

Marine Bioinformatics

in the 21st Century

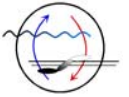
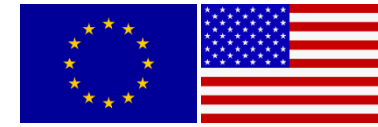


Frank Oliver Glöckner
Chair of the EC-US working group on
Marine Genomics

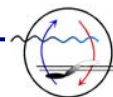


JACOBS
UNIVERSITY

EC-US Workshops

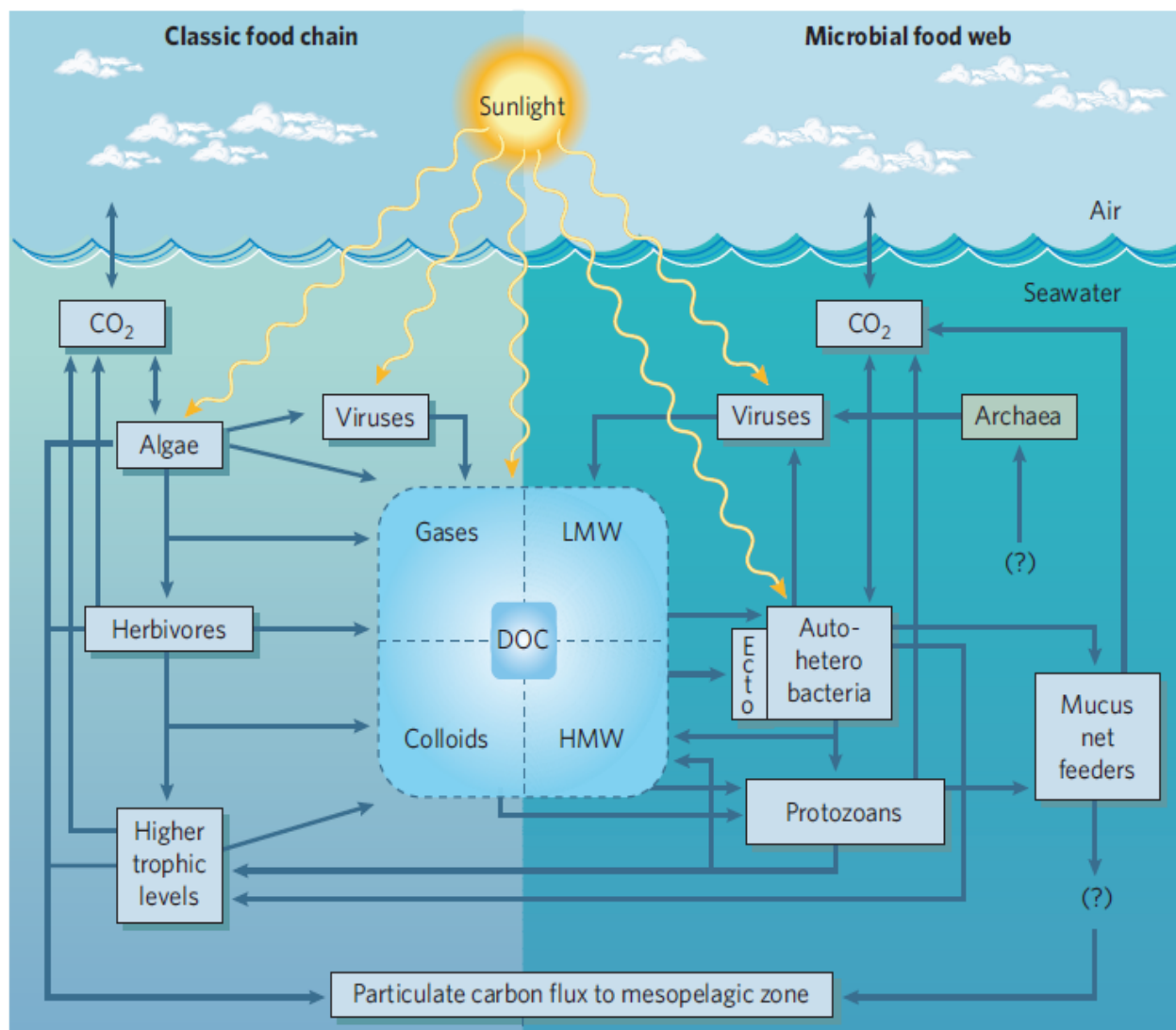


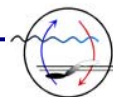
- ▶ **Genomic Approaches for Studying the Marine Environment and Resources**
Bremen, 12-13 May 2005
- ▶ **Cyberinfrastructure Resources for Genome-Enabled Research on Microbial Life and the Marine Environment**
Arlington, September 9-11, 2007
- ▶ **At the Interface of Marine Microbial Ecology and Biotechnological Applications**
Principality of Monaco, October 12-14, 2008
- ▶ **High Throughput Technologies and their Application/Influence on Marine Microbial Genomics and Biotechnology**
Washington, October 10-12, 2010



