Bioinformatics I -- Lecture 24

Metagenomics

Metagenomics --

assessing microbial biodiversity by sequencing uncultured samples

"Microbes run the world. It's that simple. Although we cannot usually see them, microbes are essential for every part of human life—indeed all life on Earth. Every process in the biosphere is touched by the seemingly endless capacity"

-- Opening lines of **"The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet", National Academy of Sciences, 2007.**

"Everything is everywhere, the environment selects"

--The Baas-Becking hypothesis

Metagenomics research

Metagenomics: Application of Genomics to Uncultured Microorganisms

Microbiology and Molecular Biology Reviews, December 2004, p. 669-685, Vol. 68, No. 4
Jo Handelsman

Abstract

Metagenomics (also referred to as environmental and community genomics) is the genomic analysis of microorganisms by direct extraction and cloning of DNA from an assemblage of microorganisms. The development of metagenomics stemmed from the ineluctable evidence that as-yet-uncultured microorganisms represent the vast majority of organisms in most environments on earth. This evidence was derived from analyses of 16S rRNA gene sequences amplified directly from the environment, an approach that avoided the bias imposed by culturing and led to the discovery of vast new lineages of microbial life. Although the portrait of the microbial world was revolutionized by analysis of 16S rRNA genes, such studies yielded only a phylogenetic description of community membership, providing little insight into the genetics, physiology, and biochemistry of the members. Metagenomics provides a second tier of technical innovation that facilitates study of the physiology and ecology of environmental microorganisms. Novel genes and gene products discovered through metagenomics include the first bacteriorhodopsin of bacterial origin; novel small molecules with antimicrobial activity; and new members of families of known proteins, such as an $Na^+(Li^+)/H^+$ antiporter, RecA, DNA polymerase, and antibiotic resistance determinants. Reassembly of multiple genomes has provided insight into energy and nutrient cycling within the community, genome structure, gene function, population genetics and microheterogeneity, and lateral gene transfer among members of an uncultured community. The application of metagenomic sequence information will facilitate the design of better culturing strategies to link genomic analysis with pure culture studies.

7,660 articles containing "metagenomics"
95% in the last 5 years
(Nov 2009, Google Scholar)

Metagenomics research

Contrasts Between Marine and Freshwater Bacterial Community Composition: Analyses of Communities in Lake George and Six Other Adirondack Lakes

Barbara A. Methe, William D. Hiorns and Jonathan P. Zehr

Limnology and Oceanography, Vol. 43, No. 2 (Mar., 1998), pp. 368-374

Abstract

The bacterial communities of seven freshwater lakes in the Adirondack Mountains of New York state were examined using culture-independent methods. β-Proteobacteria 16S rRNA sequences were recovered from all seven lakes and their presence was confirmed by direct DNA hybridization. The results are consistent with phylogenetic and in situ hybridization-based studies in other freshwater environments, but are significantly different than the results of marine oceanic studies, where β-Proteobacteria are noticeably absent. This relationship between evolutionary history and environmental distribution is striking, since these phylogenetic clades have not been correlated with consistent physiological features or biochemical capabilities, and there is no a priori reason to expect differences in phylogenetic composition between the environments. In contrast, freshwater relatives to marine phylogenetic clusters, in particular the SAR 11 cluster of the α-Proteobacteria, were identified. The data imply an underlying physiological distinction between the β- and other Proteobacteria groups and potentially an important difference between the composition of bacterial communities in marine and fresh-water environments.

1,150 articles containing "freshwater metagenomics"
(Nov 2009, Google Scholar)

Metagenomics research

The ecological role of biodiversity in agroecosystems
Agriculture, Ecosystems & Environment, Volume 74, Issues 1-3, June 1999, Pages 19-31
Miguel A. Altieri

Abstract

Increasingly research suggests that the level of internal regulation of function in agroecosystems is largely dependent on the level of plant and animal biodiversity present. In agroecosystems, biodiversity performs a variety of ecological services beyond the production of food, including recycling of nutrients, regulation of microclimate and local hydrological processes, suppression of undesirable organisms and detoxification of noxious chemicals. In this paper the role of biodiversity in securing crop protection and soil fertility is explored in detail. It is argued that because biodiversity mediated renewal processes and ecological services are largely biological, their persistence depends upon the maintenance of biological integrity and diversity in agroecosystems. Various options of agroecosystem management and design that enhance functional biodiversity in crop fields are described.

