# Overview of Molecular Subtyping of Methods for Bacterial Pathogens

Christine Keys

The U.S. Food & Drug Administration
Center for Food Safety and Applied Nutrition
Office of Regulatory Science

# Acknowledgements

Dr. Rebecca Bell (IBIS T5000)
Dr. Jie Zheng (xMAP Technology)
Dr. John Callahan (MALDI)
Alice E Hayford (Pyrosequencing)

Dr. Eric Brown Branch Chief, Molecular Methods and Subtyping Branch

Christine.Keys@FDA.HHS.GOV

# Objectives

- Methods in Use at the FDA
- Future Methods – SNP-Based
- Mass Spectrometry-Based

# How much Discrimination is the Right Amount of Discrimination?

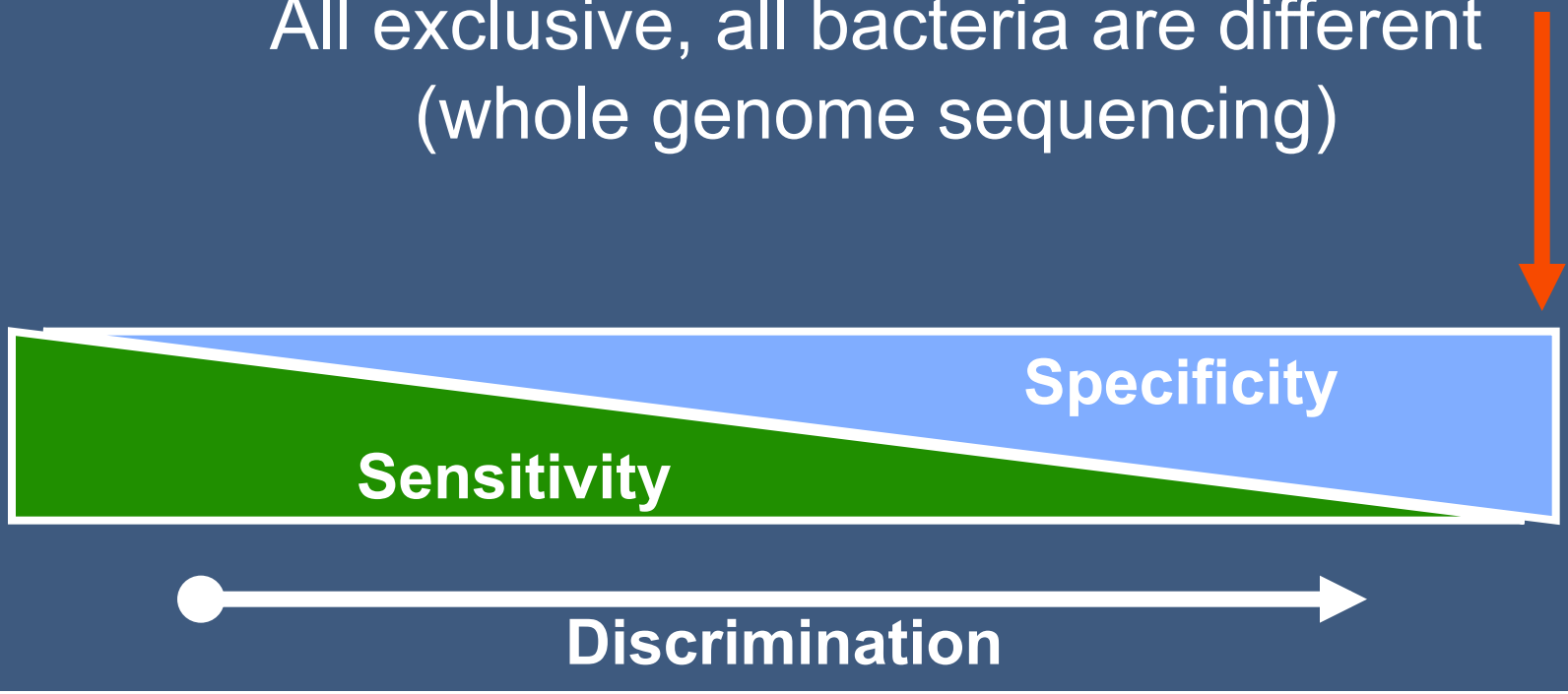
## It Depends on the Question You are Asking

All inclusive, but bacteria are not identified at all

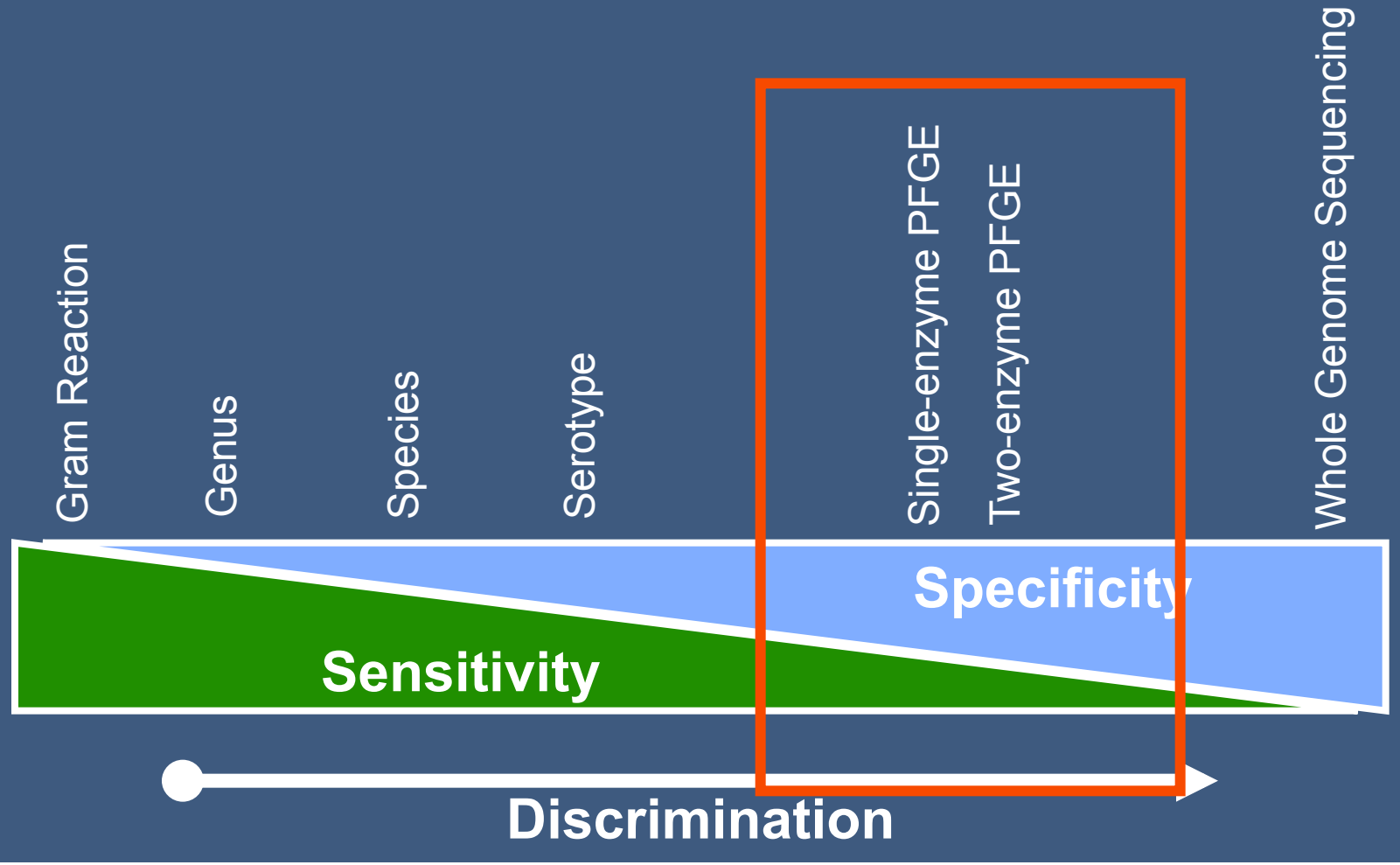
# How much Discrimination is the Right Amount of Discrimination?