The Marine Ecosystem: A Complex System

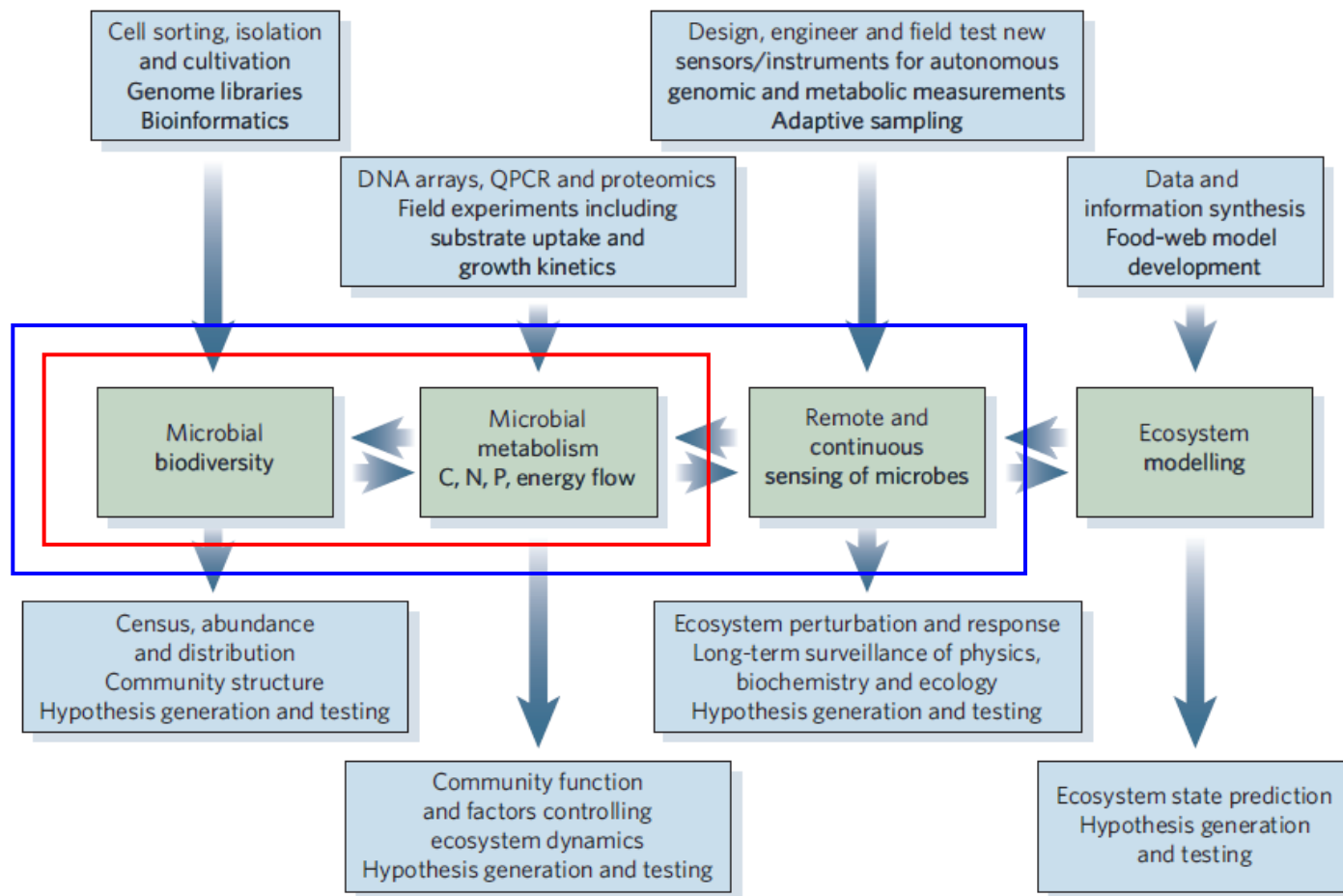
► Where are we?

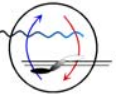




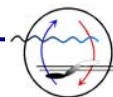
Ecosystems Biology

► What do we want?





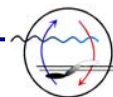


- ▶ **So what?**
- ▶ **Nothing really new...**
- ▶ **What makes the difference?**