3,260 articles containing “soil metagenomics”
(Nov 2009, Google Scholar)

Metagenomics research

Phylogenetic diversity of termite gut spirochaetes.

Environ Microbiol. 1999 Aug;1(4):331-45.

Lilburn TG, Schmidt TM, Breznak JA.

Abstract

A molecular phylogenetic analysis was done of not-yet-cultured spirochaetes inhabiting the gut of the termite, Reticulitermes flavipes (Kollar). Ninety-eight clones of near-full-length spirochaetal 16S rDNA genes were classified by ARDRA pattern and by partial sequencing. All clones grouped within the genus Treponema, and at least 21 new species of Treponema were recognized within R. flavipes alone. Analysis of 190 additional clones from guts of Coptotermes formosanus Shiraki and Zootermopsis angusticollis (Hagen), as well as published data on clones from Cryptotermes domesticus (Haviland), Mastotermes darwiniensis Froggatt, Nasutitermes lujae (Wasmann) and Reticulitermes speratus(Kolbe), revealed a similar level of novel treponemal phylogenetic diversity in these representatives of five of the seven termite families. None of the clones was closely related (i.e. all bore < or = 91% sequence similarity) to any previously recognized treponeme. The data also revealed the existence of two major phylogenetic groups of treponemes: one containing all of the currently known isolates of Treponema and a large number of phylotypes from the human gingival crevice, but only a minority of the termite gut spirochaete clones; another containing the majority of termite spirochaete clones and two Spirochaeta (S. caldaria and S. stenostrepta), which, although free living, group within the genus Treponema on the basis of 16S rRNA sequence. Signature nucleotides that almost perfectly distinguished the latter group, herein referred to as the 'termite cluster', occurred at the following (E. coli numbering) positions: 289-G x C-311; A at 812; and an inserted nucleotide at 1273. The emerging picture is that the long-recognized and striking morphological diversity of termite gut spirochaetes is paralleled by their phylogenetic diversity and may reflect substantial physiological diversity as well.

231 articles containing “termite gut metagenomics”
(Nov 2009, Google Scholar)

Metagenomics research

Variations of Bacterial Populations in Human Feces Measured by Fluorescent *In Situ* Hybridization with Group-Specific 16S rRNA-Targeted Oligonucleotide Probes

Applied and Environmental Microbiology, September 1998, p. 3336-3345, Vol. 64, No. 9

Alison H. Franks, Hermie J. M. Harmsen,* Gerwin C. Raangs, Gijsbert J. Jansen, Frits Schut, and Gjalt W. Welling.

Abstract

Six 16S rRNA-targeted oligonucleotide probes were designed, validated, and used to quantify predominant groups of anaerobic bacteria in human fecal samples. A set of two probes was specific for species of the *Bacteroides fragilis* group and the species *Bacteroides distasonis*. Two others were designed to detect species of the *Clostridium histolyticum* and the *Clostridium lituseburense* groups. Another probe was designed for the genera *Streptococcus* and *Lactococcus*, and the final probe was designed for the species of the *Clostridium coccoides-Eubacterium rectale* group. The temperature of dissociation of each of the probes was determined. The specificities of the probes for a collection of target and reference organisms were tested by dot blot hybridization and fluorescent in situ hybridization (FISH). The new probes were used in initial FISH experiments to enumerate human fecal bacteria. The combination of the two *Bacteroides*-specific probes detected a mean of 5.4×10^{10} cells per g (dry weight) of feces; the *Clostridium coccoides-Eubacterium rectale* group-specific probe detected a mean of 7.2×10^{10} cells per g (dry weight) of feces. The *Clostridium histolyticum*, *Clostridium lituseburense*, and *Streptococcus-Lactococcus* group-specific probes detected only numbers of cells ranging from 1×10^{7} to 7×10^{8} per g (dry weight) of feces. Three of the newly designed probes and three additional probes were used in further FISH experiments to study the fecal flora composition of nine volunteers over a period of 8 months. The combination of probes was able to detect at least two-thirds of the fecal flora. The normal biological variations within the fecal populations of the volunteers were determined and indicated that these variations should be considered when evaluating the effects of agents modulating the flora.

Metagenomics research

Metagenomic Analyses of an Uncultured Viral Community from Human Feces
J Bacteriol. 2003 October; 185(20): 6220–6223
Mya Breitbart,[1] Ian Hewson,[2] Ben Felts,[3] Joseph M. Mahaffy,[3] James Nulton,[3] Peter Salamon,[3] and Forest Rohwer[1,4*].

