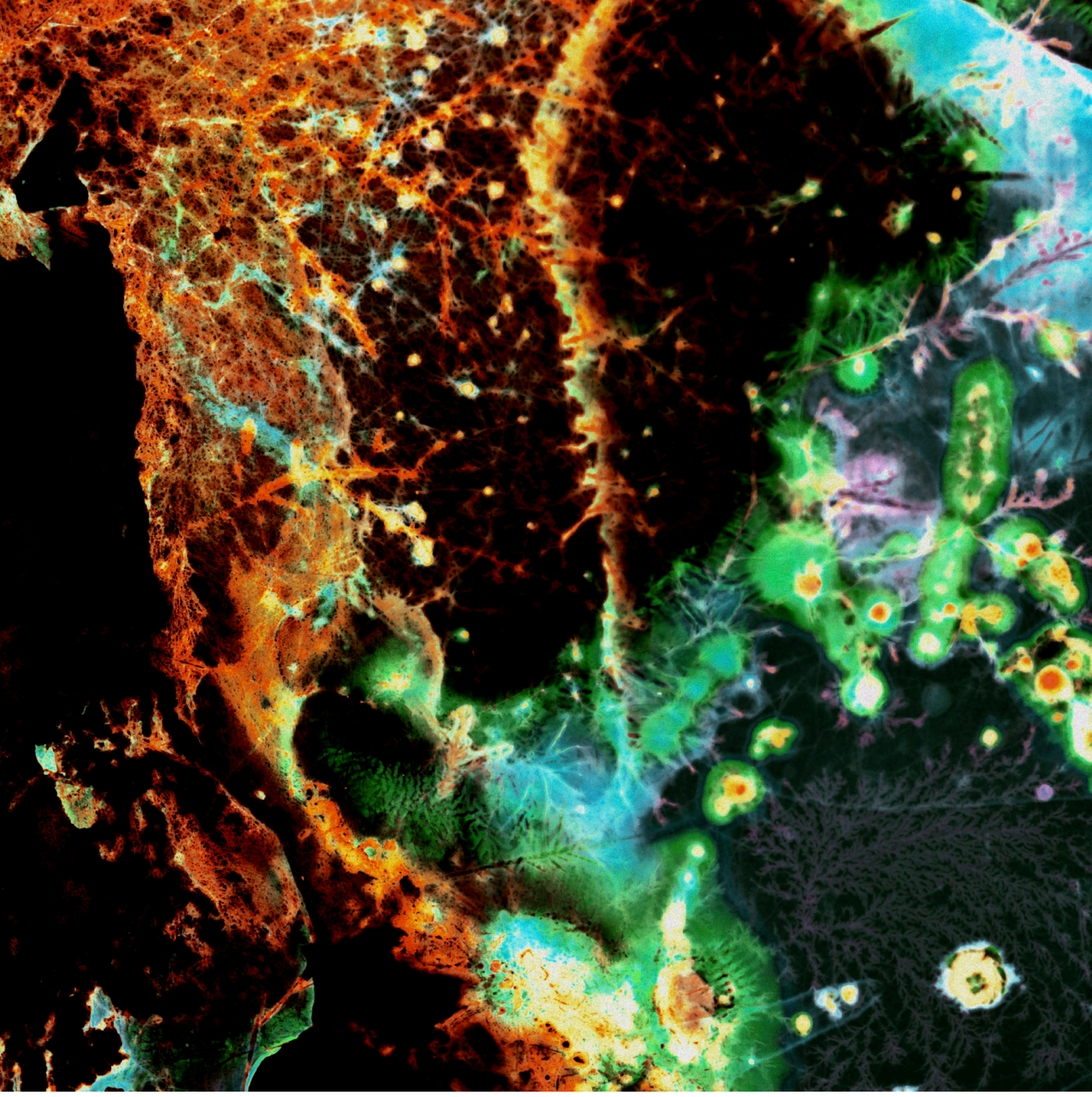


Metagenomics Research Review

An Overview of Publications Featuring Illumina® Technology

Table of Contents

This document highlights recent publications that demonstrate the use of Illumina sequencing technologies in metagenomics. To learn more about Illumina sequencing and microarray technologies, visit www.illumina.com.

Introduction

Microbial populations occur in every biological niche on earth, from the insect gut to the oceans of the world and in the sediment beneath them. Metagenomics refers to the study of genomic DNA obtained from microorganisms that cannot be cultured in the laboratory. This represents the vast majority of microorganisms on earth.[1] The new generation of sequencing technology[2], with its ability to sequence thousands of organisms in parallel, has proved to be uniquely suited to this application. As a result, the sequencing of microbial genomes has become routine. Researchers have even sequenced individual cells.

This accumulation of sequence information has greatly expanded our appreciation of the dynamic nature of microbial populations and their impact on the environment and human health. Humans actually carry ten times more bacterial cells than human cells, and 100 times more bacterial genes than our inherited human genome.[3] Microbes also hold the secrets for generating renewable biofuels and bioremediation. With this extraordinary powerful set of sequencing tools now at our disposal, it is no surprise that metagenomics has become one of the fastest growing scientific disciplines.

Reviews:

Bik, H. M., Porazinska D. L., Creer S., Caporaso J. G., Knight R., et al. (2012) Sequencing our way towards understanding global eukaryotic biodiversity. Trends Ecol Evol 27: 233–243

Caporaso, J. G., Lauber C. L., Walters W. A., Berg-Lyons D., Huntley J., et al. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J

Foster, J. A., Bunge J., Gilbert J. A. and Moore J. H. (2012) Measuring the microbiome: perspectives on advances in DNA-based techniques for exploring microbial life. Brief Bioinform

Reid, G., Younes J. A., Van der Mei H. C., Gloor G. B., Knight R., et al. (2011) Microbiota restoration: natural and supplemented recovery of human microbial communities. Nat Rev Microbiol 9: 27–38

Microbial Populations

The study of bacterial populations can usefully be divided into two functional areas: metagenomics ("Who is there?") and metatranscriptomics ("What are they doing?"). To fully answer those questions, we need to integrate the gene and transcriptome sequence analysis, and also look at individual genomes.

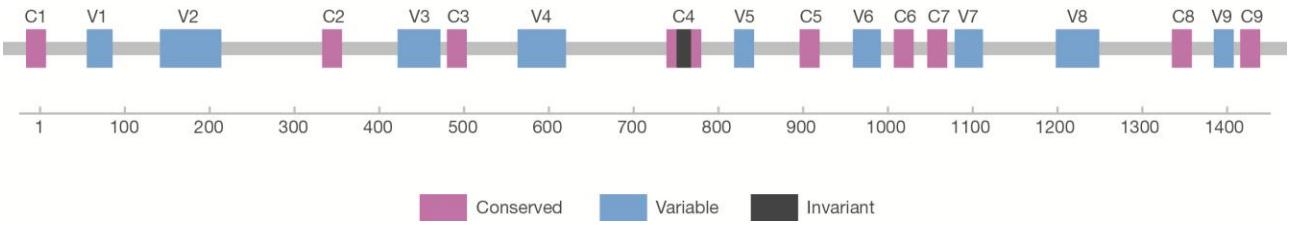
[1] Hugenholtz, P., Goebel, B. M. and Pace, N. R. (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J Bacteriol 180: 4765–4774

[2] Next Generation Sequencing (NGS) and Massively Parallel Sequencing MPS are often used interchangeably to refer to high throughput sequencing technologies. Sequencing by Synthesis (SBS) refers specifically to Illumina sequencing technology. sequencing technologies. Sequencing by Synthesis (SBS) refers specifically to Illumina sequencing technology.

[3] Barwick, B. G., Abramovitz, M., Kodani, M., Moreno, C. S., Nam, R., et al. (2010) Prostate cancer genes associated with TMPRSS2-ERG gene fusion and prognostic of biochemical recurrence in multiple cohorts. Br J Cancer 102: 570–576

16S Ribosomal RNA

The new era of metagenomics was ushered in by studies using 16S rRNA as a phylogenetic marker of microbial taxa.[4] The 16S gene occurs in all living organisms, with the notable exception of viruses and represents more than 80% of total bacterial RNA. The 16S rRNA gene structure consists of interspersed conserved and variable regions, which suits it well for PCR amplification and sequencing. Probes can be designed to hybridize to the conserved regions allowing for amplification and sequencing of the variable regions. Focusing on a small part of the microbial genome brings down the sequencing costs dramatically. This approach has been particularly effective in monitoring fluctuations in populations.[5]



Schematic representation of the 16S rRNA gene. Location of variable (blue) and conserved (purple) regions in a canonical bacterial 16S rRNA. The grey region is invariant in all bacteria.

The impact of read length, read depth, and sequencing errors of the various next-generation sequencing technologies has been extensively studied.[6] For example, Illumina technology can now sequence 250 bp paired ends, effectively interrogating 500 bases at a time. The Illumina HiSeq® and MiSeq® platforms have demonstrated highly similar results.[7]

Caporaso, J. G., Lauber C. L., Walters W. A., Berg-Lyons D., Lozupone C. A., et al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A 108 Suppl 1: 4516–4522

The authors sequenced a diverse set of 25 environmental samples and three known "mock communities" at a depth averaging 3.1 million reads per sample. They found excellent consistency in taxonomic recovery and recapture diversity patterns that were previously reported on the basis of meta analysis of many studies from the literature. The authors conclude that when sequencing a well-chosen region of the 16S gene, 75–100 bases may suffice for drawing reliable conclusions from the data.

Illumina Technology: Genome Analyzer$_{IIx}$. Primers against the V4 region of the 16S rRNA.

[4] Pace, N. R. (1997) A molecular view of microbial diversity and the biosphere. Science 276: 734-740
[5] Caporaso, J. G., Paszkiewicz, K., Field, D., Knight, R. and Gilbert, J. A. (2011) The Western English Channel contains a persistent microbial seed bank. ISME J
[6] Luo, C., Tsementzi, D., Kyrpides, N., Read, T. and Konstantinidis, K. T. (2012) Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. PLoS ONE 7: e30087
[7] Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., et al. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J

References:

Degnan, P. H. and Ochman H. (2012) Illumina-based analysis of microbial community diversity. ISME J 6: 183–194

Loman, N. J., Misra R. V., Dallman T. J., Constantinidou C., Gharbia S. E., et al. (2012) Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol

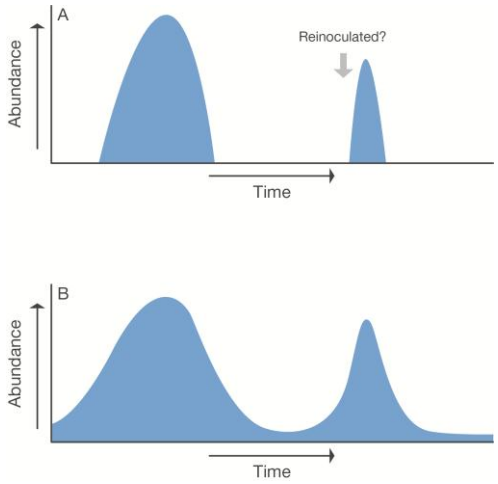
Luo, C., Tsementzi D., Kyrpides N., Read T. and Konstantinidis K. T. (2012) Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. PLoS ONE 7: e30087

Mende, D. R., Waller A. S., Sunagawa S., Jarvelin A. I., Chan M. M., et al. (2012) Assessment of metagenomic assembly using simulated next generation sequencing data. PLoS ONE 7: e31386

Ram, J. L., Karim A. S., Sendler E. D. and Kato I. (2011) Strategy for microbiome analysis using 16S rRNA gene sequence analysis on the Illumina sequencing platform. Syst Biol Reprod Med 57: 162–170

Deep Sequencing

One of the most important advantages of next-generation sequencing is the wealth of sequence information it can produce. Deep sequencing makes it possible to detect very low abundant members of complex populations. These members can act as a pool of genetically diverse organisms that will survive under changing environments or environmental stress. As a result, the ability to detect these rare populations can profoundly impact the interpretation of microbiological changes.[8] The actual read depth required will depend on the desired sensitivity, as well as the complexity of the population.

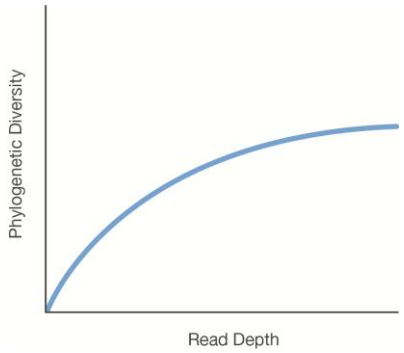


Panels A and B represent the same microbial population sampled at two different depths. In panel A, it appears that the microbes were reintroduced from an external source. However, deep sequencing in panel B reveals that the microbes were present at all points but dropped below the detection level used in panel A.

[8] Caporaso, J. G., Paszkiewicz, K., Field, D., Knight, R. and Gilbert, J. A. (2011) The Western English Channel contains a persistent microbial seed bank. ISME J

Complex Populations

The exhaustive analysis of a complex population is a significant technical challenge.[9] The primary goal is to sequence deep enough to distinguish rare members of the population from sequencing errors. A low sequencing error rate is important, as well as strict filters to remove sequencing errors.[10]



Rarefaction curve. Phylogenetic diversity increases with the read depth. The optimal read depth for discovery applications would be the read depth where phylogenetic diversity no longer increases.

Amplification and Cloning Bias

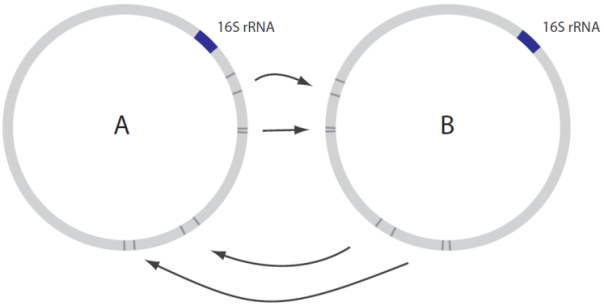
There are PCR and cloning biases inherent in 16S rRNA protocols.[11,12] The primers are designed to hybridize to conserved regions, but these regions may change over long evolutionary periods, resulting in a loss of hybridization to the probe. This will lead to an underestimation of evolutionarily distant members of the population.

