Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Biotechnology

The art and design of functional metagenomic screens

Marcus Taupp[1], Keith Mewis[2] and Steven J Hallam[1,2,3]

This article summarizes general design principles for functional metagenomics. The focus is on *Escherichia coli* as an expression host, although alternative host–vector systems are discussed in relation to optimizing gene recovery in activity-based screens. Examples of DNA isolation and enrichment approaches, library construction and phenotypic read-out are described with special emphasis on the use of high throughput technologies for rapid isolation of environmental clones encoding phenotypic traits of interest.

Addresses
[1] Department of Microbiology & Immunology, University of British Columbia, Canada
[2] Genome Science and Technology Graduate Program, University of British Columbia, Canada
[3] Graduate Program in Bioinformatics, University of British Columbia, Canada

Corresponding author: Hallam, Steven J (shallam@interchange.ubc.ca)

Current Opinion in Biotechnology 2011, **22**:1–8

This review comes from a themed issue on
Environmental biotechnology
Edited by Lindsay Eltis and Ariel Kushmaro

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2011.02.010

Introduction

Environmental genomic also known as metagenomic sequencing shows great promise in unlocking the hidden metabolic powers of uncultivated microbial communities [1]. However, accurate functional assignment is a major challenge to metabolic pathway reconstruction and gene discovery. In the absence of expression data, gene models remain hypotheses and in the absence of functional evidence, expressed proteins remain cryptic effectors of metabolism. In model systems, genetic analysis provides a successful paradigm for forward gene discovery based on phenotypic properties of organisms [2]. When coupled with clone library production, metagenomic approaches provide informational resources and genetic substrates for identification and heterologous expression of genes encoding desired phenotypic traits (reviewed in [3]. For example, a metagenomic survey of the Tammar wallaby foregut identified more than 800 putative glycoside hydrolases (GH), but no GH6 family enzymes were identified. A functional metagenomic screen of a fosmid library derived from the same source material recovered a putative β-1,4-endoglucanase belonging to the GH6 family [4]. Similarly, metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite identified more than 200 gene modules correlating to CAZy gene families [5], and recovered functional activity for 33 redundant or non-overlapping glycosyl hydrolases [6••]. Thus, functional metagenomics represents an effective adjunct to cultivation and *in silico* surveys of community metabolism founded on genetic principles and screening technologies developed for model organismal systems.

Looking in the right place

One of the keys to designing successful functional metagenomic screens lies in the availability and interpretation of environmental parameter data. Knowledge and intuition about sampling environments provides a phenotypic hierarchy useful in selecting appropriate screening targets and substrates. For example, a metagenomic library constructed from genomic DNA derived from bovine rumen was screened for hydrolytic activity [7]. Owing to the natural diet of grazing ruminants, the digestion requires hydrolytic enzymes provided by the rumen microbiota. The function-based screen revealed entirely new ester hydrolases and glycoside hydrolases as well as a novel cyclodextrinase. In another case, the knowledge of ethanol production as a common by-product in anaerobic digestors was used in a screening approach to discover genes encoding ethanol catabolising enzymes [8]. Similarly, a screen for xenobiotic degradation is more likely to succeed when clone libraries are constructed from environmental samples containing the intended screening substrate [9–11].

Careful consideration of the environment not only increases screening efficiency but opens the door to discovering novel biocatalysts with unique structural or kinetic properties. Cold-adapted and heat-adapted enzymes were isolated from habitats where those conditions naturally occur. For example, a novel low-temperature-active lipase was isolated from a metagenomic library constructed from soft-bottom sediment samples from the western Baltic Sea [12] and a thermostable esterase was isolated from a metagenomic library constructed from thermal environments in Indonesia [13]. Esterases adapted to high hydrostatic pressure and salinity have also been isolated from a metagenomic library constructed from environmental DNA derived from the brine interface of a deep-sea hypersaline anoxic basin [14].

Maximizing genomic potential

Low sample biomass compounded by the presence of inhibitory substances can impair metagenomic library

Please cite this article in press as: Taupp M, et al. The art and design of functional metagenomic screens, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.02.010

construction. Enrichment cultures offer one strategy to selectively increase biomass associated with a desired phenotypic trait with concomitant reduction of inhibitory substances (Figure 1) [15]. A similar strategy using stable isotope probing (SIP) can enrich for biomass on the basis of selective substrate utilization (stable isotope labelled substrates: ^{13}C, ^{15}N, ^{18}O) (Figure 1) (reviewed in [16]). Genomic DNA obtained from a labelling experiment is subjected to density centrifugation before metagenomic library production. This additional step separates labelled from unlabelled DNA resulting in a library enriched for microbial genomes with the potential to utilize the labelled substrate or metabolic by-products of the original substrate (for visualized experiment see [17•]).

