Lawrence Berkeley National Laboratory
Recent Work

Title
Integrated OMICs analysis in IMG and IMG/M

Permalink
https://escholarship.org/uc/item/0fz7q27j

Authors
Pati, Amrita
Varghese, Neha
Chen, I-Min Amy
et al.

Publication Date
2014-03-19

Integrated OMICs analysis in IMG and IMG/M

Amrita Pati[1], Neha Varghese[1*], I-Min Amy Chen[2], Marcel Huntemann[1], Kostas Billis[1], Kostas Mavrommatis[1], Krishna Palaniappan[2], Ernest Szeto[2], Anna Ratner[2], Manoj Pillay[2], Ken Chu[2], Victor Markowitz[2], Natalia Ivanova[1], Nikos Kyrpides[1]

[1] LBNL - Department of Energy Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA USA

[2] Biological Data Management and Technology Center (CRD), Berkeley, CA USA

*To whom correspondence should be addressed: Email: apati@lbl.gov

March 21, 2014
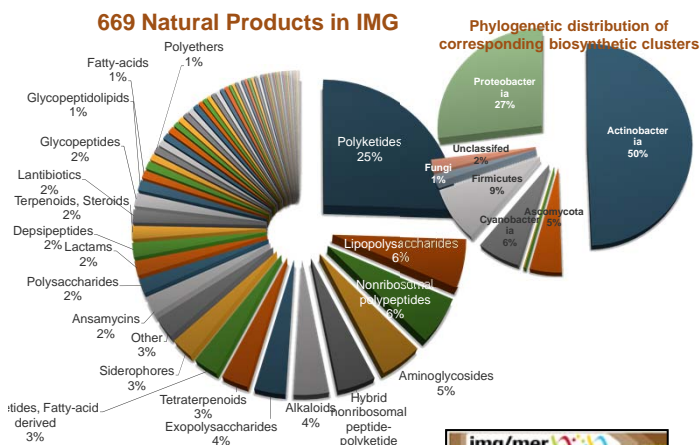
ACKNOWLEDGMENTS:

DISCLAIMER:

Integrated OMICs analysis in IMG and IMG/M

Neha Varghese[1], I-Min Amy Chen[2], Marcel Huntemann[1], Kostas Billis, Kostas Mavrommatis, Krishna Palaniappan[2], Ernest Szeto[2], Anna Ratner[2], Manoj Pillay[2], Ken Chu[2], Victor Markowitz[2], Natalia Ivanova[1], Nikos Kyrpides[1], **Amrita Pati**[1,*]

[1]DOE Joint Genome Institute, 2800 Mitchell Dr. Walnut Creek CA 94598.   [2]Biological Data Management and Technology Center, CRD, Berkeley CA   *apati@lbl.gov
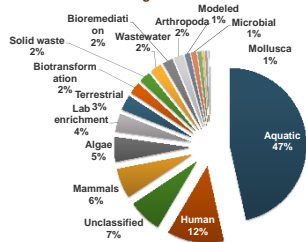
Natural Products and Biosynthetic Clusters

669 Natural Products in IMG

Pie chart labels: Polyethers 1%, Fatty-acids 1%, Glycopeptidolipids 1%, Glycopeptides 2%, Lantibiotics 2%, Terpenoids, Steroids 2%, Depsipeptides 2%, Lactams 2%, Polysaccharides 2%, Ansamycins 2%, Other 3%, Siderophores 3%, ...tides, Fatty-acid derived 3%, Tetraterpenoids 3%, Exopolysaccharides 4%, Alkaloids 4%, Hybrid nonribosomal peptide-polyketide, Aminoglycosides 5%, Nonribosomal polypeptides 6%, Lipopolysaccharides 6%, Polyketides 25%

Phylogenetic distribution of corresponding biosynthetic clusters

Proteobacteria 27%, Unclassified 2%, Fungi 1%, Firmicutes 9%, Cyanobacteria 6%, Ascomycota 5%, Actinobacteria 50%

Biosynthetic Clusters predicted using ClusterFinder

predicted in 13,929 isolate genomes: 299,413
predicted in 3,219 metagenomes: 1,932,135