Abstract

Here we present the first metagenomic analyses of an uncultured viral community from human feces, using partial shotgun sequencing. Most of the sequences were unrelated to anything previously reported. The recognizable viruses were mostly siphophages, and the community contained an estimated 1,200 viral genotypes.

609 articles containing "human feces metagenomics"
(Nov 2009, Google Scholar)

Metagenomics research

Viral metagenomics

Reviews in Medical Virology. Volume 17 Issue 2, Pages 115 - 131

Eric L. Delwart

Abstract

Characterisation of new viruses is often hindered by difficulties in amplifying them in cell culture, limited antigenic/serological cross-reactivity or the lack of nucleic acid hybridisation to known viral sequences. Numerous molecular methods have been used to genetically characterise new viruses without prior *in vitro* replication or the use of virus-specific reagents. In the recent metagenomic studies viral particles from uncultured environmental and clinical samples have been purified and their nucleic acids randomly amplified prior to subcloning and sequencing. Already known and novel viruses were then identified by comparing their translated sequence to those of viral proteins in public sequence databases. Metagenomic approaches to viral characterisation have been applied to seawater, near shore sediments, faeces, serum, plasma and respiratory secretions and have broadened the range of known viral diversity. Selection of samples with high viral loads, purification of viral particles, removal of cellular nucleic acids, efficient sequence-independent amplification of viral RNA and DNA, recognisable sequence similarities to known viral sequences and deep sampling of the nucleic acid populations through large scale sequencing can all improve the yield of new viruses. This review lists some of the animal viruses recently identified using sequence-independent methods, current laboratory and bioinformatics methods, together with their limitations and potential improvements. Viral metagenomic approaches provide novel opportunities to generate an unbiased characterisation of the viral populations in various organisms and environments.

2,070 articles containing “viral metagenomics”
(Nov 2009, Google Scholar)

Metagenomics research

Metagenomics to Paleogenomics: Large-Scale Sequencing of Mammoth DNA

Science 20 January 2006:
Vol. 311. no. 5759, pp. 392 - 394
Hendrik N. Poinar,[1,2,3*] Carsten Schwarz,[1,2] Ji Qi,[4] Beth Shapiro,[5] Ross D. E. MacPhee,[6] Bernard Buigues,[7] Alexei Tikhonov,[8] Daniel H. Huson,[9] Lynn P. Tomsho,[4] Alexander Auch,[9] Markus Rampp,[10] Webb Miller,[4] Stephan C. Schuster[4]

Abstract

We sequenced 28 million base pairs of DNA in a metagenomics approach, using a woolly mammoth (Mammuthus primigenius) sample from Siberia. As a result of exceptional sample preservation and the use of a recently developed emulsion polymerase chain reaction and pyrosequencing technique, 13 million base pairs (45.4%) of the sequencing reads were identified as mammoth DNA. Sequence identity between our data and African elephant (Loxodonta africana) was 98.55%, consistent with a paleontologically based divergence date of 5 to 6 million years. The sample includes a surprisingly small diversity of environmental DNAs. The high percentage of endogenous DNA recoverable from this single mammoth would allow for completion of its genome, unleashing the field of paleogenomics.

176 articles containing "paleogenomics metagenomics"
(Nov 2009, Google Scholar)

Metagenomics research

The Human Microbiome Project

NATURE|Vol 449|18 October 2007|
Peter J. Turnbaugh, Ruth E. Ley, Micah Hamady, Claire M. Fraser-Liggett, Rob Knight & Jeffrey I. Gordon

A strategy to understand the microbial components of the human genetic and metabolic landscape and how they contribute to normal physiology and predisposition to disease.

Roles of micriobiota

- **Harvest of otherwise inaccessible nutrients and/or sources of energy from the diet, and synthesis of vitamins**
- **Metabolism of xenobiotics, and other metabolic phenotypes**
- **Renewal of gut epithelial cells**
- **Development and activity of the immune system**