## It Depends on the Question You are Asking

All exclusive, all bacteria are different
(whole genome sequencing)

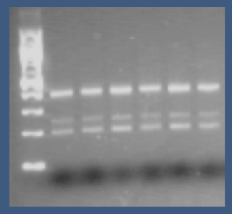


# How much Discrimination is the Right Amount of Discrimination?

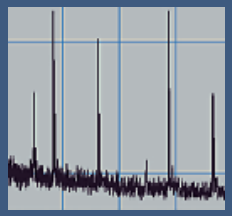


# Information vs. Throughput

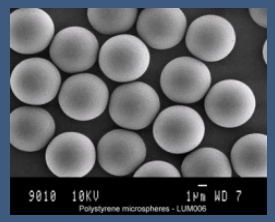
microarray



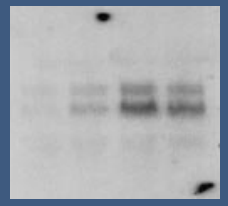
multiplex RT-PCR



Mass Spec




xMAP Technology



northern blot

Real-Time PCR

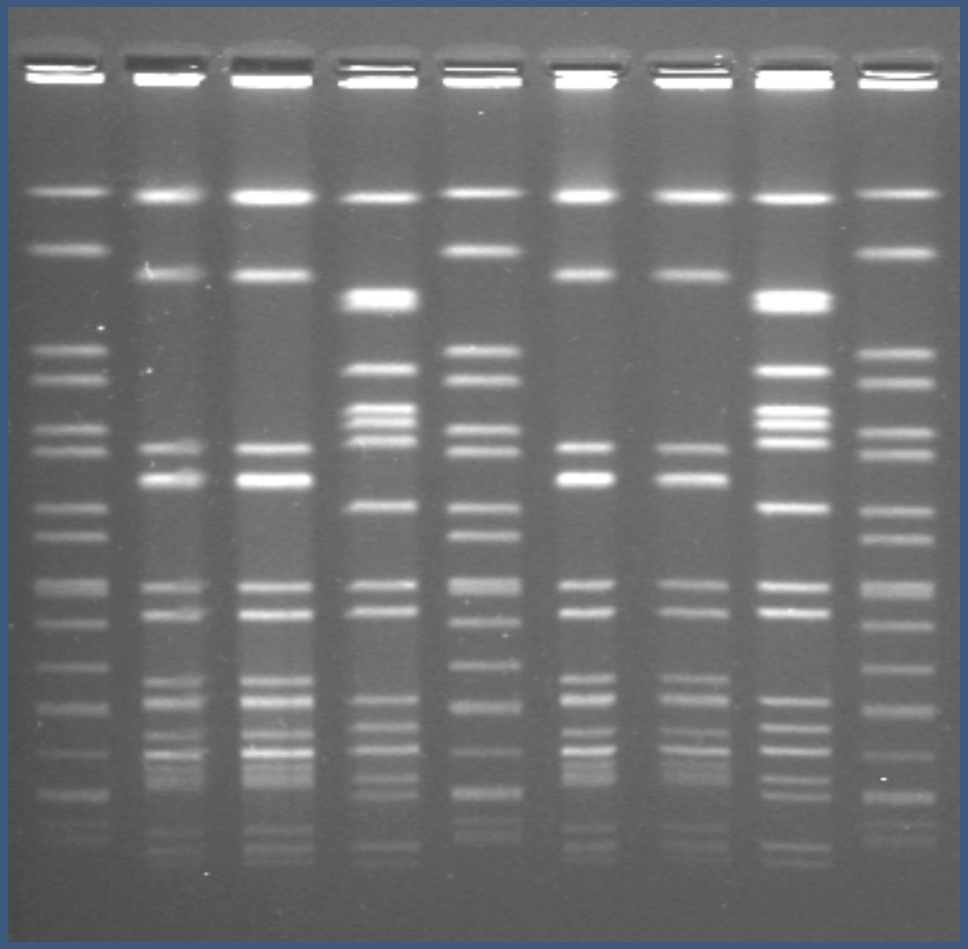
Tests per Sample

Sample throughput

# Objectives

- Methods in Use at the FDA
- Future Methods – SNP-Based
- Mass Spectrometry-Based

# PFGE
# Pulsed-Field Gel Electrophoresis



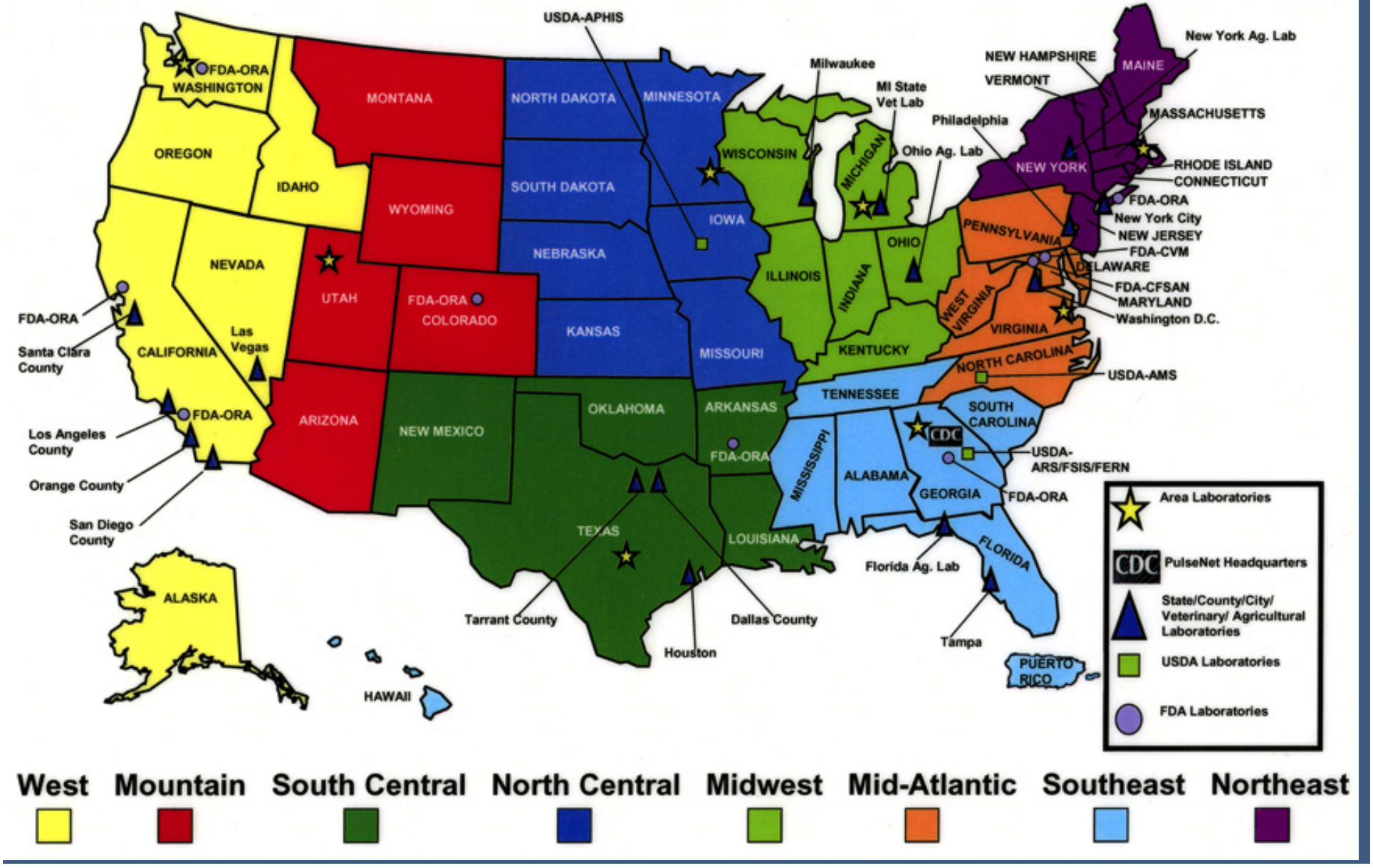
-Large DNA Fragments

-“Universal” Technique

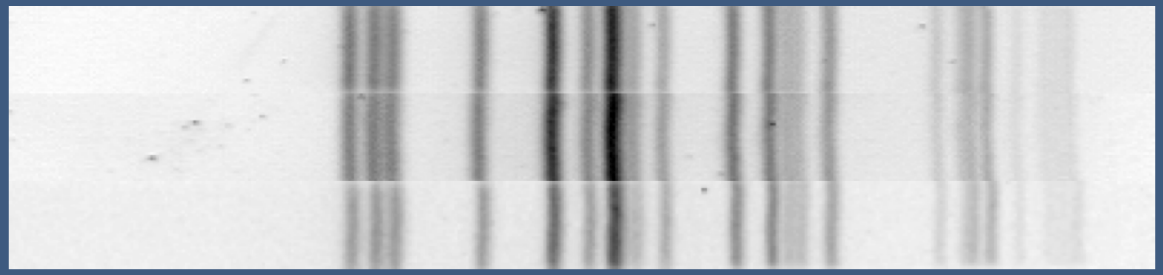
-PulseNet Standardizations

-Usually Epidemiologically Relevant

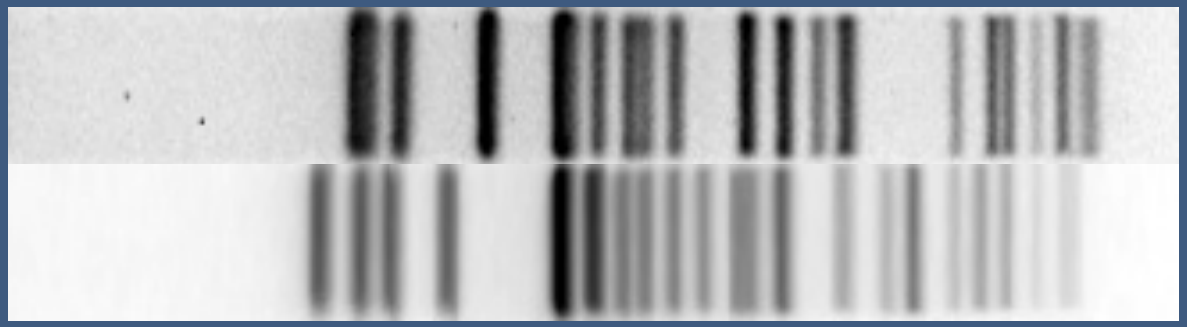
# The National Molecular Subtyping Network for Foodborne Disease Surveillance



