In/formation

a helix
winds its
mythical
way across
the space
between
what is
not and
what could
be, to hook
up with like
-minded
entities
in order
to create
even bigger
uniquely
-shaped
bodies

Jhe

Next Generation Sequencing with the Genomics Core

Wouter Bossuyt- Greg Maes

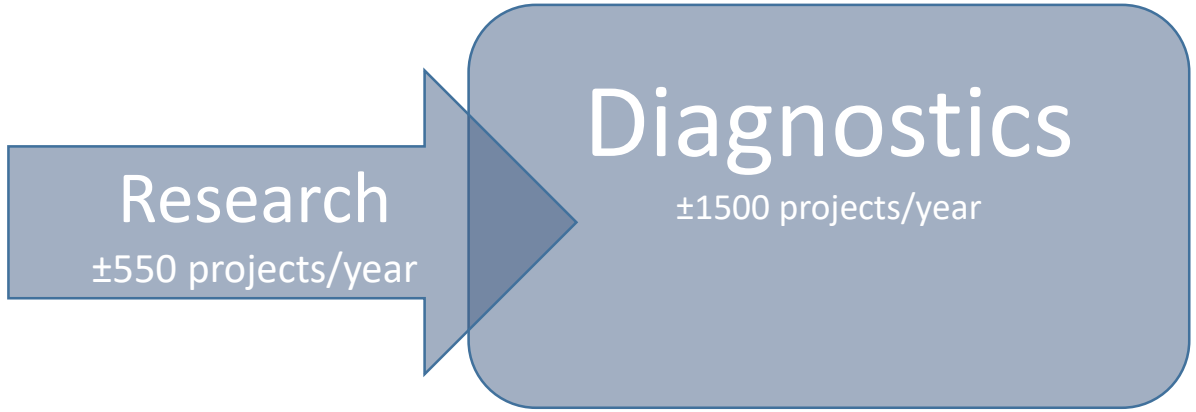
20/11/2020

Sciences Workshop Genomics Core

What’s the use of technology if it is not usable?

Mission

- Make the NGS revolution a possibility for everyone
- Introduce novel technologies as service
- Reduce cost by optimizing sequencing
- Diagnose patients



Who are we?

- 8 wet lab people
- 9 bio-informaticians
- 4 supporting people

DNA

- Amplicon, MIPs
- *de novo* assembly
- Metagenetics, -genomics
- Bacterial WGS
- Vertebrate WGS
- Long read sequencing

RNA

- Truseq stranded mRNA
- Lexogen quantseq
- IsoSeq (Pacbio)

Single cell genomics

- DepArray
- SMART-seq2
- 10x Genomics

Sequencing only

Custom Bio-informatics

Sequencing

- Illumina

Miseq

Nextseq

Hiseq 4000

Novaseq

- PacBio

Pacbio Sequel

- Nanopore

MinIon

PrometION 24

How to get started?

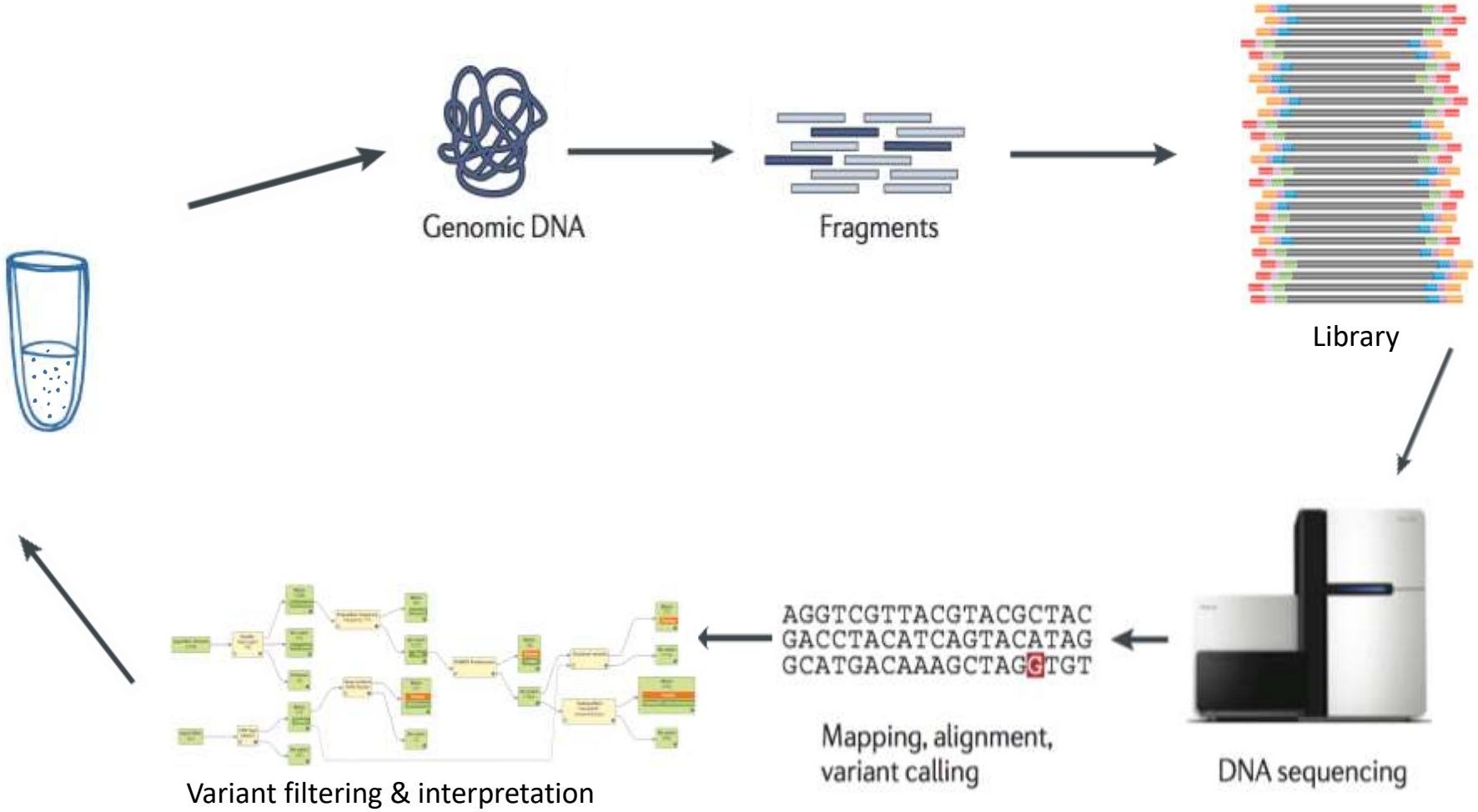
8* weeks

Library prep and sequencing

Data Analysis: What we offer

- Always **demultiplexed** (separate files for separate samples), and **mapped** for species with references genome
- Differential expression analysis included in price of RNAseq library prep
- **Custom** project possible
- **Cloud-enabling services**

