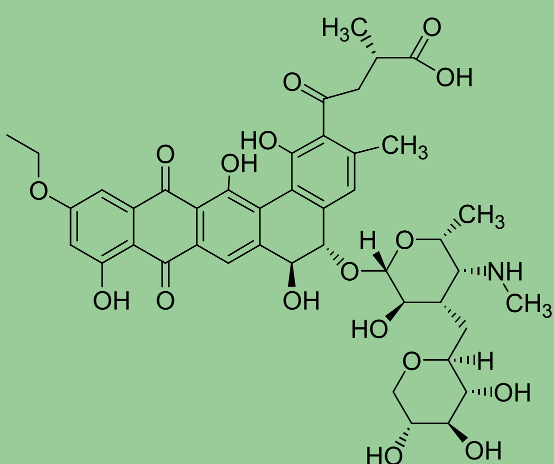
INTERNATIONAL COOPERATIVE BIODIVERSITY GROUPS
NIH NSF USDA

Bio-prospecting of uncultured microbes for novel antibiotics

Boris Wawrik, Ph.D.



THE STATE UNIVERSITY OF NEW JERSEY
RUTGERS
New Brunswick/Piscataway

Outline

- Why are novel antibiotics needed
- Emergence of antibiotic resistance
- Traditional screening approach
- Target based screening
- The promise of the uncultured
- Why is central Asia important ?
- Low tech screening
- The environmental “metagenome”
- How to make large insert libraries (BACs, Cosmids)
- Screening for genes of interest by probing and PCR
- Other potential targets in the metagenome

Why search for novel antibiotics ?

- Antibiotics were first used in chemotherapy in the 1940's
- 1950's – 2% of *S. pneumomiae* were tetracycline resistant
- 1967 – first penicillin resistant *Streptococcus* observed in Australia
- 1974 – first case in the USA
- 1998 – 34% of *S. pneumomiae* is resistant
- 1990's - >80% of *S. pneumomiae* were tetracycline resistant
- The same pattern is observed with almost every clinically important bacterial pathogen
- Since 1996 at least four patients in US hospitals have not responded to treatment by ANY antibiotic available, even Vancomycin

- Clinically important bacteria are becoming resistant to all commonly used antibiotics.
- Now more then 70% of bacteria in hospitals are resistant to at least one commonly used antibiotic
- Many strains are multiply resistant
- According to the W.H.O. the two most important reason for this are
  - over-prescription (mainly in rich countries)
  - Improper use (where people can not afford a full course of the drugs)
  - CDC reckons that in the US 1/3 of antibiotics prescriptions are unnecessary
- New resistances now emerge more quickly now than in the past
- **The rate of discovery of new antibiotics is slowing down !!!!**
- In the USA, 14,000 people die each year from drug-resistant infections picked up in hospitals

How have these pathogens become resistant ?

- Antibiotics are naturally occurring compounds produced by bacteria and fungi as chemical warfare agents – bacteria have had billions of years to adapt and develop resistance mechanisms

- Pathogenic bacteria acquire resistance mechanism by lateral gene transfer
  - Conjugation (one bacterium donates plasmid to another)
  - Transformation (bacterium picks up naked DNA from the environment)
  - Transduction (viral particle brings new genes)

- Most resistances are carried on R-plasmids (R=resistance)

- Each R-plasmid can carry multiple resistance elements

- Resistance elements are often found in transposable elements

- Resistance mechanisms include point mutations, efflux pump, enzymes that chemically change the antibiotic and make it inert

Properties of some R plasmids

Plasmid	Origin	Resistances	Size (kb)
RP1	England	CbKmTc	36
R527	Spain	CbCmGmKmSmSuTcHg	49
pMG5	Japan	AkKmSuTmBorHgPmrTer	280
pMG90	France	CbCmGmKmSmSuTcTmBorHg	150
Rms149	Germany	CbGmSmSuTra^{-}	36
pMG38	USA	CbGmKmSuTcTmHg	53
FP110	Australia	Cma*Pae*Fp110	60
pMG25	South Africa	CbCmGmKmSmSuTmBor	66
pMG69	Ireland	CbGmKmSmSuTcTmTra^{-}	47

Finding new antibiotics: The traditional approach is ‘random screening’

- Bacteria (mainly actinomycetes) are isolated and are grown and their activity spectrum is assessed
- Has NOT held pace with the rapidly emerging resistant bacteria in clinical settings over the past 20 years
- Advantages:
  - Selects for compounds that can penetrate cells
  - Compounds co-evolved with target – antimicrobial activity
  - Reproducible (you will find antibiotics producing actinomycetes in all soil)
- Disadvantages
  - Compounds are often rediscovered
  - Compounds are often toxic and non-specific
  - No rational basis for compound optimization
  - Spectrum unpredictable
  - Many different mechanisms

Two Nobel Prises for the discovery of antibiotics:

- First true antibiotic was discovered by Fleming by accident in 1928
  - penicillin – produced by the fungus *Penicillium notatum*

- First bacterial antibiotic was discovered by Waksman in 1944
  - Streptomycin – produced by *Streptomyces griseus*

Many more were discovered later -

- **Fungi**
  - Penicillin, Cephalosporin, Griseofulvin

- **Streptomyces**
  - Amphotericin B, Chloramphenicol, Erythromycin, Kanamycin, Neomycin, Nystatin, Rifampin, Streptomycin, Tetracyclin, Vancomycin, Gentamycin, Bacitracin, Polymyxin

The research dates back a little further

- **First research on this dates back to 1904 (Ehrlich and Hata)**
  - The screened chemicals for their toxicity against Syphilis
  - ‘compound 606’ = arsphenamine which is selectively toxic to Syphilis

- **Domagk at IG Farben discovered sulfonamides in 1934**
  - Screened dye chemicals for antibiotic activity
  - Discovered structural analogues of PABA (para-benzoic-acid-amine)
  - used in folic acid synthesis in bacteria -> selectively toxic (humans don’t make folic acid)

- **Approach was abandoned in favour of the random screening approach**
  - limited by the number of chemicals available

- **In recent years industry has moved towards combinatorial chemistry**
  - can generate a very large diversity of different chemical structures

Pharmaceutical companies also have moved to 'Target Based Screening'

- **A specific cellular process is targeted that is**
  - essential to bacterial growth
  - highly conserved in clinically relevant isolates
  - absent or radically different in man (comparative genome analysis)

- **Advantages**
  - Rational approach – target known
  - Very sensitive
  - Easy screening
  - Accepts combinatorial chemistry libraries in screen

- **Disadvantages**
  - molecules are not co-evolved
  - *in vitro* activity does not always translate into *in vivo* activity
  - Costly and requires genome sequencing and analysis

Prokaryotic Microorganisms Associate Program (PMAP)

- Why still target microbes in the search for antibiotics ? (has been done for more than 50 years)
  - Prokaryotes are a well known source of bioactive molecules
  - Over 17,000 active molecules have been described from prokaryotes
  - Only 1% of bacteria in the environment have been cultured
  - The remaining 99% of uncultured bacteria are thought to make many more novel secondary metabolites
  - Molecular tools are now available to access to genes of hard to culture environmental microbes

Why Central Asia ?

 Uzbekistan  Kyrgyzstan

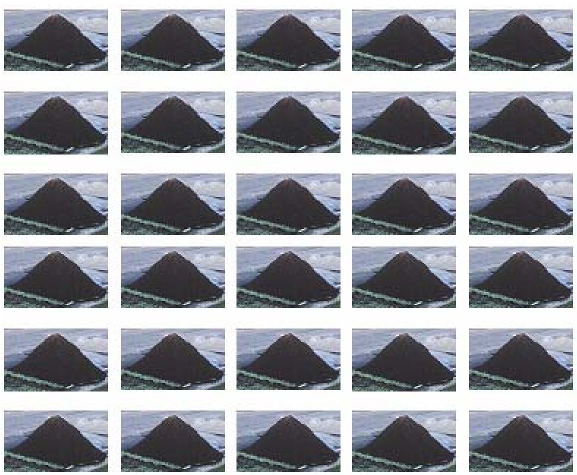
- It is assumed that different species can be found in different habitats
- Central Asia is a very unique place
- It has not been well described with respect to it’s bacterial diversity
- High bio-diversity, mixed topography and divergent habitats
- Environments include
  - desert
  - steppe
  - semiarid grassy plains
  - Aral Sea
  - Subtropical valley systems
  - Arctic in the high mountains
- -> presumably high bacterial diversity; hopefully many new species

The I.C.B.G approach to bio-prospecting bacteria

- 1st Rational approach for choosing a soil sample to work with
  - Soil with high diversity
  - Soil with unique communities
  - Soil with uncultured species
  - Soil with new secondary metabolite gene pathways
- 2nd Two approaches to antibiotics discovery:
  - Traditional screening (in a not so traditional way)
  - Molecular tools to describe the diversity of secondary metabolite pathways in a sample – Cloning these pathways for expression in other bacteria.

The average soil sample contains 10,000 different bacteria and different soil samples contain different communities

How do we rationally select a soil sample for analysis ?

