Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Environmental Sustainability

Genetic diversity assessments in the century of genome science

Tetsukazu Yahara[1], Michael Donoghue[2], Rafael Zardoya[3], Daniel P Faith[4] and Joel Cracraft[5]

Genetic elements determine phenotypes of organisms by interacting with environments. Despite genetic diversity within and between species being the fundamental basis of biological diversity, its contribution has been long neglected in biodiversity studies. This situation is rapidly changing as quantification of genetic diversity, from intraspecific up to the ecosystem level, has become more accessible owing to the development of next-generation sequencing technologies (NGSTs). Whole-genome sequencing techniques provide two specific approaches for accessing genetic diversity at large scales: metagenomics (environmental genomics) and EST (Expressed Sequence Tag) comparisons. The former has been applied successfully in the profiling of different microbial biomes, and it is particularly interesting in understanding their ecosystem structure and function. The latter is particularly useful in the studies of adaptation and the assessment of functional traits. Unquestionably, advances in the genomic sciences combined with a new generation of ecological and evolutionary science will boost new approaches to global and local assessments of biodiversity changes, and more importantly, will surely reframe the questions we are asking in biodiversity science.

Addresses
[1] Faculty of Sciences, Kyushu University, Fukuoka 812-8581, Japan
[2] Department of Ecology and Evolutionary Biology, Yale University, 21 Sachem Street, P.O. Box 208105, New Haven, CT 06520-8105, USA
[3] Museo Nacional de Ciencias Naturales, Jose Gutierrez Abascal, 2, 28006 Madrid, Spain
[4] The Australian Museum, 6 College St., Sydney, NSW 2010, Australia
[5] Department of Ornithology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, United States

Corresponding author: Yahara, Tetsukazu (tet.yahara@gmail.com)

Current Opinion in Environmental Sustainability 2010, **2**:43–49

This review comes from a themed issue on Terrestrial systems
Edited by Anne Larigauderie and Harold A. Mooney

Received 19 January 2010, accepted 18 March 2010
Available online 13th April 2010

1877-3435/$ – see front matter

DOI 10.1016/j.cosust.2010.03.008

Introduction

Although genetic diversity has long been recognized as a component of biodiversity in general, most reference to a biodiversity crisis is within the context of the loss of species and ecosystems, not genetic diversity. To take one example, among the indicators used to evaluate the CBD 2010 biodiversity targets, genetic diversity was considered important only in terms of species such as crop-relatives. This view is changing as evolutionary thinking is becoming more integrated into biodiversity studies [1], and access to large genetic datasets is becoming straightforward. It is within this context that global initiatives such as GEO BON (Global Earth Observation Biodiversity Observation Network; http://www.earthobservations.org/geobon.shtml) have called for efforts to monitor and assess genetic diversity, particularly within selected species over time [2].

Genetic diversity has long played a role in biodiversity sciences, especially in the study of population-level phenomena such as adaptation, demographic history, and extinction risk. We will suggest here that, because its importance to sustainability science will certainly grow, the analysis of genetic diversity, from within-species scales to those of ecosystems, should receive much greater attention within the biodiversity community. This new importance arises in large part from emergence of genome science and its incorporation into organismal biology. The 21st century has been characterized not only as the century of the environment but also as the century of genomics. Following the breakthrough of deciphering the human genome, whole genomes of more than 40 animals [3], 11 plants [4], and more than 1000 microbes have now been sequenced. Many more eukaryotic genome sequences are currently being determined and many more are planned (http://www.ncbi.nlm.nih.gov/sites/entrez?db=genomeprj), and indeed recently, a call was made to sequence 10 000 vertebrate species within the next decade [5]. All of these will happen because of the development of massively parallel next-generation sequencing technologies (NGSTs) [6,7•] that enable the sequencing of hundreds of megabases per instrument-run, which is precipitously reducing the cost and time of whole-genome sequencing-so much so that in 2009 a human genome was sequenced for approximately $2000 US in consumable supplies, a 200-fold decrease from the cost of the first human genome.