In addition to enrichment culture approaches, isolated environmental DNA can be subjected to whole genome amplification, that is multiple displacement amplification (MDA) to provide sufficient genetic substrate for library production (Figure 1) [18]. When applied to single cells, MDA methods are capable of providing significant coverage of individual microbial genomes [19••,20]. An alternative to WGA called linker-amplified shotgun library (LASL) [21•] also improves coverage from low biomass samples. A modification of this approach, called expressed linker-amplified shotgun libraries with topoisomerase cloning (E-LASLs) has been used to screen a metagenomic library constructed from genomic DNA derived from earthworm gut for lysin activity [22]. Recently, a combination of SIP and MDA approaches enabled metagenomic library construction and screening for methanol dehydrogenase genes from marine surface water samples highlighting the modular effectiveness of enrichment, labelling and amplification methods [23].

Selecting an expression system

Clone libraries used in functional metagenomic screens can be categorized on the basis of expression systems. Expression systems are comprised of (a) vectors with specific attributes that constrain host range, insert size, copy number, and transformation efficiency (Figure 1 and Table 1) and (b) hosts organisms that carry and express the genetic information encoded by the vector. Metagenomic libraries differ from traditional clone libraries prepared from clonal organisms in that effective genome coverage is limited by the diversity of organisms found within a given environmental sample. In most metagenomic libraries, individual clones are derived from independent donor genotypes where each donor genotype is genetically unique. While this limits coverage of individual genotypes, it increases diversity of genes recovered in the library.

Successful library production depends on isolation of high molecular weight DNA of sufficient quality for downstream enzymatic processing steps. Small insert plasmid libraries in the range of ~15 kilobase (kb) pairs are carried in high copy vectors that are transformed directly into competent *E. coli* host strains. Small insert library production requires between 3–5 μg of starting genomic DNA before size selection using either gel purification or hydrosheer methods. Transformation efficiencies of small insert libraries are high, enabling the production of libraries containing $>10^5$ clones. Although small insert libraries contain hundreds of thousands of clones with potential for constitutive or controlled gene expression, they are generally restricted to single locus activity screens.

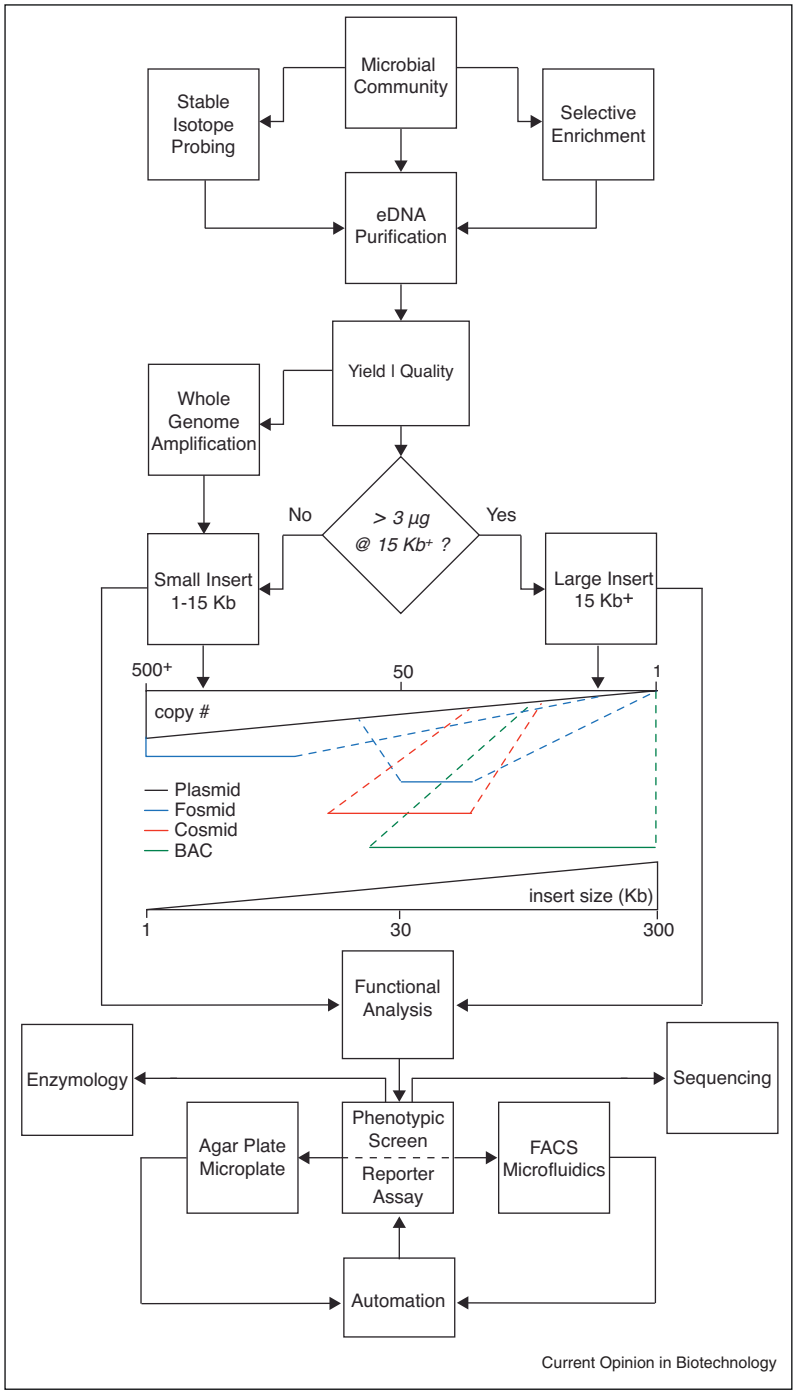
Large insert libraries including cosmids, fosmids and bacterial artificial chromosomes (BACS) utilize low copy vectors or inducible copy control vectors. For example, fosmid libraries with insert sizes ranging between 32 and 45 kb are packaged in phage lambda extracts and adsorbed directly to host cells (for visualized experiment see [24•]). Fosmid library production requires between 3–5 μg of starting genomic DNA before size selection using pulsed-field gel electrophoresis methods. Although transformation efficiency of individual fosmid clones is low, the use of phage transduction systems enables efficient library production. While large insert libraries typically contain far fewer clones than small insert libraries, their larger insert size provides extensive linkage information. Such linkage information facilitates identification of multilocus traits with increased taxonomic resolution [1,25••].

