Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Microbiology

Single cell genomics: an individual look at microbes

Ramunas Stepanauskas

Single cell genomics (SCG) uncovers hereditary information at the most basic level of biological organization. It is emerging as a powerful complement to cultivation-based and microbial community-focused research approaches. SCG has been instrumental in identifying metabolic features, evolutionary histories and inter-organismal interactions of the uncultured microbial groups that dominate many environments and biogeochemical cycles. The SCG approach also holds great promise in microbial microevolution studies and industrial bioprospecting. Methods for SCG consist of a series of integrated processes, beginning with the collection and preservation of environmental samples, followed by physical separation, lysis and whole genome amplification of individual cells, and culminating in genomic sequencing and the inference of encoded biological features.

Address
Bigelow Laboratory for Ocean Sciences, 60 Bigelow Drive, P.O. Box 380, East Boothbay, ME 04544, United States

Corresponding author: Stepanauskas, Ramunas (rstepanauskas@bigelow.org)

Current Opinion in Microbiology 2012, **15**:613–620

This review comes from a themed issue on **Genomics**

Edited by **Siv Andersson** and **Andrew Goodman**

For a complete overview see the Issue and the Editorial

Available online 29th September 2012

http://dx.doi.org/10.1016/j.mib.2012.09.001

Introduction

The introduction of DNA sequencing in phylogenetic studies in the 1970s [1] and its application to uncultured microorganisms in the 1980s and 1990s [2,3] revolutionized microbiology and revealed that the diversity of unicellular life on Earth vastly exceeds what has been discovered using classical, cultivation-based techniques. During the 2000s, large metagenomics projects started providing extensive gene content information from natural microbial communities, spanning both cultured and uncultured taxonomic groups [4–6]. Nevertheless, significant technical and conceptual challenges remain unresolved. For example, when analyzing complex microbial communities, metagenomics is not well suited to deliver unequivocal information about the organization of discovered genes within genomes, obscuring the knowledge of the metabolic potential and evolutionary histories of specific microorganisms. To bridge this gap, single cell genomics (SCG) emerged as a powerful complement to cultivation and metagenomics, by providing genomic information from individual, uncultured cells. Several prior publications reviewed progress in this rapidly advancing field [7–14]. Here I focus on the most recent research findings and technology improvements.

Recent breakthroughs enabled by SCG

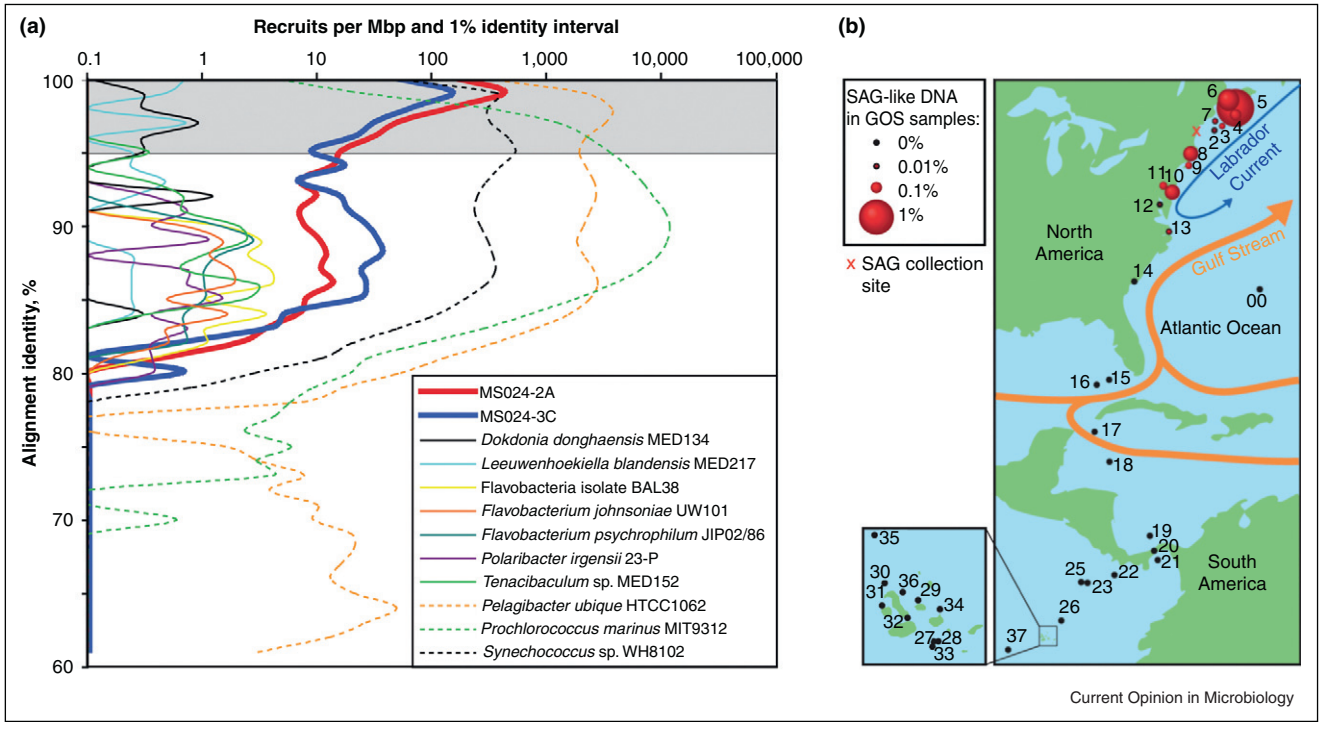
Matching phylogeny and function of the microbial 'uncultured majority'