The variable regions are different sizes and can change at different rates; for example, V6 tags appear to systematically overestimate species richness.[13] Paired-end Illumina sequencing of the larger V4 region has been successfully used to build phylogenetic trees.[14]

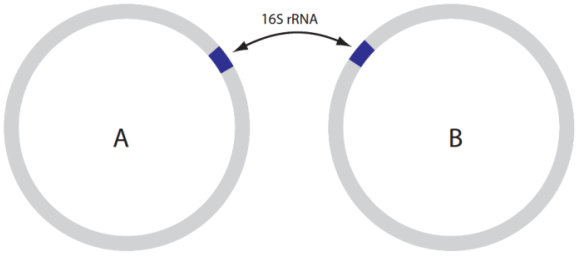
[9] Bartram, A. K., Lynch, M. D., Stearns, J. C., Moreno-Hagelsieb, G. and Neufeld, J. D. (2011) Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. Appl Environ Microbiol 77: 3846–3852
[10] Mende, D. R., Waller, A. S., Sunagawa, S., Jarvelin, A. I., Chan, M. M., et al. (2012) Assessment of metagenomic assembly using simulated next generation sequencing data. PLoS ONE 7: e31386
[11] Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., et al. (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res 21: 494–504
[12] Soergel, D. A., Dey, N., Knight, R. and Brenner, S. E. (2012) Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. ISME J
[13] Youssef, N., Sheik, C. S., Krumholz, L. R., Najar, F. Z., Roe, B. A., et al. (2009) Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. Appl Environ Microbiol 75: 5227–5236
[14] Werner, J. J., Zhou, D., Caporaso, J. G., Knight, R. and Angenent, L. T. (2011) Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. ISME J

Mosaicism

Many bacteria have a history of horizontal gene transfer. Horizontal transfer of functional genes, or even significant genomic rearrangements, may not be reported by the 16S rRNA region.[15] In addition, bacteria can tolerate transfer of complete 16S genes.[16] Any species identification using 16S gene-based probes or homology-based analysis of partial 16S sequences may lead to misidentification, because the marker may represent a transferred gene structure.[17]



Horizontal transfer of functional genes, or even significant genomic rearrangements, may not be reported by the 16S rRNA region.

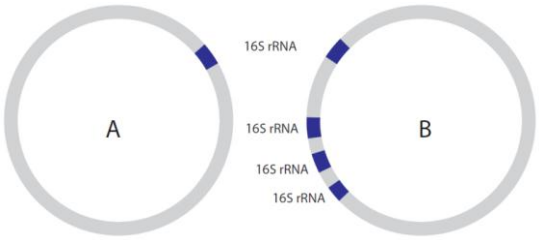


Some bacteria can tolerate the transfer of complete 16S genes, which could impact phylogenetic interpretation based on 16S rRNA metagenomics analysis.

Intragenomic Heterogenity

The copy number of rRNA operons per bacterial genome varies from 1 to 15.[18] The sequences of multi-copy rRNA can vary by as much as 6.4%.[19] This will impact abundance estimates based on the rRNA, and limit the phylogenetic resolution of species based on those sequences.

[15] Altermann, E. (2012) Tracing lifestyle adaptation in prokaryotic genomes. Front Microbiol 3: 48
[16] Asai, T., Zaporojets, D., Squires, C. and Squires, C. L. (1999) An Escherichia coli strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. Proc Natl Acad Sci U S A 96: 1971–1976
[17] Schouls, L. M., Schot, C. S. and Jacobs, J. A. (2003) Horizontal transfer of segments of the 16S rRNA genes between species of the Streptococcus anginosus group. J Bacteriol 185: 7241–7246
[18] Klappenbach, J. A., Dunbar, J. M. and Schmidt, T. M. (2000) rRNA operon copy number reflects ecological strategies of bacteria. Appl Environ Microbiol 66: 1328–1333
[19] Wang, Y., Zhang, Z. and Ramanan, N. (1997) The actinomycete Thermobispora bispora contains two distinct types of transcriptionally active 16S rRNA genes. J Bacteriol 179: 3270–3276



The copy number of rRNA operons per bacterial genome varies from 1 to 15 and the sequences of the copies can vary by as much as 6.4%.

Lack of Threshold and Identity

There is no consistent relationship between the conservation of 16S rRNA and the rest of the bacterial genome. For example the 16S rRNA genes of type strains *B. globisporus* and *B. psychrophilus* share 99.8% sequence identity, but at the genome level they exhibit only 23–50% relatedness.[20] This can be due to many factors such as mosaicism and differences in evolutionary pressure. As a result, species identification solely based on the 16S rRNA sequences may lead to misidentifications.[21]

There is no consistent relationship between the conversation of 16S rRNA and the rest of the bacterial genome. As a result, species identification solely based on the 16S rRNA sequences may lead to misidentifications.

Metagenome Sequencing

Metagenome sequencing, also called shotgun sequencing, refers to sequencing DNA fragments extracted from microbial populations. Because this approach captures the complete genomes of all the organisms in the population, mosaicism and biases have little impact.[22] The comprehensive information obtained by this approach enables the accurate inference of phylogenetic relationships to be made.

However, the most substantial advantage is the information it provides about the genes present in the bacterial population. It is possible to identify the genes present in a bacterial population without assembling the individual bacterial genomes. Functional gene groupings can be more informative and more stable than a record of bacterial species.

[20] Fox, G. E., Wisotzkey, J. D. and Jurtshuk, P., Jr. (1992) How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int J Syst Bacteriol 42: 166–170
[21] Rajendhran, J. and Gunasekaran, P. (2011) Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. Microbiol Res 166: 99–110
[22] Harris, S. R., Clarke, I. N., Seth-Smith, H. M., Solomon, A. W., Cutcliffe, L. T., et al. (2012) Whole-genome analysis of diverse Chlamydia trachomatis strains identifies phylogenetic relationships masked by current clinical typing. Nat Genet 44: 413–419

Reviews:

Gonzalez, A. and Knight R. (2012) Advancing analytical algorithms and pipelines for billions of microbial sequences. Curr Opin Biotechnol 23: 64–71

Gerlach, W. and Stoye J. (2011) Taxonomic classification of metagenomic shotgun sequences with CARMA3. Nucleic Acids Res 39: e91

Xia, L. C., Cram J. A., Chen T., Fuhrman J. A. and Sun F. (2011) Accurate genome relative abundance estimation based on shotgun metagenomic reads. PLoS ONE 6: e27992

References:

Harris, S.R., Clarke I. N., Seth-Smith H. M., Solomon A. W., Cutcliffe L. T., et al. (2012) Whole genome analysis of diverse Chlamydia trachomatis strains identifies phylogenetic relationships masked by current clinical typing. Nat Genet 44:413-419

This paper presents a detailed phylogeny based on whole-genome sequencing of representative strains of *C. trachomatis* from both trachoma and lymphogranuloma venereum (LGV) biovars. It shows that predicting phylogenetic structure using ompA, which is traditionally used to classify Chlamydia, is misleading because extensive recombination in this region masks the true relationships.

Illumina Technology: Genome Analyzer$_{II}$

Altermann, E. (2012) Tracing lifestyle adaptation in prokaryotic genomes. Front Microbiol 3: 48

Coelho, A. C., Boisvert S., Mukherjee A., Leprohon P., Corbeil J., et al. (2012) Multiple mutations in heterogeneous miltefosine-resistant Leishmania major population as determined by whole genome sequencing. PLoS Negl Trop Dis 6: e1512

Transcriptomes

The metatranscriptome is the identity and quantity of a complete set of transcripts in a population of cells. While metagenomics tells us who is there and what they are capable of based on their gene complement, metatranscriptomics tells us what they are doing at that moment. Unlike hybridization-based techniques, such as PCR, Northern blotting, or microarrays, RNA-seq sequences are matched to genes by sequence alignment. This approach offers the following advantages:

Characteristic	Application
No prior knowledge of the genome sequence is required	• Discover novel transcripts and genetic features (gene mining). • Annotate functional domains in the genome.
Accurate mapping	• The mapping of sequences with an aligner is more precise than hybridization in solution. • Transcription can be studied at a much higher resolution and specificity without interference to non-specific cross-hybridization.
Dynamic Range	• Greater dynamic range than fluorescence- based measurements. • Better discrimination at high and low levels of expression.

Gene mining has become one of the fastest growing applications of this technology. Transcription analyses of microbial populations in the rumen of various animals, termite gut, and biodigesters have led to the discovery of a staggering array of novel biocatalysts for the biofuels and related industries.[23]

Reviews:

Filiatrault, M. J. (2011) Progress in prokaryotic transcriptomics. Curr Opin Microbiol 14: 579–586

Klitgord, N. and Segre D. (2011) Ecosystems biology of microbial metabolism. Curr Opin Biotechnol 22: 541–546

Mader, U., Nicolas P., Richard H., Bessieres P. and Aymerich S. (2011) Comprehensive identification and quantification of microbial transcriptomes by genome-wide unbiased methods. Curr Opin Biotechnol 22: 32–41

Martin, F. (2011) Unearthing the truffle genome. New Phytol 189: 645–646

Trevors, J. T. (2011) Viable but non-culturable (VBNC) bacteria: Gene expression in planktonic and biofilm cells. J Microbiol Methods 86: 266–273

References:

Qi, M., Wang P., O'Toole N., Barboza P. S., Ungerfeld E., et al. (2011) Snapshot of the eukaryotic gene expression in muskoxen rumen--a metatranscriptomic approach. PLoS ONE 6: e20521

In this paper, the authors used mRNA-seq to investigate the functional diversity of the eukaryotic microorganisms and plant cell wall degrading enzymes within the rumen of muskoxen (Ovibos moschatus). They identified plant cell-wall degrading enzyme modules, including glycoside hydrolases, carbohydrate esterases, and polysaccharide lyases.

Illumina Technology: Genome Analyzer$_{II}$ Polyadenylated RNA (mRNA) was sequenced, 2.8 gigabases of sequences were obtained and 59,129 contigs assembled.

Bomar, L., Maltz M., Colston S. and Graf J. (2011) Directed culturing of microorganisms using metatranscriptomics. MBio 2: e00012–00011

The authors utilized RNA-seq to learn more about the uncultured and cultured symbionts from the digestive-tract microbiome of the medicinal leech. The expression data revealed highly expressed hydrolytic enzymes and transporters that provided critical clues for the design of a culture medium for the previously uncultured Rikenella-like symbiont. This also independently validates the pathways and the annotations the authors used to design the culture medium lyases.

Illumina Technology: Genome Analyzer and mRNA-seq kit. 34 bp read length. Approximately 15 million cDNA reads were generated.

Tisserant, E., Da Silva C., Kohler A., Morin E., Wincker P., et al. (2011) Deep RNA sequencing improved the structural annotation of the Tuber melanosporum transcriptome. New Phytol 189: 883–891

Wang, T. Y., Chen H. L., Lu M. Y., Chen Y. C., Sung H. M., et al. (2011) Functional characterization of cellulases identified from the cow rumen fungus Neocallimastix patriciarum W5 by transcriptomic and secretomic analyses. Biotechnol Biofuels 4: 24

Zhou, M., McAllister T. A. and Guan L. L. (2011) Molecular identification of rumen methanogens: Technologies, advances and prospects. Animal Feed Science and Technology 166–167: 76–86

[23] Warnecke, F. and Hess, M. (2009) A perspective: metatranscriptomics as a tool for the discovery of novel biocatalysts. J Biotechnol 142: 91–95

Plasmids (Plasmidomics)

Plasmids often serve as mediators of lateral gene transfer, a strong and sculpting evolutionary force in microbial environments.[24]

References:

Kav, A. B., Sasson G., Jami E., Doron-Faigenboim A., Benhar I., et al. (2012) Insights into the bovine rumen plasmidome. Proc Natl Acad Sci U S A 109: 5452–5457

The authors describe the rumen plasmidome as having a highly mosaic nature that can cross phyla. The functional profile of the rumen plasmidome codes for functions that are enriched in the rumen ecological niche. This could confer advantages to their hosts, suggesting that the mobile genetic elements play a role in adaptation to the environment.

Illumina Technology: Genome Analyzer$_{IIx}$ and paired-end reads.

Ho, P. L., Lo W. U., Yeung M. K., Li Z., Chan J., et al. (2012) Dissemination of pHK01-like incompatibility group IncFII plasmids encoding CTX-M-14 in Escherichia coli from human and animal sources. Vet Microbiol

Single Cell Analysis

The sequencing and analysis of the genomes of individual cells should provide the ultimate resolution of a microbial population. In addition, the analysis of the RNA from individual cells should provide insight into the activities of the cells and task-division within the population. There has been substantial progress in individual cell analysis in areas like cancer research[25,26,27,] which will inevitably find their way into metagenomics research.