**The 2006 spinach outbreak had clear associations between human, food, and animal/environmental isolates, supported by PFGE.**




**Shortly after the spinach outbreak, there were two concurrent shredded lettuce outbreaks**



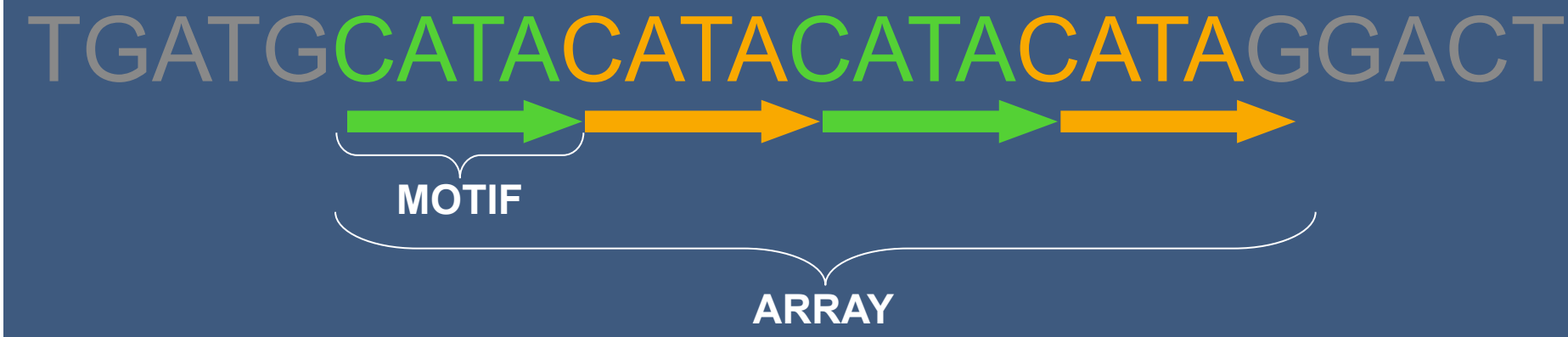
# MLVA
## Multi-Locus VNTR Analysis

## VNTR
## Variable-Number Tandem Repeat

## (Variable copy Numbers of Tandem Repeats)

# Tandem Repeat

A sequence that is made up of a tandemly repeated sequence motif
(arranged head-to-tail without interruption)




**4 bp MOTIF X 4 COPIES = 16 bp ARRAY**

Tandem Repeats may also be called:

- Simple Sequences (SS)
- Short Tandem Repeats (STR)
- Microsatellites

# Variable-Number Tandem Repeat

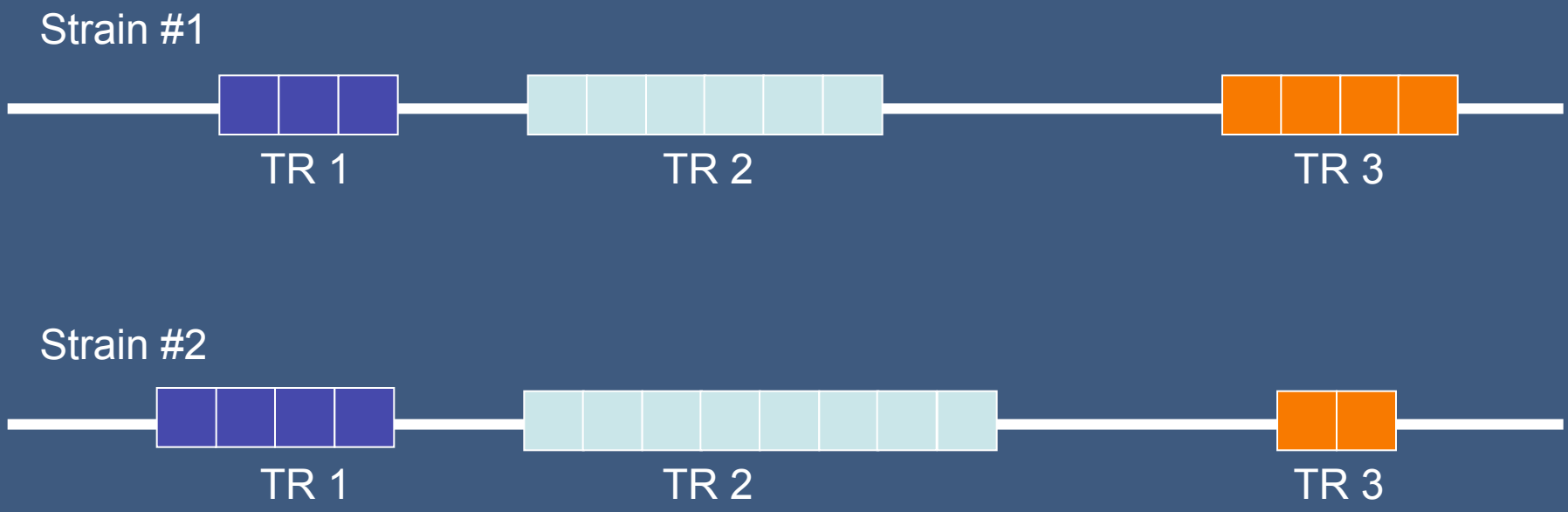
A tandem repeat that varies in the number of copies of the motif. Variation is caused when errors in copy number are made during replication.


(motif is lost)



(motif is gained)

# When we combine the analysis of multiple VNTRs throughout the genome, we create a MLVA system

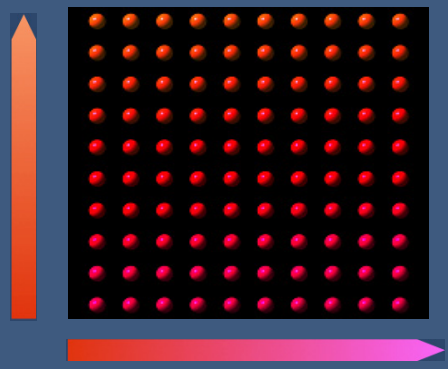


# Objectives

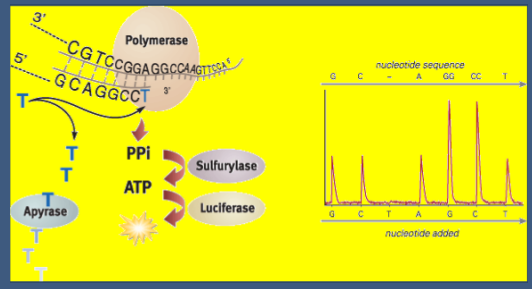
- Methods in Use at the FDA
- Future Methods – SNP-Based
- Mass Spectrometry-Based