By enabling sequencing of any genome region in an uncultured cell, SCG provides direct link information between cell's phylogenetic (e.g. SSU rRNA genes) and metabolic markers. A powerful example is the discovery by Swan *et al.* [15••] of chemolithoautotrophy pathways in uncultured Proteobacteria that constitute a major fraction of dark ocean's biomass, which may help reconcile current discrepancies in dark ocean's carbon budget. Martinez-Garcia *et al.* [16] discovered that members of the poorly understood phylum Verrucomicrobia probably are significant players in the degradation of polysaccharides. Fleming *et al.* [17] settled the 100 years-long debate about the phylogenetic position of the textbook iron oxidizer *Leptothrix ochracea*. Blainey *et al.* [18] provided further evidence for the importance of archaea in nitrogen cycling in diverse environments. Rhodopsin and bacteriochlorophyll genes were confirmed in many aquatic bacteria, indicating that photoheterotrophy is widespread among freshwater [16,19•] as well as marine [20–22] taxonomic groups. Yoon *et al.* [23••] showed that Picobiliphytes, a novel phylum of marine protists with no cultured representatives, are heterotrophic, not phototrophic as originally described [24]. These examples illustrate the power of SCG to effectively resolve the long-standing challenge of identifying the metabolic potential of the uncultured microbial groups that dominate many environments and biogeochemical cycles.

Binning of environmental omics data

SCG generates reference genomes of the uncultured microorganisms, facilitating the interpretation of community omics data sets. Woyke *et al.* [21] employed metagenomic fragment recruitment by single cell genomes to investigate biogeographic distribution of uncultured, marine Flavobacteria (Figure 1). Mason *et al.* [25••] combined SCG, metagenomics and metaproteomics to identify members of marine bacterioplankton that were involved in the degradation of hydrocarbons during the *Deepwater Horizon* oil spill. Swan *et al.* [15••] utilized a combination of SCG and other omics tools to verify the presence and expression of chemoautotrophy pathways in dark ocean bacteria. These and other examples [17,26,27]

Figure 1



The abundance and geographic distribution of surface ocean microbial DNA that is similar to genomes of two uncultured, marine Flavobacteria cells. These data demonstrate that, in contrast to all available marine Flavobacteria cultures, the two sequenced single cells represent genotypes that are numerically significant in their source environment. Single amplified genomes of Flavobacteria MS024-2A and MS024-3C from the Gulf of Maine were used to recruit metagenome fragments from the Global Ocean Sampling (GOS) expedition [6]. **(a)** Recruiting capacity of MS024-2A and MS024-3C, relative to all available cultured marine Flavobacteria, the non-marine *Flavobacterium johnsoniae*, and the three best GOS fragment recruiters *Pelagibacter*, *Prochlorococcus and Synechococcus*. **(b)** Geographic distribution of GOS metagenome fragments with >95% identity to MS024-2A. Numerals on the map indicate GOS station numbers.
Modified from [21].

illustrate how the integration of SCG with other research methods provides insights into microbial diversity, biogeography and processes that would not have been detected by any of the individual techniques alone.

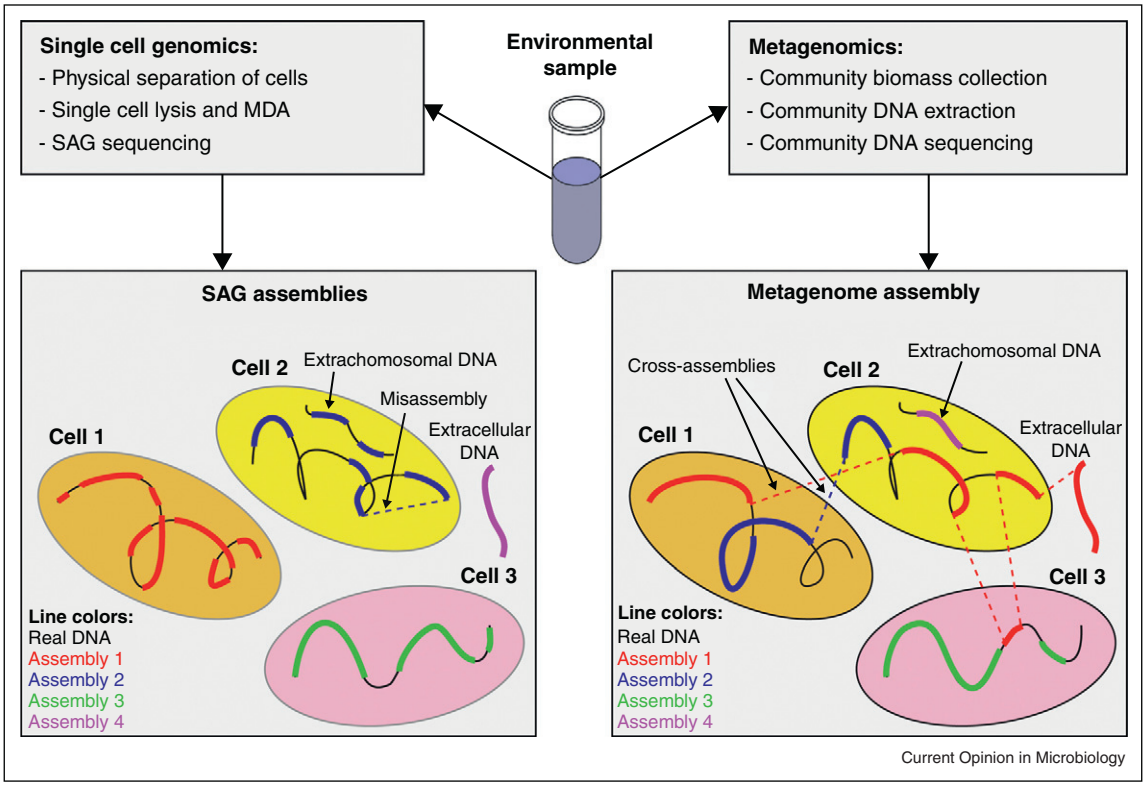
Revealing microbial interactions *in situ*