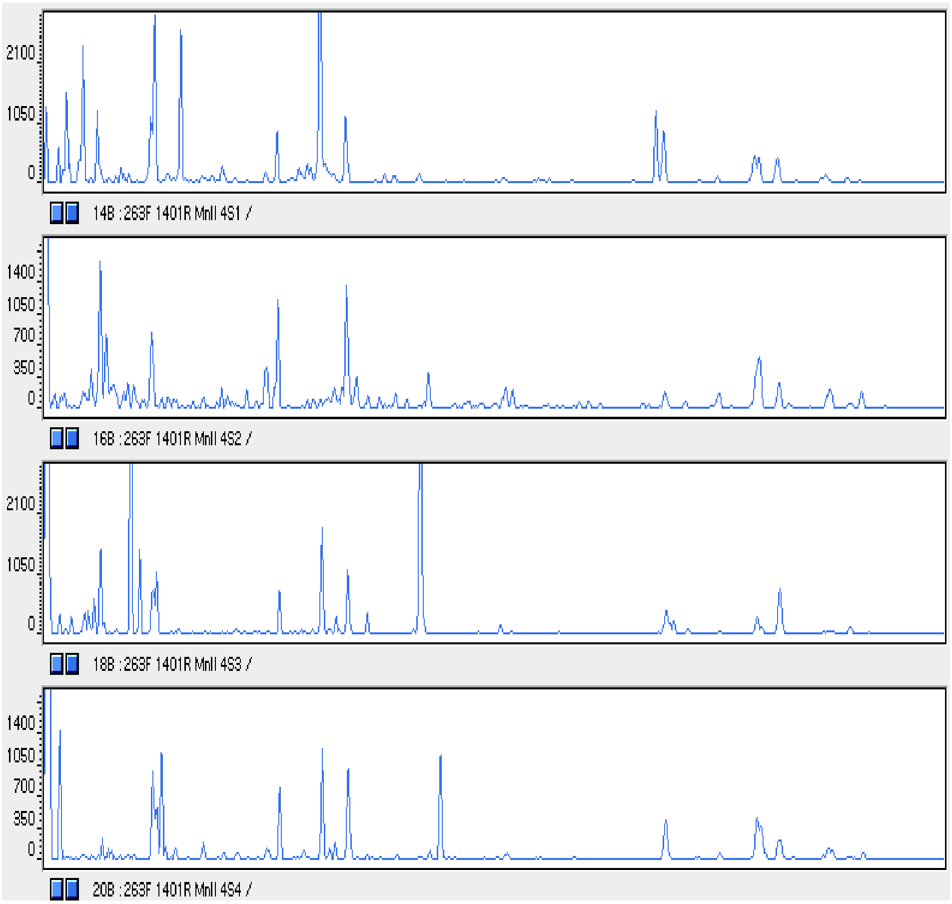


How do we find soil samples that contain unique and novel actinomycetes ?

How do we find soil samples that contain novel antibiotics genes?

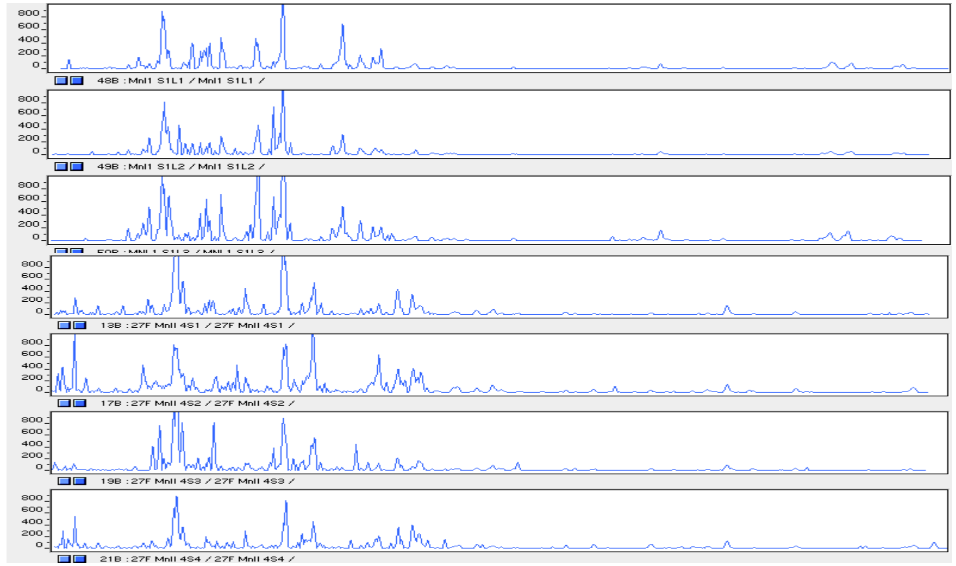
Which samples should be used for isolation and metagenomic analysis?

TRFLP analysis of soil communities

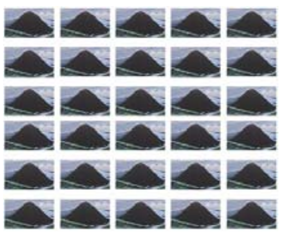


- Mixed population is amplified using primers for
  - Actinomycete 16S rRNA genes
  - Polyketide synthase genes (PKS)
- PCR product is cut with a 4bp cutting restriction endonuclease
- Different sequences will give different length fragments
- 16S tells give us a fingerprint of the actinomycete diversity
- PKS gives us fingerprint of antibiotics production in the soil

Actinomycete
16S PCR

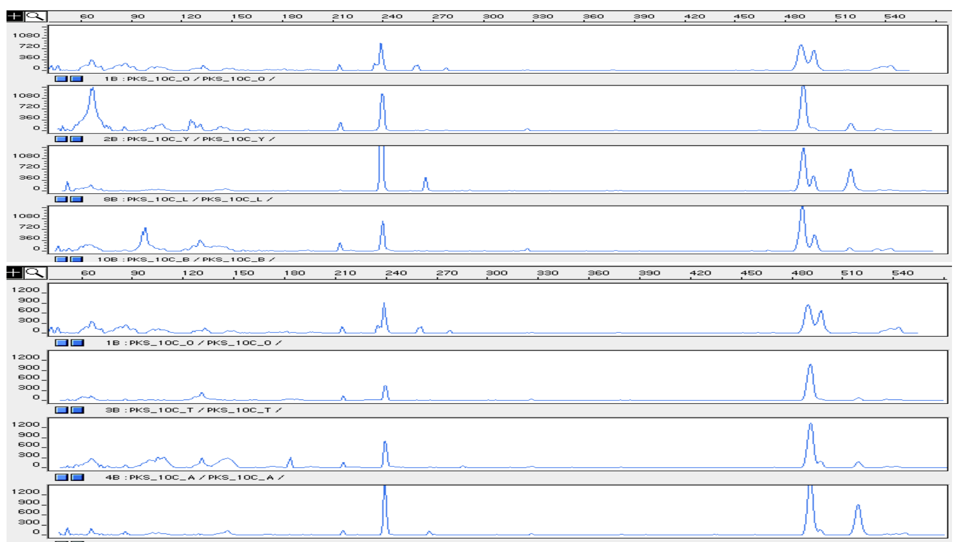


Identify samples with unique actinomycete communities



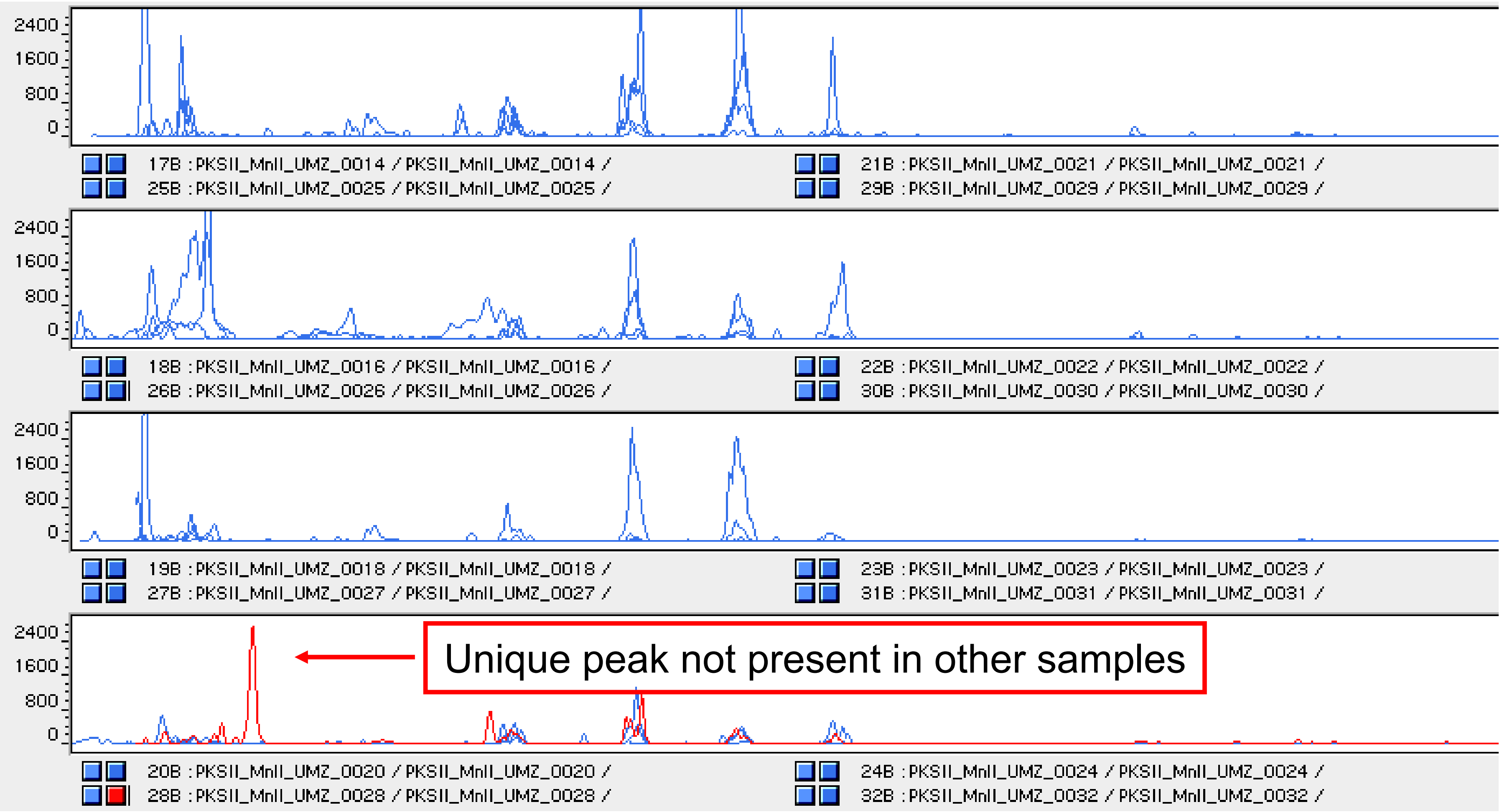
(many soil samples)