Reviews:

Walker, A. (2011) Genome watch: Singled out. Nat Rev Microbiol 9: 485

Kalisky, T., Blainey P. and Quake S. R. (2011) Genomic analysis at the single-cell level. Annu Rev Genet 45: 431–445

[24] Kav, A. B., Sasson, G., Jami, E., Doron-Faigenboim, A., Benhar, I., et al. (2012) Insights into the bovine rumen plasmidome. Proc Natl Acad Sci U S A 109: 5452–5457
[25] Islam, S., Kjallquist, U., Moliner, A., Zajac, P., Fan, J. B., et al. (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 21: 1160–1167
[26] Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., et al. (2011) Tumour evolution inferred by single-cell sequencing. Nature 472: 90–94
[27] Moon, S., Kim, Y. G., Dong, L., Lombardi, M., Haeggstrom, E., et al. (2011) Drop-on-demand single cell isolation and total RNA analysis. PLoS ONE 6: e17455

References:

Chitsaz, H., Yee-Greenbaum J. L., Tesler G., Lombardo M. J., Dupont C. L., et al. (2011) Efficient de novo assembly of single-cell bacterial genomes from short-read data sets. Nat Biotechnol 29: 915–921

The authors describe the sequencing of the genome from a single cell of an uncultivated SAR324 clade of Deltaproteobacteria, a cosmopolitan bacterial lineage in the global ocean. Metabolic reconstruction suggests that SAR324 is aerobic, motile, and chemotaxic. This approach enables acquisition of genome assemblies for individual uncultivated bacteria using only short reads, providing cell-specific genetic information absent from metagenomic studies.

Illumina Technology: Genome Analyzer 100 bp paired-end reads.

Yoon, H. S., Price D. C., Stepanauskas R., Rajah V. D., Sieracki M. E., et al. (2011) Single-cell genomics reveals organismal interactions in uncultivated marine protists. Science 332: 714–717

Viruses

In virology, next-generation sequencing has become a powerful tool that can be used to detect, identify, and quantify novel viruses in one step.[28] It is proving to be a sensitive method for detecting putative infectious agents associated with human tissues. At a modest depth of sequencing, viral transcripts can be detected at frequencies less than 1 in 1,000,000.[29] One of the happy consequences of deep sequencing is the coincidental sequencing of viral DNA or RNA, which has led to the discovery of several new viruses.[30]

Reviews:

Mokili, J. L., Rohwer F. and Dutilh B. E. (2012) Metagenomics and future perspectives in virus discovery. Curr Opin Virol 2: 63–77

Li, L. and Delwart E. (2011) From orphan virus to pathogen: the path to the clinical lab. Curr Opin Virol 1: 282–288

Rosario, K. and Breitbart M. (2011) Exploring the viral world through metagenomics. Curr Opin Virol 1: 289–297

References:

Flaherty, P., Natsoulis G., Muralidharan O., Winters M., Buenrostro J., et al. (2012) Ultrasensitive detection of rare mutations using next-generation targeted resequencing. Nucleic Acids Res 40: e2

The authors demonstrate that they can robustly detect mutations at 0.1% fractional representation. This represents accurate detection of one mutant per every 1000 wild-type alleles. The method for detecting rare variants compares the baseline error rate from multiple reference replicates to the sample error rate at each position. To demonstrate the utility of the method, they analyzed nine clinical samples of H1N1 influenza A and detected an oseltamivir (antiviral therapy) resistance mutation in the H1N1 neuraminidase gene in 0.18% of the samples.

Illumina Technology: Genome Analyzer$_{IIx}$

[28] Dunowska, M., Biggs, P. J., Zheng, T. and Perrott, M. R. (2012) Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula). Vet Microbiol 156: 418–424
[29] Moore, R. A., Warren, R. L., Freeman, J. D., Gustavsen, J. A., Chenard, C., et al. (2011) The sensitivity of massively parallel sequencing for detecting candidate infectious agents associated with human tissue. PLoS ONE 6: e19838
[30] Li, S. C., Chan, W. C., Lai, C. H., Tsai, K. W., Hsu, C. N., et al. (2011) UMARS: Un-MAppable Reads Solution. BMC Bioinformatics 12 Suppl 1: S9

DNA Viruses

Routine sequencing of DNA viruses has produced a large number of viral genomes, which highlights the remarkable variability of viruses. The differences between the genomes of laboratory strains and clinical isolates of the same virus can be substantial and underscores the need to routinely sequence clinical isolates.[31]

References:

Conway, C., Chalkley R., High A., Maclennan K., Berri S., et al. (2012) Next-generation sequencing for simultaneous determination of human papillomavirus load, subtype, and associated genomic copy number changes in tumors. J Mol Diagn 14: 104–111

This study uses next-generation sequencing to investigate viral infection in 44 head and neck tumor types from FFPE samples. The authors were able to detect HPV subtypes that would not have been detected by traditional methods. They then use 8 cell lines to show that this approach could be applied to various tumors and viruses.

Illumina Technology: Genome Analyzer with 76 bp paired-end reads.

Minot, S., Grunberg S., Wu G. D., Lewis J. D. and Bushman F. D. (2012) Hypervariable loci in the human gut virome. Proc Natl Acad Sci U S A 109: 3962–3966

RNA Viruses

The high mutation rate in RNA viruses arises from error-prone polymerases and limited RNA proofreading functions.[32] This low-replication fidelity results in RNA virus populations that have been described as quasispecies, a cloud or assemblage of wild-type (WT) and mutant genomes that exist at a mutation-selection equilibrium.[33] Recent studies have shown that virus diversity is essential for adaptive evolution and the capacity to cause disease.[34]

References:

Blasdell, K. R., Voysey R., Bulach D., Joubert D. A., Tesh R. B., et al. (2012) Kotonkan and Obodhiang viruses: African ephemeroviruses with large and complex genomes. Virology 425: 143–153

This paper describes the complete sequences of the OBOV and KOTV genomes. Genetic and serological data indicate that KOTV and OBOV should be classified as new species in the genus Ephemerovirus. This is an example of using sequencing to identify a new RNA-virus species.

Illumina Technology: Genome Analyzer with 75 bp paired-end reads.

[31] Szpara, M. L., Parsons, L. and Enquist, L. W. (2010) Sequence variability in clinical and laboratory isolates of herpes simplex virus 1 reveals new mutations. J Virol 84: 5303–5313
[32] Drake, J. W. and Holland, J. J. (1999) Mutation rates among RNA viruses. Proc Natl Acad Sci U S A 96: 13910–13913
[33] Bull, J. J., Meyers, L. A. and Lachmann, M. (2005) Quasispecies made simple. PLoS Comput Biol 1: e61
[34] Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. and Andino, R. (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439: 344–348

Al Rwahnih, M., Dolja V. V., Daubert S., Koonin E. V. and Rowhani A. (2012) Genomic and biological analysis of Grapevine leafroll-associated virus 7 reveals a possible new genus within the family Closteroviridae. Virus Res 163: 302–309

Dunowska, M., Biggs P. J., Zheng T. and Perrott M. R. (2012) Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula). Vet Microbiol 156: 418–424

Perera, O. P., Snodgrass G. L., Allen K. C., Jackson R. E., Becnel J. J., et al. (2012) The complete genome sequence of a single-stranded RNA virus from the tarnished plant bug, Lygus lineolaris (Palisot de Beauvois). J Invertebr Pathol 109: 11–19

Moore, J., Jironkin A., Chandler D., Burroughs N., Evans D. J., et al. (2011) Recombinants between Deformed wing virus and Varroa destructor virus-1 may prevail in Varroa destructor-infested honeybee colonies. J Gen Virol 92: 156–161

Virus mRNA

Sequencing viral mRNAs can provide a wealth of information about the activity of viruses, as well as their mechanisms of action.[35] This information can then be used to annotate the viral genome. Until the advent of next-generation sequencing, it was very difficult and laborious to sequence viruses. Our lack of knowledge about viral sequences represents a large gap in our understanding of microbiological population dynamics.

References:

Yang, Z., Bruno D. P., Martens C. A., Porcella S. F. and Moss B. (2011) Genome-wide analysis of the 5' and 3' ends of vaccinia virus early mRNAs delineates regulatory sequences of annotated and anomalous transcripts. J Virol 85: 5897-5909

Poxyviruses are large DNA viruses that encode a multisubunit RNA polymerase, stage-specific transcription factors, and enzymes that cap and polyadenylate mRNAs within the cytoplasm of infected animal cells. This paper determines the precise 5’ and 3’ ends of vaccinia virus (VACV) early mRNAs and then proceed to map the putative transcription start sites (TSSs) and polyadenylation sites (PASs).

Illumina Technology: Genome Analyzer

Virus Small RNAs

Small RNAs play a key role in the host-pathogen interaction during virus infections. Micro-RNAs (miRNAs) are a class of small noncoding RNAs involved in post-transcriptional regulation in organisms ranging from plants to higher mammals. Both RNA and DNA viruses use miRNA for host and viral gene regulation.

[35] Jiang, X., Jiang, H., Li, C., Wang, S., Mi, Z., et al. (2011) Sequence characteristics of T4-like bacteriophage IME08 benome termini revealed by high throughput sequencing. Virol J 8: 194

References:

Meyer, C., Grey F., Kreklywich C. N., Andoh T. F., Tirabassi R. S., et al. (2011) Cytomegalovirus microRNA expression is tissue specific and is associated with persistence. J Virol 85: 378–389

The authors report the sequencing of the small RNAs in rat cytomegalovirus-infected fibroblasts and persistently infected salivary glands. They identified 24 unique miRNAs that mapped to hairpin structures found within the viral genome.

Illumina Technology: Genome Analyzer

Yang, X., Wang Y., Guo W., Xie Y., Xie Q., et al. (2011) Characterization of small interfering RNAs derived from the geminivirus/betasatellite complex using deep sequencing. PLoS ONE 6: e16928

Host-Pathogen Interactions

Viral metagenomics is expanding our current knowledge of virus–host interactions by uncovering genes that may allow viruses to manipulate their hosts in unexpected ways. The intrinsic potential for virus discovery through viral metagenomics can help advance a wide array of disciplines, including evolutionary biology, pathogen surveillance, and biotechnology.[36]

Review:

Hann, D. R., Gimenez-Ibanez S. and Rathjen J. P. (2010) Bacterial virulence effectors and their activities. Curr Opin Plant Biol 13: 388–393

References:

Chang, S. T., Sova P., Peng X., Weiss J., Law G. L., et al. (2011) Next-generation sequencing reveals HIV-1-mediated suppression of T cell activation and RNA processing and regulation of noncoding RNA expression in a CD4+ T cell line. MBio 2

The authors sequenced the total mRNA from infected cells and detected differences in the expression of both host and viral mRNAs. Viral reads represented a large portion of the total mapped sequencing reads: approximately 20% at 12 h post infection (hpi) and 40% at 24 hpi. This study clearly demonstrates the value of a sequencing approach to host-pathogen interactions.

Illumina Technology: Genome Analyzer$_{IIx}$, Illumina FastTrack services

Avrani, S., Wurtzel O., Sharon I., Sorek R. and Lindell D. (2011) Genomic island variability facilitates Prochlorococcus-virus coexistence. Nature 474: 604–608

This study shows how next-generation sequencing technology makes it possible to study environmental interactions that were previously unimaginable. This general approach can also be applied to any organism-virus interactions, such as plants. Genome analysis of 77 sub-strains, selected for resistance to ten viruses, reveal mutations primarily in non-conserved, horizontally transferred genes that localize to a single hyper variable genomic island. Mutations affected viral attachment to the cell surface and imposed a fitness cost to the host.

Illumina Technology: Genome Analyzer

[36] Rosario, K. and Breitbart, M. (2011) Exploring the viral world through metagenomics. Curr Opin Virol 1: 289–297

Human Viral Pathogens

In addition to improving detection of disease-causing viruses, genomic methods have highlighted the prevalence of viruses in healthy individuals. For example, two groups from the family Picornaviridae are common on mucosal surfaces: rhinoviruses and gastrointestinal enteroviruses. In contrast to a "one-pathogen–one-disease" model, a more complex model of the human virome suggests that people are almost continually exposed to viruses, which may or may not cause symptoms. In this context the virome is an important component of the environment that can interact with host genetic traits to contribute to the pathogenesis of complex diseases.[37]

Reviews:

Foxman, E. F. and Iwasaki A. (2011) Genome-virome interactions: examining the role of common viral infections in complex disease. Nat Rev Microbiol 9: 254–264

Oliere, S., Douville R., Sze A., Belgnaoui S. M. and Hiscott J. (2011) Modulation of innate immune responses during human T-cell leukemia virus (HTLV-1) pathogenesis. Cytokine Growth Factor Rev 22: 197–210

References:

Nasu, A., Marusawa H., Ueda Y., Nishijima N., Takahashi K., et al. (2011) Genetic heterogeneity of hepatitis C virus in association with antiviral therapy determined by ultra-deep sequencing. PLoS ONE 6: e24907

In this study the authors used ultra-deep sequencing to determine the presence of viral quasispecies of genotype 1b HCV in patients receiving peg-interferon (IFN) a2b plus ribavirin (RBV). The authors demonstrate that clones resistant to direct-acting antivirals for HCV, such as viral protease and polymerase inhibitors, preexist with various abundances in all 27 treatment-naive patients. This observation implies that there is a risk that drug resistance may develop against these agents.

Illumina Technology: Genome Analyzer$_{II}$

Murray, C. L., Oh T. S. and Rice C. M. (2011) Keeping Track of Viruses. Microbial Forensics (Second Edition) 137–153

Yongfeng, H., Fan Y., Jie D., Jian Y., Ting Z., et al. (2011) Direct pathogen detection from swab samples using a new high-throughput sequencing technology. Clin Microbiol Infect 17: 241–244

Plant Viral Pathogens

Plants have a well-defined defense mechanism against invasive nucleic acids such as viral transcripts. The silencing pathway is quite sophisticated in higher eukaryotes, but the distinct steps and nature of effector complexes vary between and even within species.[38]

[37] Foxman, E. F. and Iwasaki, A. (2011) Genome-virome interactions: examining the role of common viral infections in complex disease. Nat Rev Microbiol 9: 254–264
[38] Alvarado, V. and Scholthof, H. B. (2009) Plant responses against invasive nucleic acids: RNA silencing and its suppression by plant viral pathogens. Semin Cell Dev Biol 20: 1032–1040

References:

Wylie, S. J. and Jones M. G. (2011) The complete genome sequence of a Passion fruit woodiness virus isolate from Australia determined using deep sequencing, and its relationship to other potyviruses. Arch Virol 156: 479–482

This paper reports the complete genome sequence (9,858 nucleotides) of the Passion fruit woodiness virus isolate MU-2 determined with Illumina sequencing. It is a typical potyvirus.