SCG provides access to sequences of all DNA in the analyzed cell, including chromosome(s), organelles, plasmids, food items, symbionts and pathogens (Figure 2). This enabled the recovery of a complete genome of a novel nano virus that probably infects uncultured Picobiliphytes [23••] and DNA of putative prey items and symbionts in diverse marine protists [23••,28]. Hongoh *et al.* [29] obtained complete genomes of intracellular symbionts inhabiting individual protist cells, which themselves are symbionts of termites. Woyke *et al.* [30•] and Pamp *et al.* [31] utilized SCG to study intracellular symbionts of sharpshooter insects and mouse gut cells, respectively. Tadmor *et al.* [32•] employed a high-density microfluidic device and multiplex PCR to identify phage-infected bacteria, while Martinez *et al.* [33] developed a fluorescent probe to specifically target virus-infected algal cells for SCG. These examples demonstrate a wide array of opportunities provided by SCG to microbial ecology, including studies of predation, infections and symbioses among microorganisms in their natural environment.

Contributions to microbial evolution studies

SCG has been instrumental in obtaining genomic sequences of microorganisms from several deep-branching phylogenetic groups with no cultured representatives, such as TM7 [34,35], OP11 [36], Picobilliphytes [23••] and divergent groups of aquatic Proteobacteria, Flavobacteria and Archaea [15••,18,21,37,38]. Bhattacharya *et al.* used SCG to study the origin of plastids [39]. Genomic contextual information obtained about gene families from deeply divergent, and presumably ancient lineages will aid our understanding of the early evolution of life. On the practical level, more phylogenetically balanced representation of genomes in databases, including lineages that have no cultured representatives, will improve models for computational gene annotation and taxonomic binning [40].

Figure 2



Strengths and limitations of SCG and metagenomic assemblies. Dashed lines indicate false assembly contiguity. Here cells 1 and 2 represent close relatives, while cell 3 is evolutionarily distant, but has horizontally acquired a DNA region from a relative of cells 1 and 2. The main strength of an SCG assembly is that all resulting contigs originate from DNA that was present in an individual cell, independent of the genetic complexity of the analyzed microbial community and the number of DNA molecules comprising each cell's genome. On the downside, SCG assemblies often are fragmented and incomplete, owing to the challenging nature of some cells (e.g. tough cell walls, DNA–protein interactions), whole genome amplification artifacts (e.g. uneven amplification, chimeras) and not fully optimized genome assembly algorithms. In some cases, metagenomic assemblies may result in more contiguous and complete genome recoveries than SCG assemblies, owing to a more even sequence coverage depth. However, metagenomic assemblies are consensuses from a multitude of cells and, potentially, detrital DNA that share high-homology regions but may vary in their genome-wide similarity, owing to mutations, horizontal gene transfer and recombination. Thus, metagenomic assemblies should be viewed only as approximations of the genome content of environmental microorganisms.

Arguably the most novel type of data that SCG provides is the quantitative information on genomic variability in natural microbial populations. Genome rearrangements, gene insertions, duplications and loss can be analyzed, even when multiple DNA molecules are present in a cell (e.g. chromosomes, plasmids, organelles) (Figure 2A). Such information is hard to obtain from metagenomics, where genome assemblies, when achieved [22,41,42], are consensuses from a multitude of cells that are assumed to be clonal. In reality, metagenomic assemblies may be mosaics of DNA from cells that share high-homology regions but vary in genome-wide similarity (Figure 2B), owing to mutations, horizontal gene transfer and recombination, which appear to be more frequent in nature than previously assumed [43–45]. In pilot studies, Martinez-Garcia *et al.* [19•] demonstrated the utility of SCG to detect horizontal gene transfer and recombination events in freshwater bacterioplankton, while Woyke *et al.* [30•] analyzed population structure of intracellular symbionts. Future, scaled-up SCG projects may provide the type of data that is required to develop and verify the currently contested [46–49] concepts of prokaryote diversity and diversification.

Biotechnology relevance

Biotechnology applications to date rely almost exclusively on the <1% of microbial diversity that has been amenable to cultivation, although metagenomics-based bioprospecting is gaining momentum [50–52]. By discovering entire genomes rather than individual genes of the uncultured microorganisms, SCG offers a powerful complement. Complex metabolic pathways can be reconstructed from the same cell, ensuring compatibility of the discovered genes with each other and facilitating the selection of suitable heterologous expression systems. Genomic information obtained through SCG may also

enable cultivation of microorganisms of interest. Early examples of SCG relevance in biotechnology include recoveries of polyketide biosynthesis pathways from sponge symbionts [53,54] and marine cyanobacteria [55], and genomes of uncultured microorganisms that degrade specific macromolecules [52,56••] and fix CO_2 through chemoautotrophy [15••]. Future integration of SCG-based biochemical pathway discovery with synthetic biology holds enormous potential for novel, environmentally responsible energy solutions, bioremediation of toxins, and natural products for nutritional, medicinal and other uses.

SCG methods

SCG consists of a series of integrated processes, starting with the collection and preservation of environmental samples, followed by physical separation, lysis and whole genome amplification of individual cells, then proceeding into either targeted loci or whole genome sequencing and sequence interpretation (Figure 3).