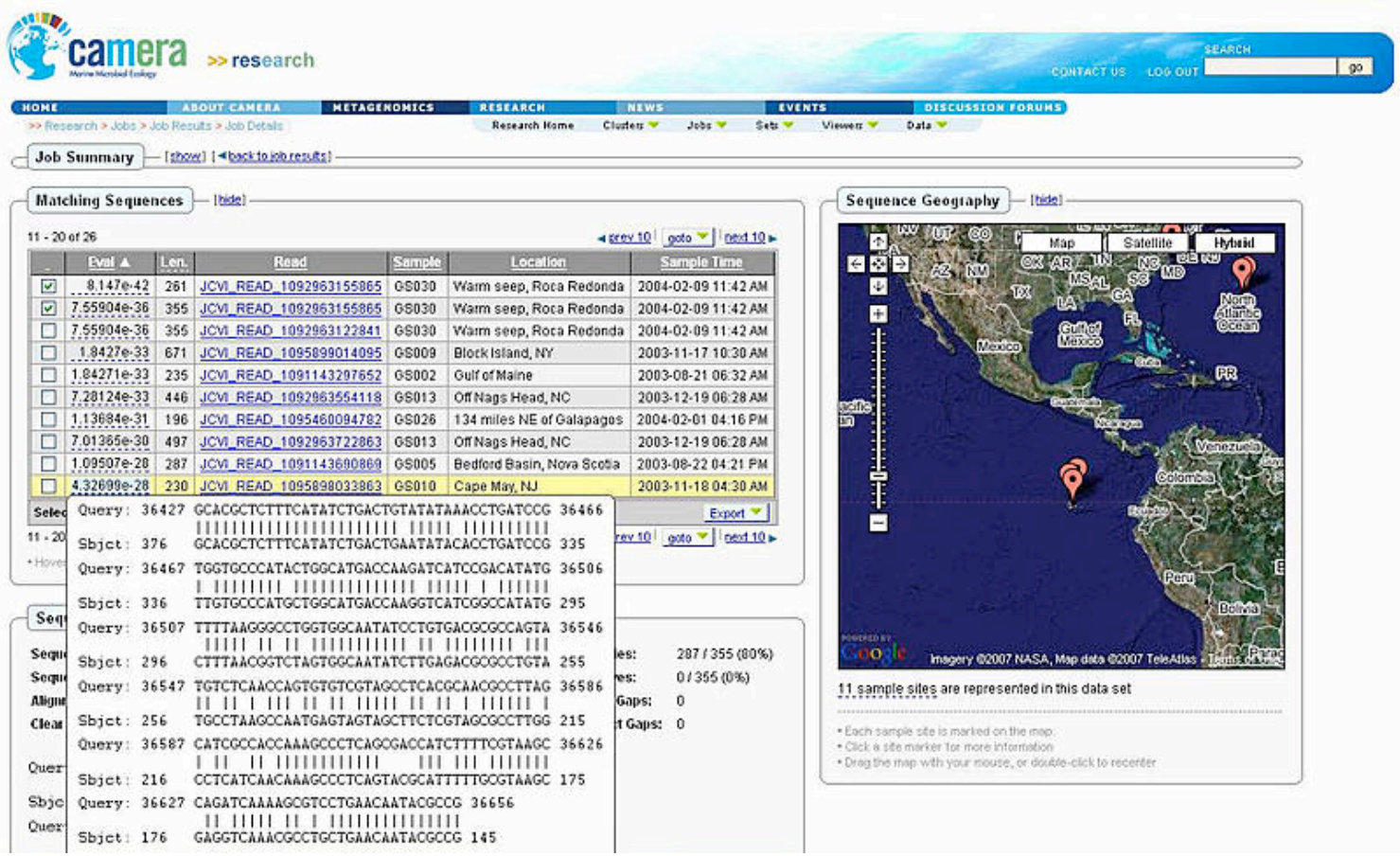
Metagenomics research

Sorcerer II

Sorcerer II Global Ocean Sampling Expedition

J Craig Ventor Institute is undertaking the global metagenomics of the world's oceans.