Polyketide synthase gene PCR

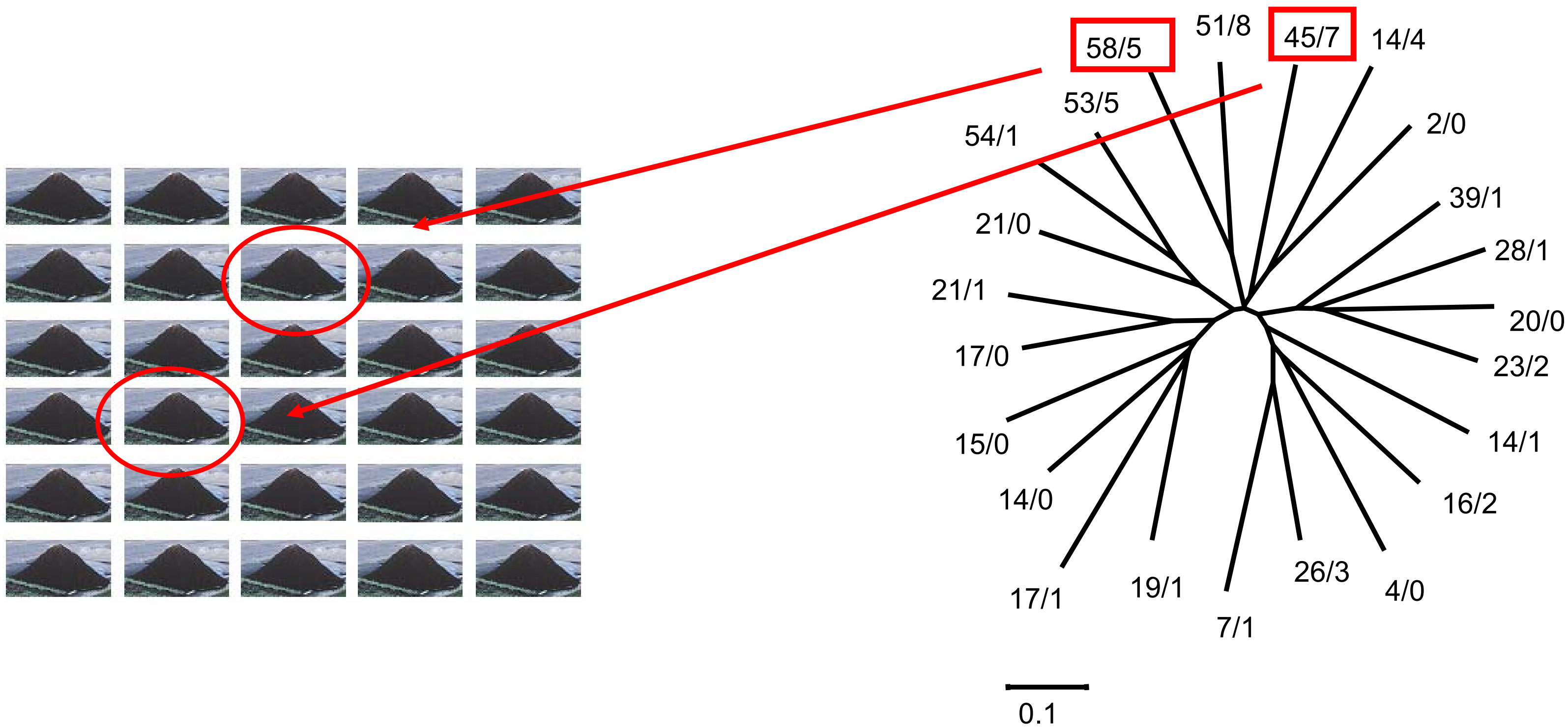


Identify samples with unique PKS signatures

Example of computational analysis



An example of Actinomycete community analysis

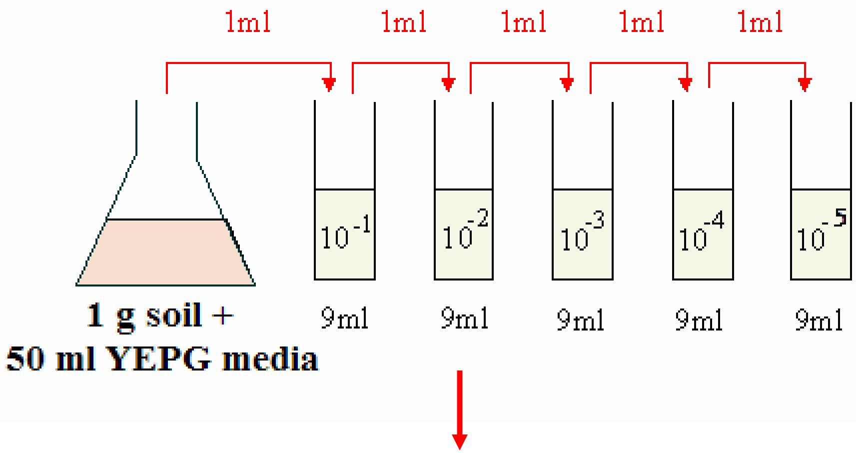


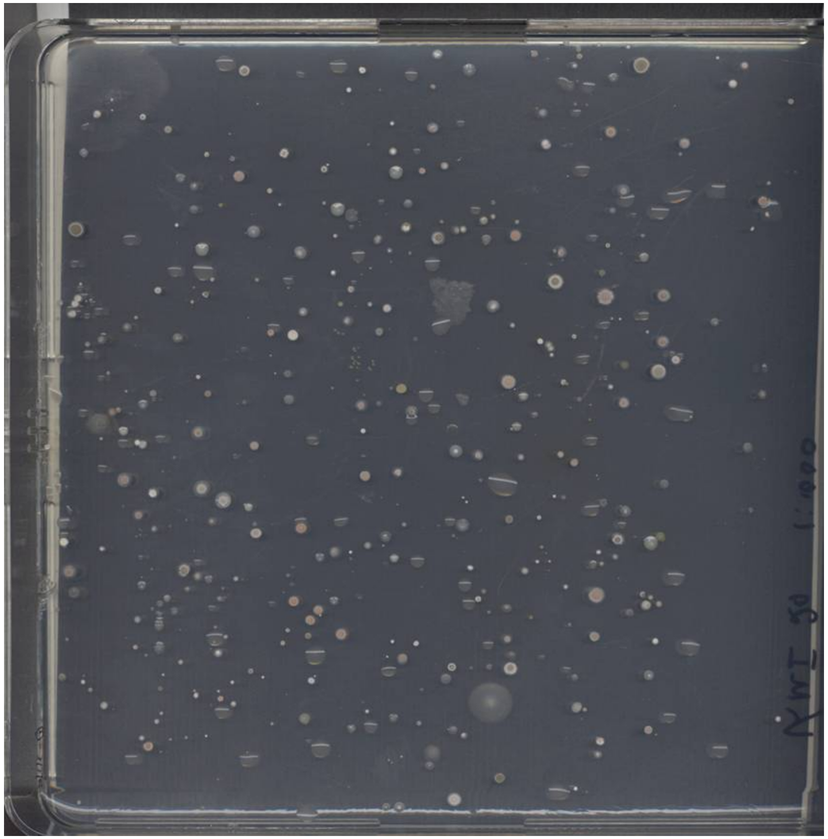
total peaks / unique peaks

Isolation of antibiotics producing actinomycetes from soil

- Traditionally isolation is performed by picking one colony at time
  - Time consuming
  - Labor intensive
  - Inefficient because antibiotics production is rare
- How do screen a large number of colony forming units for activity ?
  - without isolating thousands of individual strains
  - cheaply
  - screening the largest possible diversity of microbes
  - maybe find novel species that produce antibiotics