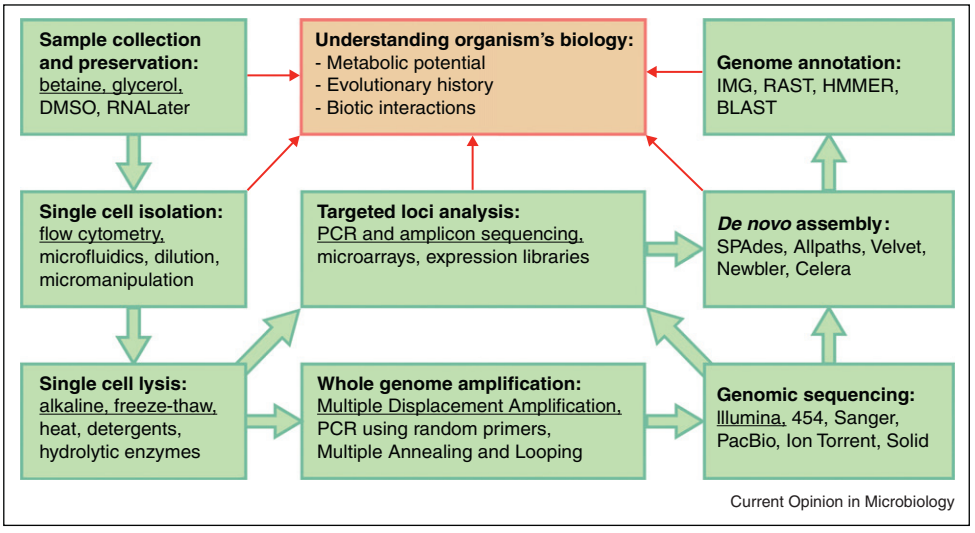
Unless analyzed immediately, environmental samples require *storage* that preserves the integrity of the cell and its DNA and does not interfere with downstream cell separation and DNA analyses. Deep freezing in the presence of glycine betaine or glycerol is the most common approach [15••,19•]. Aldehydes, which are commonly used in microbial sample preservation for microscopy, are not suitable for SCG, because they cross-link DNA and prevent its amplification.

The majority of SCG work today relies on fluorescence-activated cell sorting (FACS) for *cell separation* [15••,18,17,19•,20–22,23••,27,28,33,37–39,52–54,56••,57,58••,59•,60,61]. This well-established technology [62] offers automated, rapid delivery of individual cells into tubes or microwell plates and can be combined with a wide variety of fluorescent cell labeling techniques. Only a few picoliters of sample are sorted with each cell, minimizing the risk of contamination from extracellular DNA [63]. An alternative, micromanipulation technique has been employed in SCG as well [29,30•,55,64]. In difference to FACS, micromanipulation enables visual inspection of the analyzed cells, but it is tedious and more susceptible to contamination. Diverse microfluidic devices have also been successfully employed in SCG [18,31,32•,35,65–67,68•]. Although still lagging behind FACS in terms of cell separation versatility and throughput, further improvements and commercialization of microfluidics may reduce SCG costs and provide novel research opportunities, for example individual cell experimentation and genomics on a single lab-on-a-chip.

An ideal *cell lysis* protocol would be effective on diverse types of cells without damaging their DNA, leaving no DNA contamination and no chemicals that may inhibit downstream analyses. Most SCG work today relies on cell lysis by an alkaline solution, first described by Raghunathan *et al.* [61]. Alternative or supplementary treatments have included heat, freeze–thaw, detergents (unpublished data) and treatment with hydrolytic enzymes [15••,17,35]. Single cell lysis success rates vary widely and typically are below 40% [15••,35], so further method improvements are needed.

Whole genome amplification (WGA) is required before single cell DNA sequencing, except when direct, multiplex PCR [32•,67] is employed. Multiple displacement amplification (MDA) [69] is the most widely used technique, which produces long (average >10 kbp), overlapping

Figure 3



General workflow for microbial single cell genomics. The list of methods is not exhaustive, and the currently most popular methods are underlined.

amplicons that are suitable for whole genome sequencing and *de novo* assembly. Drawbacks of MDA include highly uneven genome coverage and chimera formation [21,61,66,70,71]. The former has been counteracted by laboratory and *in silico* DNA normalization [15••,60] and specialized *de novo* assembly software [37,72,73], while the latter is largely resolved by sufficient sequence coverage and avoidance of long mate-pair libraries [58••,60]. Polymerase chain reaction (PCR)-based techniques have also been used in single cell WGA [68•,74], producing short amplicons for genome re-sequencing. Single cell WGA products are called single amplified genomes (SAGs) [20] or plones [71] and can be further analyzed in ways that are similar to DNA extracts from pure cultures.

PCR is often used as a cost-effective option to screen large numbers of SAGs [15••,19•,20,27,28,33] or unamplified single cell DNA [32•,67] for *specific loci*. However, the immense and poorly understood genetic diversity of most microbial communities makes it difficult to design primers and probes that match all and only the target genes.

Multiple technologies are currently available for *genomic sequencing*, differing by read length, paired-end options, error types and rates, and the cost, as summarized by Loman *et al.* [75]. The choice should be guided by the intended use of the data. For example, most of the recent software development for single cell whole genome *de novo assembly* has been focused on the utilization of paired-end Illumina reads [15••,37,72,73], while Pacific Biosciences technology is increasingly used in assembly gap closure [76]. Key components of *quality assessment* include characterization of assembly fragmentation, identification of misassemblies and contaminating DNA, and estimates of the fraction of the genome that has been recovered [15••,18,21,22,30•,35,37,60]. Standard *annotation* pipelines can be employed, such as the Integrated Microbial Genomes System [77] and the Rapid Annotations using Subsystems Technology [78].