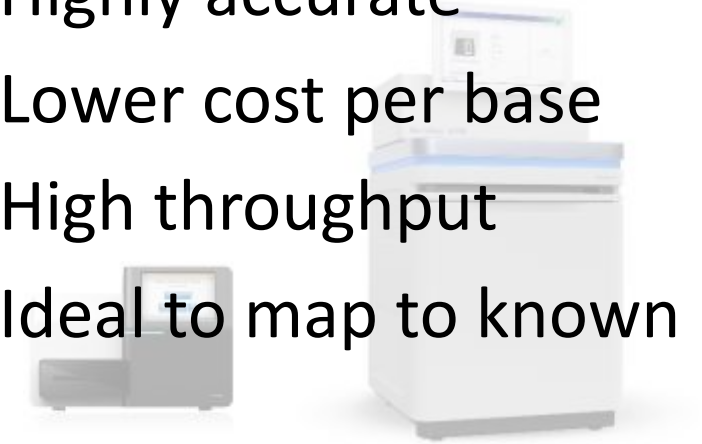
NGS workflow



Two main types of NGS

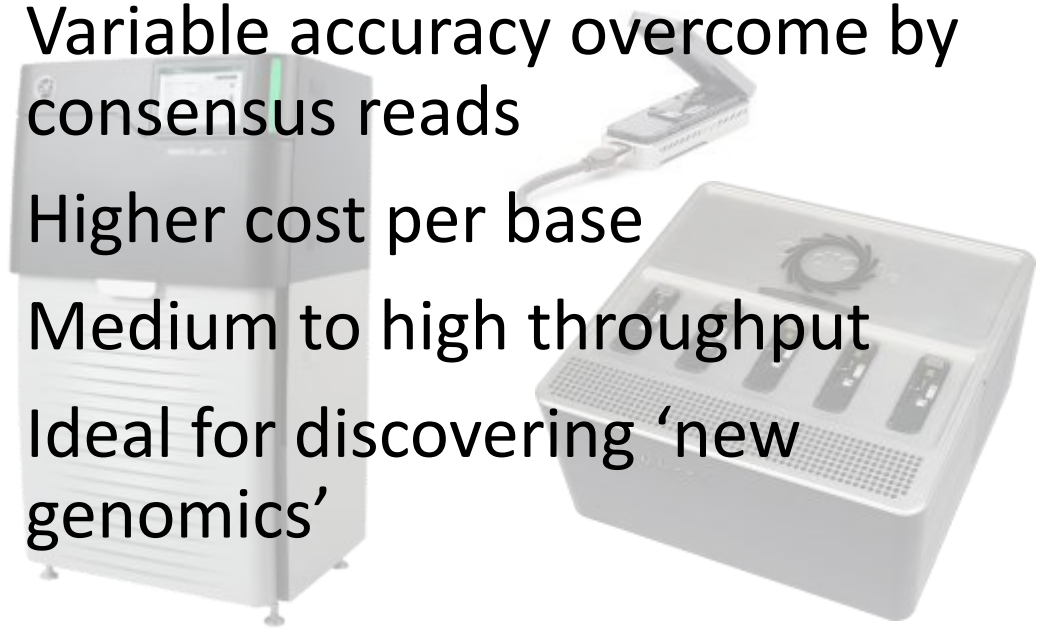
Short read: Illumina

- From from 35 to 500 bp
- Highly accurate
- Lower cost per base
- High throughput
- Ideal to map to known reference