Advances in organismal biology are accelerating rapidly because of these technological breakthroughs. In this

article, we will expand the view of genetic diversity outlined in the GEO BON Concept Document [2] and describe how traditional approaches to defining, measuring, and monitoring genetic diversity might change in the face of the new genomic science. We then consider the implications of new genomic approaches for developing future assessments of genetic diversity and suggest ways in which genomic studies might deepen our understanding of the relationship between biodiversity and ecosystem structure and function. But first we briefly describe two approaches that already are making contributions.

Measures of genetic diversity

Organisms are the carriers of genetic and developmental systems, thus population measures of genetic diversity reflect statistical sampling for some genetic attributes. Traditionally these have been allelic markers from which measures of genetic diversity such as within-population heterozygosity (H_s), total heterozygosity (H_t), and genetic differentiation between populations ($1 - H_s/H_t$, called F_{st} or G_{st} reflecting differences of exact definition) have been determined. But measuring diversity even among populations can be complex [8], and extending comparisons to measure of interspecific genetic diversity or measures of genetic diversity among species assemblages, such as habitats or ecosystems, becomes increasingly problematic. To make matters worse, DNA sequencing has revealed a panoply of genome variation beyond single nucleotide polymorphisms (SNPs), from genome rearrangements and duplications, to multiple kinds of transposable elements, all of which can affect developmental pathways and have downstream effects on phenotypes and their function. Recent findings of comparative evolutionary genomics also raise questions about the efficacy of interspecific measures of genetic diversity that go beyond simple comparisons of SNP variation in one or more shared genes. Recent whole-genome comparisons among great apes [9] and among species of *Drosophila* [10] demonstrate how difficult it will be to develop overall indices of genetic diversity even among closely related species. Although it is well known how to compare variation within specific genes or sets of genes across species, genomic comparisons reveal massive loss and gain of genes and gene families, and substantial numbers of segmental-duplication events [11] even among species as close as humans and chimpanzees, or *Drosophila melanogaster* and *D. simulans*.

What metric of 'genetic diversity' might therefore be used to measure genomic complexity among species, on the one hand, or across species assemblages, on the other? What relevance might those metrics have for assessing and monitoring species or ecosystem structure and function? These are not simple questions to answer, in part because we know so little about the comparative architecture of genomes among close-relatives and certainly within ecosystems. As comparative genomics progresses, more sophisticated metrics will be developed. Current methods are largely simple, and tree-based, and measure either a genetic 'distance' derived from some character difference, usually computed across aligned DNA sequences, or character optimizations by a parsimony criterion, such as the absence/presence/loss/rearrangements of genes among the tree of great apes [9].

Ecologists long ago developed hierarchical species-diversity measures: within-ecosystem level (α diversity), total-area level (γ diversity), and the difference between ecosystems in an area (β diversity = γ diversity/α diversity). Ideally, we would want something analogous to quantify genetic diversity at the ecosystem level. Species are traditionally counted as a measure of species diversity in a particular ecosystem, yet the assessment of genetic diversity in an ecosystem is more nuanced. Some species are closely related to each other while others have very distant relatives genetically, consequently the former may share many common genes whereas the latter may have many unique genes. As a consequence, species contribute unevenly to genetic diversity within ecological units based on their phylogenetic relatedness and thus are not random samples in a statistical sense [12,13]. One approach to describing diversity at the species level is to determine a phylogenetic tree using DNA sequences and reconstruct an evolutionary history of genetic and functional divergence based on that tree. Progress toward this is being made with microorganisms [14] but can be scaled to any living organism. The basic idea is to use phylogenetic diversity (PD) [15••,16] instead of species diversity. PD is the sum of branch lengths in a phylogenetic tree for each member of an ecosystem. Species within a community may be closely or remotely related and thus species diversity by itself is not an adequate measure of genetic diversity; PD quantifies divergence in DNA sequences among ecosystem members. PD is one metric for genetic diversity at the ecosystem level, but it applies to genomic similarities and differences among the lineages being compared, and therefore taxon sampling is an issue. It is typically based on a single gene (e.g. 16s RNA in bacteria) or small set of genes; if different genes are used, relative PD values could change. Lineage-specific genes are a further complication, which may relate to lineage-specific functions acquired through adaptation to a particular environment. Genomic comparisons will enable us to quantify how many genes are specific to a particular lineage (a set of related species); interpretations of those results again may be a function of taxon sampling (but see [15••,17]). Designing measures for diversity of those lineage-specific genes at the ecosystem level is a new challenge for biodiversity scientists. One approach might be to reconstruct gains and losses of genes on a phylogenetic tree and calculate differences of functional genes between ecosystems considering complementarity [17].