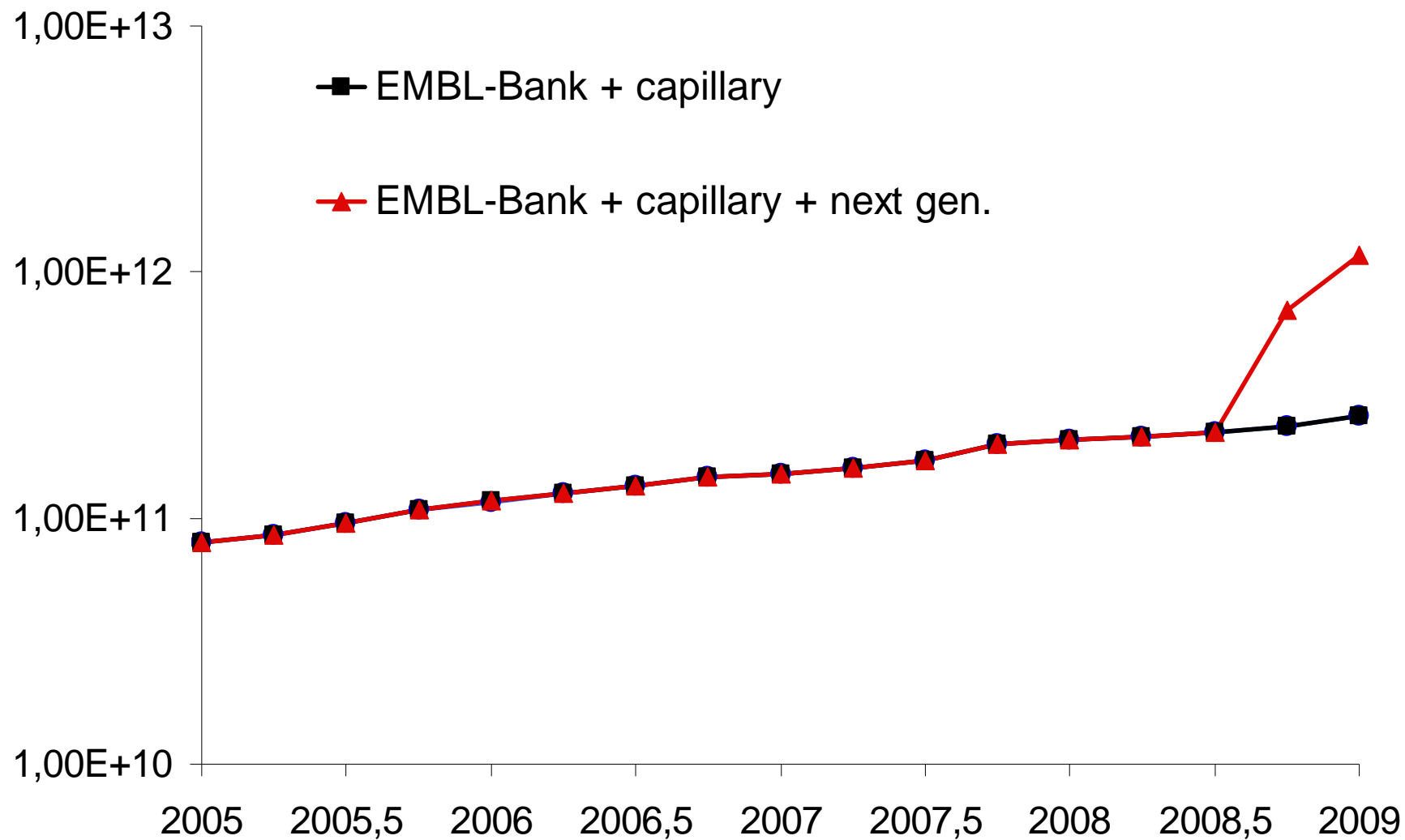


(R)Evolution in Sequencing Technology

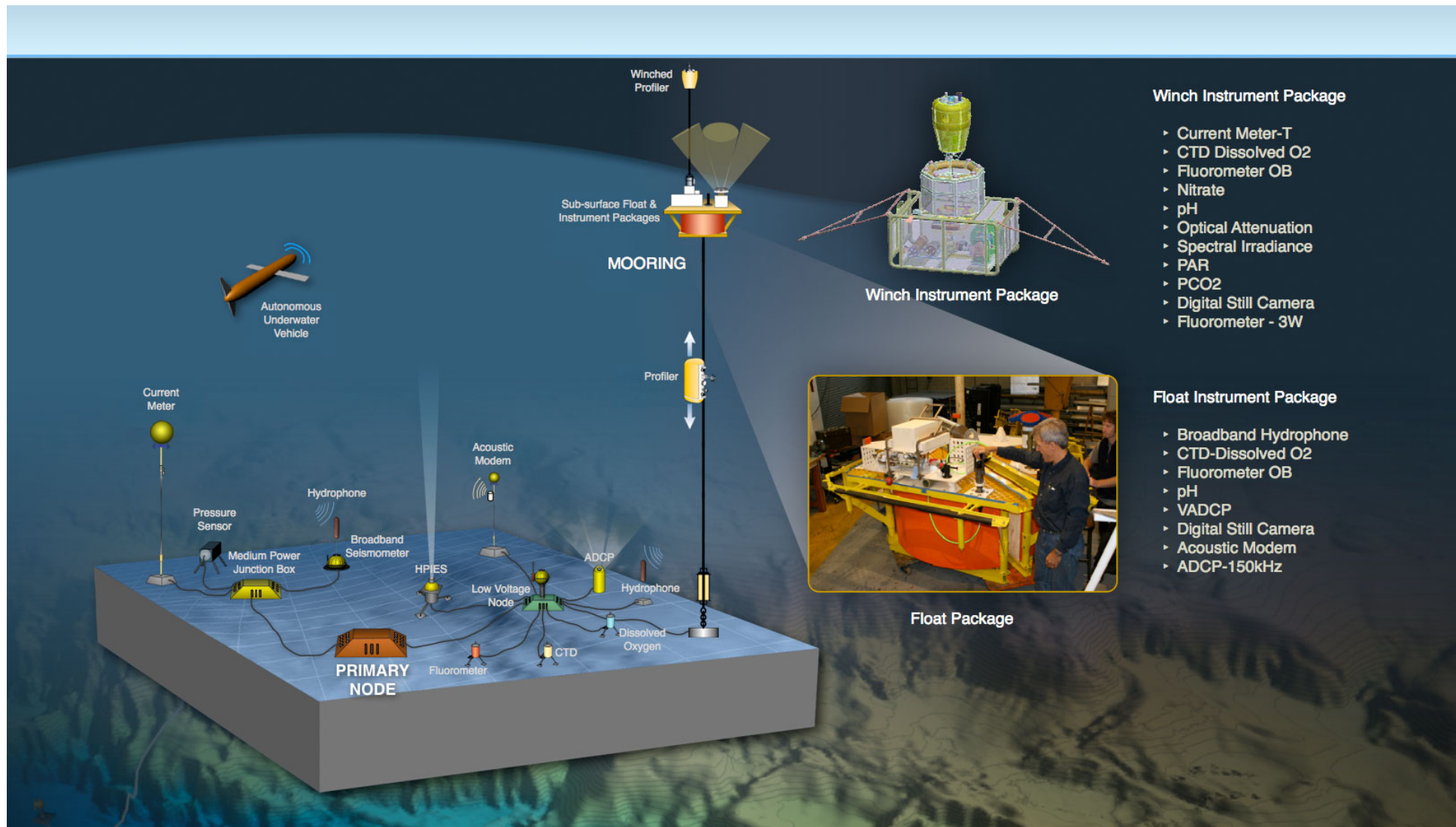
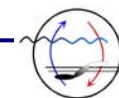
 <p>Sanger capillary sequencer</p>	 <p>454 (Roche)</p>	 <p>Illumina III</p>	 <p>SOLiD (ABI)</p>
 <p>Helicos tSMS (Helicos)</p>	 <p>Polonator (Dover)</p>	 <p>DNB Sequencing</p>	 <p>ZMW Sequencing</p>
 <p>Illumina HiSeq 2000</p>	 <p>Oxford Nanopore</p>	 <p>NAbsys</p>	 <p>US Genomics</p>
 <p>Halcyon</p>	 <p>ZS Genetics</p>	 <p>IBM DNA Transistor</p>	<p>T.B.C.</p>