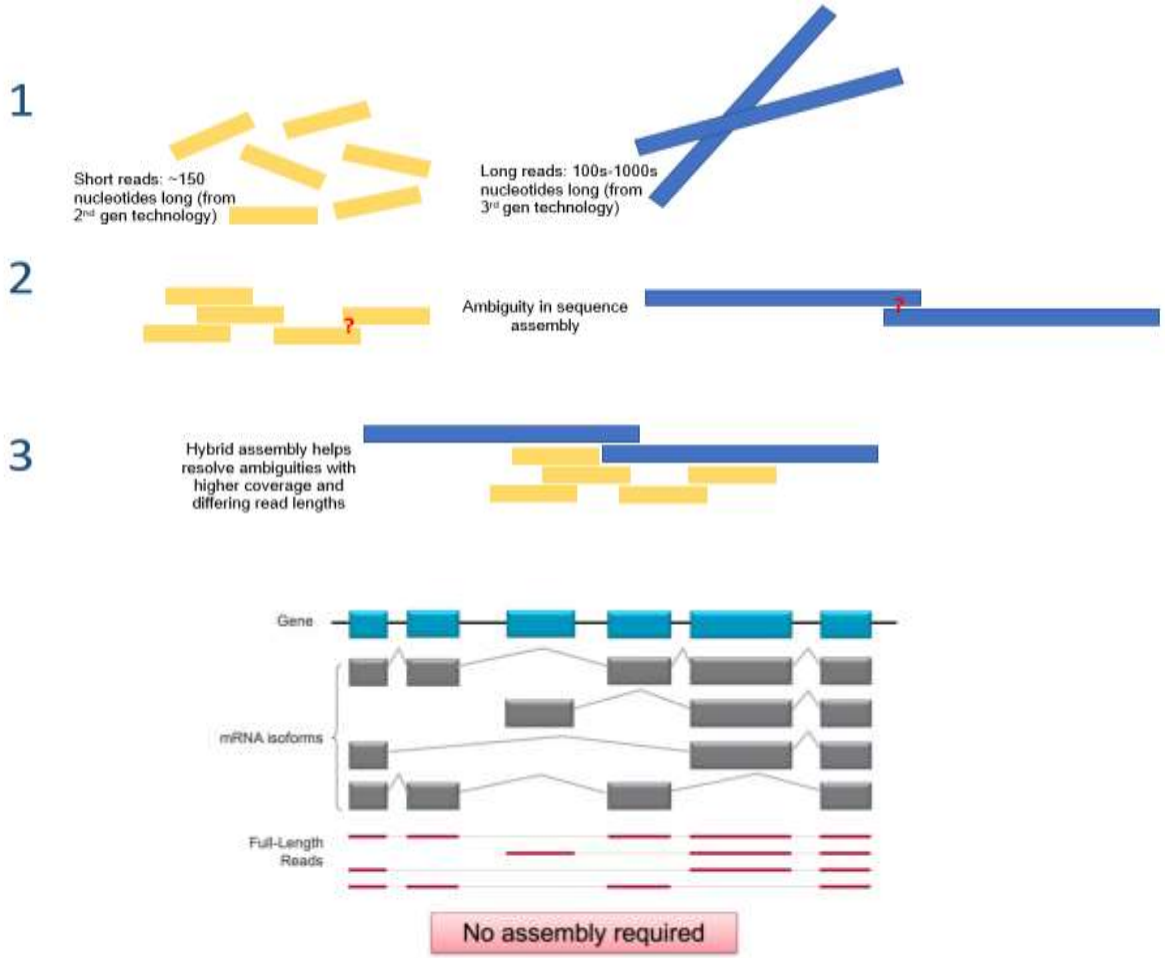
Long read: Pacbio and Nanopore

- From 500 bp to 500 kb
- Variable accuracy overcome by consensus reads
- Higher cost per base
- Medium to high throughput
- Ideal for discovering ‘new genomics’



Combining short- and long-read sequencing

1. **Genome Sequencing**
   - **Cost-efficient** *de novo* assembly for genomes or difficult regions
   - **Complementary**
   - HiC & Genome Mapping
2. **Transcriptome Sequencing**
   - Full transcripts
   - Splice variants
   - Isoforms and gene-fusions
3. **Epigenome**
   - Methylation signals



The genomics revolution in non-model organisms…

- NGS used for ecological and non-model questions
- Genetic mechanisms of ecological interactions with other species, the environment and perturbations
- New tools for old questions

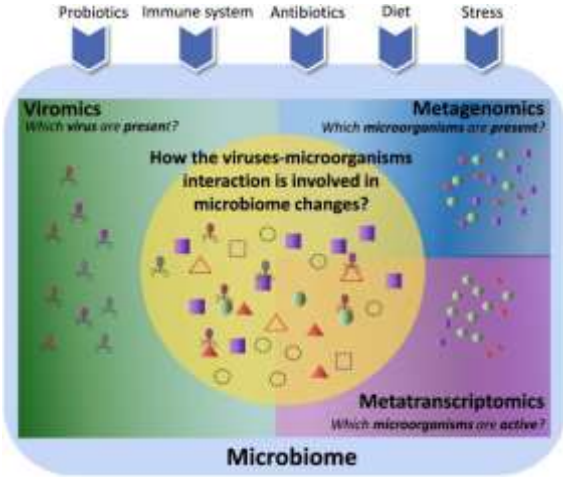
What can we do for you at GC?

- **Genetic marker** development
- **Species** identification
- **Stock** identification / traceability
- **Sequencing** and **Genotyping** service
- **Ecotoxicogenomics**
- **Breeding/Aquaculture/Wildlife** genetics
- **Bioinformatics** consulting

BIOGENOMICS
Bioinformatic and Genomic Identification - Consulting Services

Transcriptomics (short & long reads combined)

1. Transcriptome assembly of non-model organisms
   (Full transcripts IsoSeq or deep RNASeq)
2. Differential Gene Expression analysis
   (RNAseq or QuantSeq)
3. Single-cell transcriptomics
   (10X genomics)
4. Metatranscriptomics
   (functional component of microbial communities)



RNAseq: What is my scientific question?