Next generation sequencing technologies (NGSTs) and genetic diversity

Two common genomic approaches — environmental genomics (metagenomics) and the use of Expressed Sequence Tags (ESTs) — are currently being applied to numerous questions having relevance to the analysis of biodiversity and ecosystem structure and function. Both approaches provide the foundation for developing tools to assess and monitor genetic diversity, from populations to ecosystems.

Whole-genome sequencing techniques provide us with an extraordinary opportunity to compare genetic diversity not only within species but also between closely related species, giving new insights into the genetic processes of differentiation. Now intensive in-depth studies are being conducted on wild relatives of *Drosophila*, *Arabidopsis*, and other model organisms. Although the number of species for which whole genome is known remains small, we will soon see this change [5]. In the meantime, a number of investigators studying non-model organisms are using metagenomic approaches and ESTs for monitoring and assessments [18].

Here, we briefly describe these and then follow with a small number of case studies that exemplify what the future holds for assessing and monitoring genetic diversity. As the utility of metagenomics in environmental research is already comprehensively reviewed [19,20••], below, we will focus more attention upon ESTs.

Metagenomics

The goal of metagenomic approaches is to assay all the different genomes in an environmental sample simultaneously [19,20••], hence this method mostly pertains to microorganisms (from bacteria and unicellular eukaryotes up to microscopic metazoans such as nematodes) in relatively complex ecosystems. One purpose might be to inventory the species diversity present in the sample, using a given gene sequence as a marker, or to explore the functional diversity of the coding regions of genomes sampled. Simplistically, in metagenomic analysis, DNA is recovered from the environment and sequenced; sequences are assembled, compared to known sequences, and annotated computationally; and these genes samples are analyzed in terms of phylogenetic patterns, community structure, or functional and metabolic assessments are undertaken. In principle, whole genomes are discoverable, but in reality that does not happen for a number of reasons, and instead sampling captures taxonomic diversity for a large set of genes, many of which can be compared across species, thereby giving a community-level description. This approach, however, has the limitation that individual organisms are never isolated, and thus a classic taxonomic description is not feasible.

Expressed Sequence Tags (ESTs)

Groups of partial sequences of cDNAs are called ESTs; cDNAs are double-stranded DNAs complementary to mRNAs that code enzymes or other functional proteins. We can now determine more than 500 000 ESTs within a week that are expected to provide singular or unified sequences of 5000–10 000 genes. By developing ESTs for a pair of species, therefore, we can compare 5000 or more genes simultaneously. In EST production mRNA is isolated, reverse-transcribed, and resulted cDNAs are then sequenced [18]. Because ESTs represent the coding and thus functional portions of a genome (but without the non-coding regulatory sequences), those provide a picture of the functional genome of a particular organism. By mapping locations of ESTs on chromosomes using a series of crossing experiments and genetic analyses, we can obtain a more detailed picture of a genome that can aid in embarking on whole-genome analysis. This latter process is, however, time-consuming and not always necessary for biodiversity monitoring in the wild. Simple determination of ESTs for a particular organism is a fair proxy for quantifying genetic diversity within and between species at the genome level. Moreover, ESTs have also been used frequently to construct phylogenetic trees across metazoans [21].

Over the next 5–10 years, use of NGSTs will drastically transform our knowledge of genetic diversity. How will we utilize these new technologies and information in global and local assessments of biodiversity changes? Importantly, how will these new approaches reframe some of the questions we will be asking? We attempt to explore these implications through selected case studies.

