Building a meiofauna reference database in the Gulf of Mexico

9th National Summit on Coastal and Estuarine Restoration and Management
W. Kelley Thomas
Hubbard Center for Genome Studies
University of New Hampshire, Durham

Joseph Sevigny

Francesca Leasi Ph.D.
Meiofauna species as ecological indicators

- Diverse, ubiquitous and abundant
- Short generation times
- Stimulate mineralization of organic matter and bacterial production
- Trophic Links

Bonaglia et al. 2014. *Nature Communications* 5(5133)
Meiofauna species as ecological indicators

Can we use Meiofaunal communities to:

1. Investigate the changes in species composition correlated with the Deep Water Horizon Oil Spill

2. Understand the functional differences among community members based on gene catalogs

- Low dispersal potential
- High sensitivity to anthropogenic inputs
- Different responses to environmental provocation

DNA Sequencing

Traditional Taxonomy

- Time consuming and expensive
- Limited taxonomic expertise
- Phenotypic plasticity
- Cryptic species

- Metabarcoding/eDNA – single gene(s)

Positives:
- Rapid and inexpensive
- Bioinformatic infrastructure

Negatives:
- Primer biases
- Taxonomic resolution
- Poor databases
Increasing the breadth of animal genome references

- What animals should we sequence?
- A few big fish or a bunch of decent ones?
1. Sample Collection and Identification

2. Selective, whole genome sequencing

3. Initial Assembly: Quality Assessment

4. Contamination Removal

Potential deeper sequencing

5. Final Assembly and Annotations

6. Submission to Genbank and Experiments
Improving References - Challenges

BITMaB workshop: January 2018

◊ Gathering Samples
  ◊ Too few taxonomists
    ◊ Too many meiofauna
  ◊ Identification and preservation
  ◊ DNA quantity and quality
  ◊ Selection for sequencing

◊ Sequencing and Assembly
  ◊ Every genome is essentially unknown
  ◊ Recalcitrant genome structure
    ◊ Repeat structure
    ◊ Heterozygosity

◊ Metagenomes
  ◊ Bacteria and other associates
Improving References - Challenges

- Gathering references.
  - Too few taxonomists
    - Too many meiofauna
  - DNA quantity and quality
  - Selection for sequencing

- Sequencing and Assembly
  - Every genome is essentially unknown
  - Often recalcitrant genome structure
    - Repeat structure
    - Heterozygosity

- Metagenomes
  - Bacteria and other associates

Improving References - Challenges

- Gathering references.
  - Too few taxonomists
    - Too many meiofauna
  - DNA quantity and quality
  - Selection for sequencing

- Sequencing and Assembly
  - Every genome is essentially unknown
  - Often recalcitrant genome structure
    - Repeat structure
    - Heterozygosity

- Metagenomes
  - Bacteria and other associates
Progress of Genome Sequencing

- Total Samples: 302
- Genome Assembly: 217
- Draft Genomes: 119
- Awaiting Sequencing: 54

- Inadequate DNA: 31
- Abandoned: 98
- High Contamination: 44
  - No Target: 11
  - Multi-Species: 12
  - Poor Assembly: 31

- Nemertea: 58
- Annelida: 32
- Mollusca: 14
- Gastrotricha: 4
- Platyhelminthes: 5
- Nematoda: 4
- Echinodermata: 1
- Gnathostomulida: 1
How we are using these genomes for Metagenomics?

- References for Functional Metagenomics
- References for Phylogenetic Diversity Assessment
- Population and Community Structure Assessment
  - Mitochondrial metagenomics and eDNA analysis of meiofauna
mtDNA: the workhorse of animal eDNA analysis

Robust locus for population and Species level analysis

- High rate of mutation (>10 x nuclear genes)
- High resolution below species level
- Many copies per cell
- The same ~37 genes critical to life
- Reveals phylogeographic structure
- History of use in species level regulatory processes.
- Has been challenging to find conserved primers for diverse animal phyla
7,873 complete animal mitochondrial genomes

Mitochondrial Genomes in RefSeq

112,258 proteins

Enriching (meta) genomic libraries for animal mitochondrial genomes: Hybridization Capture

Millions of ssDNA probes representing all known mtDNA sequences
Yes, process now makes it cost effective to generate whole mitochondrial genomes from eDNA samples. MitoCap probes info is available by request.
Acknowledgments

Joseph Sevigny
Francesca Leasi, Ph. D
Anthony Westbrook
Krystalynne Morris
Stephen Simpson
Jeff Hall, Ph. D
Krystalle Diaz

Holly Bik, Ph. D (co-PI)
- University of California, Riverside, CA

Paul Montagna, Ph. D (co-PI)
- Harte Research Institute, Corpus Christi, TX
Michael Reuscher, Ph. D
Mellisa Rohal
Travis Washburn

Jon Norenburg, Ph.D - Smithsonian Institution
BUSCO – Genome completeness assessment using near-universal single-copy orthologs.

N50 (Contiguity)
Mitochondrial Genomes: the barcodes of life

**THE MIGRATION OF ANATOMICALLY MODERN HUMANS**

Evidence from fossils, ancient artifacts, and genetic analyses combine to tell a compelling story.

Early humans may have taken the southern route out of Africa. The southern route appears easiest, especially given the theory that the early humans were spread during the last ice age. Still, crossing the Sahara desert is an arduous journey. Studies have shown that the route is a harsh landscape, with limited water sources and harsh climates.

The northern route is more challenging, with higher elevations and difficult terrain. Nonetheless, it may have been the preferred route for some early humans due to its proximity to the Mediterranean Sea, which would have provided a more hospitable environment for migration.

Researchers have also studied mitochondrial genomes from ancient human remains to understand the genetic diversity and migration patterns of early humans. These analyses have provided insights into the evolutionary history of humans and their migration out of Africa.