Illumina Technology: Genome Analyzer$_{IIx,}$ 76-bp 31M reads

Luan, J. B., Li J. M., Varela N., Wang Y. L., Li F. F., et al. (2011) Global analysis of the transcriptional response of whitefly to tomato yellow leaf curl China virus reveals the relationship of coevolved adaptations. J Virol 85: 3330–3340

Insect Viral Pathogens

A large number of insect viruses with small RNA genomes and morphological resemblance to vertebrate picornaviruses have been characterized. While some of these viruses may cause only latent infections without much adverse effect on the host, others may be debilitating lethal infections in the host.[39]

References:

Perera, O. P., Snodgrass G. L., Allen K. C., Jackson R. E., Becnel J. J., et al. (2012) The complete genome sequence of a single-stranded RNA virus from the tarnished plant bug, Lygus lineolaris (Palisot de Beauvois). J Invertebr Pathol 109: 11-19

The authors generated a complete genome sequence of a single-stranded RNA virus by sequencing cDNA prepared from infected insects. High similarity to the honey bee sacbrood virus (SBV) genome and similarities in the genome organization and amino acid sequence with the viruses of the family Iflaviridae suggested that LyLV-1 was a novel member of this family.

Illumina Technology: Genome Analyzer$_{II}$ 36 bp reads

Moore, J., Jironkin A., Chandler D., Burroughs N., Evans D. J., et al. (2011) Recombinants between Deformed wing virus and Varroa destructor virus-1 may prevail in Varroa destructor-infested honeybee colonies. J Gen Virol 92: 156-161

The authors identified novel recombinants between deformed wing virus (DWV) and Varroa destructor virus-1 (VDV-1). This recombinant is better adapted to transmission between V. destructor and honeybees than the parental DWV.

Illumina Technology: Genome Analyzer$_{II}$ 72 bp paired-end run

Zioni, N., Soroker V. and Chejanovsky N. (2011) Replication of Varroa destructor virus 1 (VDV-1) and a Varroa destructor virus 1-deformed wing virus recombinant (VDV-1-DWV) in the head of the honey bee. Virology 417: 106–112

[39] Perera, O. P., Snodgrass, G. L., Allen, K. C., Jackson, R. E., Becnel, J. J., et al. (2012) The complete genome sequence of a single-stranded RNA virus from the tarnished plant bug, Lygus lineolaris (Palisot de Beauvois). J Invertebr Pathol 109: 11–19

Bacteriophages

Bacteriophages represent an absolute majority of all organisms in the biosphere. The genetic diversity of the population is very high and it appears that phages have been actively evolving for billions of years. Frequent horizontal genetic exchange results in pervasive mosaicism in their architectures.

Bacteriophages play an important role in modulating bacterial populations. With the advent of next-generation sequencing, the coming years of phage genome exploration promise to be especially revealing.[40]

Reviews:

Hatfull, G. F. and Hendrix R. W. (2011) Bacteriophages and their Genomes. Curr Opin Virol 1: 298–303

References:

Seed, K. D., Bodi K. L., Kropinski A. M., Ackermann H. W., Calderwood S. B., et al. (2011) Evidence of a dominant lineage of Vibrio cholerae-specific lytic bacteriophages shed by cholera patients over a 10-year period in Dhaka, Bangladesh. MBio 2: e00334-00310

Factors that affect *V. cholerae* in the environment can impact the occurrence of cholera outbreaks. One of these factors is thought to be the presence of bacterial viruses, or bacteriophages. Bacteriophages that prey on *V. cholerae* in the environment, and potentially in humans, have not been extensively genetically characterized. In this paper the authors isolated and sequenced the genomes of bacteriophages from cholera patient stool samples collected over a 10-year period in Dhaka Bangladesh, a region that suffers from regular cholera outbreaks.

Illumina Technology: Genome Analyzer paired-end libraries with barcodes.

Sabehi, G., Shaulov L., Silver D. H., Yanai I., Harel A., et al. (2012) A novel lineage of myoviruses infecting cyanobacteria is widespread in the oceans. Proc Natl Acad Sci U S A 109: 2037–2042

Jiang, X., Jiang H., Li C., Wang S., Mi Z., et al. (2011) Sequence characteristics of T4-like bacteriophage IME08 benome termini revealed by high throughput sequencing. Virol J 8: 194

Kent, B. N., Salichos L., Gibbons J. G., Rokas A., Newton I. L., et al. (2011) Complete bacteriophage transfer in a bacterial endosymbiont (Wolbachia) determined by targeted genome capture. Genome Biol Evol 3: 209–218

Santos, T. M. and Bicalho R. C. (2011) Complete genome sequence of vB_EcoM_ECO1230-10: a coliphage with therapeutic potential for bovine metritis. Vet Microbiol 148: 267–275

Symbiosis

The evolution of intimate symbiosis requires the coordination of genome content and gene expression between the distinct partner genomes. This coordination allows the fusion of capabilities of each organism into a single integrated metabolism.

[40] Hatfull, G. F. and Hendrix, R. W. (2011) Bacteriophages and their Genomes. Curr Opin Virol 1: 298–303

Reviews:

Keeney, K. M. and Finlay B. B. (2011) Enteric pathogen exploitation of the microbiota-generated nutrient environment of the gut. Curr Opin Microbiol 14: 92–98

Leggat, W., Yellowlees D. and Medina M. (2011) Recent progress in Symbiodinium transcriptomics. Journal of Experimental Marine Biology and Ecology 408: 120–125

Veneault-Fourrey, C. and Martin F. (2011) Mutualistic interactions on a knife-edge between saprotrophy and pathogenesis. Current opinion in plant biology 14: 444–450

References:

Rosenthal, A. Z., Matson E. G., Eldar A. and Leadbetter J. R. (2011) RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. ISME J 5: 1133-1142

In this paper the authors report the interactions of two closely related Treponema spirochetes in the termite gut. The two spirochetes possess complementary H_2 physiologies: one produces hydrogen while the other consumes it. *In vitro*, these two species markedly enhanced each other's growth; co-cultivation causes comprehensive changes in global gene expression. The expression of well over a 100 genes in each species was changed ≥2-fold, with over a dozen changed ≥10-fold. Some activities beneficial to the host were preferentially expressed during consortial growth. The results point to the intricate networks of metabolic and gene interactions that drive lignocellulose degradation in the termite gut microbiota.

Illumina Technology: Genome Analyzer with standard Illumina RNA-Seq sample preparation protocol

Hansen, A. K. and Moran N. A. (2011) Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. Proc Natl Acad Sci U S A 108: 2849–2854

In the phloem sap diet of aphids the 10 essential amino acids are scarce and are supplied by the obligate bacterial endosymbiont (Buchnera), which lives inside specialized cells called bacteriocytes. Although Buchnera's genome encodes most genes for essential amino acid biosynthesis, several genes in essential amino acid pathways are missing, as are most genes for production of nonessential amino acids. The authors compared pea aphid gene expression between bacteriocytes and other body tissues and found that 26 genes underlying amino acid biosynthesis were up-regulated in bacteriocytes.

Illumina Technology: Genome Analyzer RNA-Seq 74 bp reads

Burke, G. R. and Moran N. A. (2011) Massive Genomic Decay in Serratia symbiotica, a Recently Evolved Symbiont of Aphids. Genome Biol Evol 3: 195–208

Diguistini, S., Wang Y., Liao N. Y., Taylor G., Tanguay P., et al. (2011) Genome and transcriptome analyses of the mountain pine beetle-fungal symbiont Grosmannia clavigera, a lodgepole pine pathogen. Proc Natl Acad Sci U S A 108: 2504–2509

Kent, B. N., Salichos L., Gibbons J. G., Rokas A., Newton I. L., et al. (2011) Complete Bacteriophage Transfer in a Bacterial Endosymbiont (Wolbachia) Determined by Targeted Genome Capture. Genome Biol Evol 3: 209–218

Nowack, E. C., Vogel H., Groth M., Grossman A. R., Melkonian M., et al. (2011) Endosymbiotic gene transfer and transcriptional regulation of transferred genes in Paulinella chromatophora. Mol Biol Evol 28: 407–422

Shinzato, C., Shoguchi E., Kawashima T., Hamada M., Hisata K., et al. (2011) Using the Acropora digitifera genome to understand coral responses to environmental change. Nature 476: 320–323

Wang, Z., Kadouri D. E. and Wu M. (2011) Genomic insights into an obligate epibiotic bacterial predator: Micavibrio aeruginosavorus ARL-13. BMC Genomics 12: 453

Human Health

Humans carry ten times more bacterial cells than human cells, and 100 times more bacterial genes than our inherited human genome.[41] The impact on human health is so significant that the human micro biome can be considered an additional organ, or organs. Host-gene-microbial interactions are major determinants for the development of some multifactorial chronic disorders.[42]

Reviews:

Kuczynski, J., Lauber C. L., Walters W. A., Parfrey L. W., Clemente J. C., et al. (2012) Experimental and analytical tools for studying the human microbiome. Nat Rev Genet 13: 47–58

Bruls, T. and Weissenbach J. (2011) The human metagenome: our other genome? Hum Mol Genet 20: R142–148

Laing, C. R., Zhang Y., Thomas J. E. and Gannon V. P. (2011) Everything at once: comparative analysis of the genomes of bacterial pathogens. Vet Microbiol 153: 13–26

Virgin, H. W. and Todd J. A. (2011) Metagenomics and personalized medicine. Cell 147: 44–56

The Human Gut

If expanded, the surface of the small intestine alone can reach roughly the size of a tennis court, or 100 times the area of the skin.[43] Nearly 100 trillion bacteria are harbored within the human gastrointestinal tract, with archea, fungi, and viruses representing smaller numbers of the gut microbial community. Disturbances in the populations of the gut microbiota have been shown to be associated with diseases such as obesity,[44] diabetes[45,46], and inflammatory bowel disease.[47]

References:

Gisselsson, D., Lindgren D., Mengelbier L. H., Ora I. and Yeger H. (2010) Genetic bottlenecks and the hazardous game of population reduction in cell line based research. Exp Cell Res 316: 3379–3386

[41] Barwick, B. G., Abramovitz, M., Kodani, M., Moreno, C. S., Nam, R., et al. (2010) Prostate cancer genes associated with TMPRSS2-ERG gene fusion and prognostic of biochemical recurrence in multiple cohorts. Br J Cancer 102: 570–576
[42] Vaarala, O., Atkinson, M. A. and Neu, J. (2008) The "perfect storm" for type 1 diabetes: the complex interplay between intestinal microbiota, gut permeability, and mucosal immunity. Diabetes 57: 2555–2562
[43] Hooper, L. V. and Macpherson, A. J. (2010) Immune adaptations that maintain homeostasis with the intestinal microbiota. Nat Rev Immunol 10: 159–169
[44] Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., et al. (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 444: 1027–1031
[45] Wen, L., Ley, R. E., Volchkov, P. Y., Stranges, P. B., Avanesyan, L., et al. (2008) Innate immunity and intestinal microbiota in the development of Type 1 diabetes. Nature 455: 1109–1113
[46] Vaarala, O., Atkinson, M. A. and Neu, J. (2008) The "perfect storm" for type 1 diabetes: the complex interplay between intestinal microbiota, gut permeability, and mucosal immunity. Diabetes 57: 2555–2562
[47] Ott, S. J., Musfeldt, M., Wenderoth, D. F., Hampe, J., Brant, O., et al. (2004) Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. Gut 53: 685–693

The gene catalogue of the human gut microbiome[48]	
Number of genes in the study	3,300,000
Number of genes shared by less than 20% of individuals	2,376,000
Core genes (shared by at least 50% of individuals)	294,000
Average number of genes carried per individual	536,000
Number of shared genes (with at least one other individual)	204,000
Bacteria	99.1%
Archaea	0.8%
Eukaryotic or viral origin	0.1%
Number of species present in each individual	160

Reviews:

Gonzalez, A., Clemente J. C., Shade A., Metcalf J. L., Song S., et al. (2011) Our microbial selves: what ecology can teach us. EMBO Rep 12: 775–784

Karlsson, F. H., Nookaew I., Petranovic D. and Nielsen J. (2011) Prospects for systems biology and modeling of the gut microbiome. Trends Biotechnol 29: 251–258

Walter, J., Britton R. A. and Roos S. (2011) Host-microbial symbiosis in the vertebrate gastrointestinal tract and the Lactobacillus reuteri paradigm. Proc Natl Acad Sci U S A 108 Suppl 1: 4645–4652

Williams, I. R. (2011) Deep impact: deciphering mucosal microbiomes using next-generation sequencing approaches. Mucosal Immunology 4: 586–587

References:

Castellarin, M., Warren R. L., Freeman J. D., Dreolini L., Krzywinski M., et al. (2012) Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. Genome Res 22: 299–306

The authors carried out total RNA-seq of 11 colorectal tumor samples and 11 matched controls. The genome sequences of *Fusobacterium nucleatum* subsp. *nucleatum* were enriched in the tumor samples. *Fusobacterium nucleatum* is rare in the normal gut and usually associated with dental plaque and periodontitis. The results were validated by quantitative PCR analysis from a total of 99 subjects ($p = 2.5 \times 10^{-6}$). The authors proceeded to culture and sequence a strain and also developed a PCR screen.

Illumina Technology: Genome Analyzer$_{IIx,}$ for RNA-Seq and HiSeq for bacterial sequencing.