The success of genome recovery from single cells varies widely, from 0% to a finished genome, and depends on cell's intrinsic properties (cell wall structure, DNA packaging) and on all the components of the SCG pipeline (Figure 3). For example, environmental sample preservation has a significant impact on cell lysis efficiency, which may impact WGA evenness and genome assembly quality (unpublished data). Natural [30•] and artificially induced [57] polyploidy can improve single cell genome

Figure 4

Current Opinion in Microbiology

Bigelow Laboratory Single Cell Genomics Center, the first user facility providing SCG services to the broad scientific community (www.bigelow.org/scgc). Photo by Christopher Barnes.

recovery. Early SCG attempts on environmental microorganisms were hampered by significant DNA contamination [35,64,79], which may come from the sample itself, reagents, consumables and handling. More recently, contamination issue has been virtually resolved by the introduction of clean techniques in flow cytometry [15••,20,21,56••,60], the use of laser tweezers in microfluidics [18], and decontamination of MDA reagents before use [15••,56••,58••,65].

User facilities for SCG

High-quality SCG results currently require expensive instrumentation (e.g. cell sorters, robotic liquid handlers, DNA sequencers), specialized infrastructure (clean rooms, high-performance IT) and a concerted effort of scientific personnel with skills spanning microbiology, flow cytometry, microfluidics, robotic liquid handling, DNA sequencing and bioinformatics. This is not attainable by most individual research groups, nor would it be cost-effective. To address this challenge and to make SCG more accessible to the broad scientific community, my colleagues and I established the first core facility specializing in this field, the Bigelow Laboratory's Single Cell Genomics Center (SCGC; Figure 4; bigelow.org/scgc). Since its launch in 2009, the SCGC has already processed over 400 000 individual microbial cells from the ocean, soils, deep subsurface, organismal microbiomes and other types of environments for users at over 50 universities, research institutes and companies. The U.S. Department of Energy Joint Genome Institute (JGI; www.jgi.doe.gov) operates another major facility for SCG, with services provided to external users. Looking into the future, miniaturization and integration of the various SCG components will facilitate SCG implementation in individual research laboratories and in the field.

Conclusions and future prospects

The power of SCG stems from its ability to read genetic information at the most basic level of biological organization. This produces direct rather than inferred information about the content and organization of microbial genomes in the environment. SCG has already demonstrated its value by revealing metabolic features and *in situ* interactions among uncultured microorganisms, which was intractable before. In addition, SCG has great potential to bring more clarity to the contested discussion about the nature of prokaryote species and the process of diversification, by providing rich information on the structure and dynamics of natural microbial populations. SCG technology will probably continue undergoing rapid improvements. More reliable cell lysis techniques and better protocols for whole genome assembly are among the most vital. Examples of promising new directions include targeted SCG using single cell physiology probes [56••], the emerging ability to sequence individual viral particles [59•], the potential to integrate SCG with single cell transcriptomics [80] and metabolomics [81], and improved computational tools for data analysis and visualization.

Acknowledgements

Beth N. Orcutt, David Emerson, Brandon K. Swan and Michael E. Sieracki are acknowledged for valuable comments. Manuscript preparation was supported by NSF grants DEB-841933 and EF-826924.

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. Woese CR, Fox GE: **Phylogenetic structure of the prokaryotic domain: the primary kingdoms**. *Proc Natl Acad Sci USA* 1977, **74**:5088-5090.

2. Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR: **Rapid determination of 16S ribosomal-RNA sequences for phylogenetic analyses**. *Proc Natl Acad Sci USA* 1985, **82**:6955-6959.

3. Stein JL, Marsh TL, Wu KY, Shizuya H, Delong EF: **Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon**. *J Bacteriol* 1996, **178**:591-599.

4. Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu DY, Paulsen I, Nelson KE, Nelson W *et al.*: **Environmental genome shotgun sequencing of the Sargasso Sea**. *Science* 2004, **304**:66-74.

5. Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC *et al.*: **Comparative metagenomics of microbial communities**. *Science* 2005, **308**:554-557.

6. Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S, Wu D, Eisen JA, Hoffman JM, Remington K *et al.*: **The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific**. *PLoS Biol* 2007, **5**:e77.

7. Lasken RS: **Single-cell genomic sequencing using Multiple Displacement Amplification**. *Curr Opin Microbiol* 2007, **10**:510-516.

8. Binga EK, Lasken RS, Neufeld JD: **Something from (almost) nothing: the impact of multiple displacement amplification on microbial ecology**. *ISME J* 2008, **2**:233-241.

9. Ishoey T, Woyke T, Stepanauskas R, Novotny M, Lasken RS: **Genomic sequencing of single microbial cells from environmental samples**. *Curr Opin Microbiol* 2008, **11**:198-204.

10. Kalisky T, Quake SR: **Single-cell genomics**. *Nat Methods* 2011, **8**:311-314.

11. Kalisky T, Blainey P, Quake SR: **Genomic analysis at the single-cell level**. *Annu Rev Genet* 2011, **45**:431-445.

12. deJager V, Siezen RJ: **Single-cell genomics: unravelling the genomes of unculturable microorganisms**. *Microb Biotechnol* 2011, **4**:431-437.

13. Kamke J, Bayer K, Woyke T, Hentschel U: **Exploring symbioses by single cell genomics**. *Biol Bull* 2012, **223**:30-43.

14. Lasken RS: **Genomic sequencing of uncultured microorganisms from single cells**. *Nat Rev Microbiol* 2012, **10**:631-640.

15. •• Swan BK, Martinez-Garcia M, Preston CM, Sczyrba A, Woyke T, Lamy D, Reinthaler T, Poulton NJ, Masland EDP, Gomez ML *et al.*: **Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean**. *Science* 2011, **333**:1296-1300.

This publication describes an SCG-enabled discovery of chemolithoautotrophy potential in several uncultured Proteobacteria lineages that comprise a major fraction of biomass in the dark, oxygenated ocean, providing a new perspective on carbon cycling in the ocean's largest habitat.

