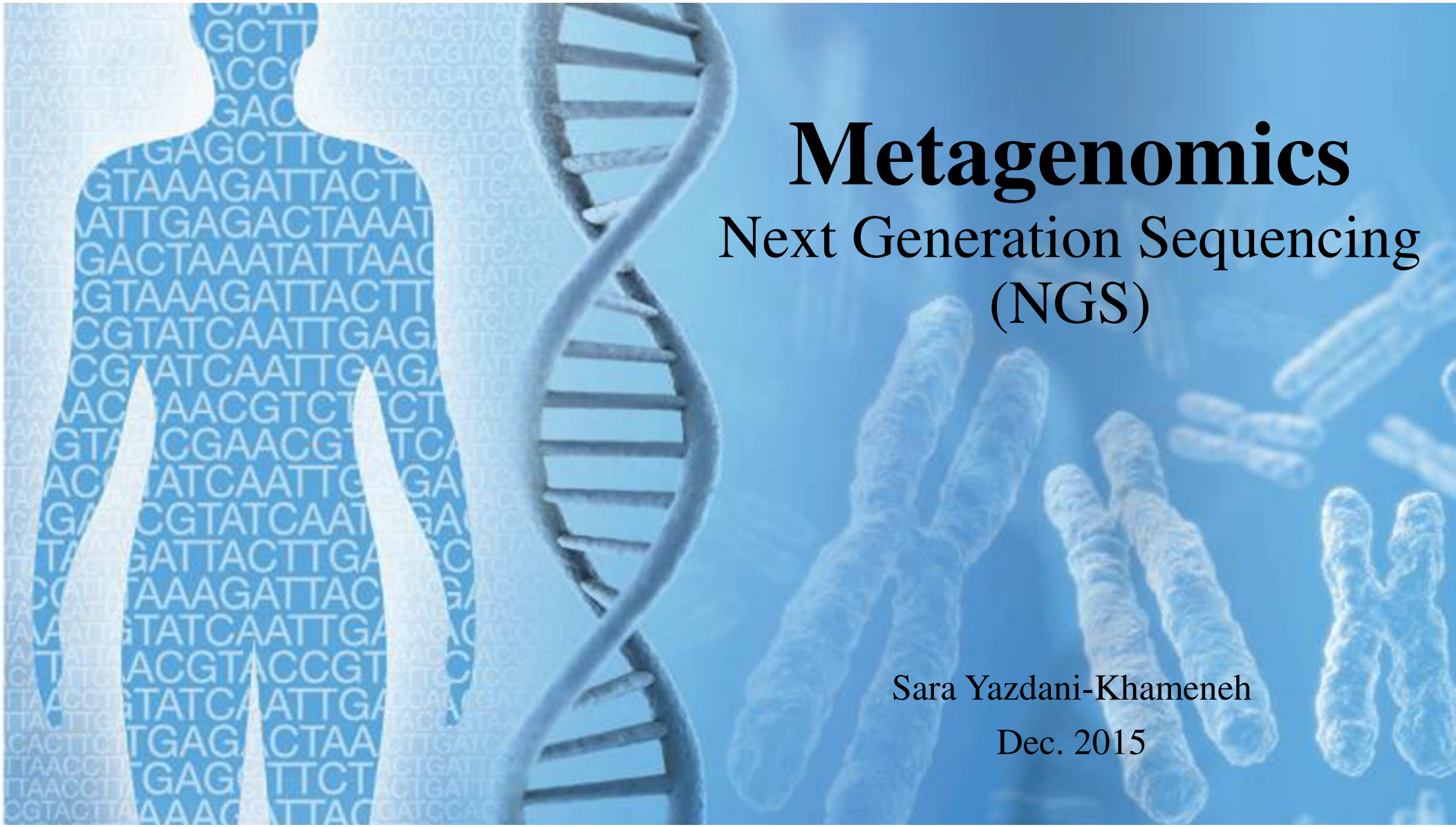


Metagenomics

Next Generation Sequencing (NGS)