Next Generation Sequencing



Data Acquisition

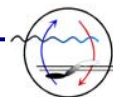


Winch Instrument Package

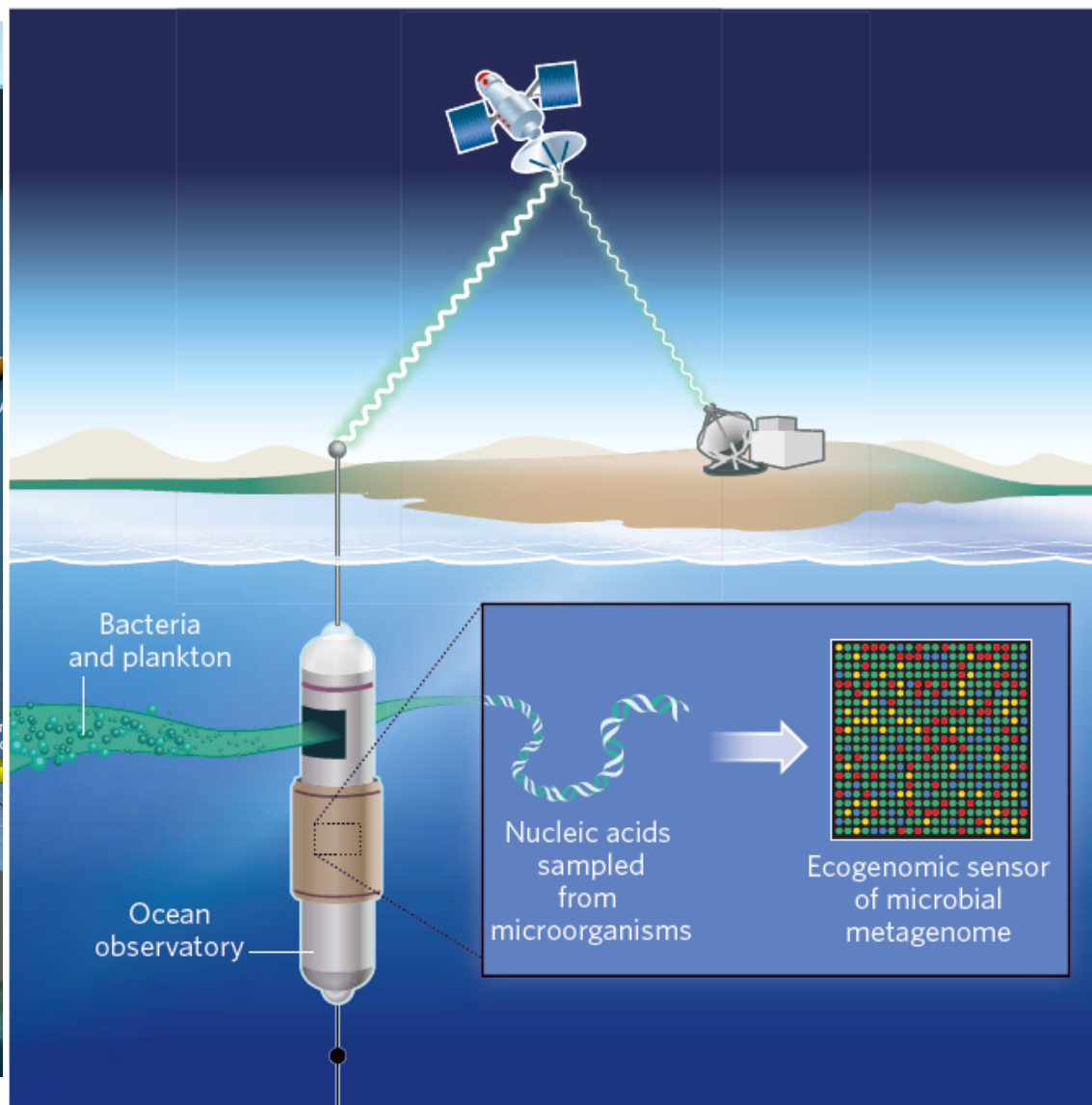
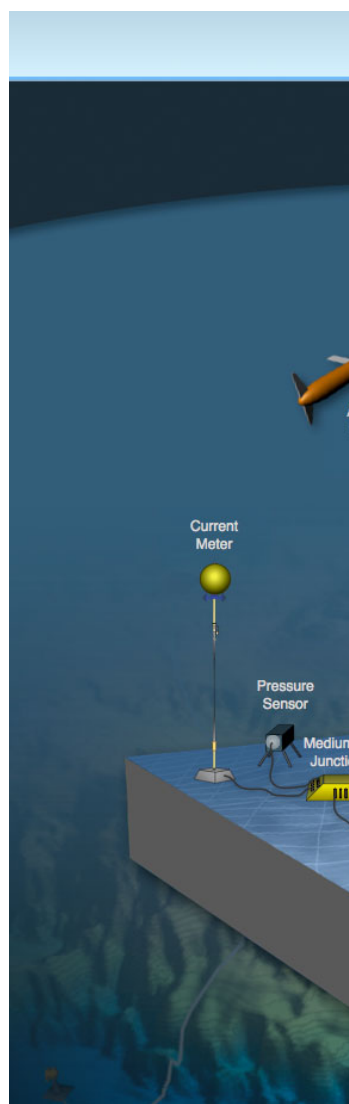
- Current Meter-T
- CTD Dissolved O2
- Fluorometer OB
- Nitrate
- pH
- Optical Attenuation
- Spectral Irradiance
- PAR
- PCO2
- Digital Still Camera
- Fluorometer - 3W

Float Instrument Package

- Broadband Hydrophone
- CTD-Dissolved O2
- Fluorometer OB
- pH
- VADCP
- Digital Still Camera
- Acoustic Modem
- ADCP-150kHz