Case studies

Metagenomic analysis: ecosystem structure and function

The study of the linkages between intraspecific and interspecific genomic diversity and the state and rate of ecosystem processes is still in its infancy [22]. The premier question is how might genetic studies contribute to a deeper understanding of the relationship between the genetic component of biodiversity and ecosystem processes? New fundamental science has been developed in recent years on ecosystem processes in marine microbial systems using metagenomic techniques. New approaches such as whole-genome shotgun (WGS) sequencing of oceanic waters by numerous investigators are revealing countless new genes and gene families [20••,23]. These, in turn, are leading to new inferences about carbon and energy flows, the discovery of new metabolic pathways, and for the first time are permitting large-scale descriptions of the functional landscape of microorganisms within an ecosystem [20••]. Parallel advances are occurring as metagenomic techniques are also accessing the functional structure of soil and fresh-water ecosystems,

vertebrate gut microbiomes, coral–microbial associations, and even the symbiotic relationships between termites and their microbiota [24,25•].

ESTs and the analysis of adaptive variation

The species flocks of cichlid fishes in the East African Tanganyika, Malawi and Victoria lake system are a prime example of adaptive radiation, showing astonishing morphological, ecological, and behavioral diversity [26,27]. The whole lineage in the rivers and lakes of East Africa, containing more than 2000 species, has an evolutionary history of a few million years. The cichlids in Lake Victoria apparently survived desiccation in refugia [28]. This iconic pattern of evolution and diversity is being threatened under rapid environmental changes including eutrophication and species invasion [26]. Thus, genetic assessments of East African cichlids are critically important for developing adequate planning of conservation management.

International efforts to characterize the cichlid genome have been made over many years and will soon be rewarded with the completion of four genomes (http://cichlid.umd.edu/CGCindex.html). Among these efforts, Salzburger *et al.* [29••] recently determined approximately 12000 ESTs of a haplochromine cichlid *Astatotilapia burtonii* and 8636 ORFs (Open Reading Frame; a sequence frame corresponding to a gene) were identified. By comparing these with the whole genome database of a puffer fish *Takifugu rubripes*, they could determine the proportion of genes unique to *A. burtonii*. When a rather relaxed threshold of similarity search is used, 3460 ORFs (40%) had similar sequences in the puffer fish genome whereas 60% were not shared. By using such techniques, we can determine how many unique genes are expected to be lost if a particular lineage of wild organisms would become extinct, and those estimates become more precise as closely related species are sampled. This approach provides a potential measure for estimating gene diversity loss under the various biodiversity loss scenarios.

Large numbers of ESTs are also useful to identify genes that evolved faster or slower in a particular lineage and to demonstrate footprints of positive natural selection in particular genes. By comparing sequences of cichlids with those of three other fish species and human, Salzburger *et al.* [29••] identified 49 genes that have a significantly faster rate of evolution in the cichlids and 213 genes that have a significantly slower rate. They also showed that four genes among the more slowly evolving ones compared to the other fish species have replacement/silent substitution ratios that indicated a signature of adaptive evolution. Variability in the amino acid substitution sites of these genes will provide a framework for examining functional genetic variability within and among cichlid species.

The approach described above is applicable to most eukaryotes in the wild and its utility has dramatically increased owing to the development of new generation sequencers. Making only two runs of the 454 sequencing, Vera *et al.* [30••] recently determined 608 053 ESTs of a butterfly *Melitaea cinxia* that has been intensively studied by ecologists. From these ESTs, they found approximately 9300 genes. In addition to these genes, they identified approximately 6000 unknown sequences with a strong signal of gene expression that may include functional genes specific to *M. cinxia*.

The findings summarized above show that we can now compare very large numbers of functional genes between species in the wild within relative short periods of time. During the next decade, the EST comparison will be applied to many organisms and our knowledge of genetic diversity in wild organisms will be significantly advanced.