Quantify

- ✓ Difference between conditions
- ✓ Change over time

Identify

- ✓ De novo transcript assembly
- ✓ Fusion genes
- ✓ Isoform

Expression profile seq

- Partial transcript
- Few reads

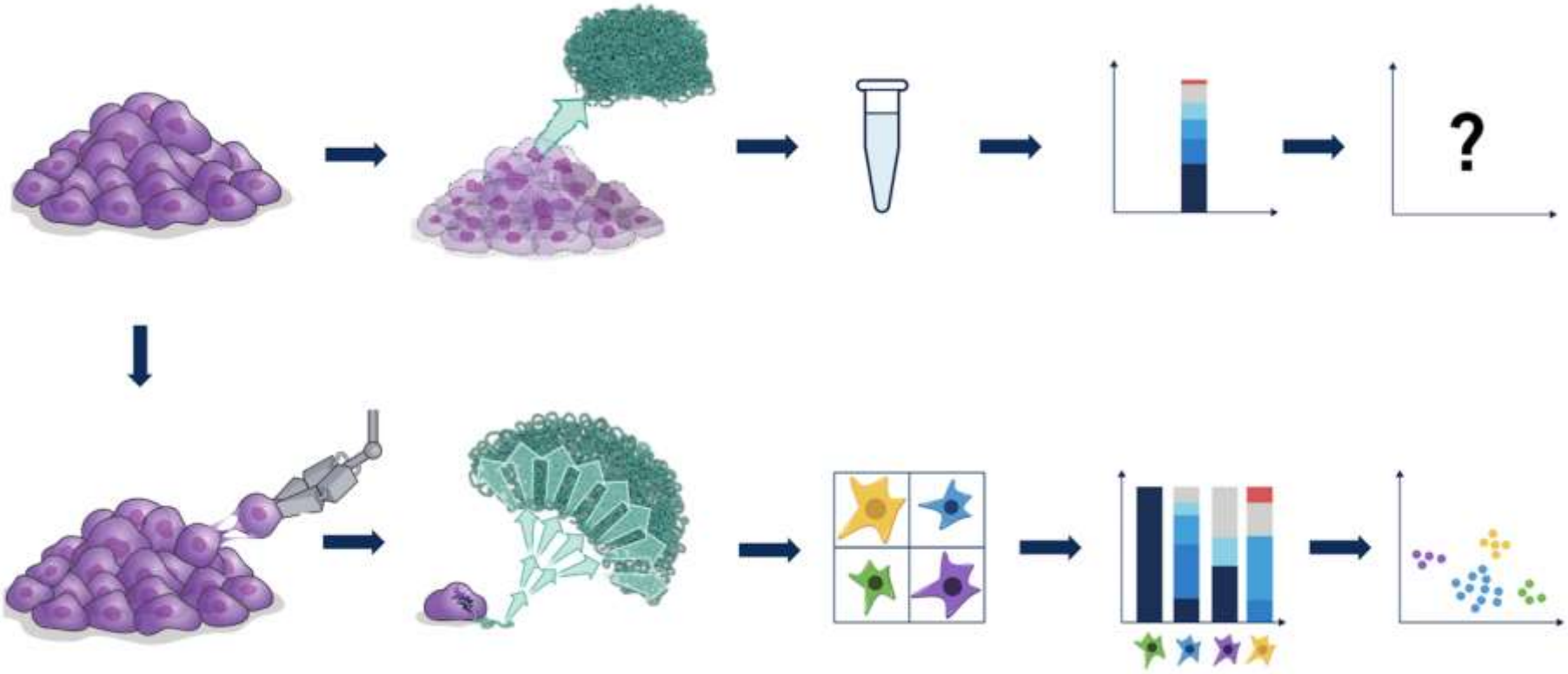
Whole transcriptome seq

- Full transcript
- Many reads

RNAseq: different library prep methods

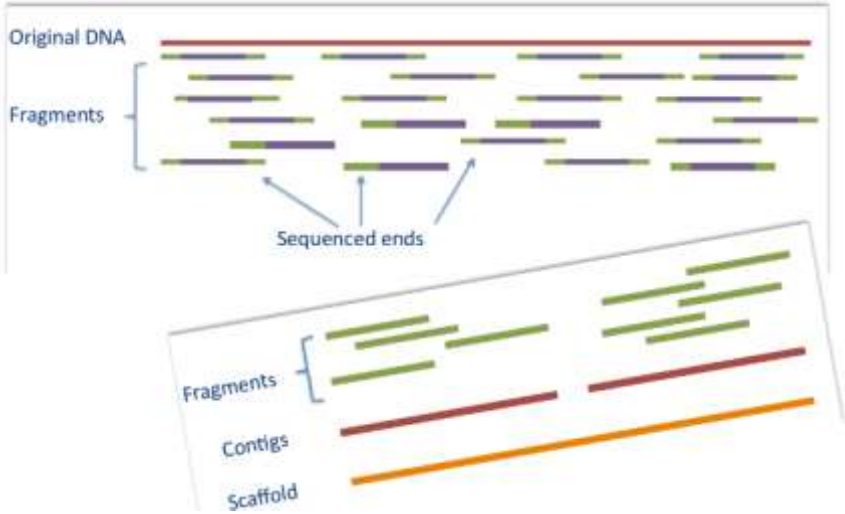
	Differential expression	Whole transcript, fusion, isoforms	Small RNA	Illumina compatible
Lexogen QuantSeq 3' mRNA	✓			✓
Lexogen Small RNA seq	✓		✓	✓
Illumina TruSeq stranded mRNA	✓	✓		✓
Illumina TruSeq stranded total RNA	✓	(✓)	✓	✓
IsoSeq	(✓)	✓		

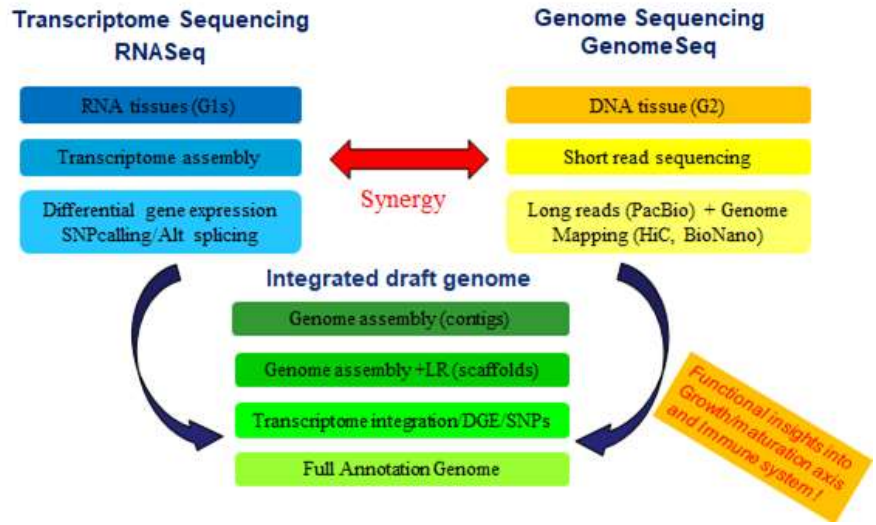
Single cell RNA-seq as a complementary technique to bulk RNA-seq



Whole Genome *de novo* assemblies

1. Organelle (circular) assemblies (mtDNA; Chloroplast DNA)
2. Microbial genome assembly (single strains or metagenomes)
3. Metazoa genomes (combination of sequencing technologies)