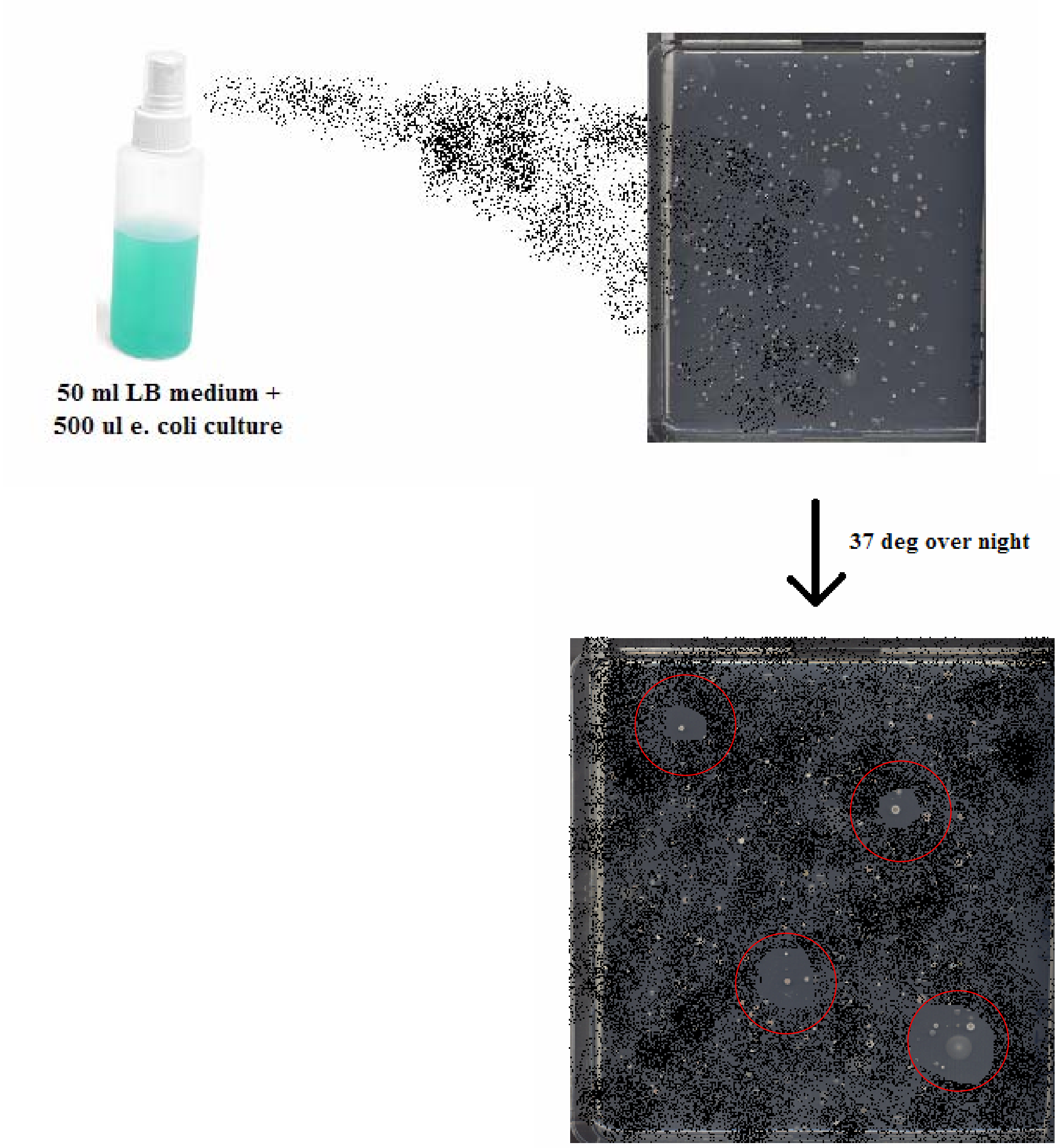
The traditional approach





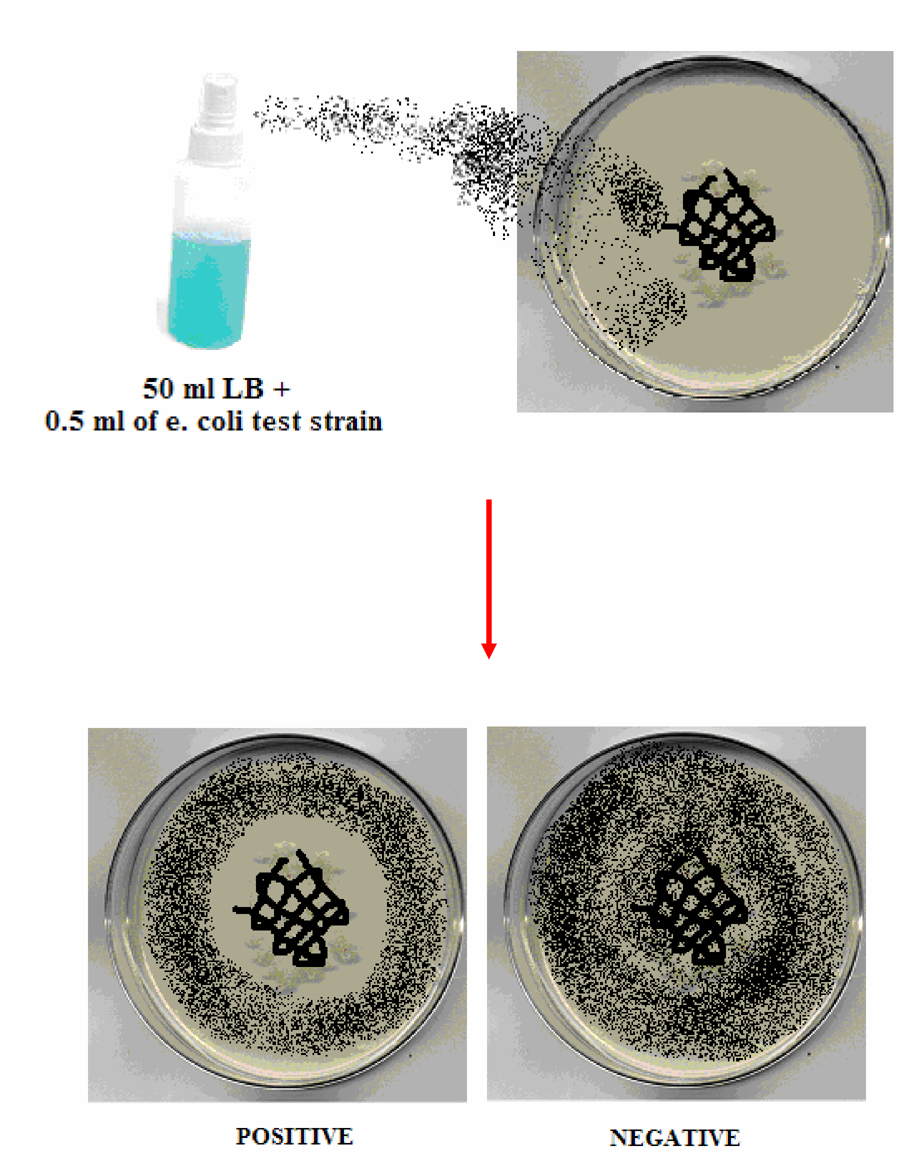
500-1000 colonies

The not so traditional, traditional screen



Isolate colonies that inhibit the growth of *E. coli*

Confirmation of positives / screening against other tester strains



Challenge-Strains currently used by us

- *E. coli* J53-AZ
  - azide resistant

- *E. coli* J53-pCFF34
  - Ak Km Nm Su Tc Tm Bla

- *E. coli* J53-pMG223
  - Ak Km Nm Su Tc Tm Bla Cm Gm Sm Sp Tp

Ak = amikacin; Cm = chloramphenicol; Km = kanamycin; Nm = neomycin; Sm = streptomycin; Sp= spectinomycin; Su = sulfonamide; Tc = tetracycline; Tm = tobramycin; Tp = trimethoprim.

Both strains make > extended-spectrum beta-lactamases (ESBLs) active on > most beta-lactam antibiotics except carbapenems and > cephamycins.

E.coli is intrinsically resistant to > erythromycin and vancomycin.

