether changes in the human microbiome can be correlated with changes in human health
 - to develop the technological tools to support these goals
 - to address the ethical, legal and social implications of human microbiome research

Nature (2007) October 449:803-

Gut microbiome

The human intestine is home to ~100 trillion microorganisms of at least 400 species
10¹¹~10¹² /ml bacteria in colon
Many gastrointestinal (GI) diseases are expected to be associated with disruption of

host-bacterial interactions

PLOS PATHOGENS (2008)