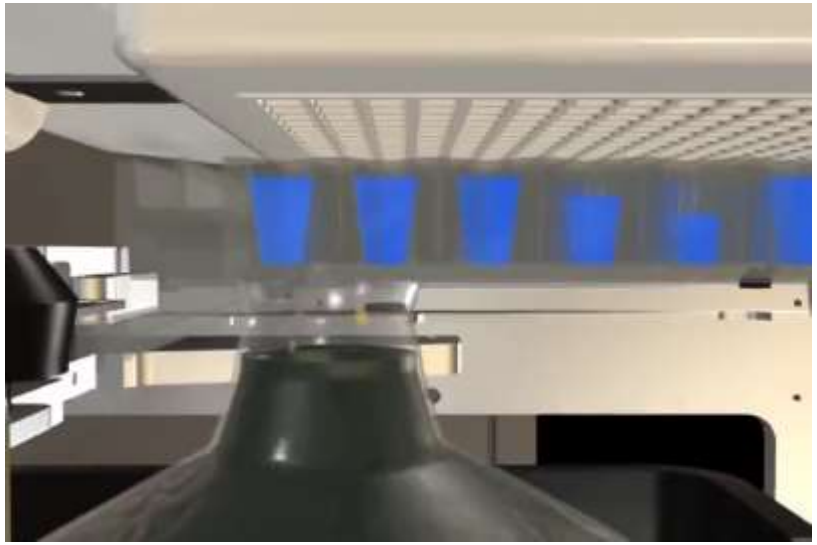


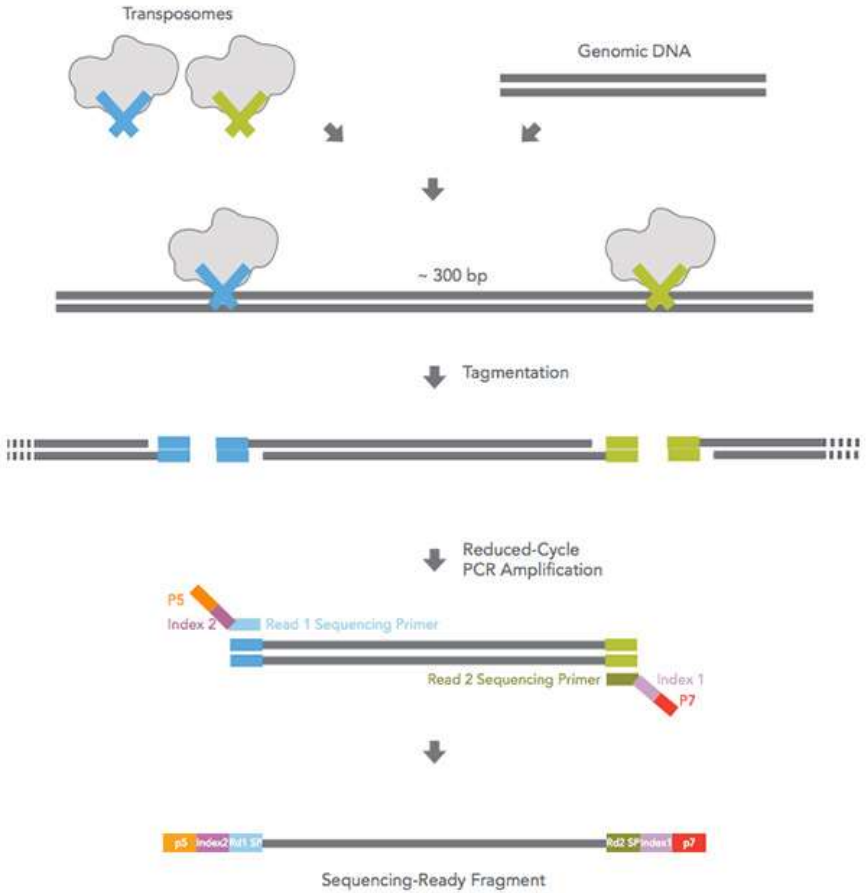
Whole genome sequencing eukaryotes

- Short reads only:
  - SNV and CNV detection
  - For known reference genomes
  - Validated pipeline for human genome
- Long reads (supplemented with short reads)
  - For SV detection (and SNVs and CNVs)
  - Sequencing difficult regions
  - De novo assembly of genomes
  - Further tools developed for new assembly strategies

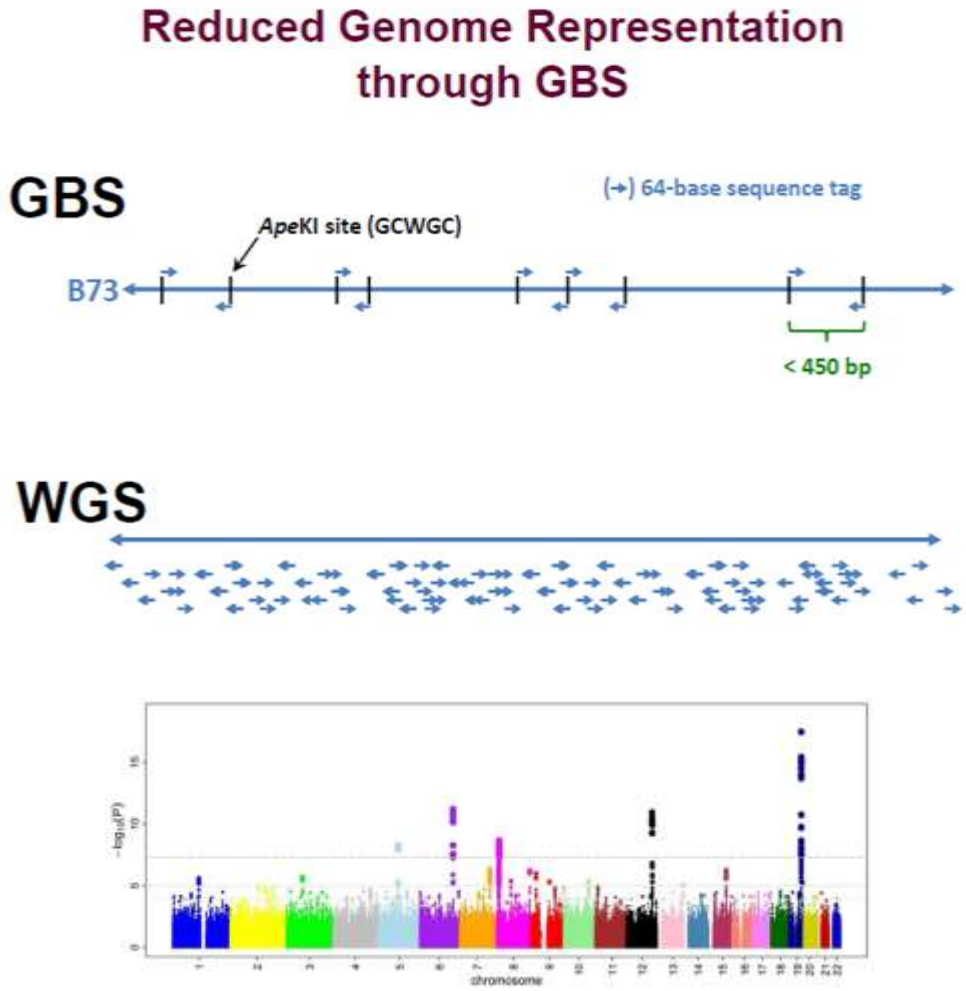
Whole genome sequencing bacteria

- Nextera XT on Echo Labcyte
  - Proven method
  - Scaled to very low volumes
  - In 384-well plates





Genotyping-by-Sequencing (GBS and RAD-Seq)