Data Acquisition

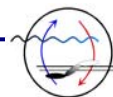


Winch Instrument Package

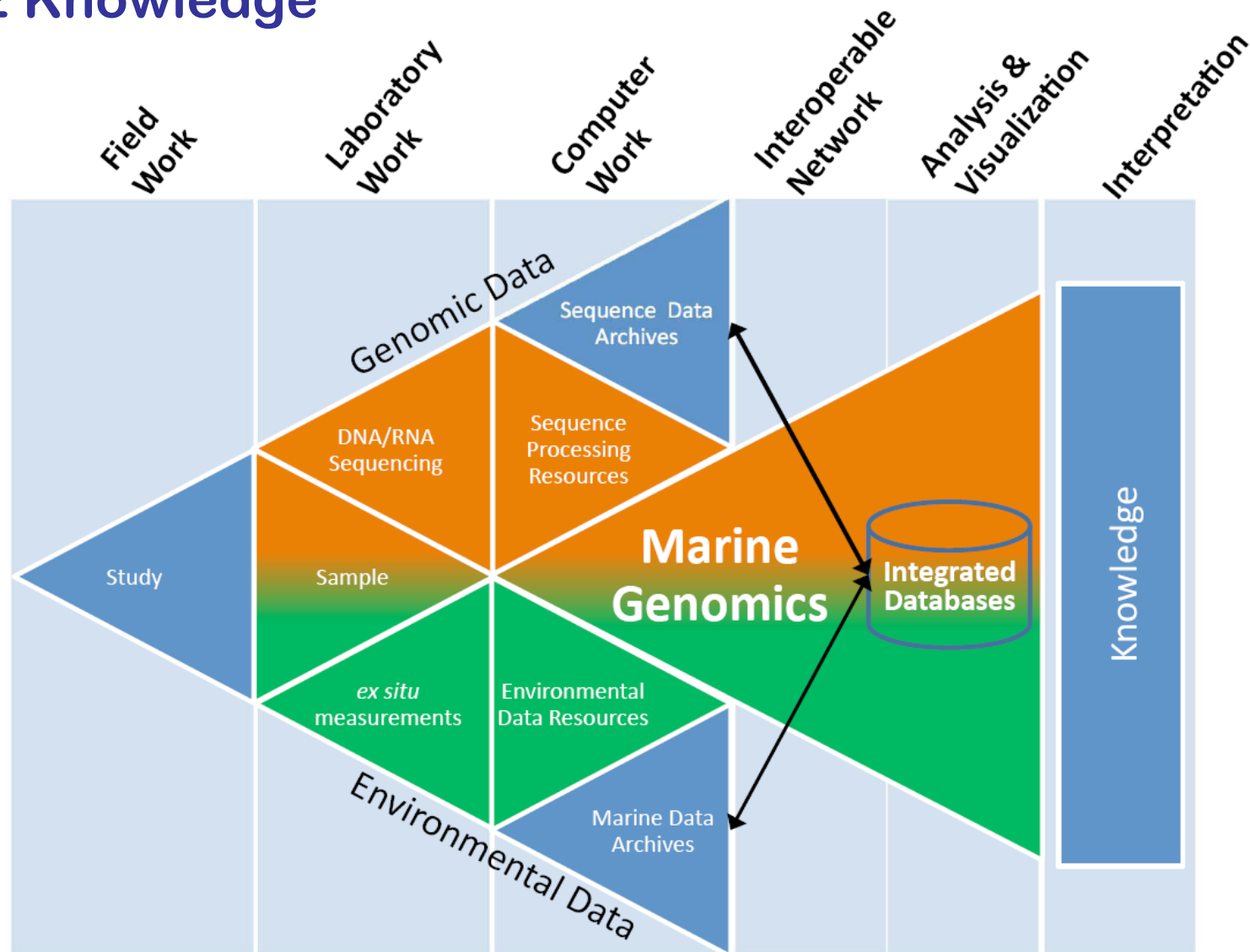
- ▶ Current Meter-T
- ▶ CTD Dissolved O₂
- ▶ Fluorometer OB
- ▶ Nitrate
- ▶ pH
- ▶ Optical Attenuation
- ▶ Spectral Irradiance
- ▶ PAR
- ▶ PCO₂
- ▶ Digital Still Camera
- ▶ Fluorometer - 3W

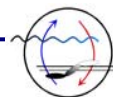
Float Instrument Package

- ▶ Broadband Hydrophone
- ▶ CTD-Dissolved O₂
- ▶ Fluorometer OB
- ▶ pH
- ▶ VADCP
- ▶ Digital Still Camera
- ▶ Acoustic Modem
- ▶ ADCP-150kHz



Data 2 Knowledge

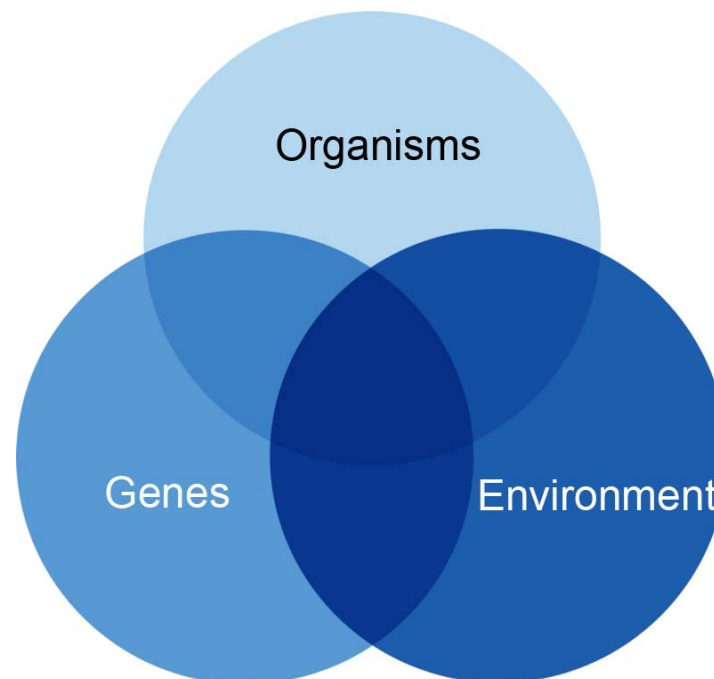




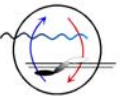
Transferring Data into Biological Knowledge