16. Garcia SL, McMahon T, Martinez-Garcia M, Srivastava A, Sczyrba A, Stepanauskas R, Grossart HP, Woyke T, Warnecke F: **Metabolic potential of a single cell belonging to one of the most abundant lineages in freshwater bacterioplankton**. *ISME J* 2012. advance online publication.

17. Fleming EJ, Langdon AE, Martinez-Garcia M, Stepanauskas R, Poulton NJ, Masland EDP, Emerson D: **What's new is old: resolving the identity of *Leptothrix ochracea* using single cell genomics, pyrosequencing and FISH**. *PLoS ONE* 2011, **6**:e17769.

18. Blainey PC, Mosier AC, Potanina A, Francis CA, Quake SR: **Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis**. *PLoS ONE* 2011, **6**:e16626.

19. • Martinez-Garcia M, Swan BK, Poulton NJ, Lluesma Gomez M, Masland D, Sieracki ME, Stepanauskas R: **High throughput single cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton**. *ISME J* 2012, **6**:113-123.
This SCG-based study demonstrates that photoheterotrophy pathways are encoded by many of the cosmopolitan taxonomic groups of freshwater bacterioplankton. Evidence is also provided for frequent inter-phyla horizontal gene transfer and recombination.

20. Stepanauskas R, Sieracki ME: **Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time**. *Proc Natl Acad Sci USA* 2007, **104**:9052-9057.

21. Woyke T, Xie G, Copeland A, Gonzalez JM, Han C, Kiss H, Saw J, Senin P, Yang C, Chatterji S *et al.*: **Assembling the marine metagenome, one cell at a time**. *PLoS ONE* 2009, **4**:e5299.

22. Dupont CL, Rusch DB, Yooseph S, Lombardo MJ, Alexander Richter R, Valas R, Novotny M, Yee-Greenbaum J, Selengut JD, Haft DH *et al.*: **Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage**. *ISME J* 2012, **6**:1186-1199.

23. •• Yoon HS, Price DC, Stepanauskas R, Rajah VD, Sieracki ME, Wilson WH, Yang EC, Duffy S, Bhattacharya D: **Single-cell genomics reveals organismal interactions in uncultivated marine protists**. *Science* 2011, **332**:714-717.
This is the first publication describing genomic sequencing of individual eukaryote cells. It shows that Picobiliphytes, a novel phylum of marine protists with no cultured representatives, are heterotrophic, not phototrophic as originally described. In addition to Picobiliphyte genomic DNA, the recovered sequences also included a complete genome of a novel nanovirus that probably infects Picobiliphytes, and DNA of putative Picobiliphyte prey items.

24. Not F, Valentin K, Romari K, Lovejoy C, Massana R, Tobe K, Vaulot D, Medlin LK: **Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes**. *Science* 2007, **315**:253-255.

25. •• Mason OU, Hazen TC, Borglin S, Chain PSG, Dubinsky EA, Fortney JL, Han J, Holman HYN, Hultman J, Lamendella R *et al.*: **Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill**. *ISME J* 2012, **6**:1715-1727.
The authors used a combination of single cell genomics, metagenomics and metatranscriptomics to discover a rapid response by specific Gammaproteobacteria to aliphatic hydrocarbons after the *Deepwater Horizon* oil spill in the Gulf of Mexico.

26. Eloe EA, Fadrosh DW, Novotny M, Zeigler Allen L, Kim M, Lombardo MJ, Yee-Greenbaum J, Yooseph S, Allen EE, Lasken R *et al.*: **Going deeper: metagenome of a hadopelagic microbial community**. *PLoS ONE* 2011, **6**:e20388.

27. Heywood JL, Sieracki ME, Bellows W, Poulton NJ, Stepanauskas R: **Capturing diversity of marine heterotrophic protists: one cell at a time**. *ISME J* 2011, **5**:674-684.

28. Martinez-Garcia M, Brazel D, Poulton NJ, Swan BK, Gomez ML, Masland D, Sieracki ME, Stepanauskas R: **Unveiling *in situ* interactions between marine protists and bacteria through single cell sequencing**. *ISME J* 2012, **6**:703-707.

29. Hongoh Y, Sharma VK, Prakash T, Noda S, Taylor TD, Kudo T, Sakaki Y, Toyoda A, Hattori M, Ohkuma M: **Complete genome of the uncultured Termite Group 1 bacteria in a single host protist cell**. *Proc Natl Acad Sci USA* 2008, **105**:5555-5560.

30. • Woyke T, Tighe D, Mavromatis K, Clum A, Copeland A, Schackwitz W, Lapidus A, Wu D, McCutcheon JP, McDonald BR *et al.*: **One bacterial cell, one complete genome**. *PLoS ONE* 2010, **5**:e10314.
This is the first demonstration of a complete genome recovery from an individual microbial cell, albeit with the aid of the analyzed cell's unusually high polyploidy and small genome size.

31. Pamp SJ, Harrington ED, Quake SR, Relman DA, Blainey PC: **Single-cell sequencing provides clues about the host interactions of segmented filamentous bacteria (SFB)**. *Genome Res* 2012, **22**:1107-1119.

32. • Tadmor AD, Ottesen EA, Leadbetter JR, Phillips R: **Probing individual environmental bacteria for viruses by using microfluidic digital PCR**. *Science* 2011, **333**:58-62.
The authors employed a high-throughput microfluidics platform to sequence marker genes of phages and their host bacteria from individual, uncultured bacteria cells residing in the termite hindgut. This study demonstrated phage host specificity in the analyzed system.

33. Martinez JM, Poulton NJ, Stepanauskas R, Sieracki ME, Wilson WH: **Targeted sorting of single virus-infected cells of the coccolithophore *Emiliania huxleyi***. *PLoS ONE* 2011, **6**:e22520.