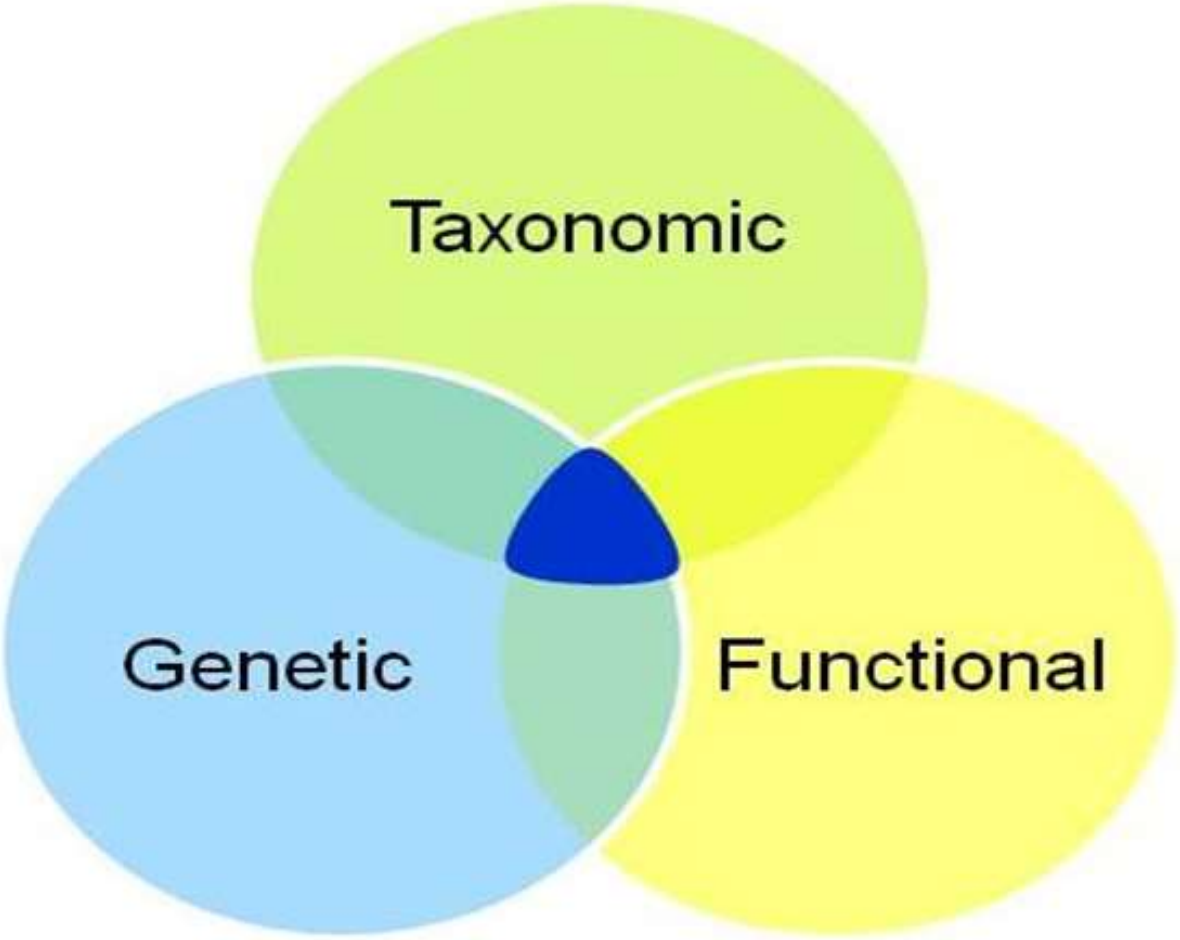
1. Population genomics and connectivity in North Sea fish populations (FWO)
2. Phylogenomics of cichlids in lake Tanganika (BRAIN)
3. GWAS and QTL Mapping of breeding traits in Orchids (O&O project with industry)
4. Population genomics and connectivity around Antarctic (BELSPO)
5. Fisheries forensics and authenticity testing
6. (dd)RAD-seq analyses for population structure, farm characterization and escapee traceability (Sea Bass, bream, turbot) (EU project)