Sara Yazdani-Khameneh

Dec. 2015

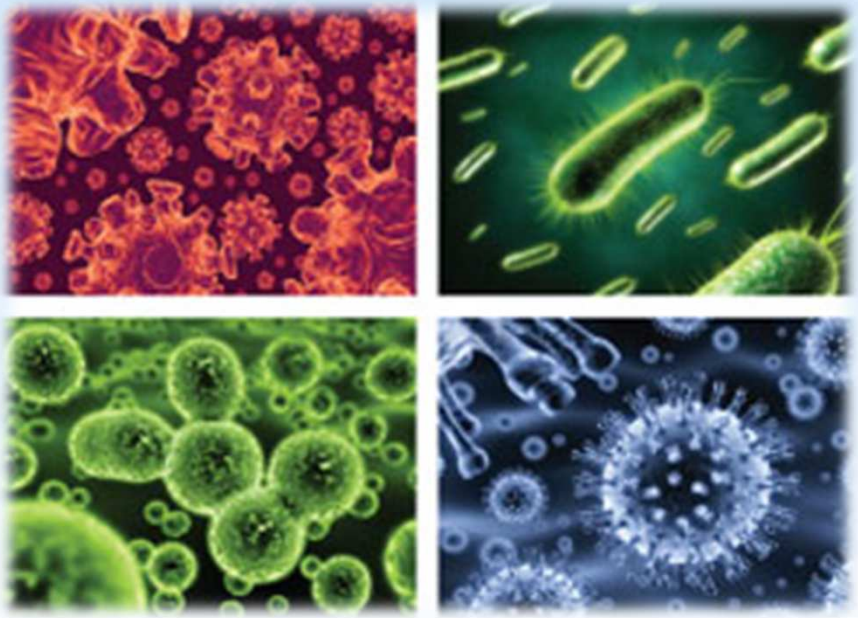
Microorganisms

- **Represent the most important and diverse group of organisms**
- **Widely distributed in many environmental habitats**
- **Play major roles in ecosystems functioning**

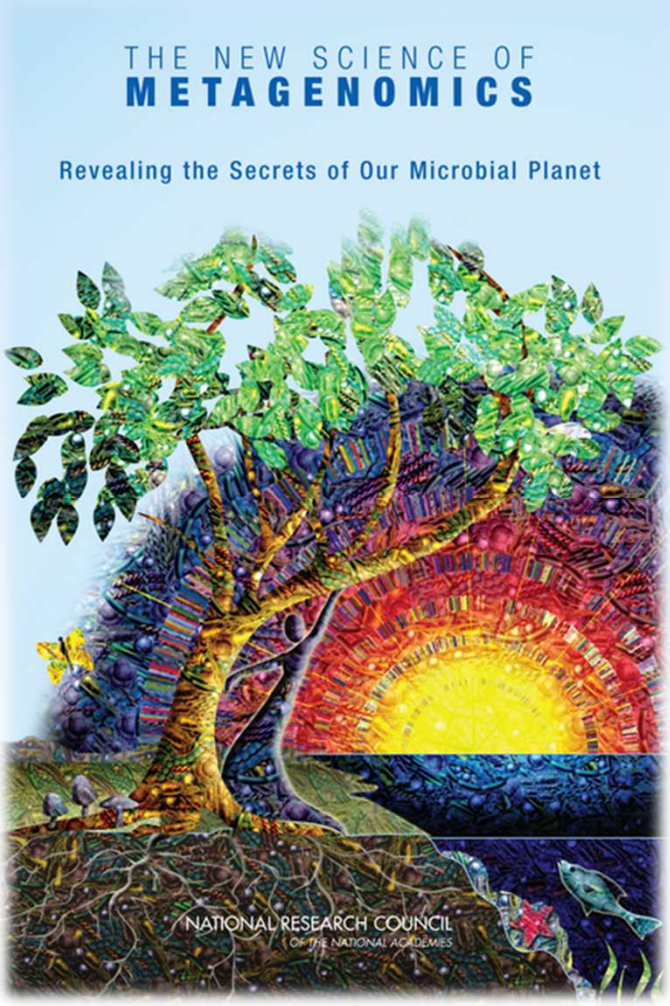
Sequencing

First DNA sequencing methods

Next generation sequencing methods

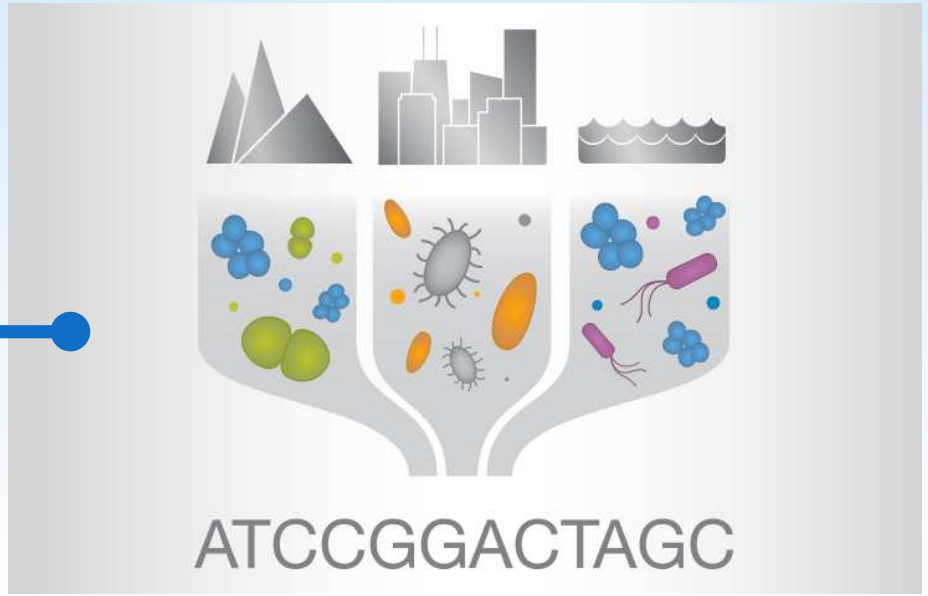


But



What is the solution ?