Metagenomics database for the Sorcerer II expedition



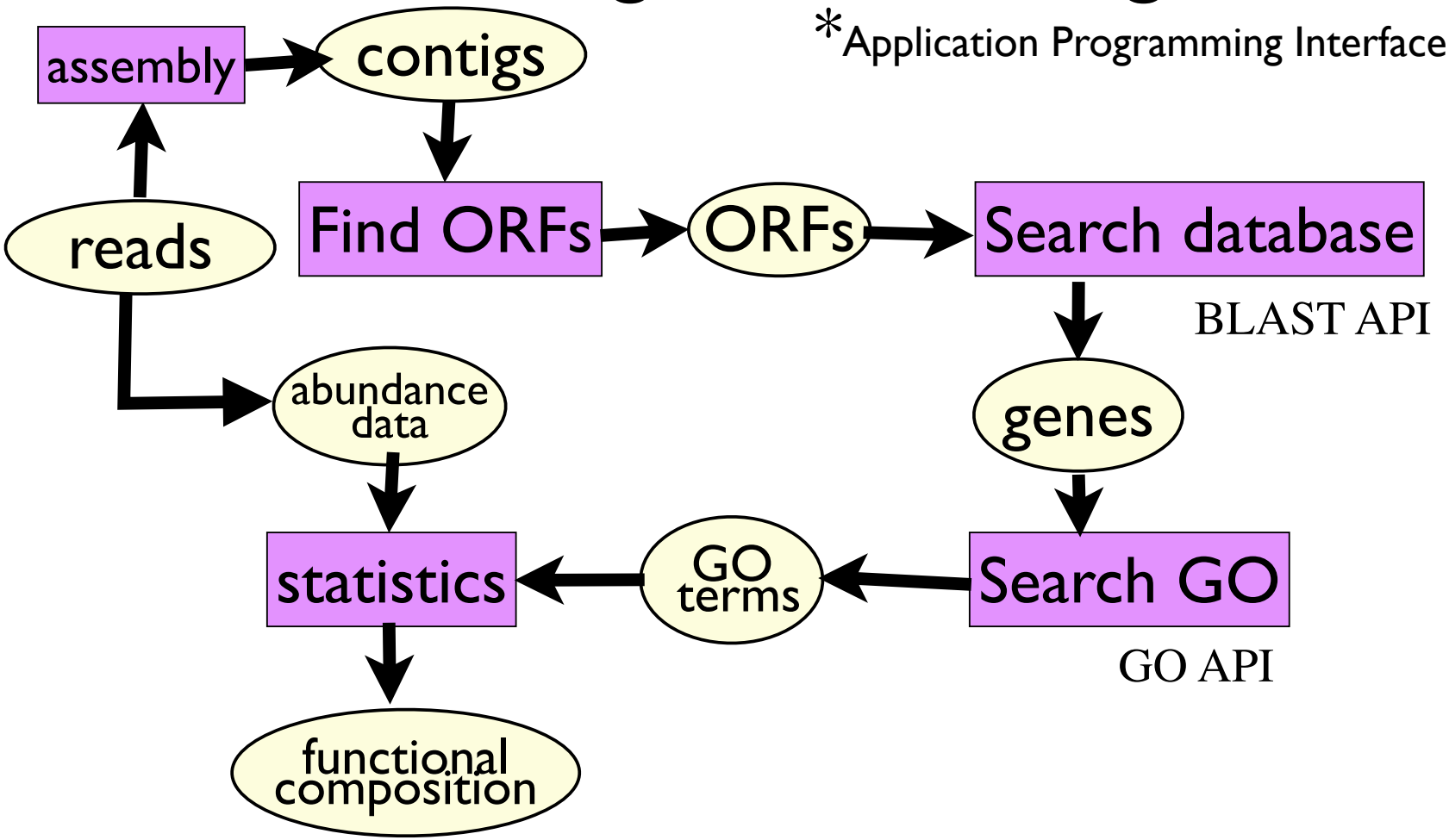
Database gives location, date, sample ID and DNA sequence. BLAST searchable. A map shows locations of hits.