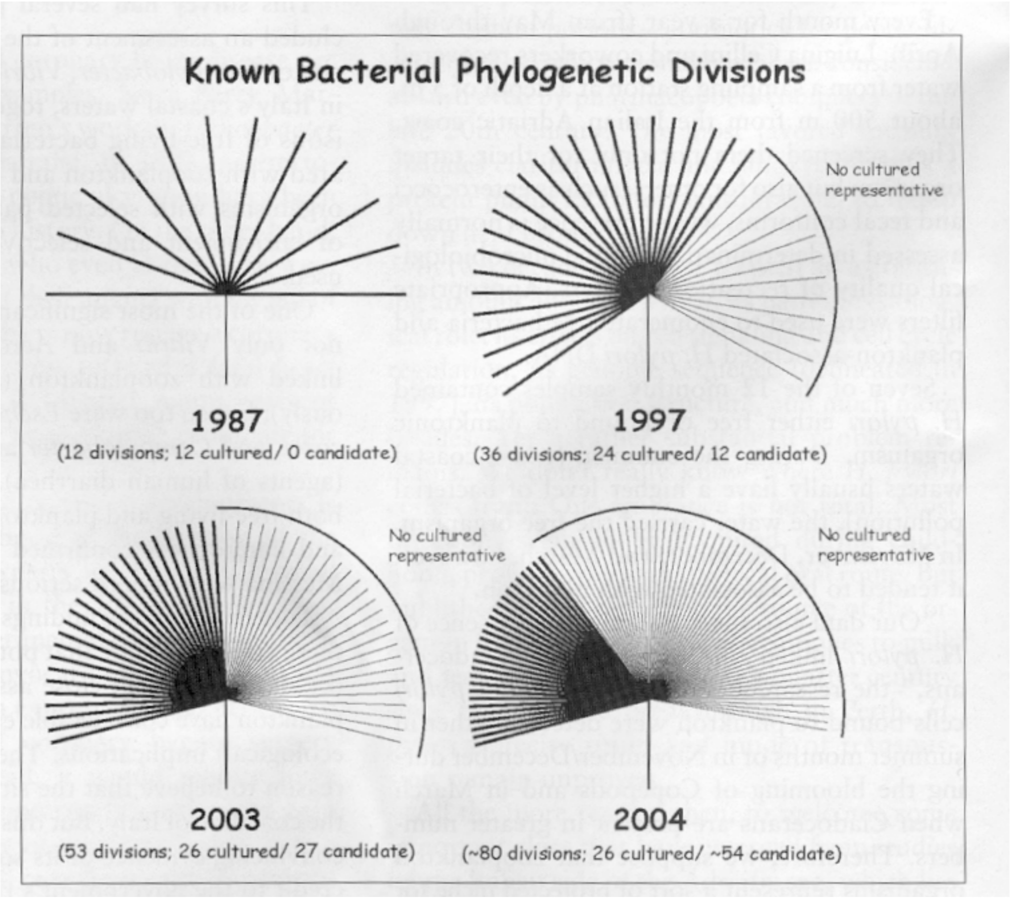
Antibiotics Screen Summary Table

Date	Soil sample	Media type	colonies	Assay strain	positives	2nd screen	positives	3rd screen	positives	total positive
2/15/2005	KMI-00020	YEP-G	1200	*e. coli*	10	*e.coli* PCFF34	2	*e. coli* PMG223	3	2
3/30/2005	KMI-00001	YEP-G	165	*e. coli*	4	*e.coli* PCFF34	0	*e. coli* PMG223	0	0
3/30/2005	KMI-00001	YEP-P	160	*e. coli*	3	*e.coli* PCFF34	0	*e. coli* PMG223	1	0
3/30/2005	KMI-00001	YEP-L	136	*e. coli*	3	*e.coli* PCFF34	2	*e. coli* PMG223	1	1
3/30/2005	KMI-00001	YEP-B	108	*e. coli*	1	*e.coli* PCFF34	1	*e. coli* PMG223	1	1
SUM			1769		21		5		6	4

pCFF34 Ak Km Nm Su Tc Tm Bla
pMG223 Ak Km Nm Su Tc Tm Bla Tp Cm Gm Sm Sp

Ak = amikacin; Cm = chloramphenicol; Km = kanamycin; Nm = neomycin; Sm = streptomycin; Sp= spectinomycin; Su = sulfonamide; Tc = tetracycline; Tm = tobramycin; Tp = trimethoprim.
Both strains make > extended-spectrum beta-lactamases (ESBLs) active on > most beta-lactam antibiotics except carbapenems and > cephamycins.
E.coli is intrinsically resistant to erythromycin and vancomycin.

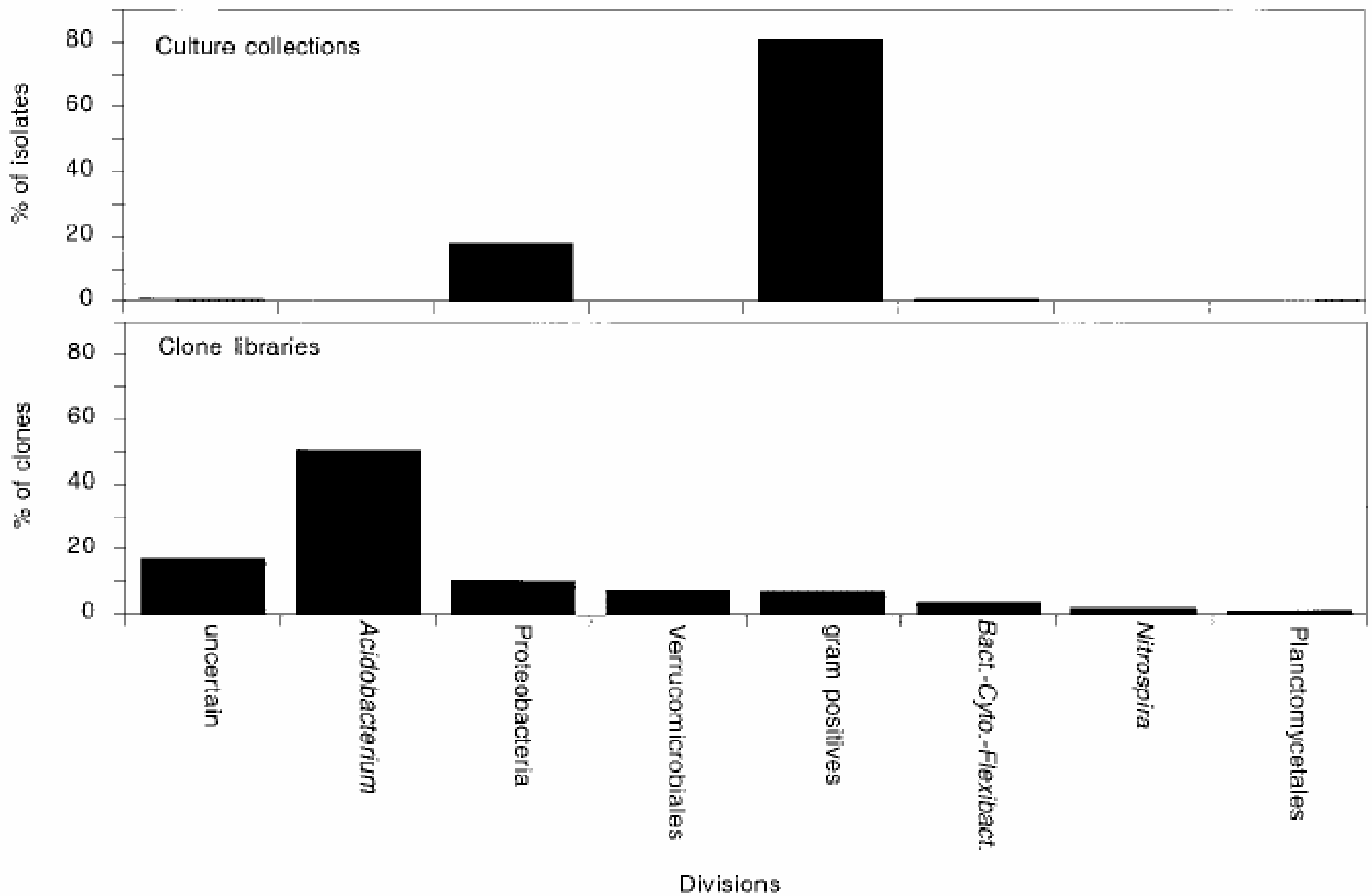
Of the 10,000 bacteria in soil, the majority has not been cultured



J. L. Fox, ASM News – January 2005

- 1987
  - 12 divisions known
  - 12 cultured divisions
  - 0 uncultured
- 2004
  - 80 divisions known
  - 26 cultured members
  - 54 uncultured

When you culture using traditional methods you do not find the diversity you see using molecular tools :



Dunbar et al., AEM No. 4, Vol 65, 1999, pp. 1662-69

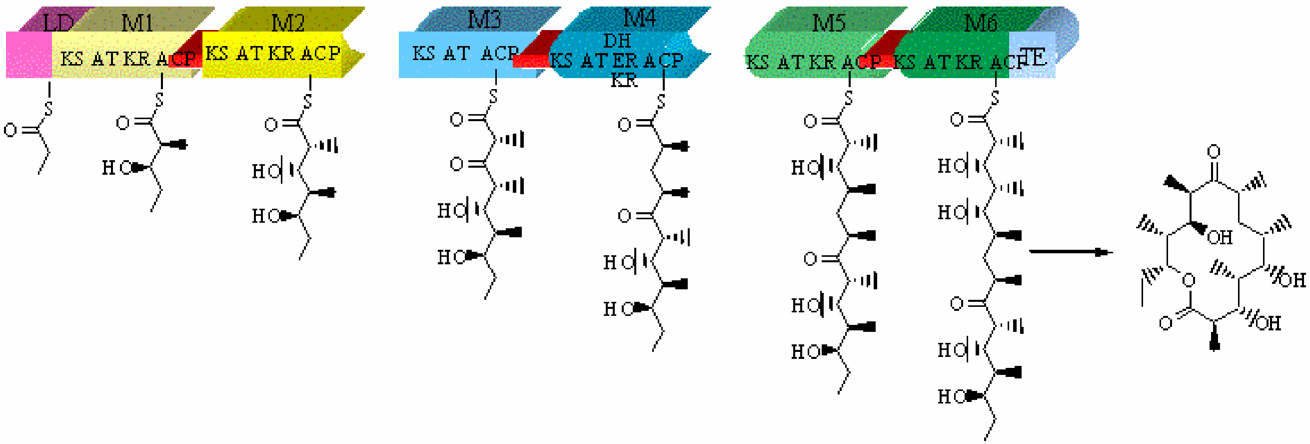
Molecular screening methods

- Is there a way to screen a soil sample for the presence of novel antibiotics without culturing or screening thousands of isolates ?
  - PCR allows us to search for the presence of known genes
  - Degenerate PCR primers allow us to amplify genes from bacteria that have not been cultured
  - Most clinically important antibiotics in actinomycetes are made by polyketide synthases (PKS genes)
  - We can make PCR primers for PKS genes and amplify DNA from soil
  - Our goal is to find PKS gene sequences that are not in the database