What is Metagenomics ?



Metagenomics

Meta-analysis: Statistically combining separate analyses.

Genomics: Comprehensive analysis of organisms' genetic material.

History

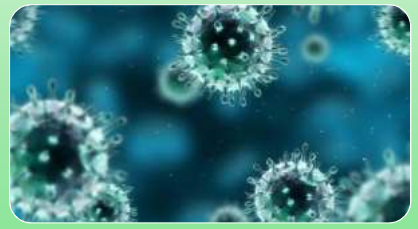
- **Late 17th century, Anton van Leeuwenhoek :**
  - First metagenomicist who directly studied organisms from pond water and his own teeth.

- **1920's**:
  - Cell culture evolved, moved away from early metagenomics.
  - If an organism could not be cultured, it could not be classified.

- **1980's**:
  - **Discrepancies observed:**
  - (1) Number of organisms under microscope in conflict with amount on plates.
  - (2) Cellular activities in situ conflicted with activities in culture.
  - (3) Cells are viable but unculturable.

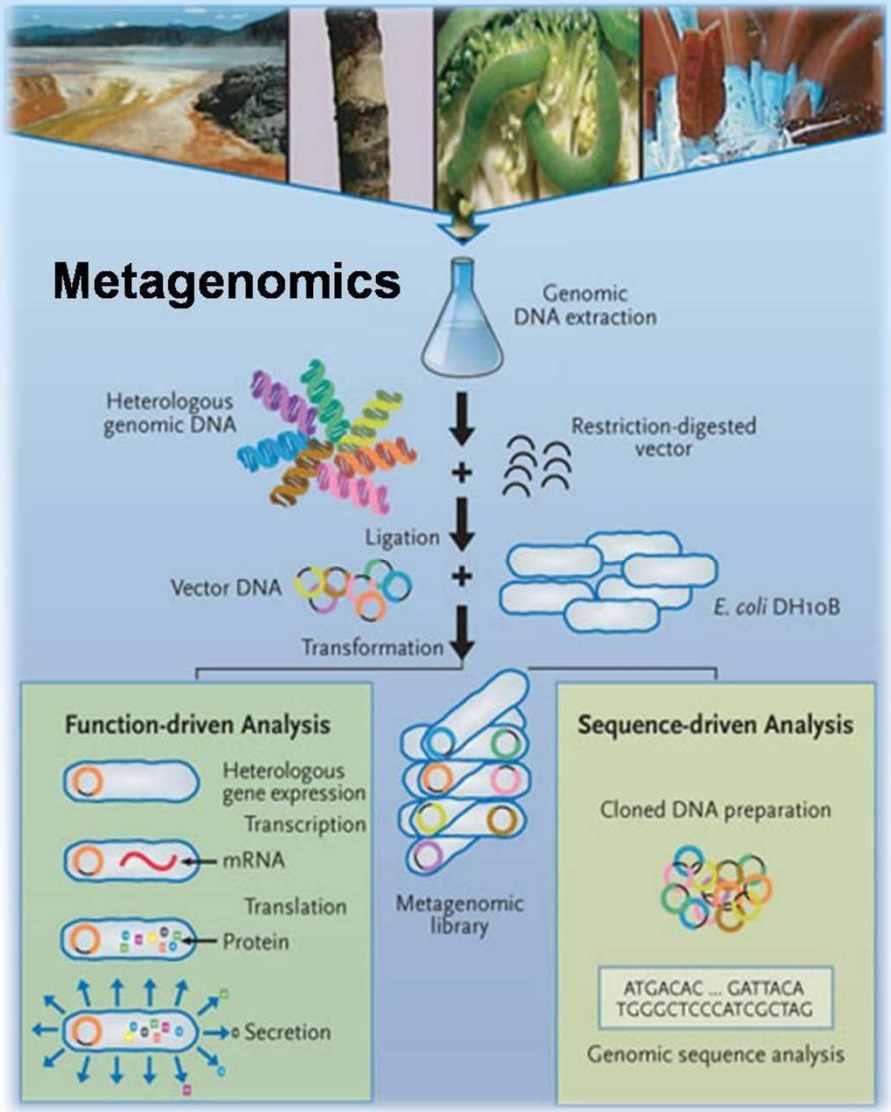
Why is Metagenomics Important ?