[48] Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464: 59–65

Kuwahara, T., Ogura Y., Oshima K., Kurokawa K., Ooka T., et al. (2011) The lifestyle of the segmented filamentous bacterium: a non-culturable gut-associated immunostimulating microbe inferred by whole-genome sequencing. DNA Res 18: 291–303

Segmented filamentous bacteria (SFBs) are non-culturable, clostridia-related Gram-positive bacteria with a unique morphology and tight attachment to intestinal epithelial cells. The authors report the complete genome sequence of mouse SFBs. The genome lacks genes for the biosynthesis of almost all amino acids, vitamins/cofactors and nucleotides, but contains a full set of genes for sporulation/germination and for chemotaxis/flagella-based motility. These findings suggest a triphasic lifestyle of the SFB, which comprises two types of vegetative (swimming and epicellular parasitic) phases and a dormant (spore) phase. This is an example of the rich biological information that can be obtained through genome sequencing.

Illumina Technology: Genome Analyzer sequences generated in the MetaHIT project[49]

Minot, S., Grunberg S., Wu G. D., Lewis J. D. and Bushman F. D. (2012) Hypervariable loci in the human gut virome. Proc Natl Acad Sci U S A 109: 3962–3966

Gong, J. and Yang C. Advances in the methods for studying gut microbiota and their relevance to the research of dietary fiber functions. Food Research International

Arumugam, M., Raes J., Pelletier E., Le Paslier D., Yamada T., et al. (2011) Enterotypes of the human gut microbiome. Nature 473: 174–180

IBD and Crohn's Disease

Crohn's disease and ulcerative colitis are the major types of inflammatory bowel disease (IBD). Crohn's disease is characterized by chronic inflammation that can affect any part of the gastrointestinal (GI) tract, whereas ulcerative colitis is limited to the colon. Symptoms include abdominal pain, diarrhea, weight loss and ulceration of the GI tract. It has been suggested that the gut microbiota is involved in IBD progression, depending on the host genotype[50] and immune system[51].

Reviews:

Man, S. M., Kaakoush N. O. and Mitchell H. M. (2011) The role of bacteria and pattern-recognition receptors in Crohn's disease. Nat Rev Gastroenterol Hepatol 8: 152–168

Ohnmacht, C., Marques R., Presley L., Sawa S., Lochner M., et al. (2011) Intestinal microbiota, evolution of the immune system and the bad reputation of pro-inflammatory immunity. Cell Microbiol 13: 653–659

[49] Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464: 59–65
[50] Festen, E. A., Goyette, P., Green, T., Boucher, G., Beauchamp, C., et al. (2011) A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. PLoS Genet 7: e1001283
[51] Karlsson, F. H., Nookaew, I., Petranovic, D. and Nielsen, J. (2011) Prospects for systems biology and modeling of the gut microbiome. Trends Biotechnol 29: 251–258

References:

Deshpande, N. P., Kaakoush N. O., Mitchell H., Janitz K., Raftery M. J., et al. (2011) Sequencing and validation of the genome of a Campylobacter concisus reveals intra-species diversity. PLoS ONE 6: e22170

This paper reports the genome of a *C. concisus* (UNSWCD) from a biopsy of a child with Crohn's disease. The authors observe substantial functional differences between UNSWCD and the reference strain (BAA-1457). While 1593 genes were conserved across UNSWCD and BAA-1457, 138 genes from UNSWCD and 281 from BAA-1457 were unique when compared against the other. Many of the observed differences are consistent with UNSWCD having adapted to greater surface interaction with host cells.

Illumina Technology: Genome Analyzer$_{IIx,}$ paired-end 2x 102 bp to 2,500x coverage.

Biswas, A., Petnicki-Ocwieja T. and Kobayashi K. S. (2012) Nod2: a key regulator linking microbiota to intestinal mucosal immunity. J Mol Med (Berl) 90: 15–24

Type 1 Diabetes

Recent studies have suggested that bacteria may play a role in the development of Type 1 diabetes (T1D), a complex, multifactorial autoimmune disease.[52]

Reviews:

Virgin, H. W. and Todd J. A. (2011) Metagenomics and personalized medicine. Cell 147: 44–56

References:

Brown, C. T., Davis-Richardson A. G., Giongo A., Gano K. A., Crabb D. B., et al. (2011) Gut microbiome metagenomics analysis suggests a functional model for the development of autoimmunity for type 1 diabetes. PLoS ONE 6: e25792

The authors analyzed stool samples from four pairs of matched T1D case-control subjects collected at the time of the development of T1D-associated autoimmunity. The data suggest that increased adhesion and flagella synthesis in autoimmune subjects may be involved in triggering a T1D associated autoimmune response. Extensive differences in metabolic potential indicate that autoimmune subjects have a functionally aberrant microbiome.

Keene, K. L., Quinlan A. R., Hou X., Hall I. M., Mychaleckyj J. C., et al. (2012) Evidence for two independent associations with type 1 diabetes at the 12q13 locus. Genes Immun 13: 66–70

Bradfield, J. P., Qu H. Q., Wang K., Zhang H., Sleiman P. M., et al. (2011) A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. PLoS Genet 7: e1002293

Cotsapas, C., Voight B. F., Rossin E., Lage K., Neale B. M., et al. (2011) Pervasive sharing of genetic effects in autoimmune disease. PLoS Genet 7: e1002254

[52] Virgin, H. W. and Todd, J. A. (2011) Metagenomics and personalized medicine. Cell 147: 44–56

Other Human Microbiota

While the gut microbiome has received much attention, there are an increasing number of studies on other human microbiota, such as vaginal and skin microbiota. The ambitious Human Microbiome Project[53] sampled several different sites on the human body, including nasal passages, oral cavities, skin, gastrointestinal tract, and urogenital tract. The project has increased awareness of the complexity and importance of the metagenome in human health.

Reviews:

Kong, H. H. (2011) Skin microbiome: genomics-based insights into the diversity and role of skin microbes. Trends Mol Med 17: 320–328

Proctor, L. M. (2011) The Human Microbiome Project in 2011 and beyond. Cell Host Microbe 10: 287–291

Reid, G., Younes J. A., Van der Mei H. C., Gloor G. B., Knight R., et al. (2011) Microbiota restoration: natural and supplemented recovery of human microbial communities. Nat Rev Microbiol 9: 27–38

References:

Macklaim, J. M., Gloor G. B., Anukam K. C., Cribby S. and Reid G. (2011) At the crossroads of vaginal health and disease, the genome sequence of Lactobacillus iners AB-1. Proc Natl Acad Sci U S A 108 Suppl 1: 4688–4695

Lactobacillus iners is present in the vagina of persons deemed healthy, infected with BV, or has just been subjected to antimicrobial therapy. This remarkable ability to survive under a range of conditions suggests that *L. iners* is not an aberrant microbiota, but an important member of the host's defenses. In this role it may serve as a persistent mutualistic lactobacilli involved in restoration and maintenance of the normal microbiota.

Illumina Technology: Genome Analyzer 12 million short (35 bp) paired-end reads were used to supplement the assembly.

Pride, D. T., Sun C. L., Salzman J., Rao N., Loomer P., et al. (2011) Analysis of streptococcal CRISPRs from human saliva reveals substantial sequence diversity within and between subjects over time. Genome Res 21: 126–136

Probiotics

Probiotics are live microorganisms that can confer a health benefit to the host. Recent advances in next-generation sequencing have advanced our understanding of probiotic functionality and the specific interactions between probiotics and their human hosts. Although not all probiotics use the same mechanisms to confer benefits to hosts, some specific mechanisms of action have been revealed through genomic investigations.[54] This should eventually lead to a more informed design of pre/probiotics, cosmetics and bio-active healthcare products[55]

[53] http://commonfund.nih.gov/hmp/
[54] Baugher, J. L. and Klaenhammer, T. R. (2011) Invited review: Application of omics tools to understanding probiotic functionality. J Dairy Sci 94: 4753–4765
[55] Gonzalez, A., Clemente, J. C., Shade, A., Metcalf, J. L., Song, S., et al. (2011) Our microbial selves: what ecology can teach us. EMBO Rep 12: 775–784

Reviews:

Baugher, J. L. and Klaenhammer T. R. (2011) Invited review: Application of omics tools to understanding probiotic functionality. J Dairy Sci 94: 4753–4765

Reid, G., Younes J. A., Van der Mei H. C., Gloor G. B., Knight R., et al. (2011) Microbiota restoration: natural and supplemented recovery of human microbial communities. Nat Rev Microbiol 9: 27–38

References:

Chen, C., Ai L., Zhou F., Wang L., Zhang H., et al. (2011) Complete genome sequence of the probiotic bacterium Lactobacillus casei LC2W. J Bacteriol 193: 3419–3420

Macklaim, J. M., Gloor G. B., Anukam K. C., Cribby S. and Reid G. (2011) At the crossroads of vaginal health and disease, the genome sequence of Lactobacillus iners AB-1. Proc Natl Acad Sci U S A 108 Suppl 1: 4688–4695

Saulnier, D. M., Santos F., Roos S., Mistretta T. A., Spinler J. K., et al. (2011) Exploring metabolic pathway reconstruction and genome-wide expression profiling in Lactobacillus reuteri to define functional probiotic features. PLoS ONE 6: e18783

Wang, Y., Chen C., Ai L., Zhou F., Zhou Z., et al. (2011) Complete genome sequence of the probiotic Lactobacillus plantarum ST-III. J Bacteriol 193: 313–314

Wang, Y., Chen Y., Shen Q. and Yin X. (2011) Molecular cloning and identification of the laspartomycin biosynthetic gene cluster from Streptomyces viridochromogenes. Gene 483: 11–21

Soil

It has been estimated that there are between 100,000 and 1,000,000 different microbial species per gram of soil[56], but only a very small portion can be cultured in the laboratory.

Reviews:

Maron, P. A., Mougel C. and Ranjard L. (2011) Soil microbial diversity: Methodological strategy, spatial overview and functional interest. C R Biol 334: 403–411

References:

Mackelprang, R., Waldrop M. P., DeAngelis K. M., David M. M., Chavarria K. L., et al. (2011) Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. Nature 480: 368–371

The authors use shotgun metagenomic sequencing to determine the impact of thaw on microbial phylogenetic and functional genes, and relate these data to measurements of methane emissions.

Illumina Technology: Genome Analyzer$_{II}$ with 113 bp paired-end sequencing

Bartram, A. K., Lynch M. D., Stearns J. C., Moreno-Hagelsieb G. and Neufeld J. D. (2011) Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. Appl Environ Microbiol 77: 3846–3852

[56] Curtis, T. P. and Sloan, W. T. (2005) Microbiology. Exploring microbial diversity--a vast below. Science 309: 1331–1333

Ma, M., Wang C., Ding Y., Li L., Shen D., et al. (2011) Complete genome sequence of Paenibacillus polymyxa SC2, a strain of plant growth-promoting Rhizobacterium with broad-spectrum antimicrobial activity. J Bacteriol 193: 311–312

Veneault-Fourrey, C. and Martin F. (2011) Mutualistic interactions on a knife-edge between saprotrophy and pathogenesis. Curr Opin Plant Biol 14: 444–450

Bioremediation

Bioremediation is an ecofriendly and cost-competitive strategy for eliminating xenobiotic compounds from polluted environments. Next-generation sequencing is providing crucial insights in the molecular and biological mechanisms involved in bioremediation of environmental pollutants. These insights will improve microbial bioremediation strategies, monitoring their progress and determining their success.[57]

References:

Zhan, Y., Yan Y., Zhang W., Chen M., Lu W., et al. (2012) Comparative analysis of the complete genome of an Acinetobacter calcoaceticus strain adapted to a phenol-polluted environment. Res Microbiol 163: 36–43

The authors report the complete genome sequence of *Acinetobacter calcoaceticus* PHEA-2, a non-pathogenic phenol-degrading bacterium previously isolated from industrial wastewater of an oil refinery in China. They conclude that many genes associated with environmental adaptation were acquired by horizontal gene transfer, including an 8-kb phenol degradation gene cluster.

Illumina Technology: Genome Analyzer

Li, K., Wang S., Shi Y., Qu J., Zhai Y., et al. (2011) Genome sequence of Paracoccus sp. Strain TRP, a Chlorpyrifos Biodegrader. J Bacteriol 193: 1786–1787, Zhang, T., Zhang X. X. and Ye L. (2011) Plasmid metagenome reveals high levels of antibiotic resistance genes and mobile genetic elements in activated sludge. PLoS ONE 6: e26041

The paper describes the draft genome of *Paracoccus sp*. strain TRP. This bacterium was isolated from activated sludge and is capable of completely biodegrading the pesticides chlorpyrifos and 3,5,6-trichloro-2-pyridinol. The draft genome sequence could be used to predict genes for xenobiotic biodegradation and metabolism.