# SNP Discovery and Analysis

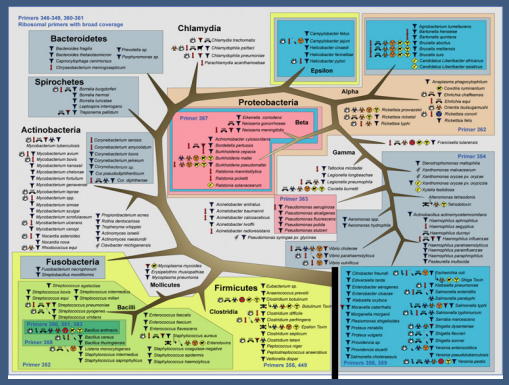
Xmap Technology →



Pyrosequencing →

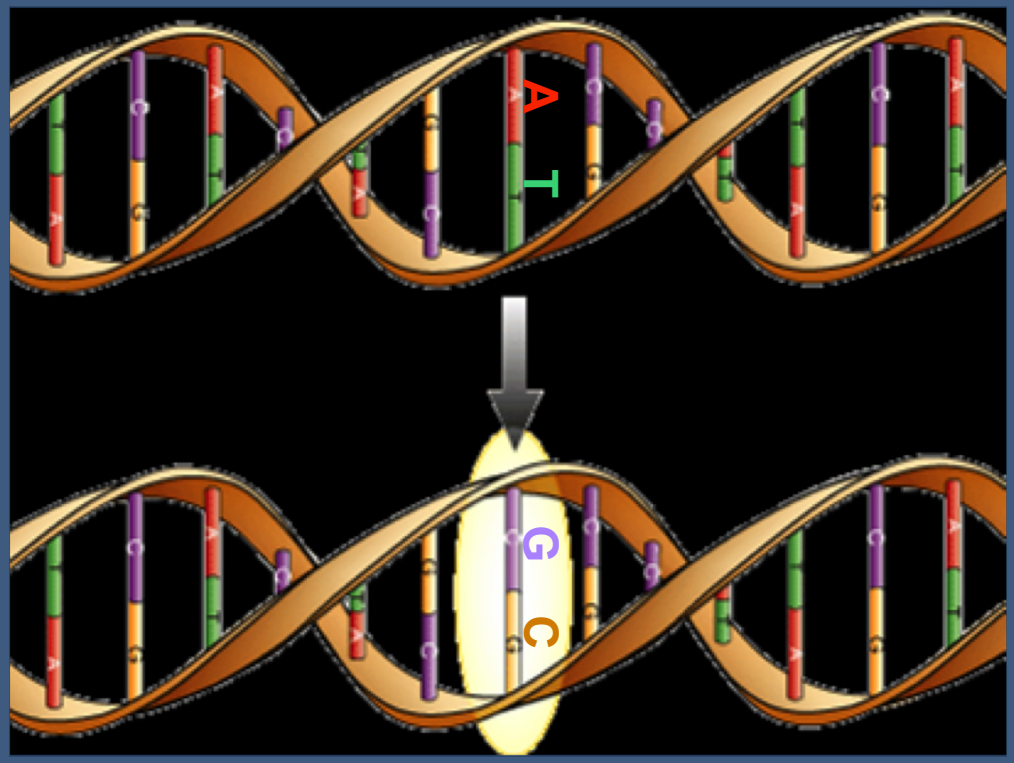



IBIS T5000 →



# SNP

## Single Nucleotide Polymorphism




**General Population Allele** ⟶ TCACACTGGATCA

**Discriminating Allele** ⟶ TCACACGGGATCA

# Advantages to using xMAP technology

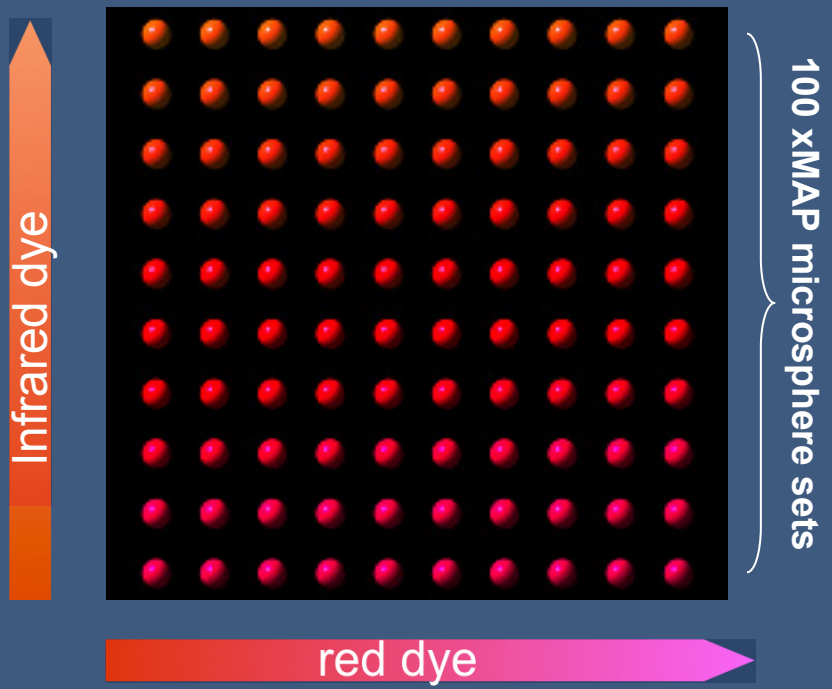
- Nearly solution phase - Fast, Reproducible

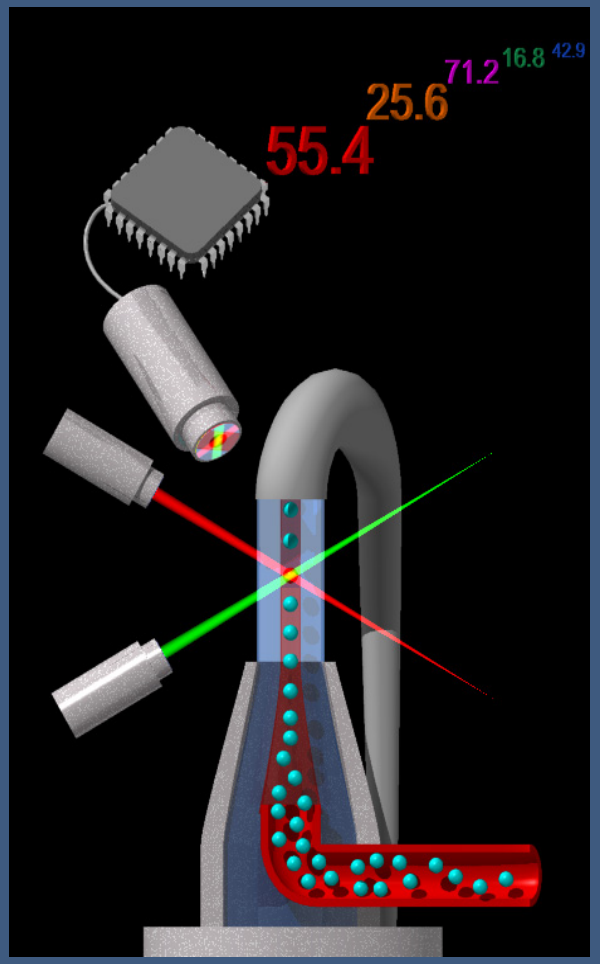
- Multiplex – up to 100 targets

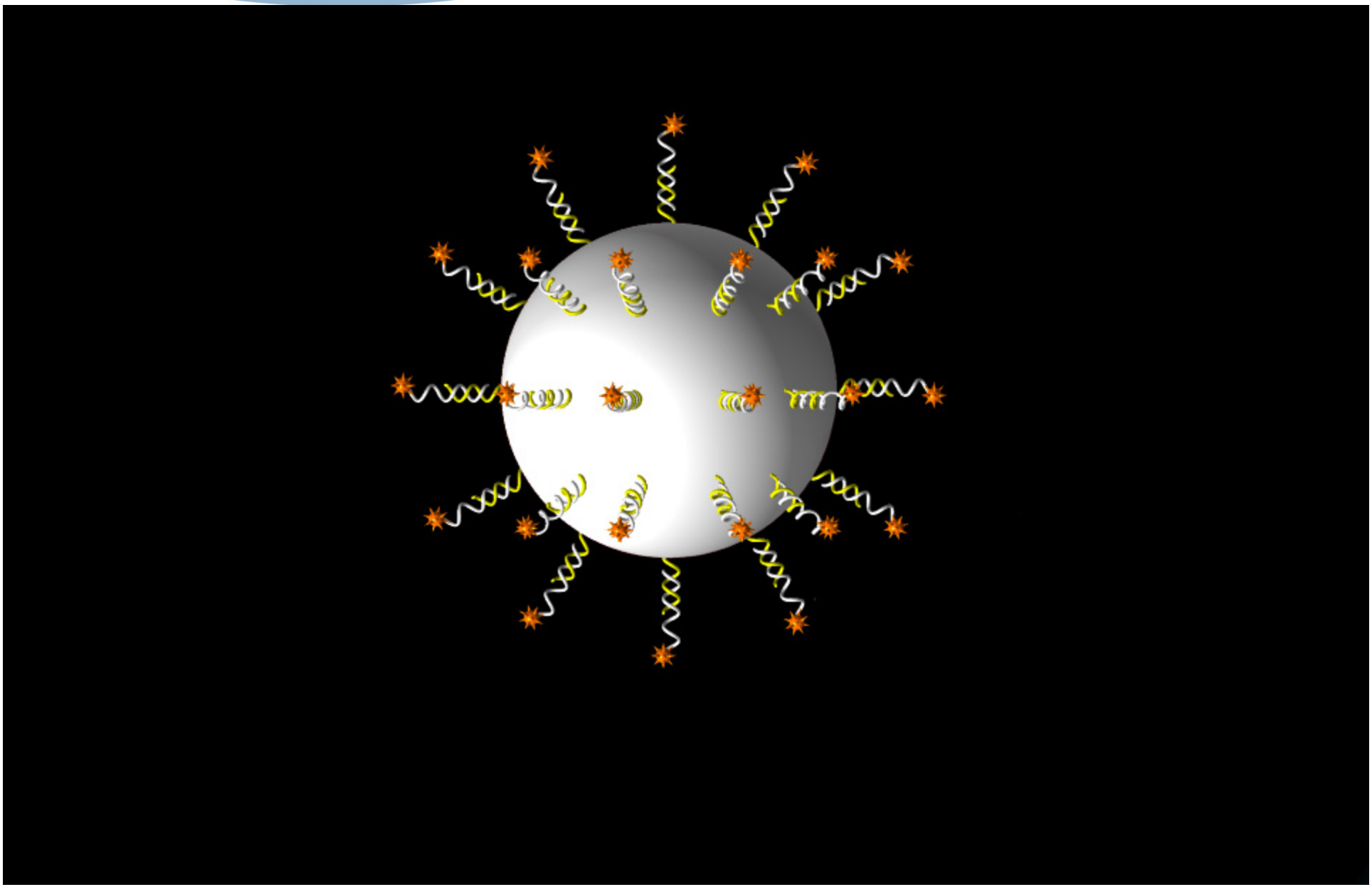
- Once an assay is designed, simple to run