EST comparisons of intraspecific genetic diversity

Single species are often distributed across large environmental gradients and show adaptive differentiation of phenotypes. Genetic backgrounds of this adaptive differentiation have been of long-term interest in ecological genetics. Whole-genome sequencing in some model organisms has greatly improved our ability to detect genetic elements associated with adaptive differentiation of phenotypes [31–33]. In *Arabidopsis*, for example, fitness effects associated with the major gene for flowering time, *FRIGIDA*, are well documented by using 136 accessions collected across Europe [34]. Similar approaches can be applied to various organisms by determining a large number of ESTs as in *M. cinxia* and designing primers for detecting variation of candidate genes that are expected to determine a particular phenotype.

An advantage of this approach is its use of knowledge about functional genes previously documented in model organisms and its facilitation of identifying candidate genes for further analysis [33,35]. If one is interested in flower color variation, for example, ESTs from cDNA can be prepared from flower buds and then gene sequences associated with flower-color biosynthetic pathways can be determined. From there, one can look for amino acid replacements, design primers for that gene, and examine associations of allelic differences with flower color phenotypes.

A disadvantage of this approach is that we still have limited knowledge about the genetic background of most phenotypes. At present, it is not easy to nominate candidate genes for functional leaf traits, for example. In *Populus*, the whole genome has been sequenced and many ESTs are available, but researchers are still looking for genes that are associated with community phenotypes [36•]. However, progress of molecular biology on functional traits is rapid and we can be optimistic that many

candidate genes will be available soon. Another disadvantage of this approach is more fundamental. Many alleles of many functional loci are fixed within a species [37], therefore only a small fraction of genetic diversity can be examined if the analysis is confined to a species. Genetic diversity within a species is of course important because loss of such diversity would result in loss of evolutionary potential to adapt to environmental changes. On the other hand, most phenotypic diversity is found not within a species but between species. Thus, global genetic diversity assessments should pay more attention to genetic diversity between species.

ESTs and the assessment of functional traits

To predict the response of a local ecosystem to environmental changes such as global warming or human land use and habitat fragmentation, we need to understand how each functional type of organism responds to these changes. Hence, there is an increasing interest in functional trait diversity within and between ecosystems — terrestrial [38,39], freshwater, and marine [40•]. For plants, a global database of more than 1480 functional traits has been developed (http://www.try-db.org/index.php?n=Main.HomePage) and a new generation of dynamic global vegetation models is utilized.

NGSTs will revolutionize how we identify and assess functional traits. Currently, ESTs provide another source for functional traits called a transcript profile [41•]. If mRNA isolated from a particular organ under a specific environment is directly used for cDNA formation, the resultant frequency distribution of ESTs provides an estimate of the relative abundances of those transcripts. This profile is a highly plastic trait especially in plants. If a plant is stressed by drought or extreme temperature, different sets of mRNAs increase or decrease. Thus, comparison of ESTs between stressed and non-stressed samples of the same species is often used to identify stress-induced genes [35]. As far as we know, no attempts have been made to compare ESTs among remotely related species under the same controlled stress, and this approach should be useful for examining how different species within a community respond to a particular change in the environment at the functional gene level. For marine microbes, DeLong *et al.* [42] found that genomic profiles of microbial communities vary considerably along the sea depth gradient, reflecting adaptation to different light environments. Similar studies for eukaryotes along latitudinal and altitudinal gradients would provide significant insight into diversity for functional response of dominant species to environmental changes.

Conclusion

Recently, GEO BON [2] has called for efforts to monitor and assess genetic diversity in addition to diversity in the levels of species and ecosystem. This is a timely call because the development of technologies including NGSTs has enabled us to compare many genes simultaneously within and between species. Of course, comparisons using a few genes or a few neutral genetic markers among many samples are still useful to provide evidence for genetic and PD analyses. However we can now adopt a complementary approach to compare many genes between a limited number of samples, for instance, between a threatened species and its relative. This approach is also useful to examine the genetic bases of adaptive phenotypic change by detecting non-synonymous substitutions and by linking genetic diversity with ecosystem function.

It is clear that new technologies such as NGSTs will provide a wealth of new tools to assess and monitor genetic diversity at various spatial and temporal scales. As the efficiency, expense, and processing time of these approaches continues to decrease, it will become more feasible to apply these tools on a regular basis for monitoring the microbial biodiversity and ecosystem function in 'real time' across numerous ecosystems. Many of these technologies will be portable to nonmicrobial organisms and systems, and comparisons across ecosystems of their 'genetic architecture' will become increasingly possible.