Organisms can be studied directly in their environments bypassing the need to isolate each species



There are significant advantages for viral metagenomics, because of difficulties cultivating the appropriate host

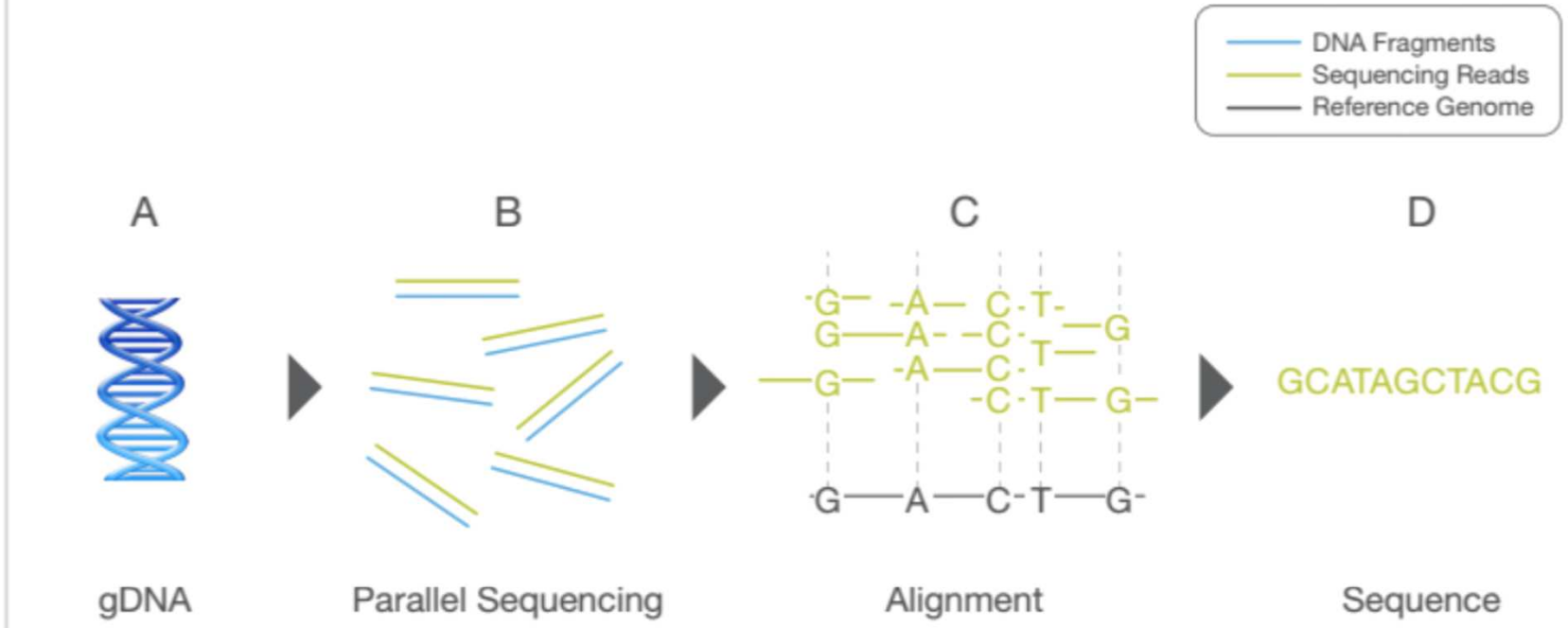
Genomic information has advanced research in a diverse array of fields, including forensic science and biomedical research



Metagenomics steps

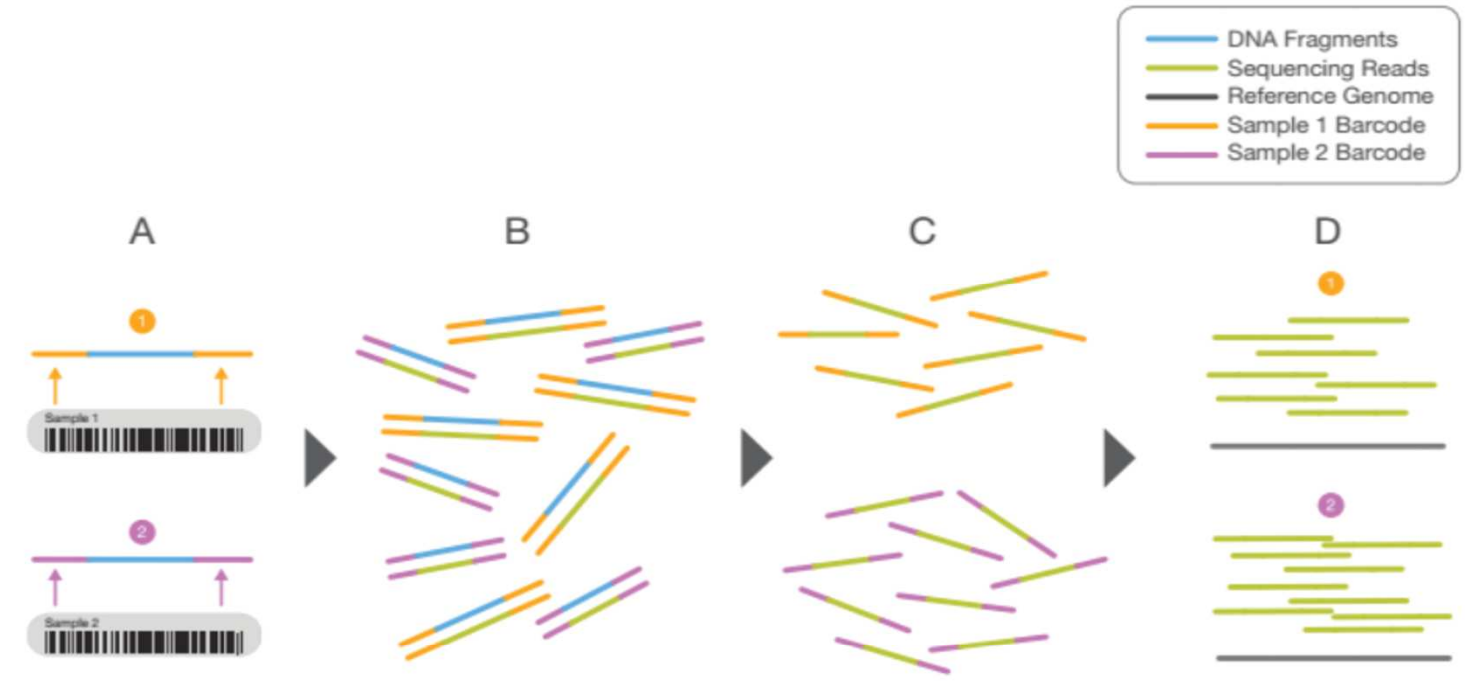
- Isolate DNA
  - Depends on sample type
- Clone DNA
- Insert into plasmid
- Develop sample library
- Screen or sequence