34. Podar M, Abulencia CB, Walcher M, Hutchison D, Zengler K, Garcia JA, Holland T, Cotton D, Hauser L, Keller M: **Targeted access to the genomes of low-abundance organisms in complex microbial communities**. *Appl Environ Microbiol* 2007, **73**:3205-3214.

35. Marcy Y, Ouverney C, Bik EM, Losekann T, Ivanova N, Martin HG, Szeto E, Platt D, Hugenholtz P, Relman DA *et al.*: **Dissecting biological ''dark matter'' with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth**. *Proc Natl Acad Sci USA* 2007, **104**:11889-11894.

36. Youssef NH, Blainey PC, Quake SR, Elshahed MS: **Partial genome assembly for a candidate division OP11 single cell from an anoxic spring (Zodletone spring Oklahoma)**. *Appl Environ Microbiol* 2011, **77**:7804-7814.

37. Chitsaz H, Yee-Greenbaum JL, Tesler G, Lombardo MJ, Dupont CL, Badger JH, Novotny M, Rusch DB, Fraser LJ, Gormley NA *et al.*: **Efficient *de novo* assembly of single-cell bacterial genomes from short-read data sets**. *Nat Biotechnol* 2011, **10**:915-921.

38. Ghai R, Pasic L, Fernandez AB, Martin-Cuadrado AB, Mizuno CM, McMahon KD, Papke RT, Stepanauskas R, Rodriguez-Brito B, Rohwer F *et al.*: **New abundant microbial groups in aquatic hypersaline environments**. *Sci Rep* 2011, **1**:srep00135.

39. Bhattacharya D, Price DC, Yoon HS, Yang EC, Poulton NJ, Andersen RA, Das SP: **Single cell genome analysis supports a link between phagotrophy and primary plastid endosymbiosis**. *Sci Rep* 2012, **2**:e356.

40. Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, Kunin V, Goodwin L, Wu M, Tindall BJ *et al.*: **A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea**. *Nature* 2009, **462**:1056-1060.

41. Narasingarao P, Podell S, Ugalde JA, Brochier-Armanet C, Emerson JB, Brocks JJ, Heidelberg KB, Banfield JF, Allen EE: ***De novo* metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities**. *ISME J* 2012, **6**:81-93.

42. Iverson V, Morris RM, Frazar CD, Berthiaume CT, Morales RL, Armburst EV: **Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota**. *Science* 2012, **335**:587-590.

43. Fraser C, Hanage WP, Spratt BG: **Recombination and the nature of bacterial speciation**. *Science* 2007, **315**:476-480.

44. Vergin KL, Tripp HJ, Wilhelm LJ, Denver DR, Rappe MS, Giovannoni SJ: **High intraspecific recombination rate in a native population of *Candidatus* Pelagibacter ubique (SAR11)**. *Environ Microbiol* 2007, **9**:2430-2440.

45. Shapiro BJ, Friedman J, Cordero OX, Preheim SP, Timberlake SC, Szabo G, Polz MF, Alm EJ: **Population genomics of early events in the ecological differentiation of bacteria**. *Science* 2012, **335**:48-51.

46. Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, Stackebrandt E, Van de Peer Y, Vandamme P, Thompson FL *et al.*: **Re-evaluating prokaryotic species**. *Nat Rev Microbiol* 2005, **3**:733-739.

47. Woese CR, Goldenfeld N: **How the microbial world saved evolution from the scylla of molecular biology and the charybdis of the modern synthesis**. *Microbiol Mol Biol Rev* 2009, **73**:14-21.

48. Doolittle RF, Zhaxybayeva O: **On the origin of prokaryotic species**. *Genome Res* 2009, **19**:744-756.

49. Lawrence JG, Retchless AC: **The myth of bacterial species and speciation**. *Biol Philos* 2010, **25**:569-588.

50. Schloss PD, Handelsman J: **Biotechnological prospects from metagenomics**. *Curr Opin Biotechnol* 2003, **14**:303-310.

51. Lorenz P, Eck J: **Metagenomics and industrial applications**. *Nat Rev Microbiol* 2005, **3**:510-516.

52. Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T *et al.*: **Metagenomic discovery of biomass-degrading genes and genomes from cow rumen**. *Science* 2011, **331**:463-467.

53. Siegl A, Hentschel U: **PKS and NRPS gene clusters from microbial symbiont cells of marine sponges by whole genome amplification**. *Environ Microbiol Rep* 2010, **2**:507-513.

54. Siegl A, Kamke J, Hochmuth T, Piel J, Richter M, Liang C, Dandekar T, Hentschel U: **Single-cell genomics reveals the lifestyle of Poribacteria, a candidate phylum symbiotically associated with marine sponges**. *ISME J* 2011, **5**:61-70.

55. Grindberg RV, Ishoey T, Brinza D, Esquenazi E, Coates RC, Liu W-T, Gerwick L, Dorrestein PC, Pevzner P, Lasken R *et al.*: **Single cell genome amplification accelerates identification of the apratoxin biosynthetic pathway from a complex microbial assemblage**. *PLoS ONE* 2011, **6**:E18565.

56. •• Martinez-Garcia M, Brazel DM, Swan BK, Arnosti C, Chain PSG, Reitenga KG, Xie G, Poulton NJ, Lluesma Gomez M, Masland DED *et al.*: **Capturing single cell genomes of active polysaccharide degraders: an unexpected contribution of Verrucomicrobia**. *PLoS ONE* 2012, **7**:e35314.

In this study, a novel combination of high-throughput single cell physiology and genomics techniques was employed. The authors discovered that the poorly understood bacteria phylum Verrucomicrobia may be important in polysaccharide degradation in marine and freshwater environments.

57. Dichosa AEK, Fitzsimons MS, Lo CC, Weston LL, Preteska LG, Snook JP, Zhang X, Gu W, McMurry K, Green LD *et al.*: **Artificial polyploidy improves bacterial single cell genome recovery**. *PLoS ONE* 2012, **7**:e37387.