Earth has been characterized as a 'living planet,' filled with biological diversity that itself has shaped the history of the physical planet itself. All organisms arose and have been shaped through numerous evolutionary processes. Darwin eloquently captured this view at the close of *The Origin of Species* by his metaphor of biological diversity as 'an entangled bank,' clothed with many plants and animals dependent on each other in a complex manner. Now we know that behind Darwin's entangled bank, is a plethora of genomic elements, which themselves are entangled in complex genomic and developmental networks, and that the manifestations of these networks underlie the networks of interacting species. Even though we do not know the precise directions biodiversity science will take, it is evident that NGSTs along with new associated bioinformatic tools will revolutionize the study of these entangled networks within individual organisms and thus the environments they comprise. But biological diversity is rapidly being lost. We are loosing not only ecosystem services but also, through the genomes of organisms, evolutionary or evosystem services [43•]. There is an imperative therefore to understand this component of diversity and the consequences of its loss. This presents a major challenge for the interdisciplinary science that links evolution, genomics, and ecology.

Acknowledgements

We acknowledge the support of the organizers of the second Open Science Conference of DIVERSITAS, held in Cape Town in 2009. The ideas described in this paper have been developed through discussion in the Open Science Conference and also with the scientific committee members of the bioGENESIS core project. We also thank Makiko Mimura, Shin-ichi Morinaga, Chris Wood and two anonymous reviewers for their thoughtful comments on the manuscript.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Donoghue MJ, Yahara T, Conti E, Cracraft J, Crandall KA, Faith DP, Hauser C, Hendry AP, Joly C, Kogure K, *et al.*: **bioGENESIS: providing an evolutionary framework for biodiversity science**. *DIVERSITAS Report* 2009, **6**:1–52. http://www.diversitas-international.org/uploads/File/biogenesis_sp_final.pdf.

2. GEO BON: *The GEO Biodiversity Observation Network Concept Document*. Group on Earth Observation; 2008 In: *http://www.earthobservations.org/documents/cop/bi_geobon/200811_geobon_concept_document.pdf*.

3. Ponting CP: **Functional repertoires of metazoan genomes**. *Nat Rev Genet* 2009, **9**:689-698.

4. Yang X, Kalluri UC, DiFazio SP, Wullschleger SD, Tschaplinski TJ, Cheng MG, Tuskan GA: **Popular genomics: state of the science**. *Crit Rev Plant Sci* 2009, **28**:285-308.

5. G10KCOS: **Genome 10K: a proposal to obtain whole genome sequence of 10000 vertebrate species**. *J Heredity* 2009, **100**:659-674.

6. Shendure J, Ji H: **Next-generation DNA sequencing**. *Nat Biotechnol* 2008, **26**:1135-1145.

7. • Metzker ML: **Sequencing technologies — the next generation**. *Nat Rev Genet* 2010, **11**:31-46.
This is an updated and most comprehensive review of the next-generation sequencing technologies. Performances of available technologies are compared. In addition, the author describes the utility of the third-generation sequencing (Real-time sequencing) that is expected to improve our sequencing efficiency dramatically again.

8. Holsinger KE, Weir BS: **Genetics of geographically structured populations: defining, estimating and interpreting F_{ST}**. *Nat Rev Genet* 2009, **10**:639-650.

9. Demuth JP, De Bie TD, Stejich JE, Cristianini N, Hahn MW: **The evolution of mammalian gene families**. *PLoS ONE* 2006, **1(1)**:e85 doi: 10.1371/journal.pone.0000085.

10. Drosophila 12 Genomes Consortium: **Evolution of genes and genomes on the Drosophila phylogeny**. *Nature* 2007, **450**:203-218.

11. Marques-Bonet T, Kidd JM, Ventura M, Graves TA, Cheng Z, Hillier LW, Jiang Z, Baker C, Malfavon-Borja R, Fulton LA *et al.*: **A burst of segmental duplications in the genome of the African Great Ape ancestor**. *Nature* 2009, **457**:877-881.