Illumina Technology: Genome Analyzer

Albertsen, M., Hansen L. B., Saunders A. M., Nielsen P. H. and Nielsen K. L. (2011) A metagenome of a full-scale microbial community carrying out enhanced biological phosphorus removal. ISME J

Li, S. G., Tang Y. Q., Nie Y., Cai M. and Wu X. L. (2011) Complete genome sequence of Polymorphum gilvum SL003B-26A1T, a crude oil-degrading bacterium from oil-polluted saline soil. J Bacteriol 193: 2894–2895

[57] Desai, C., Pathak, H. and Madamwar, D. (2010) Advances in molecular and "-omics" technologies to gauge microbial communities and bioremediation at xenobiotic/anthropogen contaminated sites. Bioresour Technol 101: 1558–1569

Marine Environment

Marine environments, including the subsurface are believed to contain a total of approximately 3.67 × 10^{30} microorganisms.[58] With approximately 71% of the earth's surface covered by the ocean, this environment represents 80% of life on earth, and an enormous pool of potential microbial biodiversity and exploitable biotechnology.[59]

Reviews:

Kennedy, J., O'Leary N. D., Kiran G. S., Morrissey J. P., O'Gara F., et al. (2011) Functional metagenomic strategies for the discovery of novel enzymes and biosurfactants with biotechnological applications from marine ecosystems. J Appl Microbiol 111: 787–799

Leggat, W., Yellowlees D. and Medina M. (2011) Recent progress in Symbiodinium transcriptomics. Journal of Experimental Marine Biology and Ecology 408: 120–125

References:

Avrani, S., Wurtzel O., Sharon I., Sorek R. and Lindell D. (2011) Genomic island variability facilitates Prochlorococcus-virus coexistence. Nature 474: 604–608

Collins, A. J. and Nyholm S. V. (2011) Draft genome of Phaeobacter gallaeciensis ANG1, a dominant member of the accessory nidamental gland of Euprymna scolopes. J Bacteriol 193: 3397–3398

Fan, L., Bo S., Chen H., Ye W., Kleinschmidt K., et al. (2011) Genome sequence of Bacillus subtilis subsp. spizizenii gtP20b, isolated from the Indian ocean. J Bacteriol 193: 1276–1277

Jang, Y., Oh H. M., Kang I., Lee K., Yang S. J., et al. (2011) Genome sequence of strain IMCC3088, a proteorhodopsin-containing marine bacterium belonging to the OM60/NOR5 clade. J Bacteriol 193: 3415–3416

Lecroq, B., Lejzerowicz F., Bachar D., Christen R., Esling P., et al. (2011) Ultra-deep sequencing of foraminiferal microbarcodes unveils hidden richness of early monothalamous lineages in deep-sea sediments. Proc Natl Acad Sci U S A 108: 13177–13182

Oh, H. M., Lee K., Jang Y., Kang I., Kim H. J., et al. (2011) Genome sequence of strain IMCC9480, a xanthorhodopsin-bearing betaproteobacterium isolated from the Arctic Ocean. J Bacteriol 193: 3421

Petersen, J. M., Zielinski F. U., Pape T., Seifert R., Moraru C., et al. (2011) Hydrogen is an energy source for hydrothermal vent symbioses. Nature 476: 176–180

Roh, H., Uguru G. C., Ko H. J., Kim S., Kim B. Y., et al. (2011) Genome sequence of the abyssomicin- and proximicin-producing marine actinomycete Verrucosispora maris AB-18-032. J Bacteriol 193: 3391–3392

Trevors, J. T. (2011) Viable but non-culturable (VBNC) bacteria: Gene expression in planktonic and biofilm cells. J Microbiol Methods 86: 266–273

Yoon, H. S., Price D. C., Stepanauskas R., Rajah V. D., Sieracki M. E., et al. (2011) Single-cell genomics reveals organismal interactions in uncultivated marine protists. Science 332: 714–717

[58] Whitman, W. B., Coleman, D. C. and Wiebe, W. J. (1998) Prokaryotes: the unseen majority. Proc Natl Acad Sci U S A 95: 6578–6583
[59] Kennedy, J., Marchesi, J. R. and Dobson, A. D. (2008) Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments. Microb Cell Fact 7: 27

Biofuels and Biocatalysts

Microbial enzymes have many known applications as biocatalysts. However, only a few are currently employed for biocatalysis, even though an annotated collection of more than 190 billion bases is available in metagenome sequence databases.[60] A primary focus of the biofuel field is the breakdown of lignocellulosic plant material due to its sheer abundance and low commercial value.

Reviews:

Xing, M. N., Zhang X. Z. and Huang H. (2012) Application of metagenomic techniques in mining enzymes from microbial communities for biofuel synthesis. Biotechnol Adv

Kennedy, J., O'Leary N. D., Kiran G. S., Morrissey J. P., O'Gara F., et al. (2011) Functional metagenomic strategies for the discovery of novel enzymes and biosurfactants with biotechnological applications from marine ecosystems. J Appl Microbiol 111: 787–799

Langridge, G. (2011) Plant power: converting a kingdom. Nat Rev Microbiol 9: 316

Stephenson, P. G., Moore C. M., Terry M. J., Zubkov M. V. and Bibby T. S. (2011) Improving photosynthesis for algal biofuels: toward a green revolution. Trends Biotechnol 29: 615–623

References:

Qi, M., Wang P., O'Toole N., Barboza P. S., Ungerfeld E., et al. (2011) Snapshot of the eukaryotic gene expression in muskoxen rumen--a metatranscriptomic approach. PLoS ONE 6: e20521

In this paper a metatranscriptomic approach was used to investigate the functional diversity of the eukaryotic microorganisms within the rumen of muskoxen (Ovibos moschatus), with a focus on plant cell wall degrading enzymes. The authors identified plant cell wall degrading enzyme modules including glycoside hydrolases, carbohydrate esterases and polysaccharide lyases. These included a number of glycoside hydrolase family 6 (GH6), GH48 and swollenin modules, which have rarely been described in previous gut metagenomic studies.

Illumina Technology: Genome Analyzer$_{II}$

Rosenthal, A. Z., Matson E. G., Eldar A. and Leadbetter J. R. (2011) RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. ISME J 5: 1133–1142

In this paper the authors report the interactions of two closely related Treponema spirochetes in the termite gut. The two spirochetes possess complementary H2 physiologies: one produces hydrogen while the other consumes it. In vitro these two species markedly enhanced each other's growth and co-cultivation causes comprehensive changes in global gene expression. The expression of well over a 100 genes in each species was changed ≥2-fold, with over a dozen changed ≥10-fold. Some activities beneficial to the host were preferentially expressed during consortial growth. The results point to the intricate networks of metabolic and gene interactions that drive lignocellulose degradation in the termite gut microbiota.

Illumina Technology: Genome Analyzer with standard Illumina RNA-Seq sample preparation protocol.

[60] Fernandez-Arrojo, L., Guazzaroni, M. E., Lopez-Cortes, N., Beloqui, A. and Ferrer, M. (2010) Metagenomic era for biocatalyst identification. Curr Opin Biotechnol 21: 725–733

Hess, M., Sczyrba A., Egan R., Kim T. W., Chokhawala H., et al. (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science 331: 463–467

In this study the authors identified 27,755 putative carbohydrate-active genes and expressed 90 candidate proteins, of which 57% were enzymatically active against cellulosic substrates. They also assembled 15 uncultured microbial genomes, which were validated by complementary methods, including single-cell genome sequencing. The number of candidate carbohydrate-active genes identified in the present study was larger by a factor of 5 than the combined number of candidate carbohydrate active genes from all previous studies.

Illumina Technology: Genome Analyzer$_{IIx}$ 125 bp paired end and HiSeq 2000 100 bp paired end sequences with 200 and 300 bp inserts. Two additional libraries were built with insert sizes of 3kb and 5kb respectively to facilitate genome assembly. Both of these two "jumping" libraries were sequenced using Illumina Genome Analyzer$_{IIx}$ by 75bp pair-end sequencing.

Wang, T. Y., Chen H. L., Lu M. Y., Chen Y. C., Sung H. M., et al. (2011) Functional characterization of cellulases identified from the cow rumen fungus Neocallimastix patriciarum W5 by transcriptomic and secretomic analyses. Biotechnol Biofuels 4: 24

Li, F., Jiang P., Zheng H., Wang S., Zhao G., et al. (2011) Draft genome sequence of the marine bacterium Streptomyces griseoaurantiacus M045, which produces novel manumycin-type antibiotics with a pABA core component. J Bacteriol 193: 3417–3418

Zhou, M., McAllister T. A. and Guan L. L. (2011) Molecular identification of rumen methanogens: Technologies, advances and prospects. Animal Feed Science and Technology 166–167: 76–86

Roh, H., Uguru G. C., Ko H. J., Kim S., Kim B. Y., et al. (2011) Genome sequence of the abyssomicin- and proximicin-producing marine actinomycete Verrucosispora maris AB-18-032. J Bacteriol 193: 3391–3392

Xiong, X. H., Zhi J. J., Yang L., Wang J. H., Zhao Y., et al. (2011) Complete genome sequence of the bacterium Methylovorus sp. strain MP688, a high-level producer of pyrroloquinolone quinone. J Bacteriol 193: 1012–1013

Glossary of Terms

allochthonous	Species that originate from another source and are transient residents of the gut.
archaea:	Single-celled microorganisms with no cell nucleus or any other membrane-bound organelles within their cells. Archaea are particularly numerous in the oceans, and the archaea in plankton may be one of the most abundant groups of organisms on the planet.
autochthonous	Species that colonize and reside at high levels permanently in the gut.
axenic	Germ-free, or without other living organisms.
bacteroidetes	A Gram-negative bacterial phylum that includes Bacteroides, Porphyromonas, and Prevotella.
bacteriocytes	Specialized cells that contain endosymbiotic bacteria.
binning	Grouping sequences based on their nucleotide composition or similarity to a reference database.
contig	A set of overlapping DNA segments.
dysbiosis	Disruption in the normal homeostatic and beneficial relationship between microbes and their host, including disruptions in microbial community structure and function.
enterotype	A classification of living organisms based on its bacteriological ecosystem in the human gut microbiome.
ESS	Environmental Shotgun Sequencing
familial clustering	If a family member is diagnosed with a disease such as type 1 diabetes, ulcerative colitis, or Crohn's disease, then the risk of other first-degree family members is much greater (perhaps as much as 50-fold for some multifactorial disorders) than that for a person taken at random from the general population.
firmicutes	A Gram-positive bacterial phylum that includes the genera *Clostridium, Eubacterium, Faecalibacterium, Roseburia* and *Ruminococcus.*
metadata	Definitional data that provide information about or documentation of other data.
metagenetics	Genetic and genomic studies that consider all of the genes in the metagenome as opposed to considering, in isolation, host genes or genes that confer particular properties (e.g., virulence or commensalism) upon an individual microbe.
metagenome	The DNA obtained from uncultured microorganisms.
metagenomics	The study of genomic DNA obtained from uncultured microorganisms.
metaproteomics	The study of protein molecular data obtained from environmental samples.
metatranscriptomics	The study of transcription sequence data obtained from environmental samples.
microbiome	All genes present in the microbiota.
microbiota	A community of microorganisms harbored by a host.
ontology	A formal representation of a set of concepts and the relationships between them. Ontologies are used to create a consensual unambiguous controlled vocabulary.
OTU	Operational taxonomic unit, species distinction in microbiology. Typically, rDNA and a percent similarity threshold are used for classifying microbes within the same, or different, OTUs.
prebiotics	Food ingredient that promotes the activity and/or growth of specific bacteria so that a health benefit is provided to the host (e.g. inulin and trans- galacto-oligosaccharides). Usually, *Bifidobacteria* and *Lactobacillus* are selectively targeted for growth.
probiotics	A live microbial culture that, when supplemented in sufficient dose, provides a health benefit for the host (e.g., *L. plantarum* and *Bif. longum*).
quasispecies	a cloud or assemblage of wild-type (WT) and mutant genomes that exist at a mutation-selection equilibrium.
rarefaction curve	A curve describing the growth of a number of species discovered as a function of individuals sampled.
ribotype	A phylotypic classification based on rDNA sequences.
scaffold	A series of contigs that are in the right order but not necessarily connected in one contiguous stretch.
shadow ORF	An incorrectly identified ORF that overlaps the coding region of the true ORF.
virome	The sum of all viruses living in the tissues of the host or infecting organisms in the microbiome. These viruses may be further divided into viruses that infect members of each of the three domains of life (e.g., bacterial virome, bacterial phages, or the eukaryotic virome).
xenobiotic	chemical compound, such as a drug, pesticide, or carcinogen that is foreign to a living organism.

Bibliography

Al Rwahnih, M., Dolja, V. V., Daubert, S., Koonin, E. V. and Rowhani, A. (2012) Genomic and biological analysis of Grapevine leafroll-associated virus 7 reveals a possible new genus within the family Closteroviridae. Virus Res 163: 302–309

Albertsen, M., Hansen, L. B., Saunders, A. M., Nielsen, P. H. and Nielsen, K. L. (2011) A metagenome of a full-scale microbial community carrying out enhanced biological phosphorus removal. ISME J

Altermann, E. (2012) Tracing lifestyle adaptation in prokaryotic genomes. Front Microbiol 3: 48

Alvarado, V. and Scholthof, H. B. (2009) Plant responses against invasive nucleic acids: RNA silencing and its suppression by plant viral pathogens. Semin Cell Dev Biol 20: 1032–1040

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., et al. (2011) Enterotypes of the human gut microbiome. Nature 473: 174–180

Asai, T., Zaporojets, D., Squires, C. and Squires, C. L. (1999) An Escherichia coli strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. Proc Natl Acad Sci U S A 96: 1971–1976

Avrani, S., Wurtzel, O., Sharon, I., Sorek, R. and Lindell, D. (2011) Genomic island variability facilitates Prochlorococcus-virus coexistence. Nature 474: 604–608

Bartram, A. K., Lynch, M. D., Stearns, J. C., Moreno-Hagelsieb, G. and Neufeld, J. D. (2011) Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. Appl Environ Microbiol 77: 3846–3852

Barwick, B. G., Abramovitz, M., Kodani, M., Moreno, C. S., Nam, R., et al. (2010) Prostate cancer genes associated with TMPRSS2-ERG gene fusion and prognostic of biochemical recurrence in multiple cohorts. Br J Cancer 102: 570–576

Baugher, J. L. and Klaenhammer, T. R. (2011) Invited review: Application of omics tools to understanding probiotic functionality. J Dairy Sci 94: 4753–4765

Bik, H. M., Porazinska, D. L., Creer, S., Caporaso, J. G., Knight, R., et al. (2012) Sequencing our way towards understanding global eukaryotic biodiversity. Trends Ecol Evol 27: 233–243

Biswas, A., Petnicki-Ocwieja, T. and Kobayashi, K. S. (2012) Nod2: a key regulator linking microbiota to intestinal mucosal immunity. J Mol Med (Berl) 90: 15–24

Blasdell, K. R., Voysey, R., Bulach, D., Joubert, D. A., Tesh, R. B., et al. (2012) Kotonkan and Obodhiang viruses: African ephemeroviruses with large and complex genomes. Virology 425: 143–153