Figure 1: Conceptual Overview of Whole-Genome Resequencing



A. Extracted gDNA.
B. gDNA is fragmented into a library of small segments that are each sequenced in parallel.
C. Individual sequence reads are reassembled by aligning to a reference genome.
D. The whole-genome sequence is derived from the consensus of aligned reads.

Figure 2: Conceptual Overview of Sample Multiplexing

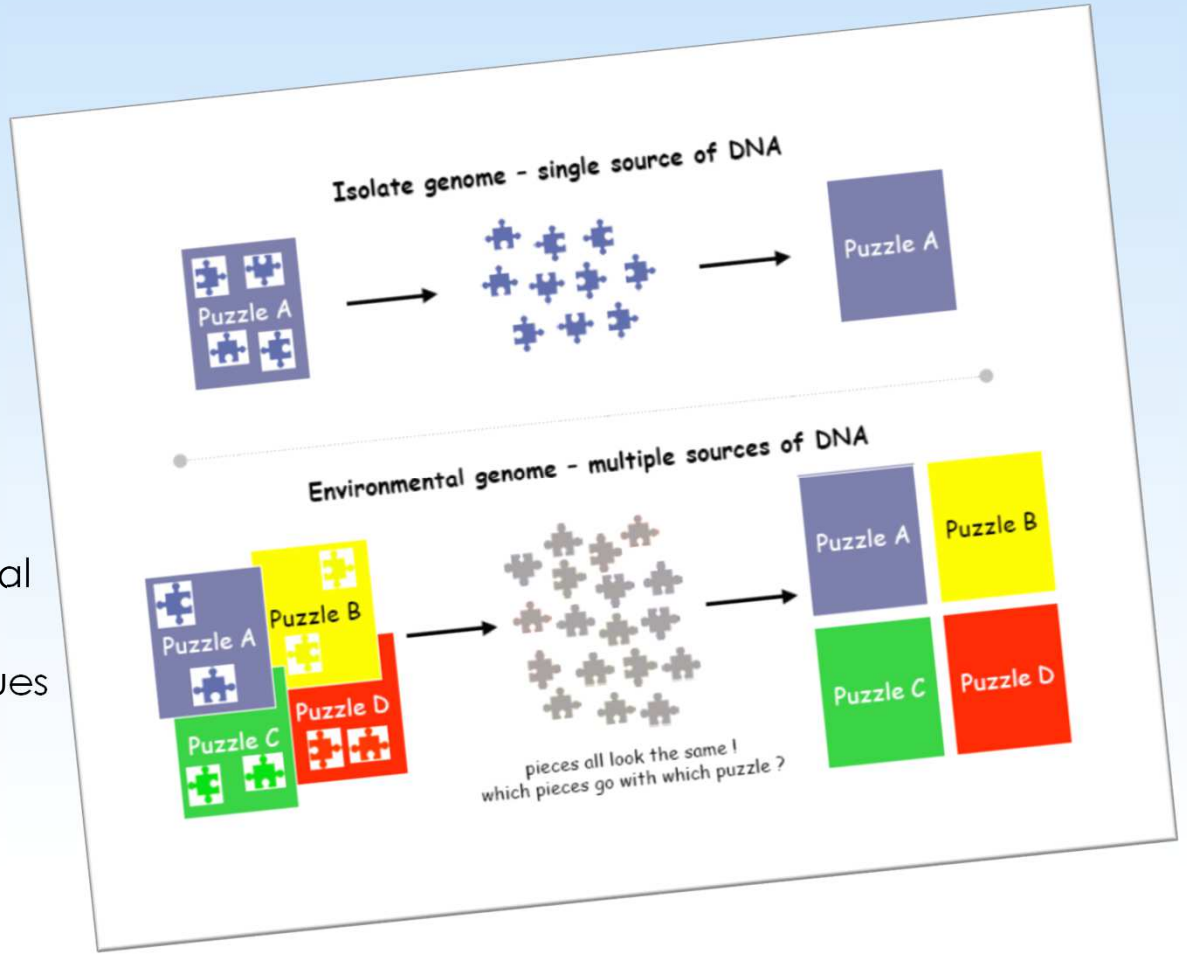


A. Two representative DNA fragments from two unique samples, each attached to a specific barcode sequence that identifies the sample from which it originated.
B. Libraries for each sample are pooled and sequenced in parallel. Each new read contains both the fragment sequence and its sample-identifying barcode.
C. Barcode sequences are used to de-multiplex, or differentiate reads from each sample.
D. Each set of reads is aligned to the reference sequence.

A2

Comparison

- **Traditional microbial genomics**
  - Sequence the genome of one organism at a time
  - Use cultures to isolate microbe of interest
- **Metagenomics**
  - Extract sequence data from microbial communities as they exist in nature
  - Bypass the need for culture techniques
    - Sequence all DNA in sample