12. Felsenstein J: **Phylogenies and the comparative method**. *Am Nat* 1985, **125**:1-15.

13. Purvis A, Lambaut A: **Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data**. *Bioinformatics* 1995, **11**:247-251.

14. Lozupone CA, Knight R: **Global patterns in bacterial diversity**. *Proc Natl Sci Acad U S A* 2007, **104**:11436-11440.

15. •• Lozupone CA, Knight R: **Species divergence and the measurement of microbial diversity**. *FEMS Microbiol Rev* 2008, **32**:557-578.
This review provides a comprehensive summary of methods to quantify genetic diversity within and between microbial communities. The methods described are applicable also to communities of macro-organisms.

16. Faith DP: **Conservation evaluation and phylogenetic diversity**. *Biol Conserv* 1992, **61**:1-10.

17. Faith DP, Lozupone CA, Nipperess D, Knight R: **The cladistic basis for the phylogenetic diversity (PD) measures links evolutionary features to environmental gradients and supports broad applications of microbial ecology's "phylogenetic beta diversity" framework**. *Int J Mol Sci* 2009, **10**:4723-4741.

18. Bouk A, Vision T: **The molecular ecologist's guide to expressed sequence tags**. *Mol Ecol* 2007, **16**:907-924.

19. Riesenfeld CS, Schloss PD, Handelsman J: **Metagenomics: genomic analysis of microbial communities**. *Ann Rev Genet* 2004, **38**:525-555.

20. •• DeLong EF: **The microbial oceans from genomes to biomes**. *Nature* 2009, **459**:200-206.
The author reviews how the latest genetic approaches to marine microbes including the new generation sequencing can provide a fresh perspective on various biological interactions and biogeochemical processes within marine ecosystems.

21. Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, Smith SA, Seaver E, Rouse GW, Obst M, Edgecombe GD *et al.*: **Broad phylogenomic sampling improves resolution of the animal tree of life**. *Nature* 2008, **452**:745-749.

22. Hillebrand H, Matthiessen B: **Biodiversity in a complex world: consolidation and progress in functional biodiversity research**. *Ecol Lett* 2009, **12**:1405-1419.

23. DeLong EF, Karl DM: **Genomic perspectives in microbial oceanography**. *Nature* 2005, **437**:336-342.

24. Warnecke F, Luginbühl P, Ivanova I, Ghassemian M, Richardson TH, Stege JT, Cayouette M, McHardy AC, Djordjevic G, Aboushadi N *et al.*: **Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite**. *Nature* 2007, **450**:560-565.

25. • Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, Furlan M, Desnues C, Haynes M, Li L *et al.*: **Functional metagenomic profiling of nine biomes**. *Nature* 2008, **452**:629-632.
The authors conducted a metagenomic comparison of approximately 15 million sequences from 45 distinct microbiomes including terrestrial, freshwater and marine samples. This large-scale study shows us how large samples can be compared now, how large microbe diversity can be discovered, and how little is known about roles of microbe diversity upon ecosystem functions.

26. Seehausen O: **African cichlid fish: a model system in adaptive radiation research**. *Proc Roy Soc B* 2006, **273**:1987-1998.

27. Seehausen O, Terai Y, Magalhaes IS, Carleton KL, Mrosso HDJ, Miyagi R, van der Sluijs I, Schneider MV, Maan ME, Tachida H *et al.*: **Speciation through sensory drive in cichlid fish**. *Nature* 2008, **455**:620-626.

28. Elmer KR, Reggio C, Wirth T, Verheyen E, Salzburger W, Meyer A: **Pleistocene desiccation in East Africa bottlenecked but did not extirpate the adaptive radiation of Lake Victoria haplochromine cichlid fishes**. *Proc Natl Acad Sci U S A* 2009, **106**:13404-13409.

29. •• Salzburger W, Renn SCP, Steinke D, Braasch I, Hofmann HA, Meyer A: **Annotation of expressed sequence tags for the East African cichlid fish *Astatotilapia burtoni* and evolutionary analyses of cichlid ORFs**. *BMC Genomics* 2008, **9**:96.
This is an excellent case study demonstrating how ESTs can be used to characterize a genome of a non-model organism and to develop markers of adaptive genetic variation under natural selection.