Bomar, L., Maltz, M., Colston, S. and Graf, J. (2011) Directed culturing of microorganisms using metatranscriptomics. MBio 2: e00012-00011

Bradfield, J. P., Qu, H. Q., Wang, K., Zhang, H., Sleiman, P. M., et al. (2011) A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. PLoS Genet 7: e1002293

Brown, C. T., Davis-Richardson, A. G., Giongo, A., Gano, K. A., Crabb, D. B., et al. (2011) Gut microbiome metagenomics analysis suggests a functional model for the development of autoimmunity for type 1 diabetes. PLoS ONE 6: e25792

Bruls, T. and Weissenbach, J. (2011) The human metagenome: our other genome? Hum Mol Genet 20: R142–148

Bull, J. J., Meyers, L. A. and Lachmann, M. (2005) Quasispecies made simple. PLoS Comput Biol 1: e61

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., et al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A 108 Suppl 1: 4516–4522

Caporaso, J. G., Paszkiewicz, K., Field, D., Knight, R. and Gilbert, J. A. (2011) The Western English Channel contains a persistent microbial seed bank. ISME J

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., et al. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J

Castellarin, M., Warren, R. L., Freeman, J. D., Dreolini, L., Krzywinski, M., et al. (2012) Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. Genome Res 22: 299–306

Chang, S. T., Sova, P., Peng, X., Weiss, J., Law, G. L., et al. (2011) Next-generation sequencing reveals HIV-1-mediated suppression of T cell activation and RNA processing and regulation of noncoding RNA expression in a CD4+ T cell line. MBio 2:

Chen, C., Ai, L., Zhou, F., Wang, L., Zhang, H., et al. (2011) Complete genome sequence of the probiotic bacterium Lactobacillus casei LC2W. J Bacteriol 193: 3419–3420

Chitsaz, H., Yee-Greenbaum, J. L., Tesler, G., Lombardo, M. J., Dupont, C. L., et al. (2011) Efficient de novo assembly of single-cell bacterial genomes from short-read data sets. Nat Biotechnol 29: 915–921

Coelho, A. C., Boisvert, S., Mukherjee, A., Leprohon, P., Corbeil, J., et al. (2012) Multiple mutations in heterogeneous miltefosine-resistant Leishmania major population as determined by whole genome sequencing. PLoS Negl Trop Dis 6: e1512

Collins, A. J. and Nyholm, S. V. (2011) Draft genome of Phaeobacter gallaeciensis ANG1, a dominant member of the accessory nidamental gland of Euprymna scolopes. J Bacteriol 193: 3397–3398

Conway, C., Chalkley, R., High, A., Maclennan, K., Berri, S., et al. (2012) Next-generation sequencing for simultaneous determination of human papillomavirus load, subtype, and associated genomic copy number changes in tumors. J Mol Diagn 14: 104–111

Cotsapas, C., Voight, B. F., Rossin, E., Lage, K., Neale, B. M., et al. (2011) Pervasive sharing of genetic effects in autoimmune disease. PLoS Genet 7: e1002254

Curtis, T. P. and Sloan, W. T. (2005) Microbiology. Exploring microbial diversity--a vast below. Science 309: 1331–1333

Degnan, P. H. and Ochman, H. (2012) Illumina-based analysis of microbial community diversity. ISME J 6: 183–194

Desai, C., Pathak, H. and Madamwar, D. (2010) Advances in molecular and "-omics" technologies to gauge microbial communities and bioremediation at xenobiotic/anthropogen contaminated sites. Bioresour Technol 101: 1558–1569

Deshpande, N. P., Kaakoush, N. O., Mitchell, H., Janitz, K., Raftery, M. J., et al. (2011) Sequencing and validation of the genome of a Campylobacter concisus reveals intra-species diversity. PLoS ONE 6: e22170

Drake, J. W. and Holland, J. J. (1999) Mutation rates among RNA viruses. Proc Natl Acad Sci U S A 96: 13910–13913

Dunowska, M., Biggs, P. J., Zheng, T. and Perrott, M. R. (2012) Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula). Vet Microbiol 156: 418–424

Fan, L., Bo, S., Chen, H., Ye, W., Kleinschmidt, K., et al. (2011) Genome sequence of Bacillus subtilis subsp. spizizenii gtP20b, isolated from the Indian ocean. J Bacteriol 193: 1276–1277

Fernandez-Arrojo, L., Guazzaroni, M. E., Lopez-Cortes, N., Beloqui, A. and Ferrer, M. (2010) Metagenomic era for biocatalyst identification. Curr Opin Biotechnol 21: 725–733

Festen, E. A., Goyette, P., Green, T., Boucher, G., Beauchamp, C., et al. (2011) A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. PLoS Genet 7: e1001283

Filiatrault, M. J. (2011) Progress in prokaryotic transcriptomics. Curr Opin Microbiol 14: 579–586

Flaherty, P., Natsoulis, G., Muralidharan, O., Winters, M., Buenrostro, J., et al. (2012) Ultrasensitive detection of rare mutations using next-generation targeted resequencing. Nucleic Acids Res 40: e2

Foster, J. A., Bunge, J., Gilbert, J. A. and Moore, J. H. (2012) Measuring the microbiome: perspectives on advances in DNA-based techniques for exploring microbial life. Brief Bioinform

Fox, G. E., Wisotzkey, J. D. and Jurtshuk, P., Jr. (1992) How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int J Syst Bacteriol 42: 166–170

Foxman, E. F. and Iwasaki, A. (2011) Genome-virome interactions: examining the role of common viral infections in complex disease. Nat Rev Microbiol 9: 254–264

Gerlach, W. and Stoye, J. (2011) Taxonomic classification of metagenomic shotgun sequences with CARMA3. Nucleic Acids Res 39: e91

Gong, J. and Yang, C. (2012) Advances in the methods for studying gut microbiota and their relevance to the research of dietary fiber functions. Food Research International

Gonzalez, A., Clemente, J. C., Shade, A., Metcalf, J. L., Song, S., et al. (2011) Our microbial selves: what ecology can teach us. EMBO Rep 12: 775–784

Gonzalez, A. and Knight, R. (2012) Advancing analytical algorithms and pipelines for billions of microbial sequences. Curr Opin Biotechnol 23: 64–71

Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., et al. (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454–pyrosequenced PCR amplicons. Genome Res 21: 494–504

Hann, D. R., Gimenez-Ibanez, S. and Rathjen, J. P. (2010) Bacterial virulence effectors and their activities. Curr Opin Plant Biol 13: 388–393

Harris, S. R., Clarke, I. N., Seth-Smith, H. M., Solomon, A. W., Cutcliffe, L. T., et al. (2012) Whole-genome analysis of diverse Chlamydia trachomatis strains identifies phylogenetic relationships masked by current clinical typing. Nat Genet 44: 413–419

Hatfull, G. F. and Hendrix, R. W. (2011) Bacteriophages and their Genomes. Curr Opin Virol 1: 298–303

Hess, M., Sczyrba, A., Egan, R., Kim, T. W., Chokhawala, H., et al. (2011) Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science 331: 463–467

Ho, P. L., Lo, W. U., Yeung, M. K., Li, Z., Chan, J., et al. (2012) Dissemination of pHK01-like incompatibility group IncFII plasmids encoding CTX-M-14 in Escherichia coli from human and animal sources. Vet Microbiol
Hooper, L. V. and Macpherson, A. J. (2010) Immune adaptations that maintain homeostasis with the intestinal microbiota. Nat Rev Immunol 10: 159–169

Hugenholtz, P., Goebel, B. M. and Pace, N. R. (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J Bacteriol 180: 4765–4774

Islam, S., Kjallquist, U., Moliner, A., Zajac, P., Fan, J. B., et al. (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 21: 1160–1167

Jang, Y., Oh, H. M., Kang, I., Lee, K., Yang, S. J., et al. (2011) Genome sequence of strain IMCC3088, a proteorhodopsin-containing marine bacterium belonging to the OM60/NOR5 clade. J Bacteriol 193: 3415–3416

Jiang, X., Jiang, H., Li, C., Wang, S., Mi, Z., et al. (2011) Sequence characteristics of T4-like bacteriophage IME08 benome termini revealed by high throughput sequencing. Virol J 8: 194

Kalisky, T., Blainey, P. and Quake, S. R. (2011) Genomic analysis at the single-cell level. Annu Rev Genet 45: 431–445

Karlsson, F. H., Nookaew, I., Petranovic, D. and Nielsen, J. (2011) Prospects for systems biology and modeling of the gut microbiome. Trends Biotechnol 29: 251–258

Kav, A. B., Sasson, G., Jami, E., Doron-Faigenboim, A., Benhar, I., et al. (2012) Insights into the bovine rumen plasmidome. Proc Natl Acad Sci U S A 109: 5452–5457

Keene, K. L., Quinlan, A. R., Hou, X., Hall, I. M., Mychaleckyj, J. C., et al. (2012) Evidence for two independent associations with type 1 diabetes at the 12q13 locus. Genes Immun 13: 66–70

Kennedy, J., Marchesi, J. R. and Dobson, A. D. (2008) Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments. Microb Cell Fact 7: 27

Kennedy, J., O'Leary, N. D., Kiran, G. S., Morrissey, J. P., O'Gara, F., et al. (2011) Functional metagenomic strategies for the discovery of novel enzymes and biosurfactants with biotechnological applications from marine ecosystems. J Appl Microbiol 111: 787–799

Kent, B. N., Salichos, L., Gibbons, J. G., Rokas, A., Newton, I. L., et al. (2011) Complete bacteriophage transfer in a bacterial endosymbiont (Wolbachia) determined by targeted genome capture. Genome Biol Evol 3: 209–218

Klappenbach, J. A., Dunbar, J. M. and Schmidt, T. M. (2000) rRNA operon copy number reflects ecological strategies of bacteria. Appl Environ Microbiol 66: 1328–1333

Klitgord, N. and Segre, D. (2011) Ecosystems biology of microbial metabolism. Curr Opin Biotechnol 22: 541–546

Kong, H. H. (2011) Skin microbiome: genomics-based insights into the diversity and role of skin microbes. Trends Mol Med 17: 320–328

Kuczynski, J., Lauber, C. L., Walters, W. A., Parfrey, L. W., Clemente, J. C., et al. (2012) Experimental and analytical tools for studying the human microbiome. Nat Rev Genet 13: 47–58

Kuwahara, T., Ogura, Y., Oshima, K., Kurokawa, K., Ooka, T., et al. (2011) The lifestyle of the segmented filamentous bacterium: a non-culturable gut-associated immunostimulating microbe inferred by whole-genome sequencing. DNA Res 18: 291–303

Laing, C. R., Zhang, Y., Thomas, J. E. and Gannon, V. P. (2011) Everything at once: comparative analysis of the genomes of bacterial pathogens. Vet Microbiol 153: 13–26

Langridge, G. (2011) Plant power: converting a kingdom. Nat Rev Microbiol 9: 316

Lecroq, B., Lejzerowicz, F., Bachar, D., Christen, R., Esling, P., et al. (2011) Ultra-deep sequencing of foraminiferal microbarcodes unveils hidden richness of early monothalamous lineages in deep-sea sediments. Proc Natl Acad Sci U S A 108: 13177–13182

Leggat, W., Yellowlees, D. and Medina, M. (2011) Recent progress in Symbiodinium transcriptomics. Journal of Experimental Marine Biology and Ecology 408: 120–125

Li, F., Jiang, P., Zheng, H., Wang, S., Zhao, G., et al. (2011) Draft genome sequence of the marine bacterium Streptomyces griseoaurantiacus M045, which produces novel manumycin-type antibiotics with a pABA core component. J Bacteriol 193: 3417–3418

Li, K., Wang, S., Shi, Y., Qu, J., Zhai, Y., et al. (2011) Genome sequence of Paracoccus sp. Strain TRP, a Chlorpyrifos Biodegrader. J Bacteriol 193: 1786–1787

Li, L. and Delwart, E. (2011) From orphan virus to pathogen: the path to the clinical lab. Curr Opin Virol 1: 282–288

Li, S. C., Chan, W. C., Lai, C. H., Tsai, K. W., Hsu, C. N., et al. (2011) UMARS: Un-MAppable Reads Solution. BMC Bioinformatics 12 Suppl 1: S9

Li, S. G., Tang, Y. Q., Nie, Y., Cai, M. and Wu, X. L. (2011) Complete genome sequence of Polymorphum gilvum SL003B-26A1T, a crude oil-degrading bacterium from oil-polluted saline soil. J Bacteriol 193: 2894–2895

Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., et al. (2012) Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol

Luan, J. B., Li, J. M., Varela, N., Wang, Y. L., Li, F. F., et al. (2011) Global analysis of the transcriptional response of whitefly to tomato yellow leaf curl China virus reveals the relationship of coevolved adaptations. J Virol 85: 3330–3340

Luo, C., Tsementzi, D., Kyrpides, N., Read, T. and Konstantinidis, K. T. (2012) Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. PLoS ONE 7: e30087

Ma, M., Wang, C., Ding, Y., Li, L., Shen, D., et al. (2011) Complete genome sequence of Paenibacillus polymyxa SC2, a strain of plant growth-promoting Rhizobacterium with broad-spectrum antimicrobial activity. J Bacteriol 193: 311–312

Mackelprang, R., Waldrop, M. P., DeAngelis, K. M., David, M. M., Chavarria, K. L., et al. (2011) Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. Nature 480: 368–371

Macklaim, J. M., Gloor, G. B., Anukam, K. C., Cribby, S. and Reid, G. (2011) At the crossroads of vaginal health and disease, the genome sequence of Lactobacillus iners AB-1. Proc Natl Acad Sci U S A 108 Suppl 1: 4688–4695