    - Select DNA based on universal sequences



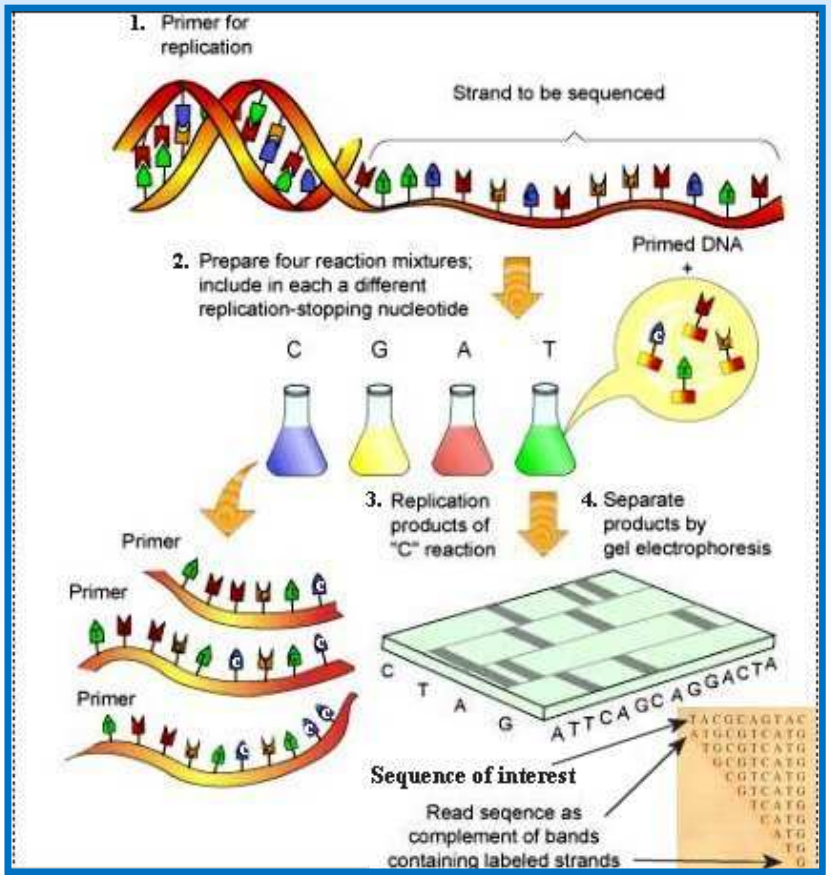
Slide 13

A2 Traditional genomics focuses on the sequencing and analysis of the genomes of individual organisms. When applied to microbes, it typically involves culturing the organism of interest followed by sequencing. Metagenomics is a new area of microbial genomics that aims to sequence the full or partial genomes of all members of a microbial community (also called a consortium). The term microbial community refers to the complex microbial ecosystems that exist almost everywhere in nature. For example, a project in soil metagenomics might extract DNA from a soil sample in a corn field and attempt to sequence all the DNA found in the sample. By directly sequencing the DNA, researchers bypass the need to culture organisms. Since only a very small minority of single-cell organisms have been successfully cultured in the laboratory, metagenomics becomes a very powerful technique for sequencing genes from organisms that can not be cultured. Alternatively, homologous genes from a variety of organisms in the microbial community can be selectively sequenced via PCR using tags that exist in known organisms.