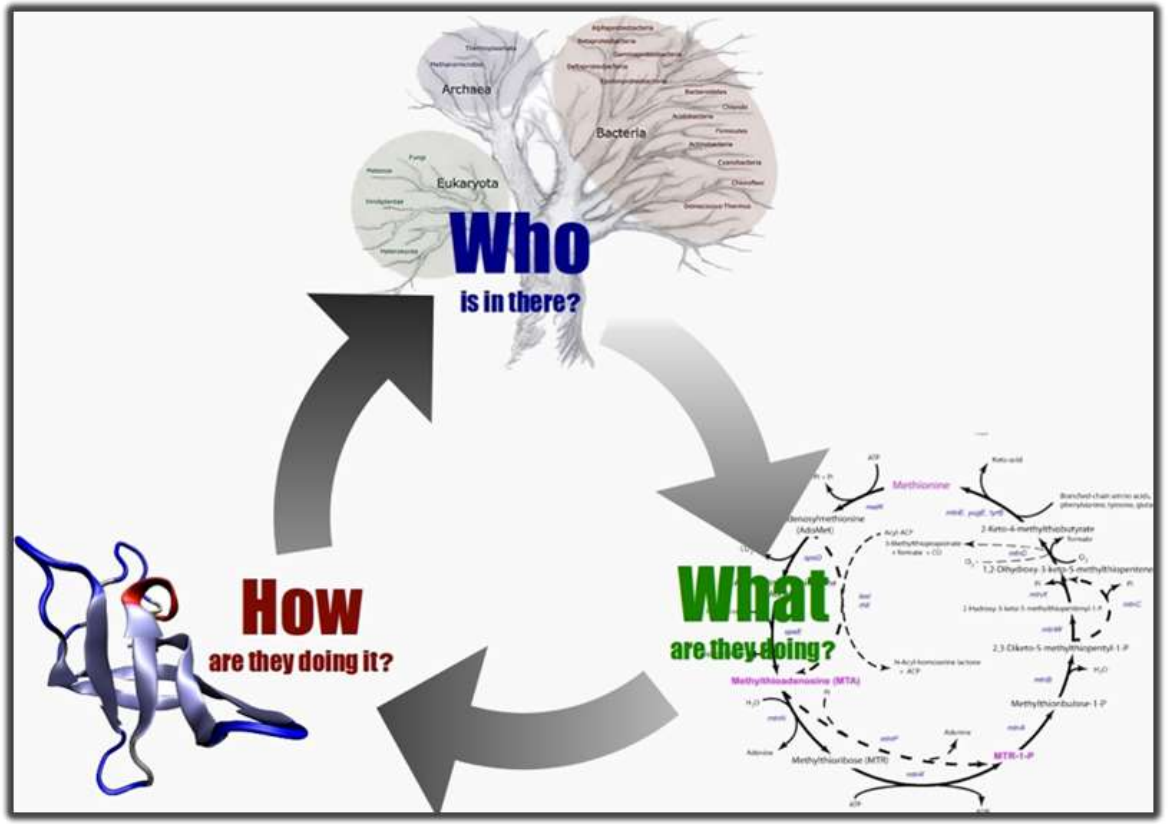
BIOGENOMICS

Starting from a DNA soup…

- Looking what biological groups, species, resistance genes, or variants are present in a DNA sample is called **metagenomics (https://www.ebi.ac.uk/metagenomics/)** and **metabarcoding (http://www.boldsystems.org/)**
- “Sample the entire zoo in a blender to discover which species are present”
- Targeted or shotgun approaches

Different technologies, different questions





Different technologies, different questions

- Whole genome approaches
  - Which genes are present?
  - What functions do genes have (transcriptomics)?
  - How does the gene pool react to certain conditions?
- Targeted approaches
  - Which species are present?
  - Which specific gene variants are present?

RESEARCH ARTICLE

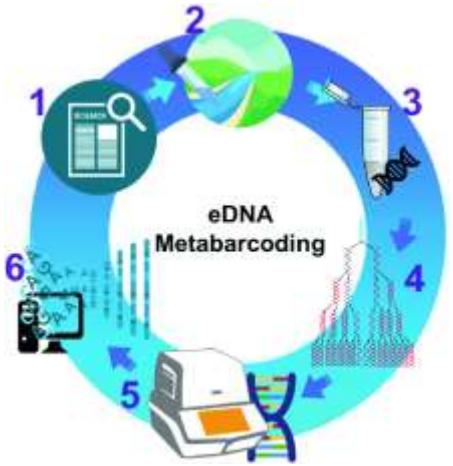
Environmental Genome Shotgun Sequencing of the Sargasso Sea

J. Craig Venter,[1]* Karin Remington,[1] John F. Heidelberg,[1] Aaron L. Halpern,[2] Doug Rusch,[2] Jonathan A. Eisen,[3] Dongying Wu,[3] Ian Paulsen,[3] Karen E. Nelson,[3] William Nelson,[3] Derrick E. Fouts,[3] Samuel Levy,[2] Anthony H. Knap,[4] Michael W. Lomas,[4] Ken Nealson,[5] Owen White,[3] Jeremy Peterson,[3] Jeff Hoffman,[1] Rachel Parsons,[6] Holly Baden-Tillson,[1] Cynthia Pfannkoch,[1] Yu-Hui Rogers,[4] Hamilton O. Smith[1]

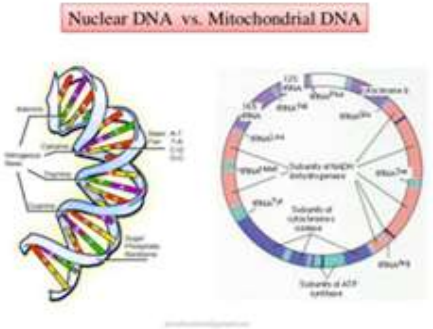
We have applied "whole-genome shotgun sequencing" to microbial populations collected en masse on tangential flow and impact filters from seawater samples collected from the Sargasso Sea near Bermuda. A total of 1.045 billion base pairs of nonredundant sequence was generated, annotated, and analyzed to elucidate the gene content, diversity, and relative abundance of the organisms within these environmental samples. These data are estimated to derive from at least 1800 genomic species based on sequence relatedness, including 148 previously unknown bacterial phylotypes. We have identified over 1.2 million previously unknown genes represented in these samples, including more than 782 new rhodopsin-like photoreceptors. Variation in species present and stoichiometry suggests substantial oceanic microbial diversity.

Metabarcoding (COI)

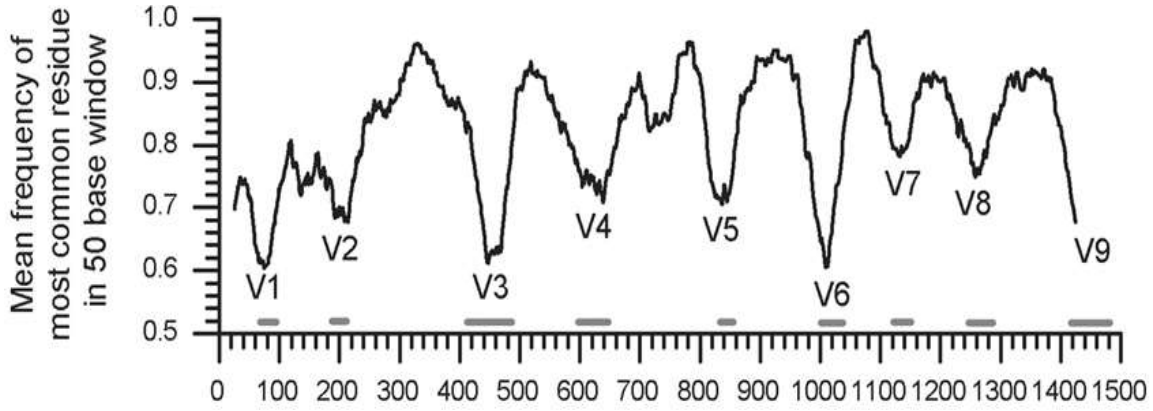
- Identifying species through sequencing unique but conserved markers
  - Conserved enough to be present in all species
  - Diverse enough to be different for each species
- From single sample identification
- …to eDNA analysis