▶ Need to

- Standardize
- Store
- Analyse
- Compare
- Integrate
- Visualize
- Interpret

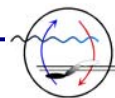


the data and make them accessible for scientists,
public and policy makers



Bioinformatics in 21st Century

- ▶ **The major high level recommendations emerging from the EC-US workshops on Marine Genomics**
- ▶ The most significant challenges that the Task Force could help to address
- ▶ The highest priorities of the community for the next steps based on these

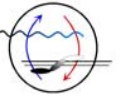


Recommendations from the EC-US workshops



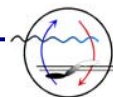
▶ Bremen, Arlington, Monaco

- Improved assembly and annotation tools
- Standards for data analysis and contextual data
- Integrative bioinformatic approaches
 - ◆ on the level of the ecosystem
 - ◆ on the level of (meta)genomics, pure cultures and single cell genomics
- Functional assessment of genes with unknown functions
- Specialized e-infrastructure for marine (meta)genomics
- Training in environmental bioinformatics

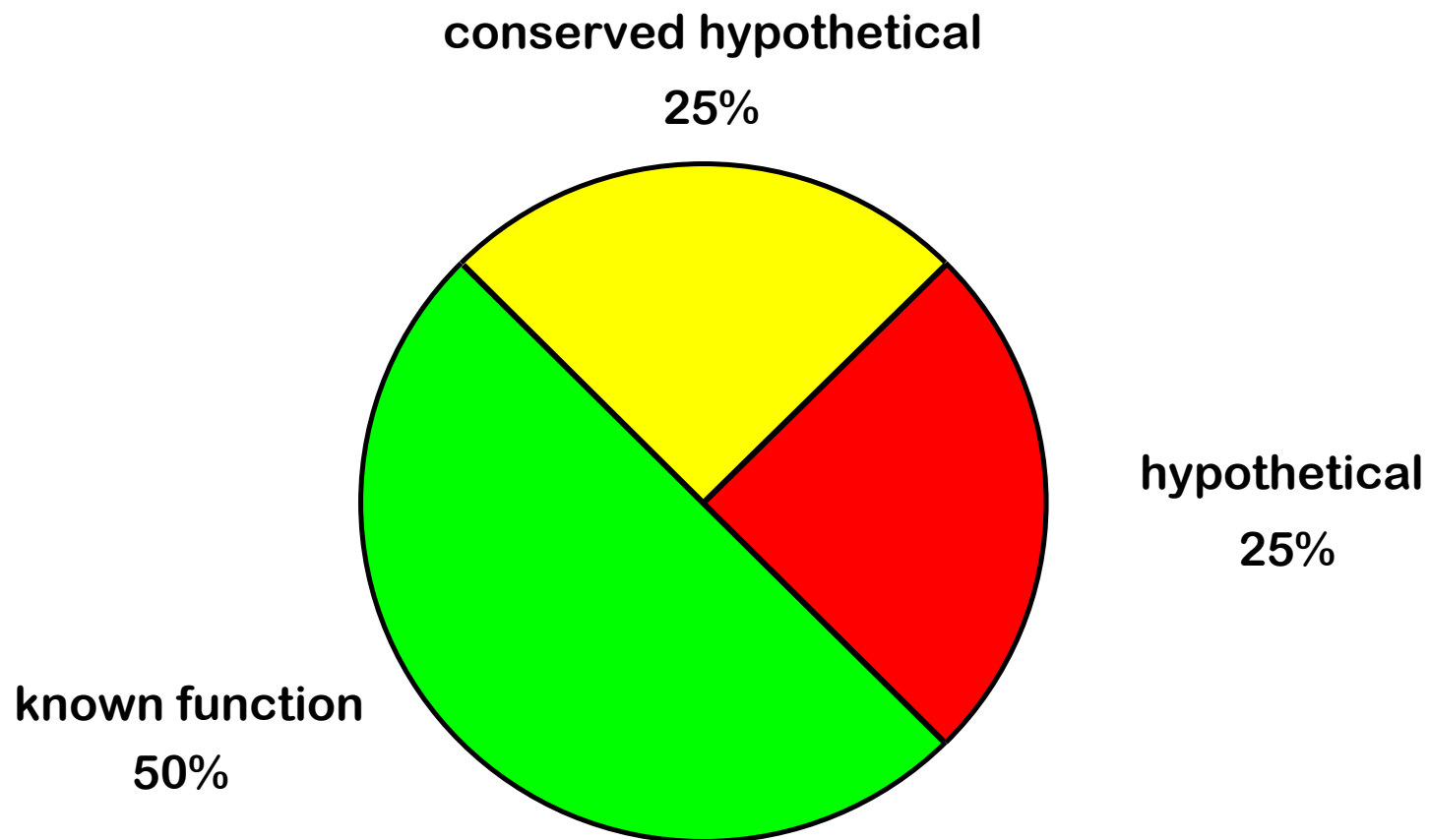


Bioinformatics in 21st Century

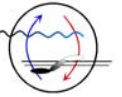
- ▶ The major high level recommendations emerging from the workshop
- ▶ **The most significant challenges that the Task Force could help to address**
- ▶ The highest priorities of the community for the next steps based on these