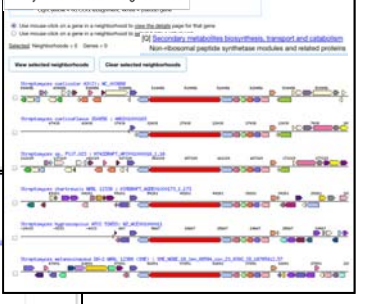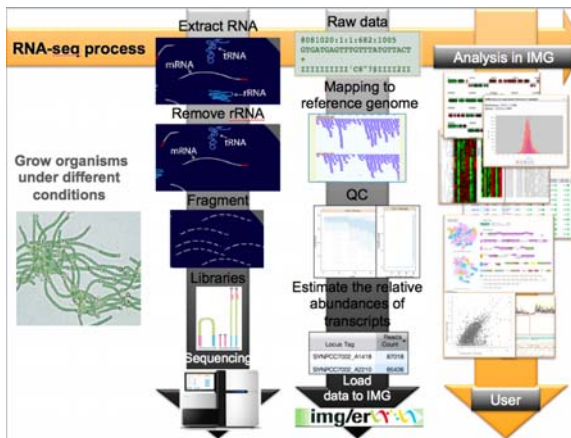### Habitat distribution of 4,082 putative full-length PKS/NRPS pathways in metagenomes

Pie chart labels: Bioremediation 2%, Arthropoda 1%, Microbial 1%, Modeled, Solid waste 2%, Wastewater 2%, Mollusca 1%, Biotransformation 2%, Terrestrial 3%, Lab enrichment 4%, Algae 5%, Aquatic 47%, Mammals, Human 12%, Unclassified 7%

Biosynthetic Cluster Detail

Biosynthetic Cluster Neighborhood

Transcriptomes and Proteomes

RNA-seq process: Extract RNA, Remove rRNA, Fragment, Libraries, Sequencing → Grow organisms under different conditions. Raw data → Mapping to reference genome → QC → Estimate the relative abundances of transcripts → Load data to IMG → User. Analysis in IMG.

RNASeq Expression Studies

Datasets in study

Gene expression quantification

Pathways for Sample

2-sample comparison – regression analysis

2-Oxocarboxylic acid metabolism (64)
ABC transporters (63)
Adherens junction (14)
Adipocytokine signaling pathway (7)

Function-based sample comparisons

RNASeq Data for Gene

Gene centric expression from different samples

Methylomes

DNA Methylation → SMRT® DNA Sequencing → Analysis of Polymerase Kinetics. Forward Strand / Reverse Strand.

Phylogenetic distribution of motifs

Methylome listing in IMG

native C2 | native C2 PB | native XL | native C2 old

Motif count by chemistry type

Genomewise impact of methylations on genes involved in various functions

Methylomics Experiment

Motifs on Sequence

Location of methylated motifs on sequence

Motif String: GATC
Center Pos: 1
Fraction: 0.7924572
N detected: 55472
Group Tag: GATC
Mean Score: 45.61139
Modification Type: m6A
DNA Coord: 799

Motif Summary
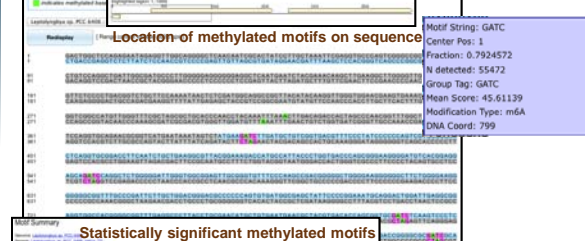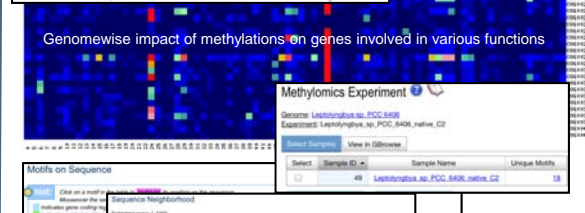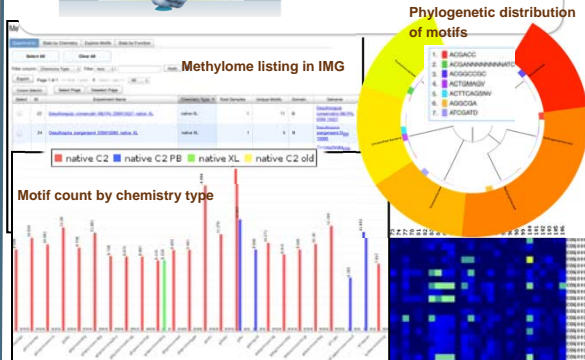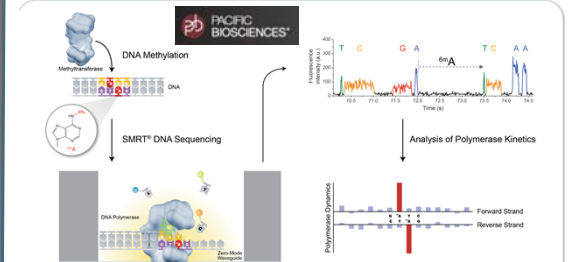
Statistically significant methylated motifs

U.S. DEPARTMENT OF ENERGY | Office of Science