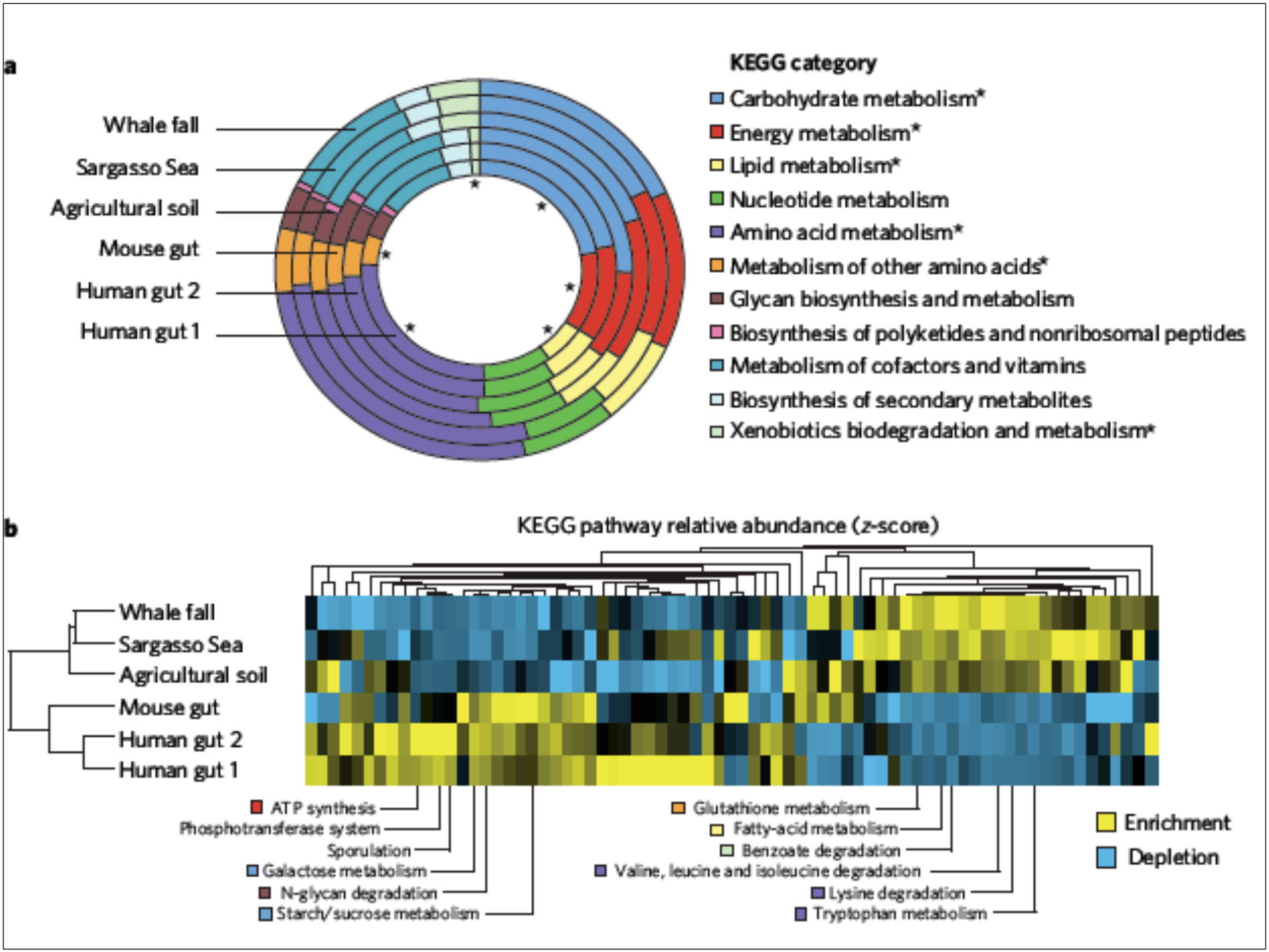
Algorithms for Metagenomics

- Multiple genome assembly from short reads
  - pyrosequencing produces short reads
- Functional metagenomics
  - functions carried out by communities
  - comparative functional metagenomics
- Microbial tree of life
  - microbiome

Functional Metagenomics using APIs*

*Application Programming Interface




a
Whale fall
Sargasso Sea
Agricultural soil
Mouse gut
Human gut 2
Human gut 1
KEGG category
Carbohydrate metabolism*
Energy metabolism*
Lipid metabolism*
Nucleotide metabolism
Amino acid metabolism*
Metabolism of other amino acids*
Glycan biosynthesis and metabolism
Biosynthesis of polyketides and nonribosomal peptides
Metabolism of cofactors and vitamins
Biosynthesis of secondary metabolites
Xenobiotics biodegradation and metabolism*
b
KEGG pathway relative abundance (z-score)
Whale fall
Sargasso Sea
Agricultural soil
Mouse gut
Human gut 2
Human gut 1
ATP synthesis
Phosphotransferase system
Sporulation
Galactose metabolism
N-glycan degradation
Starch/sucrose metabolism
Glutathione metabolism
Fatty-acid metabolism
Benzoate degradation
Valine, leucine and isoleucine degradation
Lysine degradation
Tryptophan metabolism
Enrichment
Depletion

In class exercise: functional metagenomics

- Download reads from file
- In amigo.geneontology.org, run BLAST search. Choose closest homolog.
- Display GO terms for Biological Process
  - What is the function?