Genes with Unknown Functions

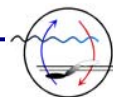


- ▶ **Hypothetical protein:** Predicted gene with no similarity in public databases – unique for the sample/organism
- ▶ **Conserved hypothetical protein:** Predicted gene with similarity to another gene(s) but lacking functional assignment

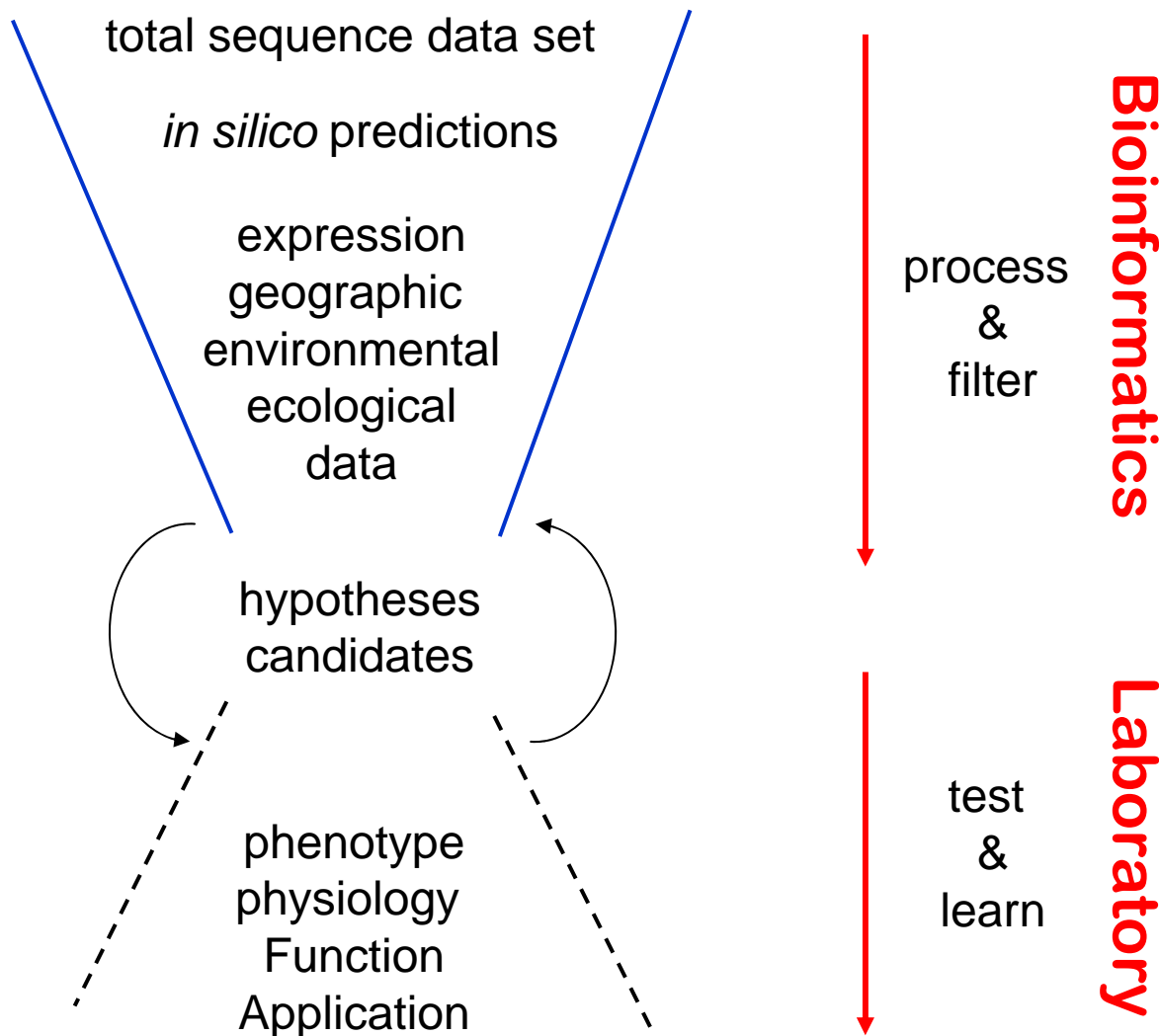
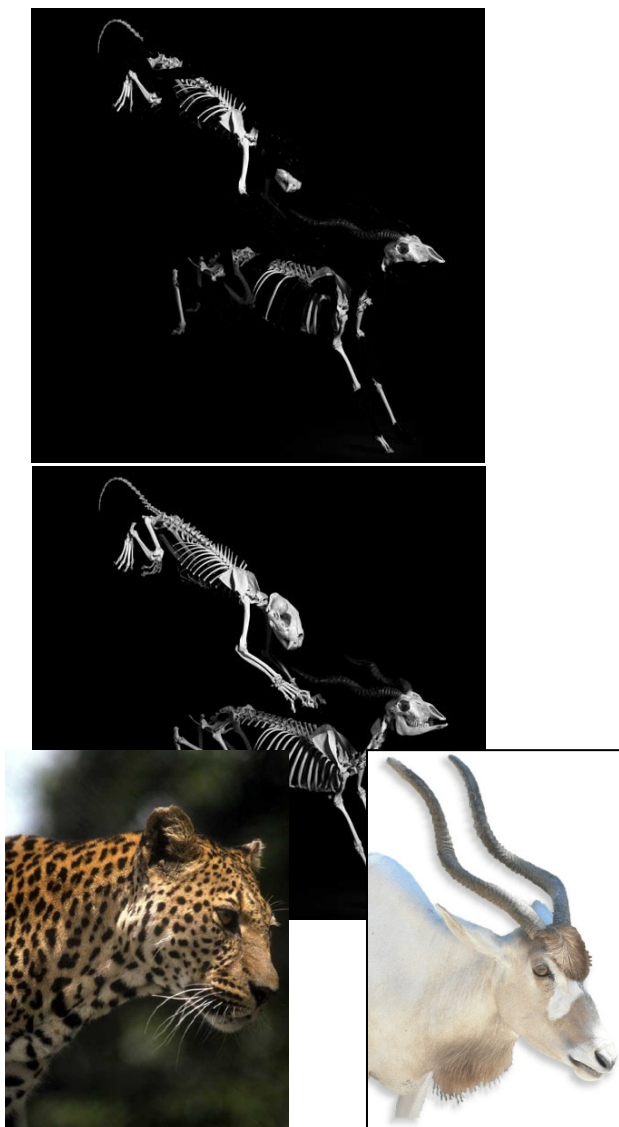