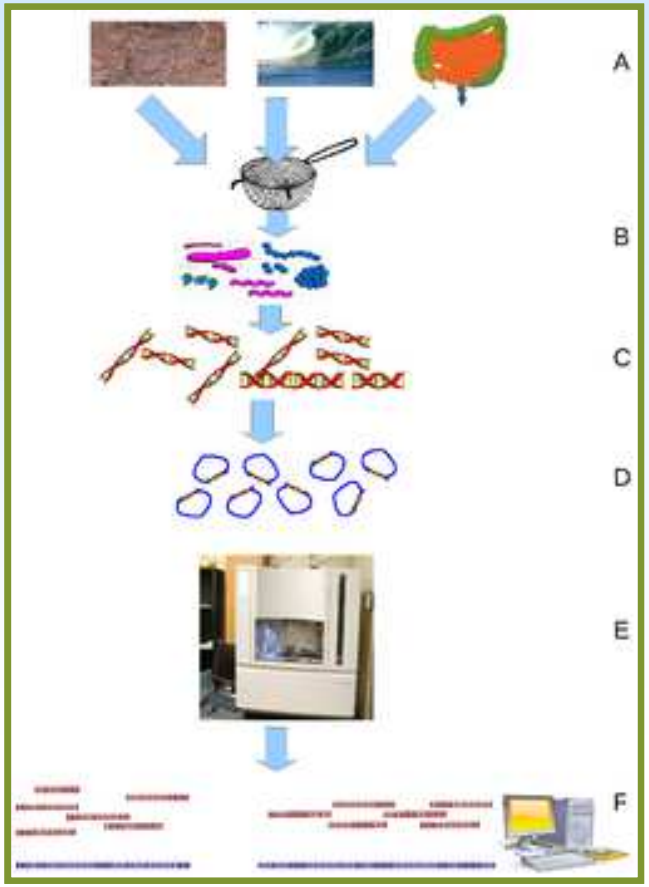
Author; 12/22/2015

A3

Sanger Sequencing (CE)



Metagenomics



Slide 14

A3 برای توضیحات به قسمت عکسها و عکس جدول مقایسه متاژومیکس و سنجر مراجعه شود

Author; 12/25/2015

Human genome project

Metagenomics	Sanger sequencing method
5 Person	1 Person
A single run	10 years
1 week	3 years
Less than $ 5,000	About 3 billion USD

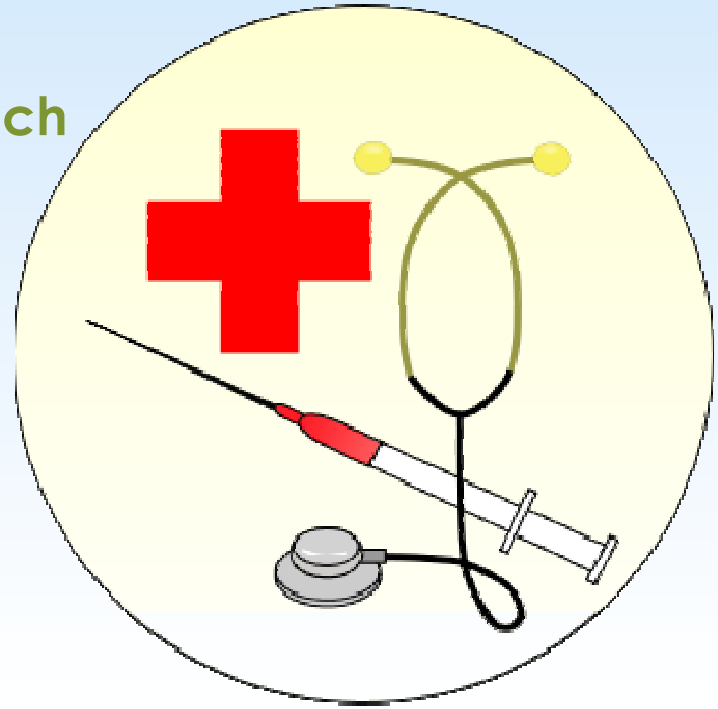
Advantages of metagenomics

- Provides a much cheaper and higher throughput alternative to sequencing DNA than traditional Sanger sequencing. Whole small genomes can now be sequenced in a day.
- Affordable and efficient for quickly interrogating particular genomic regions of interest
- Provides a deeper coverage of genomic regions of interest
- High-throughput sequencing of the human genome facilitates the discovery of genes and regulatory elements associated with disease.
- Can be utilized in deciding a therapeutic plan of action for both germline and somatic cancers
- Detects and quantifies low-frequency variants such as rare drug-resistant viral mutations (e.g., HIV, HBV or microbial pathogens)