A large number of secondary metabolites are synthesized by polyketide synthases

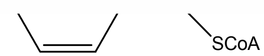
- In the environment polyketides are often involved in
  - self-defense
  - aggression (chemical warfare)
  - communication
  - development
- Polyketides are pharmacologically important as
  - anti-bacterial drugs (e.g. erythromycin)
  - anti-fungal drugs (e.g. simocyclinone)
  - Cholesterol control (e.g. lovostatin)
  - tumor treatment (e.g. avermectin)
  - immunosuppresants (e.g. rapamycin)

Modular Type I PKS

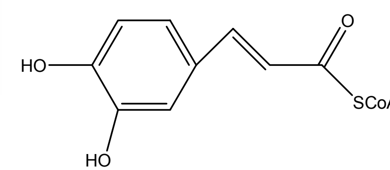


acyl-transferases (AT)
acyl carrier proteins (ACP)
b-keto-acyl synthases (KS)
b-keto reductases (KR)
dehydratases (DH
enoyl reductases (ER)
thiolesterase (TE)

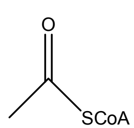
Examples of chain building units :



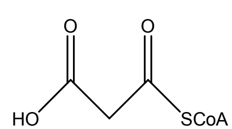
Benzoyl CoA



Cafeoyl CoA

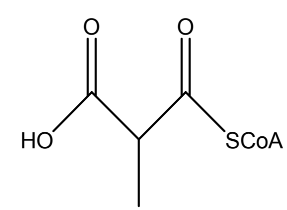


Acetyl CoA



Malonyl CoA

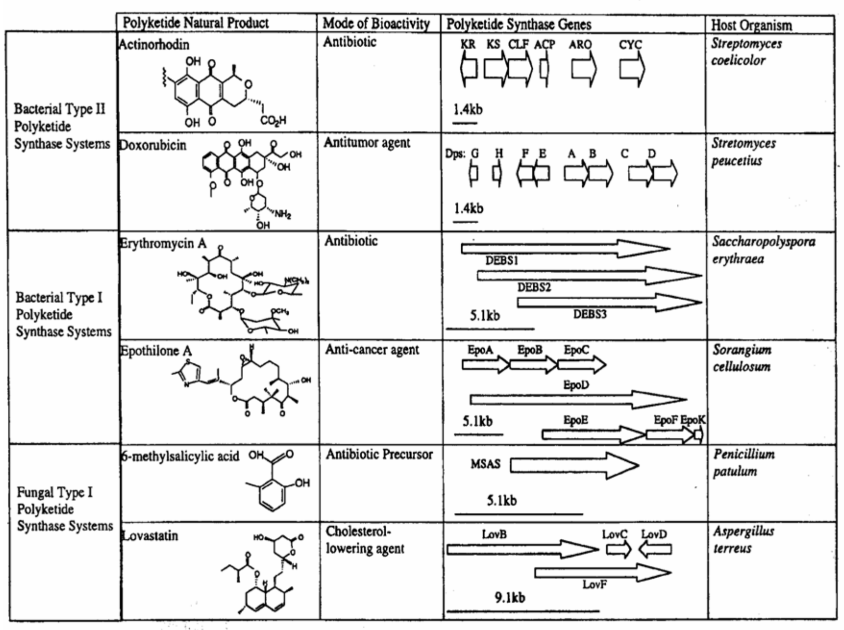
Octanoyl CoA



Methylmalonyl CoA

Types of PKS genes

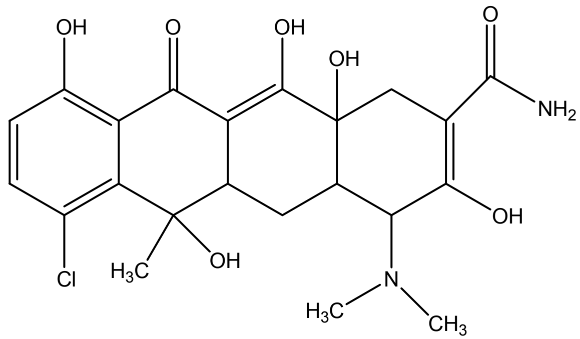
- Type I (bacteria)
  - Large multifunctional enzyme
  - Analogous to vertebrate fatty acid synthesis
  - Multi-domain - Carrying all the active sites needed to produce final product
- Type II (bacteria)
  - Multiple mono-functional enzymes
- Type III (plants and bacteria)
  - chalcones and stilbenes in plants
  - polyhydroxy phenols in bacteria



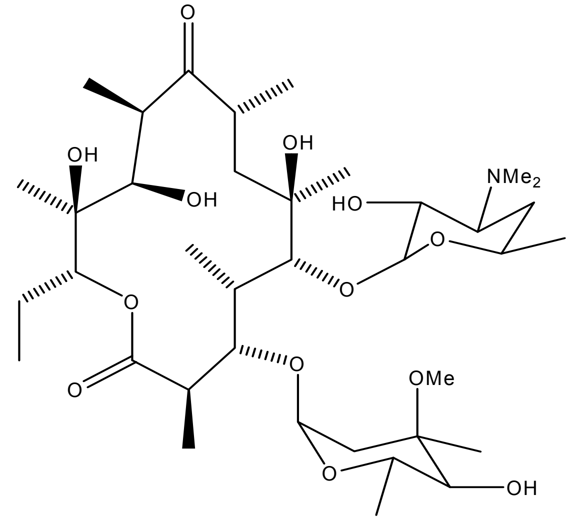
	Polyketide Natural Product	Mode of Bioactivity	Polyketide Synthase Genes	Host Organism
Bacterial Type II Polyketide Synthase Systems	Actinorhodin	Antibiotic	KR KS CLF ACP ARO CYC; 1.4kb	*Streptomyces coelicolor*
	Doxorubicin	Antitumor agent	Dps: G H F E A B C D; 1.4kb	*Stretomyces peucetius*
Bacterial Type I Polyketide Synthase Systems	Erythromycin A	Antibiotic	DEBS1 DEBS2 DEBS3; 5.1kb	*Saccharopolyspora erythraea*
	Epothilone A	Anti-cancer agent	EpoA EpoB EpoC EpoD EpoE EpoF EpoK; 5.1kb	*Sorangium cellulosum*
Fungal Type I Polyketide Synthase Systems	6-methylsalicylic acid	Antibiotic Precursor	MSAS; 5.1kb	*Penicillium patulum*
	Lovastatin	Cholesterol-lowering agent	LovB LovC LovD LovF; 9.1kb	*Aspergillus terreus*