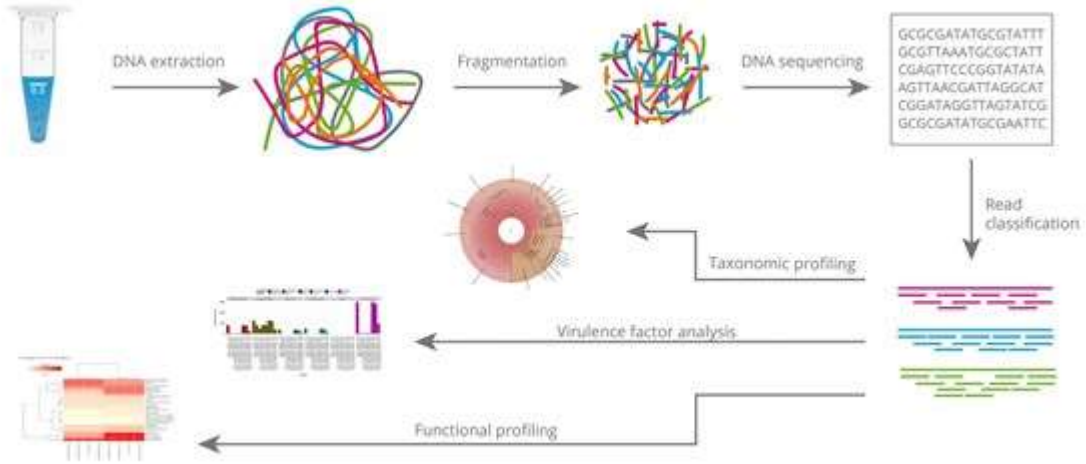


- Samples:
  - Bulk (mix, pollen, scat)
  - eDNA (water, soil)
- PCR and NGS dependent
- Targets 1 of a few genes



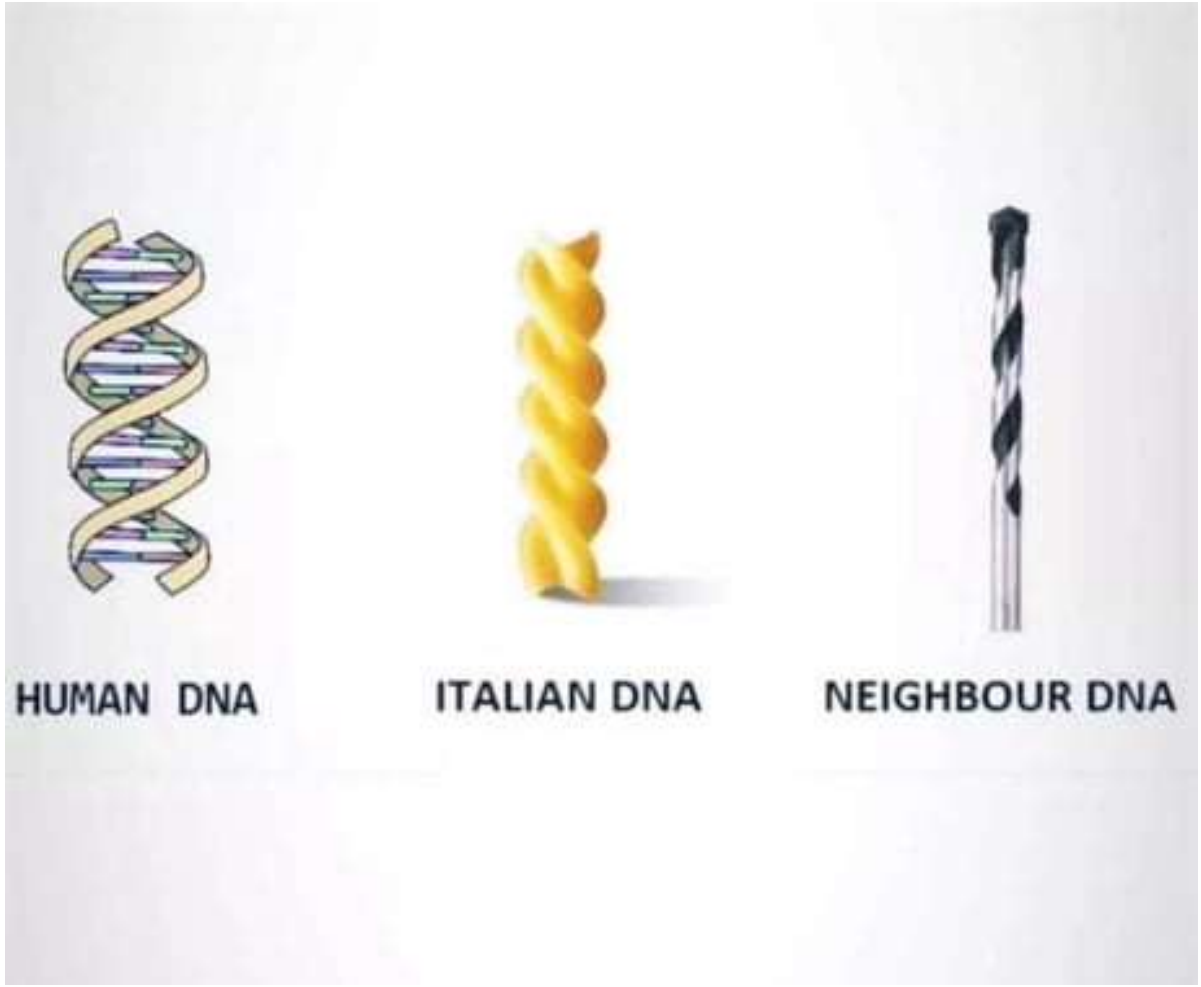
Metagenomics (16s amplicon sequencing & WGS)





1. Monitoring drinking water quality (De Watergroep + PIDPA) using Miseq + PacBio
2. Skin and mucus microbiome (fish mucus microbial diversity, immunity interaction metatranscripome)
3. Gut microbiome analysis and link to diet composition (metabarcoding) and environmental factors
4. Microbiome of Ostracoda living in extreme environments (Pollution, temperature, salinity)
5. Comparison or contemporary and ancient fish microbiome community (archived material, link to climate change)
6. Human gut microbiome (PacBio/Long read of 1600 bp)
7. Microbiome exchange between species (conspecific vs heterospecific matings, host-species adaptation)
8. Full genome metagenomics (various projects)

Thank you!



Today's special guests

Prof. Ellen Decaestecker

- PhD in Biology (KU Leuven and U Basel (Host-parasite co-evolution)
- Postdoc at Centre for Immunity, Infection and Evolution (U. Edinburgh)
- Professor at KU Leuven/KULAK (host-microbiome research: microbiome transplants, metagenomics and host epigenetics)

Dr. Shinjini Mukherjee

- Master Biotechnology (India)
- PhD University Helsinki (bioremediation petroleum hydrocarbons)
- Marie-Curie Postdoc KU Leuven (microbial responses urbanization)
- Currently postdoc KU Leuven on eco-evolutionary microbiome dynamics in Daphnia spp.