Metagenomics application areas

Applications of NGS in medicine

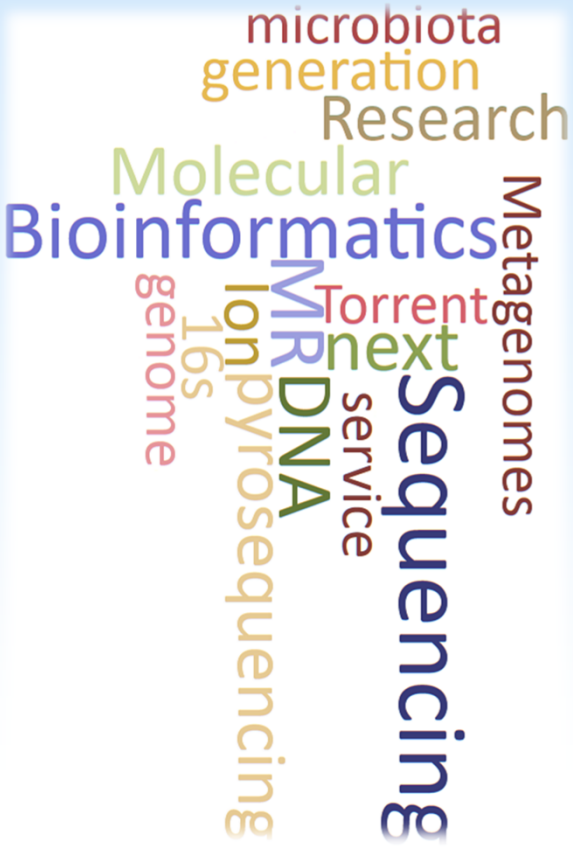
- Detecting mutations that play a role in diseases such as cancer
- Identifying genes responsible for inherited skin diseases
- Determining RNA expression levels
- Identifying novel virulence factors through sequencing of bacterial and viral species



Disadvantages

- NGS, although much less costly in time and money in comparison to first-generation sequencing, is still too expensive for many labs.

- Data analysis can be time-consuming and may require special knowledge of bioinformatics to garner accurate information from sequence data.

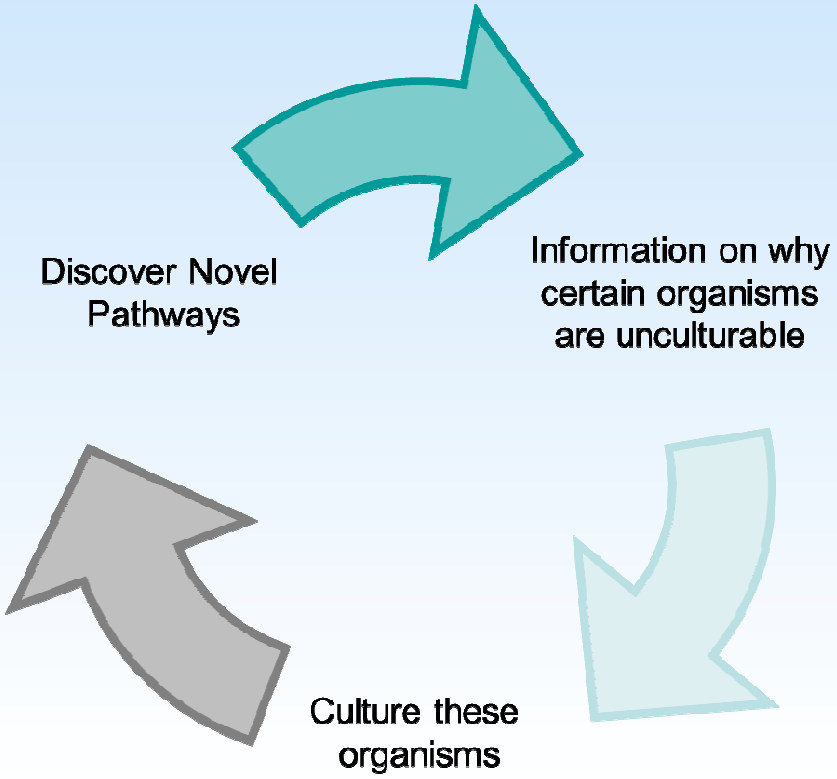


Conclusions

- Metagenomics has evolved from multiple limitations in genology and phylogeny.
- Common techniques can be used to analyze the genetic material from bacteria and organisms grown in their environment.
- Crucial symbiotic relationships are more easily studied using metagenomics through allowing the symbiont to grow in its natural environment.
- Phylogenic trees can be developed based on sequence-driven approaches
- Novel pathways will be determined using the technology required for faster analysis of a broader range of organisms

Future Directions

- New enzymes, antibiotics, and other reagents identified
- More exotic habitats can be intently studied
- Can only progress as library technology progresses, including sequencing technology
- Improved bioinformatics will quicken analysis for library profiling
- Investigating ancient DNA remnants
- Discoveries such as phylogenic tags (rRNA genes, etc) will give momentum to the growing field
- Learning novel pathways will lead to knowledge about the current nonculturable bacteria to then culture these systems



Thank You
For Your
Time