58. •• Woyke T, Sczyrba A, Lee J, Rinke C, Tighe D, Clingenpeel S, Malmstrom R, Stepanauskas R, Cheng J-F: **Decontamination of MDA reagents for single cell whole genome amplification**. *PLoS ONE* 2011, **10**:e26161.

This paper describes a simple and effective method to remove DNA contaminants from multiple displacement amplification reagents, making them suitable for single cell genomics.

59. • Allen LZ, Ishey T, Novotny MA, McLean JS, Lasken RS, Williamson SJ: **Single virus genomics: a new tool for virus discovery**. *PLoS ONE* 2011, **6**:e17722.

This is the first demonstration of DNA sequencing from individual viral particles, obtained from bacterial laboratory cultures.

60. Rodrigue S, Malmstrom RR, Berlin AM, Birren BW, Henn MR, Chisholm SW: **Whole genome amplification and *de novo* assembly of single bacterial cells**. *PLoS ONE* 2009, **4**:e6864.

61. Raghunathan A, Ferguson HR, Bornarth CJ, Song WM, Driscoll M, Lasken RS: **Genomic DNA amplification from a single bacterium**. *Appl Environ Microbiol* 2005, **71**:3342-3347.

62. Shapiro HM: *Practical Flow Cytometry*. Wiley-Liss; 2003.

63. Sieracki M, Poulton N, Crosbie N: **Automated isolation techniques for microalgae**. In *Algal Culturing Techniques.* Edited by Andersen R. Elsevier Academic; 2005:101-116.

64. Kvist T, Ahring BK, Lasken RS, Westermann P: **Specific single-cell isolation and genomic amplification of uncultured microorganisms**. *Appl Microbiol Biotechnol* 2007, **74**:926-935.

65. Blainey PC, Quake SR: **Digital MDA for enumeration of total nucleic acid contamination**. *Nucleic Acids Res* 2011, **39**:e19.

66. Marcy Y, Ishoey T, Lasken RS, Stockwell TB, Walenz BP, Halpern AL, Beeson KY, Goldberg SMD, Quake SR: **Nanoliter reactors improve multiple displacement amplification of genomes from single cells**. *PLoS Genet* 2007, **3**:1702-1708.

67. Ottesen EA, Hong JW, Quake SR, Leadbetter JR: **Microfluidic digital PCR enables multigene analysis of individual environmental bacteria**. *Science* 2006, **314**:1464-1467.

68. • Leung K, Zahn H, Leaver T, Konwar KM, Hanson NW, Page AP, Lo CC, Chain PS, Hallam SJ, Hansen CL: **A programmable droplet-based microfluidic device applied to multiparameter analysis of single microbes and microbial communities**. *Proc Natl Acad Sci USA* 2012, **109**:7665-7670.

This publication reports a significant advance in the development of microfluidics technology for single cell genomics and other microbiology applications. The described device combines microvalve and droplet technologies and introduces automated control for enhanced throughput.

69. Dean FB, Hosono S, Fang LH, Wu XH, Faruqi AF, Bray-Ward P, Sun ZY, Zong QL, Du YF, Du J *et al.*: **Comprehensive human genome amplification using multiple displacement amplification**. *Proc Natl Acad Sci USA* 2002, **99**:5261-5266.

70. Lasken RS, Stockwell TB: **Mechanism of chimera formation during the Multiple Displacement Amplification reaction**. *BMC Biotechnol* 2007, **7**:19.

71. Zhang K, Martiny AC, Reppas NB, Barry KW, Malek J, Chisholm SW, Church GM: **Sequencing genomes from single cells by polymerase cloning**. *Nat Biotechnol* 2006, **24**:680-686.

72. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD *et al.*: **SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing**. *J Comput Biol* 2012, **19**:455-477.

73. Peng Y, Leung HCM, Yiu SM, Chin FYL: **IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth**. *Bioinformatics* 2012, **28**:1420-1428.

74. Zhang L, Cui X, Schmitt K, Hubert R, Navidi W, Arnheim N: **Whole genome amplification from a single cell: implications for genetic analysis**. *Proc Natl Acad Sci USA* 1992, **89**:5847-5851.

75. Loman NJ, Constantinidou C, Chan JZ, Halachev M, Sergeant M, Penn CW, Robinson ER, Pallen MJ: **High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity**. *Nat Rev Microbiol* 2012, **10**:599-606.

76. Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, Wang Z, Rasko DA, McCombie WR, Jarvis ED *et al.*: **Hybrid error correction and *de novo* assembly of single-molecule sequencing reads**. *Nat Biotechnol* 2012, **30**:693-700.

77. Markowitz VM, Chen IA, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Anderson I, Lykidis A, Mavromatis K *et al.*: **The integrated microbial genomes system: an expanding comparative analysis resource**. *Nucleic Acids Res* 2010, **38**:D382-D390.

78. Aziz RK, Bartels D, Best AA, Dejongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M *et al.*: **The RAST server: rapid annotations using subsystems technology**. *BMC Genomics* 2008, **8**:9.

79. Lasken RS, Raghunathan A, Kvist T, Ishoey T, Westermann P, Ahring BK, Boissy R: **Multiple displacement amplification from single bacterial cells**. In *Whole Genome Amplification: Methods Express Series*. Edited by Hughes S, Lasken RS. Scion Publishing Ltd.; 2005:119-147.

80. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A *et al.*: **mRNA-Seq whole-transcriptome analysis of a single cell**. *Nat Methods* 2009, **6**:377-382.

81. Heinemann M, Zenobi R: **Single cell metabolomics**. *Curr Opin Biotechnol* 2011, **22**:26-31.