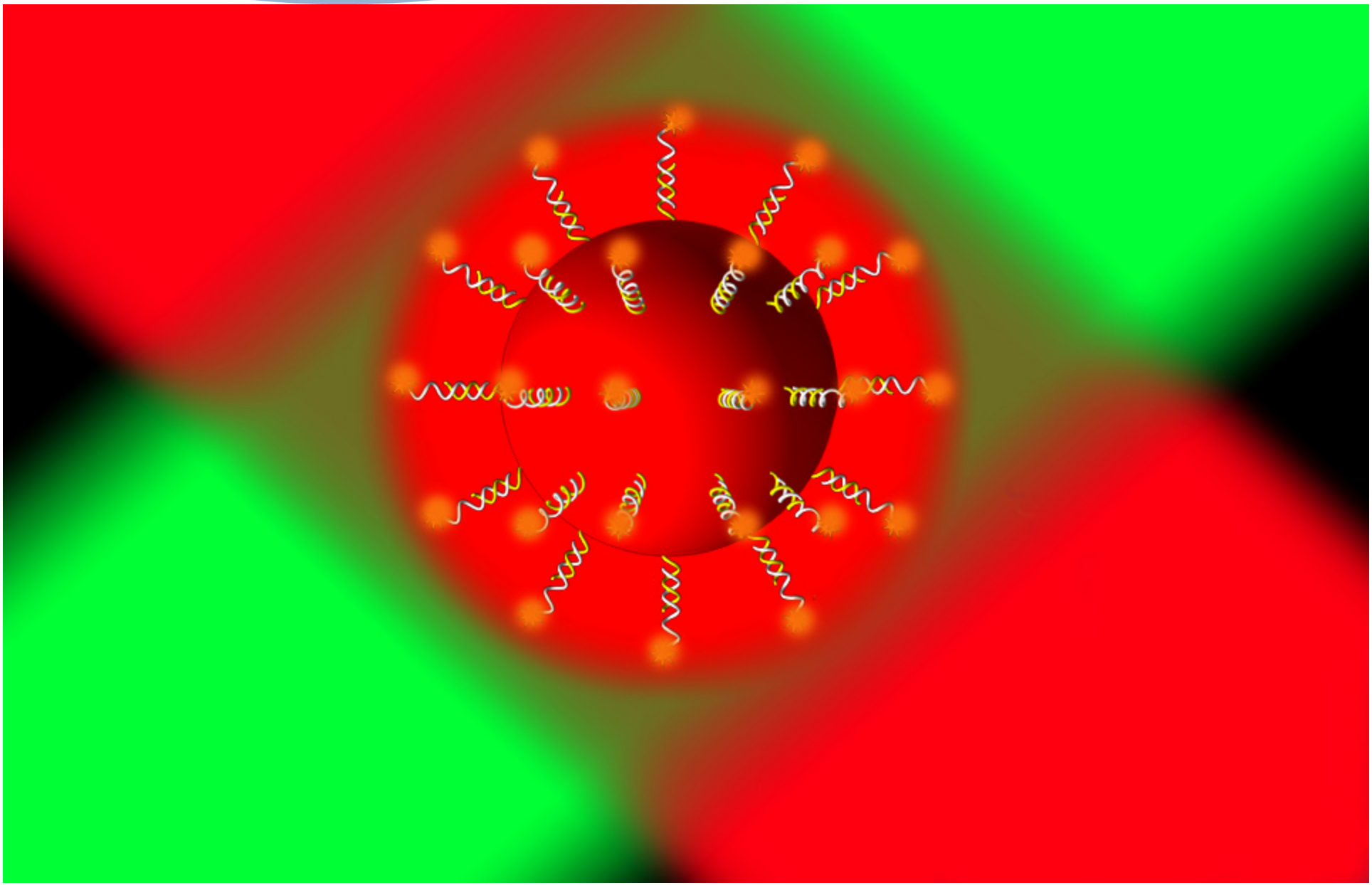
- Can use a variety of capture reagents: antibodies, peptides, or oligonucleotides

# xMAP Technology (Luminex/BioPlex)







# Pyrosequencing

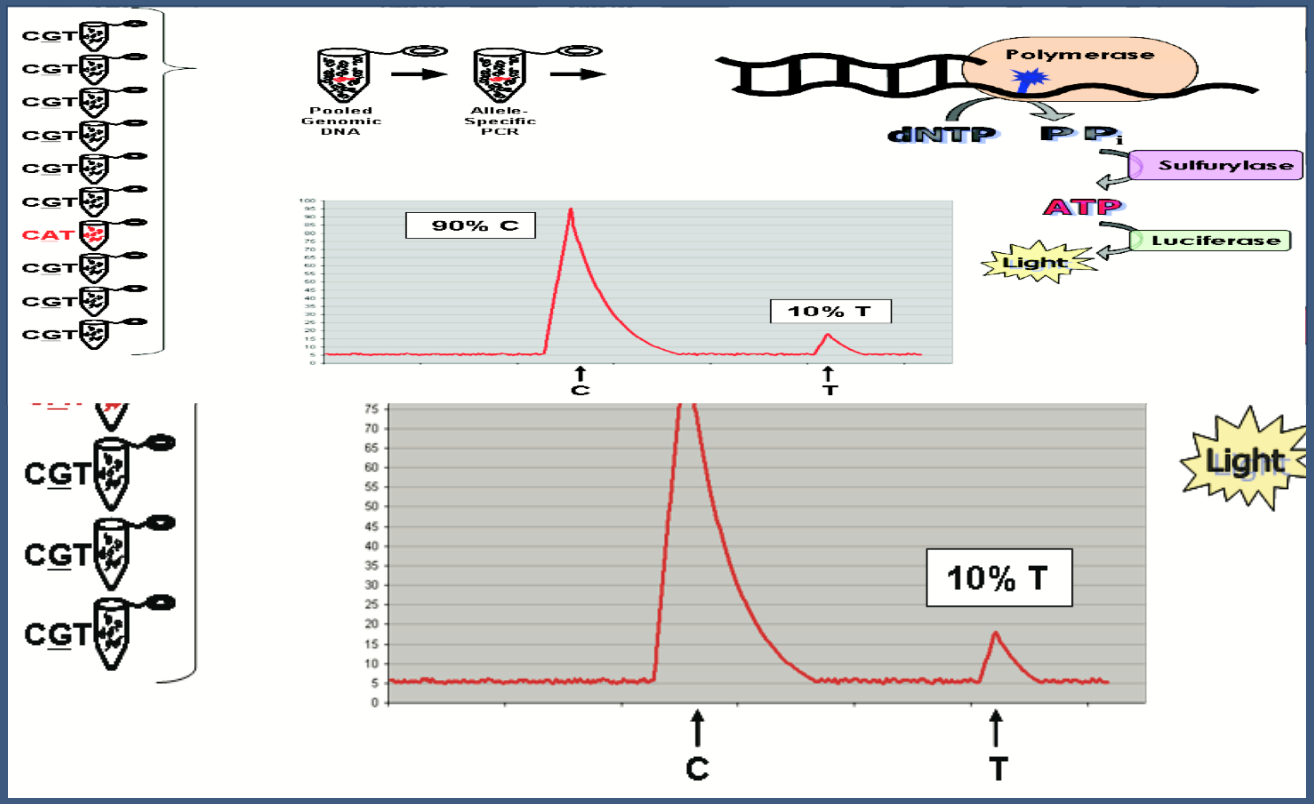
Pyrosequencing is a real-time DNA sequencing technique for rapid analysis of short sequences. Primarily used for SNP discovery and analysis, the technology quantitatively measures allele frequency in a heterogeneous population.