Pfeiffer et Khosla, Microb. Molec. Biol. Rev., Mar. 2001 p.106-118

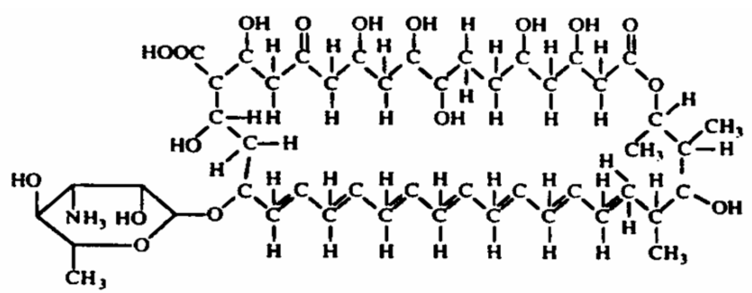
Examples of polyketide antibiotics



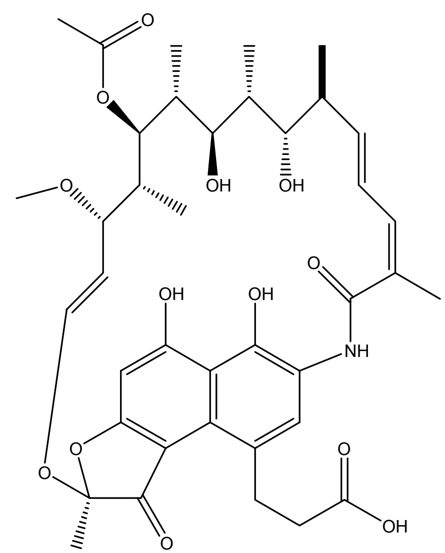
(A) tetracycline



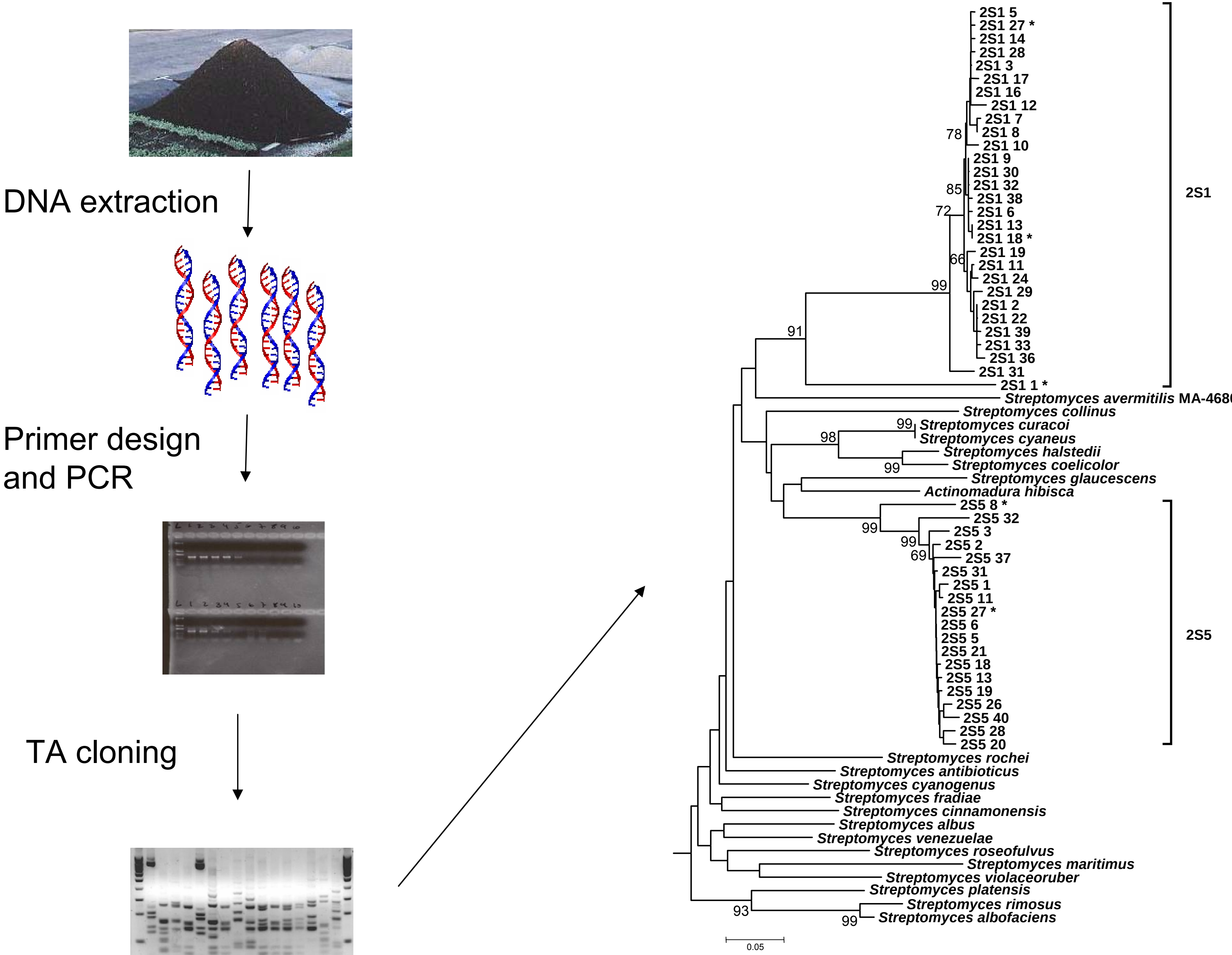
(B) erythromycin

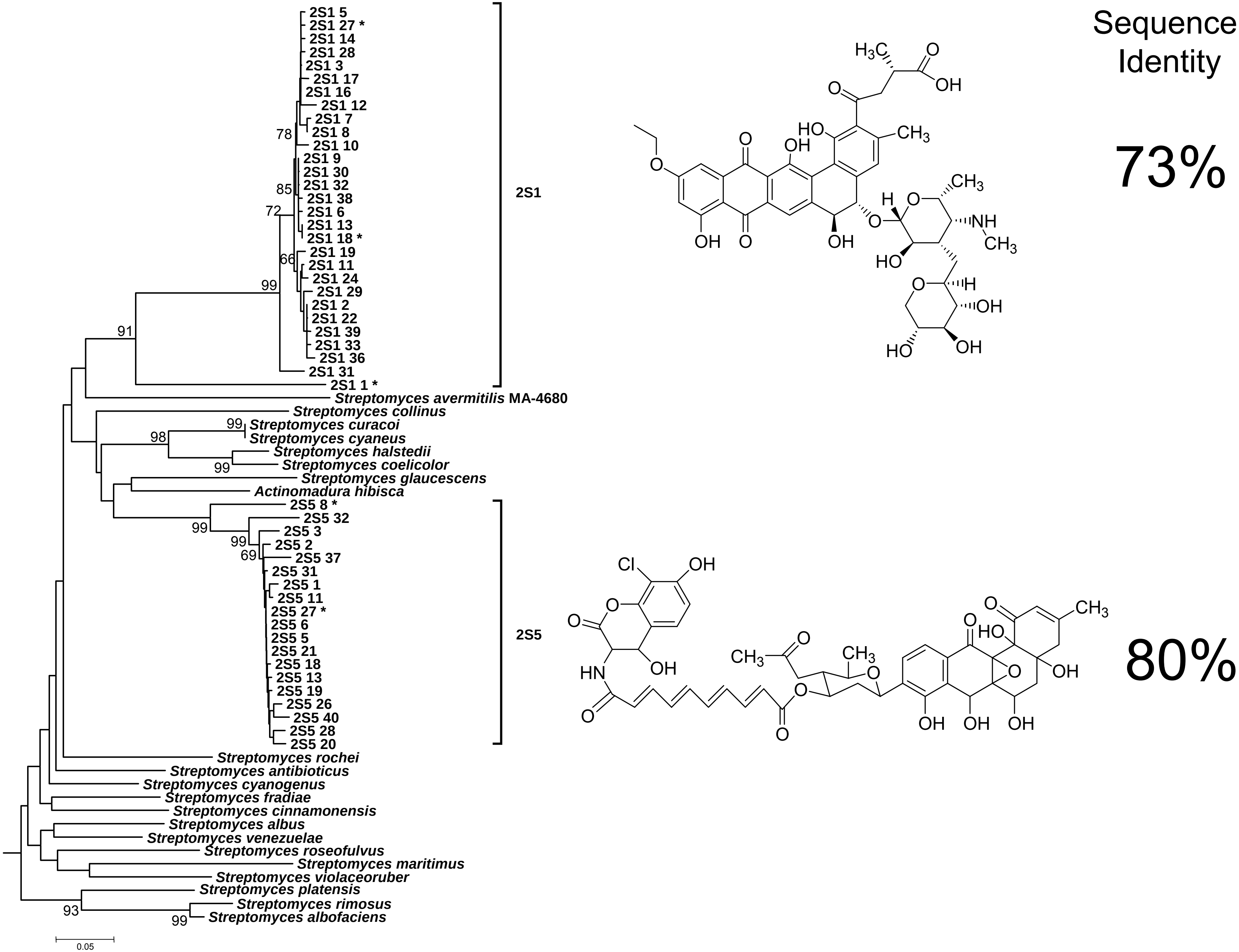


(C) Amphotericin B



(D) Rifamycin B


DNA extraction
Primer design
and PCR
TA cloning
2S1 5
2S1 27 *
2S1 14
2S1 28
2S1 3
2S1 17
2S1 16
2S1 12
2S1 7
2S1 8
2S1 10
2S1 9
2S1 30
2S1 32
2S1 38
2S1 6
2S1 13
2S1 18 *
2S1 19
2S1 11
2S1 24
2S1 29
2S1 2
2S1 22
2S1 39
2S1 33
2S1 36
2S1 31
2S1 1 *
2S1
Streptomyces avermitilis MA-4680
Streptomyces collinus
Streptomyces curacoi
Streptomyces cyaneus
Streptomyces halstedii
Streptomyces coelicolor
Streptomyces glaucescens
Actinomadura hibisca
2S5 8 *
2S5 32
2S5 3
2S5 2
2S5 37
2S5 31
2S5 1
2S5 11
2S5 27 *
2S5 6
2S5 5
2S5 21
2S5 18
2S5 13
2S5 19
2S5 26
2S5 40
2S5 28
2S5 20
2S5
Streptomyces rochei
Streptomyces antibioticus
Streptomyces cyanogenus
Streptomyces fradiae
Streptomyces cinnamonensis
Streptomyces albus
Streptomyces venezuelae
Streptomyces roseofulvus
Streptomyces maritimus
Streptomyces violaceoruber
Streptomyces platensis
Streptomyces rimosus
Streptomyces albofaciens
78
85
72
66
99
91
99
98
99
99
99
69
93
99
0.05