30. •• Vera JC, Wheat CW, Fescemyer HW, Frilander, Crawford DL, Hanski I, Marden JH: **Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing**. *Mol Ecol* 2008, **17**:1636-1647.
This is the first application of the new-generation sequencing technology to develop ESTs for non-model organisms. By comparing with [29••], we can learn how dramatically NGST increased the efficiency of characterizing a genomic profile of a non-model organism.

31. Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh YP, Hahn MH, Nista PM, Jones CD, Kern AD, Dewey CN *et al.*: **Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans***. *PLoS Biol* 2007, **5(11)** e310.

32. Li YF, Costello JC, Holloway AK, Hahn MW: **Reverse ecology and the power of population genomics**. *Evolution* 2008, **62**:2984-2994.

33. Ellgren H, Sheldon BC: **Genetic bases of fitness differences in natural populations**. *Nature* 2008, **452**:169-175.

34. Korves TM, Schmid KJ, Caicedo AL, Mays C, Stinchcombe JR, Purugganan MD, Schmitt J: **Fitness effects associated with the major flowering time gene FRIGIDA in *Arabidopsis thaliana* in the field**. *Am Nat* 2007, **169**:E141-E157.

35. Reusch TBH, Wood TE: **Molecular ecology of global change**. *Mol Ecol* 2007, **16**:3973-3992.

36. • Whitham TG, DiFazio SP, Schweitzer JA, Shuster SM, Allan GJ, Bailey JK, Woolbright SA: **Extending genomics to natural communities and ecosystems**. *Science* 2008, **320**:492-495.
The authors have developed a methodology called 'community genetics' that links genetic diversity with ecosystem functions. In this paper, the authors extended their concept and provided a perspective to link genomics with ecosystem processes.

37. Lexer C, Rosenthal DM, Raymond O, Donovan LA, Rieseberg LH: **Genetics of species differences in the wild annual sunflowers *Helianthus annuus* and *H. petiolaris***. *Genetics* 2005, **169**:2225-2239.

38. Westby M, Wright IJ: **Land-plant ecology on the basis of functional traits**. *Trends Ecol Evolut* 2006, **21**:262-268.

39. Ackerly DD, Cornwell WK: **A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components**. *Ecol Lett* 2007, **10**:135-145.

40. • Litchman E, Klausmeier CA: **Trait-based community ecology of phytoplankton**. *Annu Rev Ecol Evolut Syst* 2008, **39**:615-639.
This is a timely review of trait-based approaches to community ecology of phytoplankton. Because most functional traits are genetically determined, trait-based approaches provide a practical way to assess functional genetic changes in the levels of community or ecosystem.

41. • Leakey ADB, Ainsworth EA, Bernard SM, Cody Markelz RJ, Ort DR, Placella SA, Rogers A, Smith MD, Sudderth EA, Weston DJ *et al.*: **Gene expression profiling: opening the black box of plant ecosystem responses to global change**. *Global Change Biol* 2009, **15**:1201-1213.
This review provides an introduction of functional genomics to ecologists who are interested in how gene expression profiles of an individual change with environmental changes. The authors summarize advantages and constraints of various tools and approaches to compare gene expression profiles between different environments.

42. DeLong EF, Preston CM, Mincer M, Rich V, Hallam SJ, Frigaard NU, Martinez A, Sullivan MB, Edwards R, Brito BR *et al.*: **Community genomics among stratified microbial assemblages in the ocean's interior**. *Science* 2006, **311**:496-503.

43. • Faith DP, Magallón S, Hendry AP, Conti E, Yahara T, Donoghue MJ: **Evosystem services: an evolutionary perspective on the links between biodiversity and human-well-being**. *Curr Opin Environ Sustain* 2010, **2**:.
As a complement to the widely accepted concept of ecosystem services, the authors coined a new term 'evosystem services' that highlights the benefits to humans provided by past and future evolutionary processes.