With selection of unique markers, at least 500 strains can be typed daily

Advantages:

- High accuracy potential
- Ease of use
- Highly flexible
- Now emerging as a popular platform for microbial typing

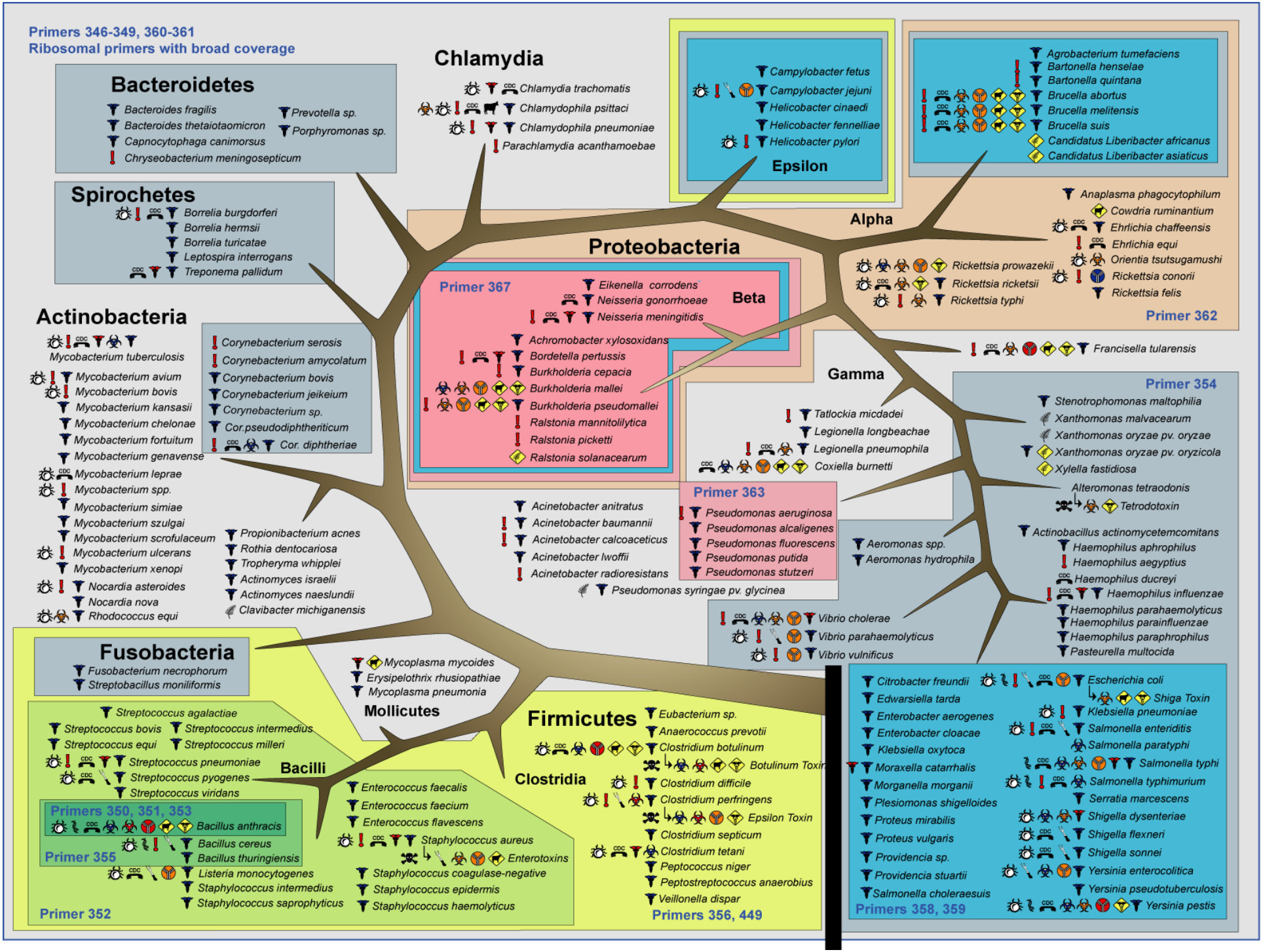
# Principle of Pyrosequencing



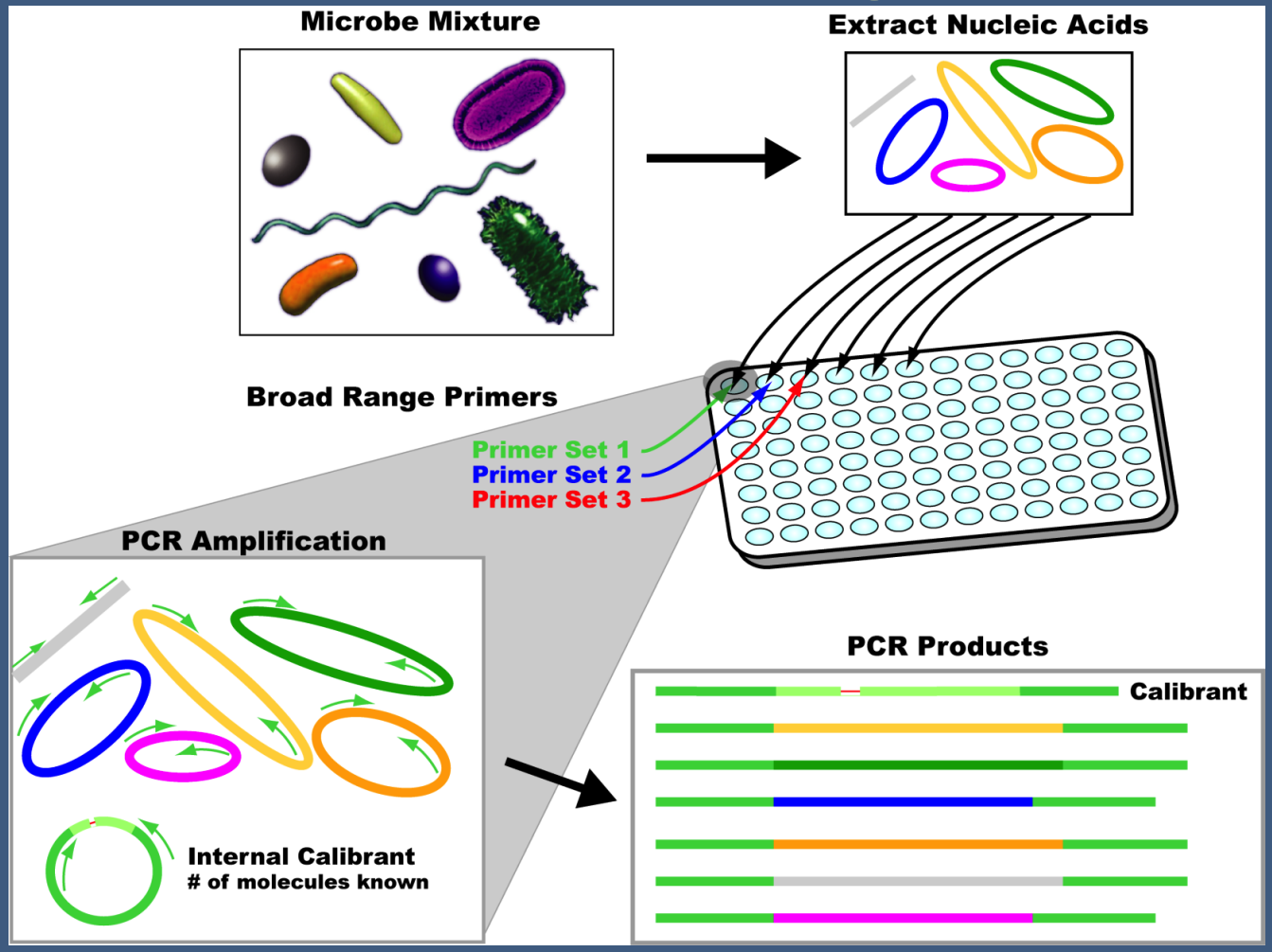

**Pyrosequencing™ is a rapid sequencing method that utilizes the pyrophosphate released upon nucleotide incorporation to generate ATP, which is used as a substrate for luciferase to emit a bioluminescent signal.**

# Highlights of the Ibis T5000

- Rapid identification (~ 5 hours)
- Broad identification of all microbes
  - Bacteria, Viruses, Fungi, Protozoa
  - No culturing
  - No DNA sequencing
- No need for *a priori* knowledge of the infectious organisms present in the sample or the suspected organisms' DNA sequence
  - Emerging infectious disease
- Mixed populations of microbes
  - Quantitative
- High resolution genotyping, strain identification and antibiotic resistance determination
- Cost effective, rapid, high throughput