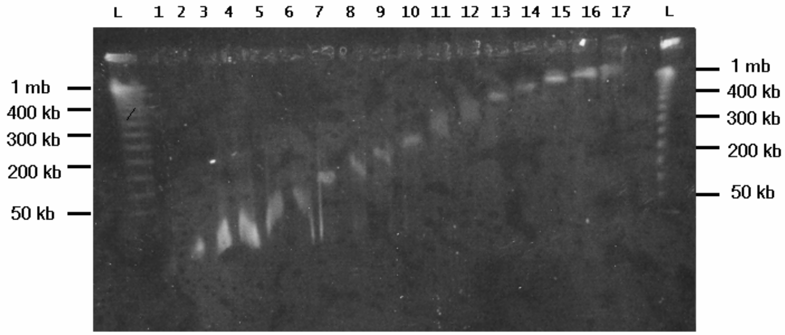

Sequence Identity
73%
80%
2S1
2S5
2S1 5
2S1 27 *
2S1 14
2S1 28
2S1 3
2S1 17
2S1 16
2S1 12
2S1 7
2S1 8
2S1 10
2S1 9
2S1 30
2S1 32
2S1 38
2S1 6
2S1 13
2S1 18 *
2S1 19
2S1 11
2S1 24
2S1 29
2S1 2
2S1 22
2S1 39
2S1 33
2S1 36
2S1 31
2S1 1 *
Streptomyces avermitilis MA-4680
Streptomyces collinus
Streptomyces curacoi
Streptomyces cyaneus
Streptomyces halstedii
Streptomyces coelicolor
Streptomyces glaucescens
Actinomadura hibisca
2S5 8 *
2S5 32
2S5 3
2S5 2
2S5 37
2S5 31
2S5 1
2S5 11
2S5 27 *
2S5 6
2S5 5
2S5 21
2S5 18
2S5 13
2S5 19
2S5 26
2S5 40
2S5 28
2S5 20
Streptomyces rochei
Streptomyces antibioticus
Streptomyces cyanogenus
Streptomyces fradiae
Streptomyces cinnamonensis
Streptomyces albus
Streptomyces venezuelae
Streptomyces roseofulvus
Streptomyces maritimus
Streptomyces violaceoruber
Streptomyces platensis
Streptomyces rimosus
Streptomyces albofaciens
0.05

What are large insert libraries and why ?

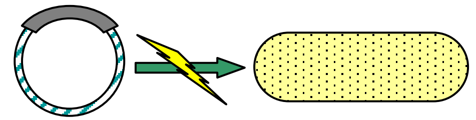
- PCR based approaches are always biased, mainly because primers preferentially amplify some sequences over other
- PCR primers can only be designed based on sequences that are already known
- Large insert libraries are generated by randomly cloning DNA from the environment
  - Large inserts >40,000 base pairs
  - 10,000 colonies are picked and screened using robots
  - Average bacterial genome = 4-8 million base pairs
  - 10,000 * 40,000 = 400,000,000 base pairs / 4 mb = 100 bacterial genomes are contained in a library of 10,000 clones
- Cloning vectors vary on application and insert size required
  - BAC 10-200 kb
  - Cosmid 35-45 kb

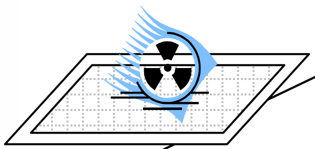
How do we gain access to all those uncultured microbes ?

We can extract HMW DNA

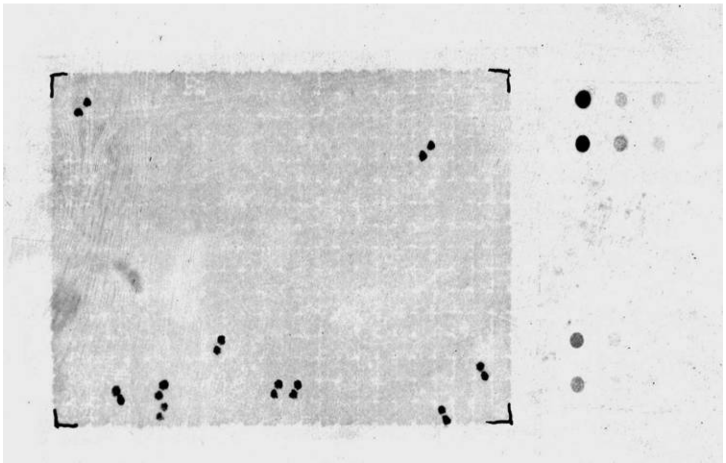


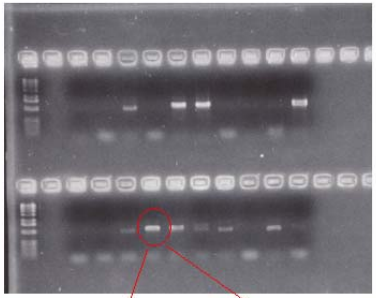
Clone into large insert (BAC or cosmid) vector





Probe for known types of genes





Screen for known genes by PCR

Plate 24 combined ROWS

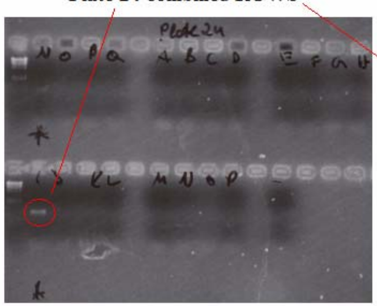
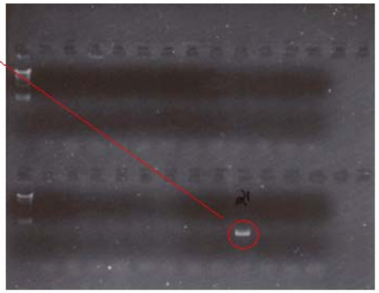
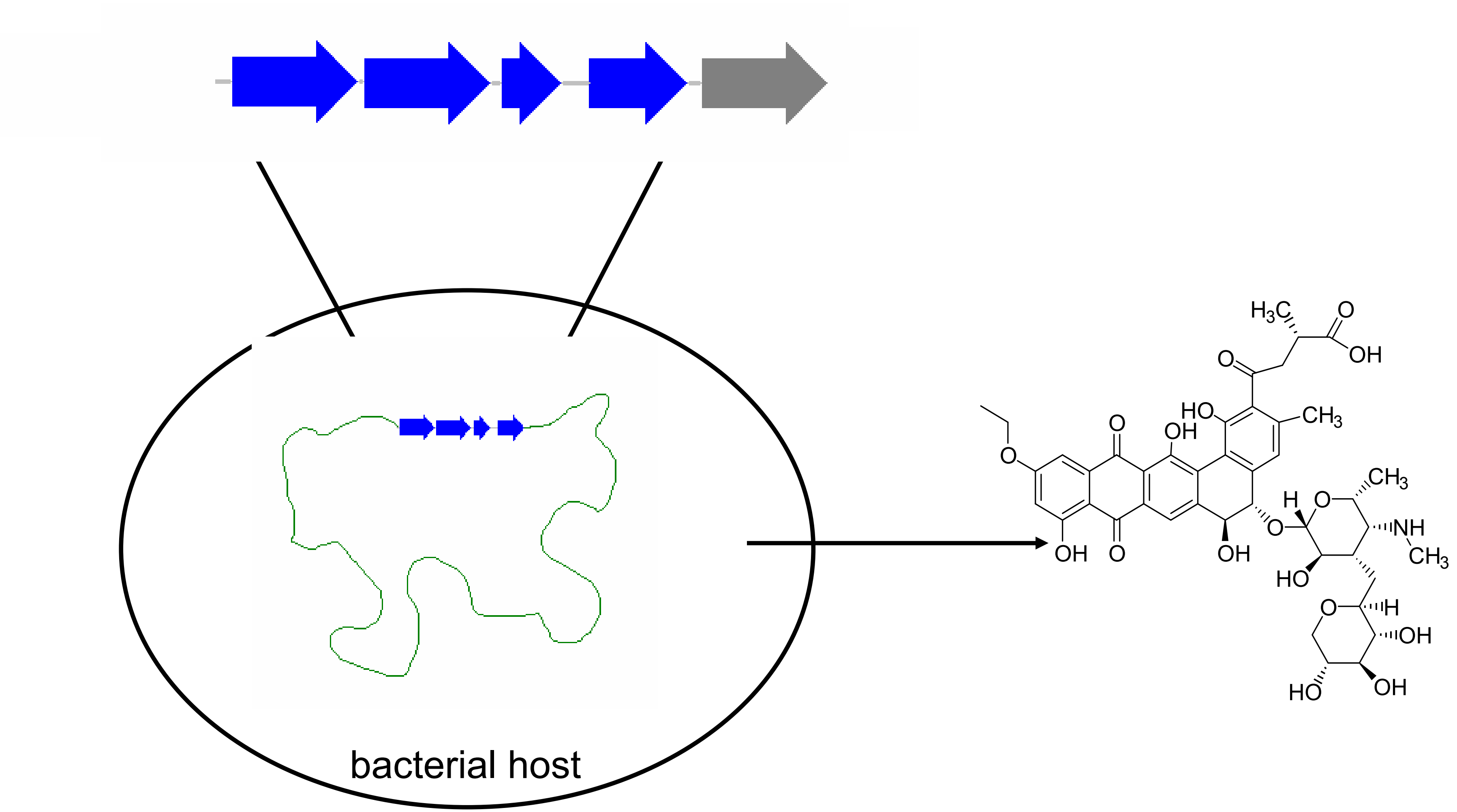


Plate 24 combined COLUMNS



The PKS gene cluster can then (maybe) be expressed in another host for secondary metabolite production



International Cooperative Biodiversity Group

United States Department of Agriculture

National Institute of Health

National Science Foundation

Rutgers University

University of Illinois