Bioinformatics in 21st Century

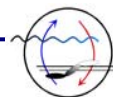
- ▶ The major high level recommendations emerging from the workshop
- ▶ The most significant challenges that the Task Force could help to address
- ▶ **The highest priorities of the community for the next steps based on these**



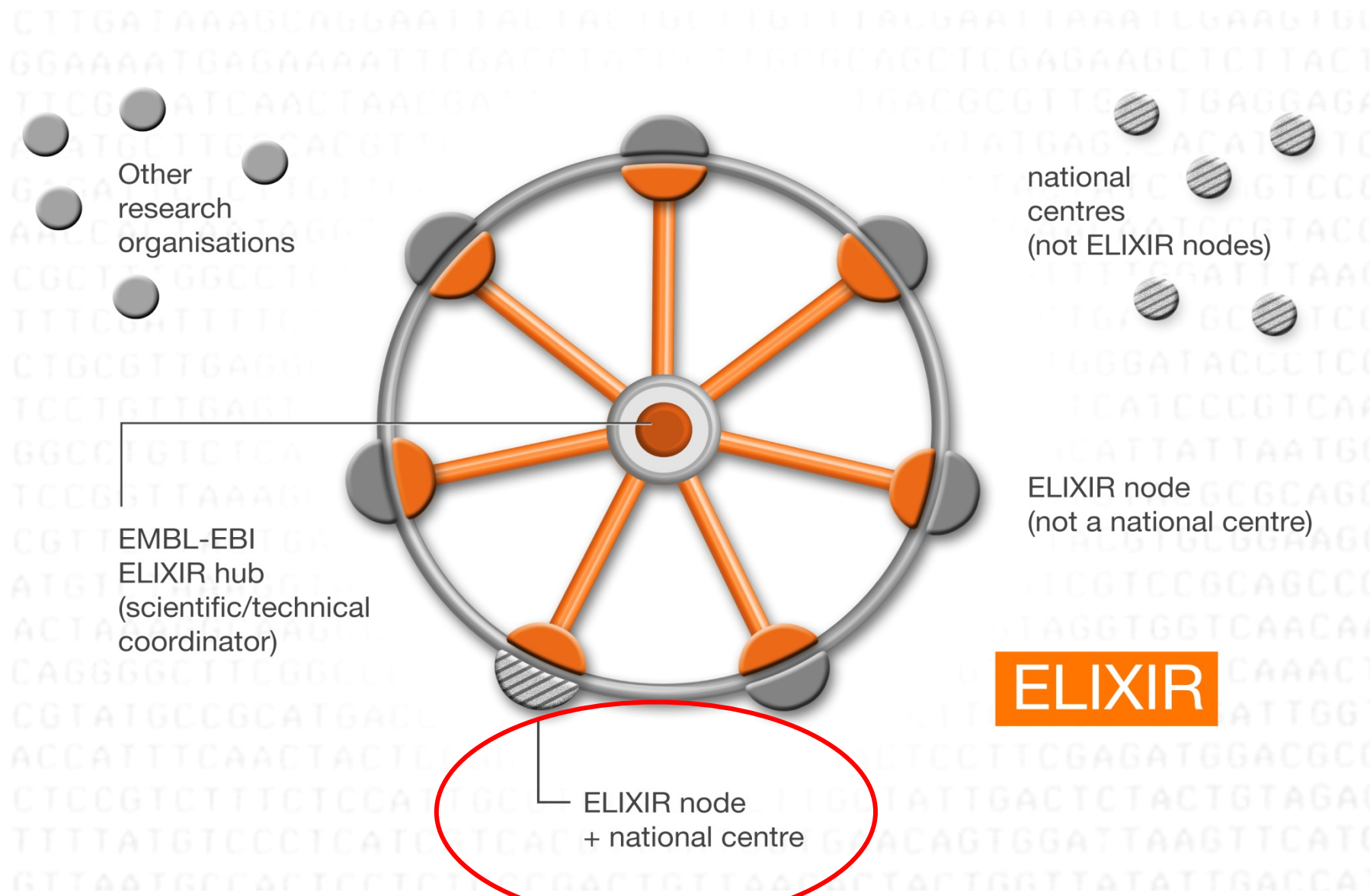
Data 2 Knowledge -> Data Integration & Cooperation

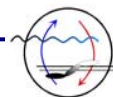


Redrawn after Frederking & Thaler, Evolution

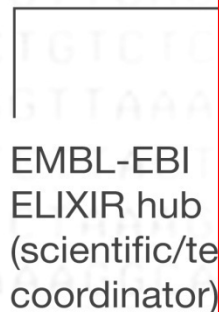


ELIXIR: A Specialised Node for Marine Bioinformatics



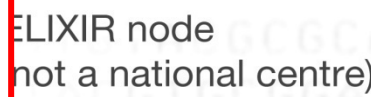


ELIXIR: A Specialised Node for Marine Bioinformatics



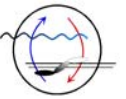
► Data management

- Specialised databases
- Data processing
- Quality management
- Standardization
- Data curation
- Data integration
- Visualization
- Data deposition
- Training



ELIXIR

+ national centre



Thanks for your attention