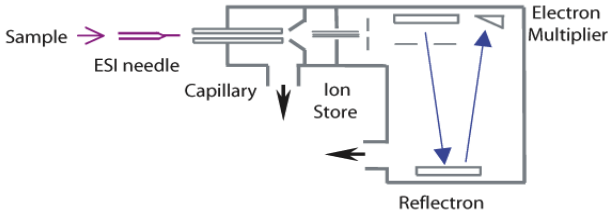

Primers 346-349, 360-361
Ribosomal primers with broad coverage
Bacteroidetes
Bacteroides fragilis
Bacteroides thetaiotaomicron
Capnocytophaga canimorsus
Chryseobacterium meningosepticum
Prevotella sp.
Porphyromonas sp.
Spirochetes
Borrelia burgdorferi
Borrelia hermsii
Borrelia turicatae
Leptospira interrogans
Treponema pallidum
Actinobacteria
Mycobacterium tuberculosis
Mycobacterium avium
Mycobacterium bovis
Mycobacterium kansasii
Mycobacterium chelonae
Mycobacterium fortuitum
Mycobacterium genavense
Mycobacterium leprae
Mycobacterium spp.
Mycobacterium simiae
Mycobacterium szulgai
Mycobacterium scrofulaceum
Mycobacterium ulcerans
Mycobacterium xenopi
Nocardia asteroides
Nocardia nova
Rhodococcus equi
Corynebacterium serosis
Corynebacterium amycolatum
Corynebacterium bovis
Corynebacterium jeikeium
Corynebacterium sp.
Cor.pseudodiphtheriticum
Cor. diphtheriae
Propionibacterium acnes
Rothia dentocariosa
Tropheryma whipplei
Actinomyces israelii
Actinomyces naeslundii
Clavibacter michiganensis
Chlamydia
Chlamydia trachomatis
Chlamydophila psittaci
Chlamydophila pneumoniae
Parachlamydia acanthamoebae
Proteobacteria
Epsilon
Campylobacter fetus
Campylobacter jejuni
Helicobacter cinaedi
Helicobacter fennelliae
Helicobacter pylori
Alpha
Agrobacterium tumefaciens
Bartonella henselae
Bartonella quintana
Brucella abortus
Brucella melitensis
Brucella suis
Candidatus Liberibacter africanus
Candidatus Liberibacter asiaticus
Anaplasma phagocytophilum
Cowdria ruminantium
Ehrlichia chaffeensis
Ehrlichia equi
Orientia tsutsugamushi
Rickettsia conorii
Rickettsia felis
Rickettsia prowazekii
Rickettsia ricketsii
Rickettsia typhi
Primer 362
Beta
Primer 367
Eikenella corrodens
Neisseria gonorrhoeae
Neisseria meningitidis
Achromobacter xylosoxidans
Bordetella pertussis
Burkholderia cepacia
Burkholderia mallei
Burkholderia pseudomallei
Ralstonia mannitolilytica
Ralstonia picketti
Ralstonia solanacearum
Gamma
Francisella tularensis
Primer 354
Stenotrophomonas maltophilia
Xanthomonas malvacearum
Xanthomonas oryzae pv. oryzae
Xanthomonas oryzae pv. oryzicola
Xylella fastidiosa
Alteromonas tetraodonis
Tetrodotoxin
Actinobacillus actinomycetemcomitans
Haemophilus aphrophilus
Haemophilus aegyptius
Haemophilus ducreyi
Haemophilus influenzae
Haemophilus parahaemolyticus
Haemophilus parainfluenzae
Haemophilus paraphrophilus
Pasteurella multocida
Tatlockia micdadei
Legionella longbeachae
Legionella pneumophila
Coxiella burnetti
Primer 363
Pseudomonas aeruginosa
Pseudomonas alcaligenes
Pseudomonas fluorescens
Pseudomonas putida
Pseudomonas stutzeri
Acinetobacter anitratus
Acinetobacter baumannii
Acinetobacter calcoaceticus
Acinetobacter lwoffii
Acinetobacter radioresistans
Pseudomonas syringae pv. glycinea
Aeromonas spp.
Aeromonas hydrophila
Vibrio cholerae
Vibrio parahaemolyticus
Vibrio vulnificus
Fusobacteria
Fusobacterium necrophorum
Streptobacillus moniliformis
Mollicutes
Mycoplasma mycoides
Erysipelothrix rhusiopathiae
Mycoplasma pneumonia
Firmicutes
Bacilli
Streptococcus agalactiae
Streptococcus bovis
Streptococcus intermedius
Streptococcus equi
Streptococcus milleri
Streptococcus pneumoniae
Streptococcus pyogenes
Streptococcus viridans
Primers 350, 351, 353
Bacillus anthracis
Primer 355
Bacillus cereus
Bacillus thuringiensis
Listeria monocytogenes
Staphylococcus intermedius
Staphylococcus saprophyticus
Primer 352
Enterococcus faecalis
Enterococcus faecium
Enterococcus flavescens
Staphylococcus aureus
Enterotoxins
Staphylococcus coagulase-negative
Staphylococcus epidermis
Staphylococcus haemolyticus
Clostridia
Eubacterium sp.
Anaerococcus prevotii
Clostridium botulinum
Botulinum Toxin
Clostridium difficile
Clostridium perfringens
Epsilon Toxin
Clostridium septicum
Clostridium tetani
Peptococcus niger
Peptostreptococcus anaerobius
Veillonella dispar
Primers 356, 449
Citrobacter freundii
Edwarsiella tarda
Enterobacter aerogenes
Enterobacter cloacae
Klebsiella oxytoca
Moraxella catarrhalis
Morganella morganii
Plesiomonas shigelloides
Proteus mirabilis
Proteus vulgaris
Providencia sp.
Providencia stuartii
Salmonella choleraesuis
Escherichia coli
Shiga Toxin
Klebsiella pneumoniae
Salmonella enteriditis
Salmonella paratyphi
Salmonella typhi
Salmonella typhimurium
Serratia marcescens
Shigella dysenteriae
Shigella flexneri
Shigella sonnei
Yersinia enterocolitica
Yersinia pseudotuberculosis
Yersinia pestis
Primers 358, 359

# Ibis Process Part 1: Sample Prep and Broad Range PCR

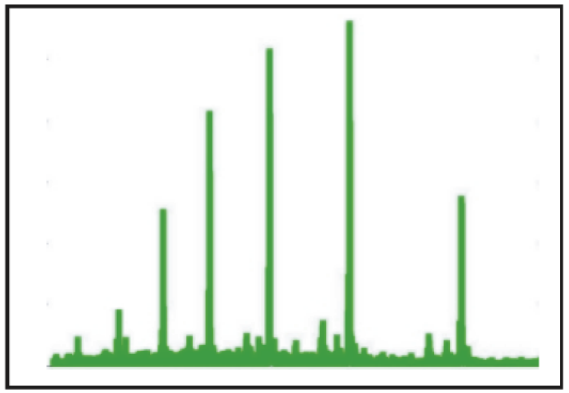


# Ibis Process Part 2: MS Analysis and Signal Processing

## Mass Spectrometer



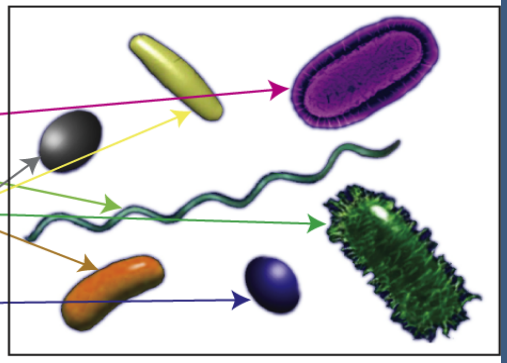

## Spectral Signal



## Signal Processing Masses to Base Compositions

| ORGANISM | Mass | Base Count |
|---|---|---|
| Bacillus anthracis | 35278.823 | A26 G34 C27 T27 |
| Borrelia burgdorferi | 33770.606 | A29 G31 C23 T26 |
| Clostridium botulinum | 35843.944 | A29 G33 C30 T24 |
| Escherichia coli | 35641.855 | A22 G39 C29 T25 |
| Staph aureus | 35240.807 | A24 G35 C30 T25 |
| Strep pneumoniae | 35270.806 | A24 G35 C28 T27 |
| Strep pyogenes (Group A Strep) | 35281.808 | A23 G37 C30 T24 |