Mader, U., Nicolas, P., Richard, H., Bessieres, P. and Aymerich, S. (2011) Comprehensive identification and quantification of microbial transcriptomes by genome-wide unbiased methods. Curr Opin Biotechnol 22: 32–41

Man, S. M., Kaakoush, N. O. and Mitchell, H. M. (2011) The role of bacteria and pattern-recognition receptors in Crohn's disease. Nat Rev Gastroenterol Hepatol 8: 152–168

Maron, P. A., Mougel, C. and Ranjard, L. (2011) Soil microbial diversity: Methodological strategy, spatial overview and functional interest. C R Biol 334: 403–411

Martin, F. (2011) Unearthing the truffle genome. New Phytol 189: 645–646

Mende, D. R., Waller, A. S., Sunagawa, S., Jarvelin, A. I., Chan, M. M., et al. (2012) Assessment of metagenomic assembly using simulated next generation sequencing data. PLoS ONE 7: e31386

Meyer, C., Grey, F., Kreklywich, C. N., Andoh, T. F., Tirabassi, R. S., et al. (2011) Cytomegalovirus microRNA expression is tissue specific and is associated with persistence. J Virol 85: 378–389

Minot, S., Grunberg, S., Wu, G. D., Lewis, J. D. and Bushman, F. D. (2012) Hypervariable loci in the human gut virome. Proc Natl Acad Sci U S A 109: 3962–3966

Mokili, J. L., Rohwer, F. and Dutilh, B. E. (2012) Metagenomics and future perspectives in virus discovery. Curr Opin Virol 2: 63–77

Moon, S., Kim, Y. G., Dong, L., Lombardi, M., Haeggstrom, E., et al. (2011) Drop-on-demand single cell isolation and total RNA analysis. PLoS ONE 6: e17455

Moore, J., Jironkin, A., Chandler, D., Burroughs, N., Evans, D. J., et al. (2011) Recombinants between Deformed wing virus and Varroa destructor virus-1 may prevail in Varroa destructor-infested honeybee colonies. J Gen Virol 92: 156–161

Moore, R. A., Warren, R. L., Freeman, J. D., Gustavsen, J. A., Chenard, C., et al. (2011) The sensitivity of massively parallel sequencing for detecting candidate infectious agents associated with human tissue. PLoS ONE 6: e19838

Murray, C. L., Oh, T. S. and Rice, C. M. (2011) Keeping Track of Viruses. Microbial Forensics (Second Edition) 137–153

Nasu, A., Marusawa, H., Ueda, Y., Nishijima, N., Takahashi, K., et al. (2011) Genetic heterogeneity of hepatitis C virus in association with antiviral therapy determined by ultra-deep sequencing. PLoS ONE 6: e24907

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., et al. (2011) Tumour evolution inferred by single-cell sequencing. Nature 472: 90–94

Oh, H. M., Lee, K., Jang, Y., Kang, I., Kim, H. J., et al. (2011) Genome sequence of strain IMCC9480, a xanthorhodopsin-bearing betaproteobacterium isolated from the Arctic Ocean. J Bacteriol 193: 3421

Ohnmacht, C., Marques, R., Presley, L., Sawa, S., Lochner, M., et al. (2011) Intestinal microbiota, evolution of the immune system and the bad reputation of pro-inflammatory immunity. Cell Microbiol 13: 653–659

Oliere, S., Douville, R., Sze, A., Belgnaoui, S. M. and Hiscott, J. (2011) Modulation of innate immune responses during human T-cell leukemia virus (HTLV-1) pathogenesis. Cytokine Growth Factor Rev 22: 197–210

Ott, S. J., Musfeldt, M., Wenderoth, D. F., Hampe, J., Brant, O., et al. (2004) Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. Gut 53: 685–693

Pace, N. R. (1997) A molecular view of microbial diversity and the biosphere. Science 276: 734–740

Perera, O. P., Snodgrass, G. L., Allen, K. C., Jackson, R. E., Becnel, J. J., et al. (2012) The complete genome sequence of a single-stranded RNA virus from the tarnished plant bug, Lygus lineolaris (Palisot de Beauvois). J Invertebr Pathol 109: 11–19

Petersen, J. M., Zielinski, F. U., Pape, T., Seifert, R., Moraru, C., et al. (2011) Hydrogen is an energy source for hydrothermal vent symbioses. Nature 476: 176–180

Pride, D. T., Sun, C. L., Salzman, J., Rao, N., Loomer, P., et al. (2011) Analysis of streptococcal CRISPRs from human saliva reveals substantial sequence diversity within and between subjects over time. Genome Res 21: 126–136

Proctor, L. M. (2011) The Human Microbiome Project in 2011 and beyond. Cell Host Microbe 10: 287–291

Qi, M., Wang, P., O'Toole, N., Barboza, P. S., Ungerfeld, E., et al. (2011) Snapshot of the eukaryotic gene expression in muskoxen rumen--a metatranscriptomic approach. PLoS ONE 6: e20521

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464: 59–65

Rajendhran, J. and Gunasekaran, P. (2011) Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. Microbiol Res 166: 99–110

Ram, J. L., Karim, A. S., Sendler, E. D. and Kato, I. (2011) Strategy for microbiome analysis using 16S rRNA gene sequence analysis on the Illumina sequencing platform. Syst Biol Reprod Med 57: 162–170

Reid, G., Younes, J. A., Van der Mei, H. C., Gloor, G. B., Knight, R., et al. (2011) Microbiota restoration: natural and supplemented recovery of human microbial communities. Nat Rev Microbiol 9: 27–38

Roh, H., Uguru, G. C., Ko, H. J., Kim, S., Kim, B. Y., et al. (2011) Genome sequence of the abyssomicin- and proximicin-producing marine actinomycete Verrucosispora maris AB-18-032. J Bacteriol 193: 3391–3392

Rosario, K. and Breitbart, M. (2011) Exploring the viral world through metagenomics. Curr Opin Virol 1: 289–297

Rosenthal, A. Z., Matson, E. G., Eldar, A. and Leadbetter, J. R. (2011) RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. ISME J 5: 1133–1142

Sabehi, G., Shaulov, L., Silver, D. H., Yanai, I., Harel, A., et al. (2012) A novel lineage of myoviruses infecting cyanobacteria is widespread in the oceans. Proc Natl Acad Sci U S A 109: 2037–2042

Santos, T. M. and Bicalho, R. C. (2011) Complete genome sequence of vB_EcoM_ECO1230-10: a coliphage with therapeutic potential for bovine metritis. Vet Microbiol 148: 267–275

Saulnier, D. M., Santos, F., Roos, S., Mistretta, T. A., Spinler, J. K., et al. (2011) Exploring metabolic pathway reconstruction and genome-wide expression profiling in Lactobacillus reuteri to define functional probiotic features. PLoS ONE 6: e18783

Schouls, L. M., Schot, C. S. and Jacobs, J. A. (2003) Horizontal transfer of segments of the 16S rRNA genes between species of the Streptococcus anginosus group. J Bacteriol 185: 7241–7246

Seed, K. D., Bodi, K. L., Kropinski, A. M., Ackermann, H. W., Calderwood, S. B., et al. (2011) Evidence of a dominant lineage of Vibrio cholerae-specific lytic bacteriophages shed by cholera patients over a 10-year period in Dhaka, Bangladesh. MBio 2: e00334–00310

Soergel, D. A., Dey, N., Knight, R. and Brenner, S. E. (2012) Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. ISME J

Stephenson, P. G., Moore, C. M., Terry, M. J., Zubkov, M. V. and Bibby, T. S. (2011) Improving photosynthesis for algal biofuels: toward a green revolution. Trends Biotechnol 29: 615–623

Szpara, M. L., Parsons, L. and Enquist, L. W. (2010) Sequence variability in clinical and laboratory isolates of herpes simplex virus 1 reveals new mutations. J Virol 84: 5303–5313

Tisserant, E., Da Silva, C., Kohler, A., Morin, E., Wincker, P., et al. (2011) Deep RNA sequencing improved the structural annotation of the Tuber melanosporum transcriptome. New Phytol 189: 883–891

Trevors, J. T. (2011) Viable but non-culturable (VBNC) bacteria: Gene expression in planktonic and biofilm cells. J Microbiol Methods 86: 266–273

Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., et al. (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 444: 1027–1031

Vaarala, O., Atkinson, M. A. and Neu, J. (2008) The "perfect storm" for type 1 diabetes: the complex interplay between intestinal microbiota, gut permeability, and mucosal immunity. Diabetes 57: 2555–2562

Veneault-Fourrey, C. and Martin, F. (2011) Mutualistic interactions on a knife-edge between saprotrophy and pathogenesis. Curr Opin Plant Biol 14: 444–450

Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. and Andino, R. (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439: 344–348

Virgin, H. W. and Todd, J. A. (2011) Metagenomics and personalized medicine. Cell 147: 44–56

Walker, A. (2011) Genome watch: Singled out. Nat Rev Microbiol 9: 485

Walter, J., Britton, R. A. and Roos, S. (2011) Host-microbial symbiosis in the vertebrate gastrointestinal tract and the Lactobacillus reuteri paradigm. Proc Natl Acad Sci U S A 108 Suppl 1: 4645–4652

Wang, T. Y., Chen, H. L., Lu, M. Y., Chen, Y. C., Sung, H. M., et al. (2011) Functional characterization of cellulases identified from the cow rumen fungus Neocallimastix patriciarum W5 by transcriptomic and secretomic analyses. Biotechnol Biofuels 4: 24

Wang, Y., Zhang, Z. and Ramanan, N. (1997) The actinomycete Thermobispora bispora contains two distinct types of transcriptionally active 16S rRNA genes. J Bacteriol 179: 3270–3276

Wang, Y., Chen, C., Ai, L., Zhou, F., Zhou, Z., et al. (2011) Complete genome sequence of the probiotic Lactobacillus plantarum ST-III. J Bacteriol 193: 313–314

Wang, Y., Chen, Y., Shen, Q. and Yin, X. (2011) Molecular cloning and identification of the laspartomycin biosynthetic gene cluster from Streptomyces viridochromogenes. Gene 483: 11–21

Warnecke, F. and Hess, M. (2009) A perspective: metatranscriptomics as a tool for the discovery of novel biocatalysts. J Biotechnol 142: 91–95

Wen, L., Ley, R. E., Volchkov, P. Y., Stranges, P. B., Avanesyan, L., et al. (2008) Innate immunity and intestinal microbiota in the development of Type 1 diabetes. Nature 455: 1109–1113

Werner, J. J., Zhou, D., Caporaso, J. G., Knight, R. and Angenent, L. T. (2011) Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. ISME J

Whitman, W. B., Coleman, D. C. and Wiebe, W. J. (1998) Prokaryotes: the unseen majority. Proc Natl Acad Sci U S A 95: 6578–6583

Williams, I. R. (2011) Deep impact: deciphering mucosal microbiomes using next-generation sequencing approaches. Mucosal Immunology 4: 586–587

Wylie, S. J. and Jones, M. G. (2011) The complete genome sequence of a Passion fruit woodiness virus isolate from Australia determined using deep sequencing, and its relationship to other potyviruses. Arch Virol 156: 479–482

Xia, L. C., Cram, J. A., Chen, T., Fuhrman, J. A. and Sun, F. (2011) Accurate genome relative abundance estimation based on shotgun metagenomic reads. PLoS ONE 6: e27992

Xing, M. N., Zhang, X. Z. and Huang, H. (2012) Application of metagenomic techniques in mining enzymes from microbial communities for biofuel synthesis. Biotechnol Adv

Xiong, X. H., Zhi, J. J., Yang, L., Wang, J. H., Zhao, Y., et al. (2011) Complete genome sequence of the bacterium Methylovorus sp. strain MP688, a high-level producer of pyrroloquinolone quinone. J Bacteriol 193: 1012–1013

Yang, X., Wang, Y., Guo, W., Xie, Y., Xie, Q., et al. (2011) Characterization of small interfering RNAs derived from the geminivirus/betasatellite complex using deep sequencing. PLoS ONE 6: e16928

Yang, Z., Bruno, D. P., Martens, C. A., Porcella, S. F. and Moss, B. (2011) Genome-wide analysis of the 5' and 3' ends of vaccinia virus early mRNAs delineates regulatory sequences of annotated and anomalous transcripts. J Virol 85: 5897–5909

Yongfeng, H., Fan, Y., Jie, D., Jian, Y., Ting, Z., et al. (2011) Direct pathogen detection from swab samples using a new high-throughput sequencing technology. Clin Microbiol Infect 17: 241–244

Yoon, H. S., Price, D. C., Stepanauskas, R., Rajah, V. D., Sieracki, M. E., et al. (2011) Single-cell genomics reveals organismal interactions in uncultivated marine protists. Science 332: 714–717

Youssef, N., Sheik, C. S., Krumholz, L. R., Najar, F. Z., Roe, B. A., et al. (2009) Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. Appl Environ Microbiol 75: 5227–5236

Zhan, Y., Yan, Y., Zhang, W., Chen, M., Lu, W., et al. (2012) Comparative analysis of the complete genome of an Acinetobacter calcoaceticus strain adapted to a phenol-polluted environment. Res Microbiol 163: 36–43

Zhang, T., Zhang, X. X. and Ye, L. (2011) Plasmid metagenome reveals high levels of antibiotic resistance genes and mobile genetic elements in activated sludge. PLoS ONE 6: e26041

Zhou, M., McAllister, T. A. and Guan, L. L. (2011) Molecular identification of rumen methanogens: Technologies, advances and prospects. Animal Feed Science and Technology 166–167: 76–86

Zioni, N., Soroker, V. and Chejanovsky, N. (2011) Replication of Varroa destructor virus 1 (VDV-1) and a Varroa destructor virus 1-deformed wing virus recombinant (VDV-1-DWV) in the head of the honey bee. Virology 417: 106–112

FOR RESEARCH USE ONLY

Pub. No. 773-2012-022 Current as of 06 July 2012

illumina®