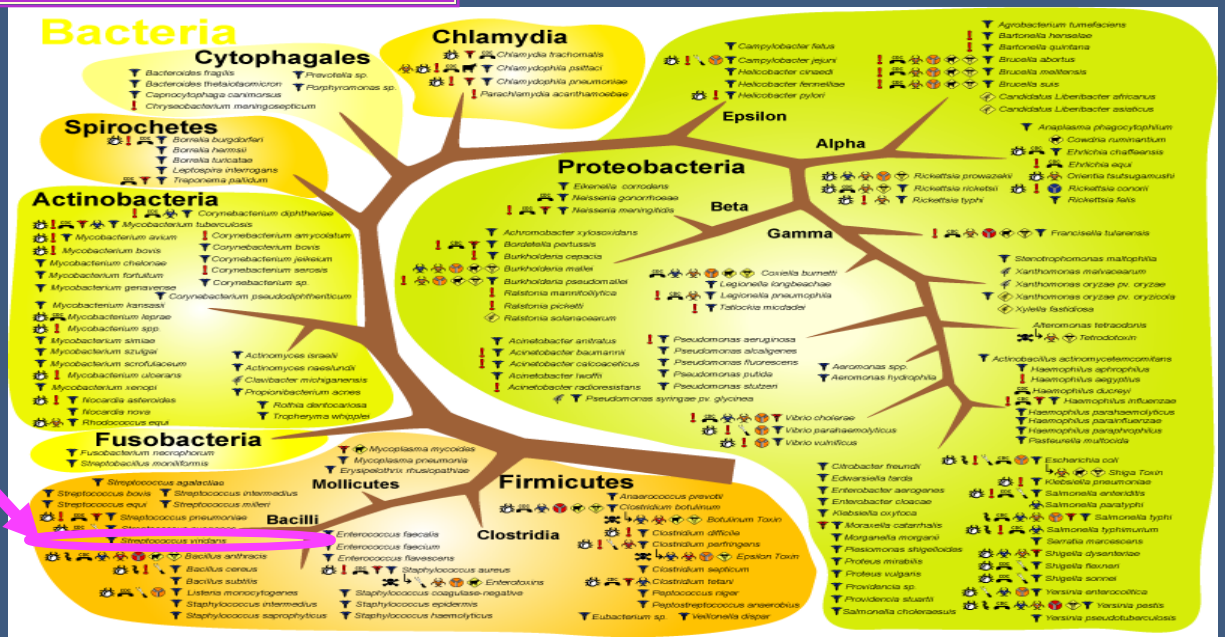
## Base Compositions Map to Microbes



# Ibis Process Part 3: Triangulation Using Multiple Primer Pairs

| Primer #1 | Mass | Base Count |
|---|---|---|
| Blue | 18234.970 | $A_{12}G_{17}C_{17}T_{13}$ |
| Blue | 17948.926 | $A_{14}G_{14}C_{12}T_{18}$ |
| Blue | 18610.017 | $A_{11}G_{19}C_{15}T_{15}$ |
| Blue | 17936.912 | $A_{11}G_{17}C_{16}T_{14}$ |
| Blue | 18877.118 | $A_{18}G_{15}C_{15}T_{13}$ |

**Bioinformatics!**

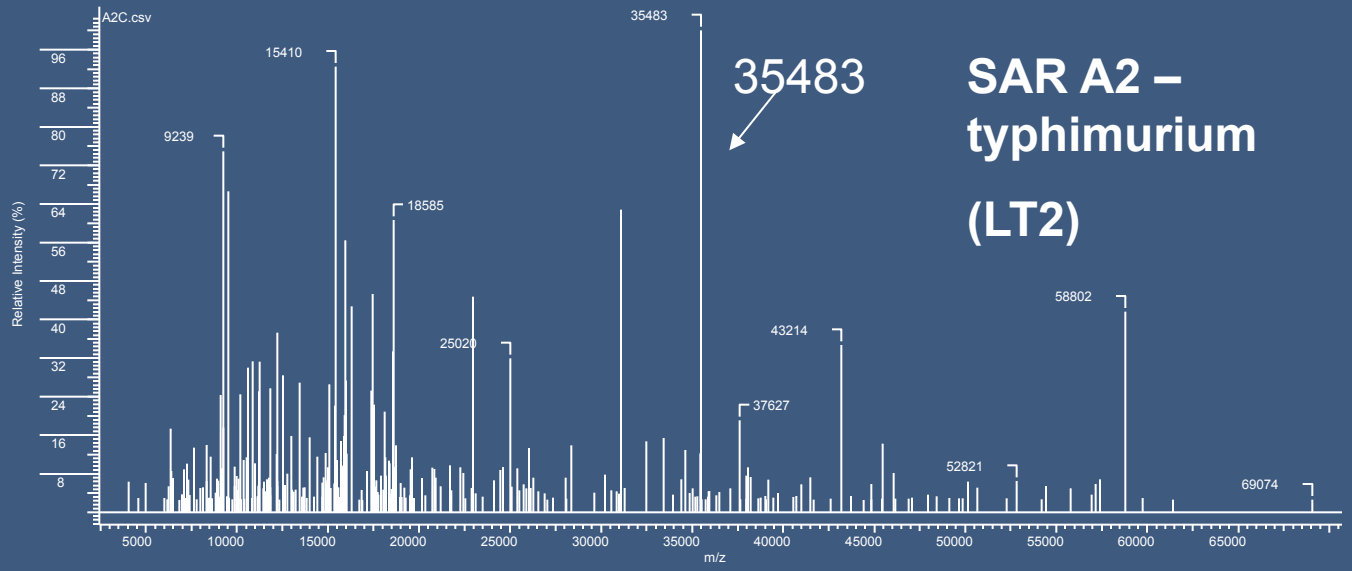


# Objectives

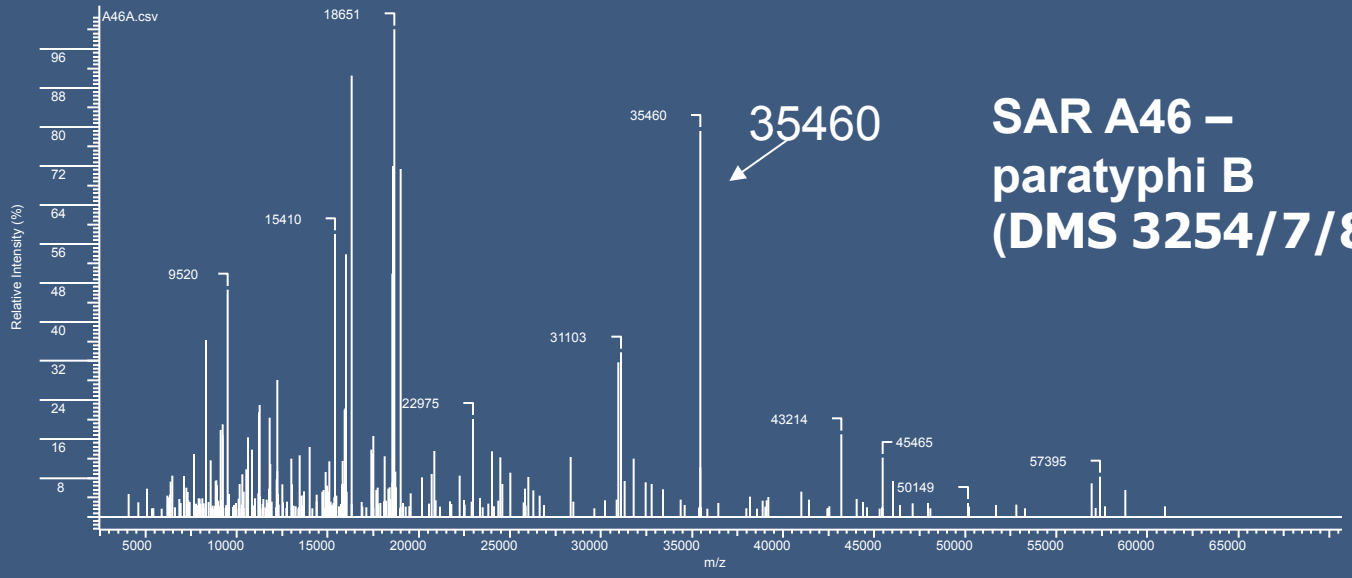
- Methods in Use at the FDA
- Future Methods – SNP-Based
- Mass Spectrometry-Based

# MS based approaches

- High mass accuracy analysis of PCR amplicons
- Fatty acids (FAME), desorption electrospray of phospholipids and metabolites
- Peptide-based methods (bottom up): digest everything with enzymes, look for peptides
- Matrix-assisted Laser Desorption/Ionization (MALDI) of whole cell
- Intact protein LC/MS of protein extract




**SAR A2 – typhimurium (LT2)**




**SAR A46 – paratyphi B (DMS 3254/7/81)**

# Advantages of MALDI

- Intact protein MS is an additional tool for differentiation of closely related *Salmonella* strains
- Provides identification of protein “targets” for detection and sub-typing of strains without genome sequencing
- Could be used in combination with “top-down methods” for rapid analysis of protein variants
- Method may be adapted to epidemiological and microbial forensics needs

# Acknowledgements

Dr. Rebecca Bell (IBIS T5000)
Dr. Jie Zheng (xMAP Technology)
Dr. John Callahan (MALDI)
Alice E Hayford (Pyrosequencing)

Dr. Eric Brown Branch Chief, Molecular Methods and Subtyping Branch

Christine.Keys@FDA.HHS.GOV

FDA
CFSAN

# Working to Keep Food and Cosmetics Safe and Promote Good Nutrition