Currently, *E. coli* is the dominant screening host for functional metagenomics and most commercially available large insert library production systems utilize *E. coli* as a replication host. Indeed, numerous *E. coli* strains have been engineered that allow for stable replication of various single-copy or multi-copy vectors, minimal recombination and protection from lytic phage (reviewed in [26]). The low copy number status of fosmids or BACs within *E. coli* promotes stable replication and minimizes overexpression of toxic genes. However, the same attributes that make large insert clones useful for genomic sequencing can become liabilities when searching for function within a metagenomic library. Copy control systems have been developed allowing for on-demand induction of fosmid or BAC inserts [27]. This approach was successfully used in functional screens for proteorhodopsin photosystems and phosphonate utilization pathways from marine metagenomic libraries [28••,29]. Martinez *et al.* observed a strong positive correlation with visible screening efficiency and increased fosmid copy number [28••]. Similar observations were made in a screen for low-temperature lipase activity [12].

The most commonly invoked limitation in using *E. coli* expression systems in functional metagenomic screens is a limitation common to all expression systems. Indeed, although *E. coli* can support the expression of genes from numerous donor genomes, an unknown quantity cannot

Please cite this article in press as: Taupp M, et al. The art and design of functional metagenomic screens, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.02.010

Figure 1



A workflow detailing processes involved in production and functional screening of metagenomic libraries. Starting with the microbial community, optional enrichment steps such as selective enrichment or stable isotope probing may be used before library construction. The yield and quality of DNA determines the insert size and copy number of the library. Functional analysis using a phenotypic screen or assay transpires, often using automation and high throughput methodology such as microfluidics or FACS. Following functional identification, samples can be sequenced or characterized following standard biochemical methods.

Please cite this article in press as: Taupp M, et al. The art and design of functional metagenomic screens, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.02.010

Table 1

Common expression hosts used in functional metagenomic screens circa 2008 to present

Host bacterium	Phylogenetic affiliation	Gram stain	Mode of DNA delivery	Phenotypic trait in screen	Additional remarks	Reference
Agrobacterium tumefaciens	α-proteobacteria	negative	transformation	pigmentation, morphology		[35••]
Caulobacter vibrioides	α-proteobacteria	negative	transformation			[35••]
Rhizobium leguminosarum	α-proteobacteria	negative	conjugation	alcohol/aldehyde dehydrogenase		[8]
Burkholderia graminis	β-proteobacteria	negative	transformation	antibiosis		[35••]
Ralstonia metallidurans	β-proteobacteria	negative	electroporation	pigmentation, antibiosis		[35••]
Escherichia coli	γ-proteobacteria	negative	transformation, transduction	see references		reviewed in [3,53]
Pseudomonas fluorescens	γ-proteobacteria	negative	conjugation	not available	proteomic analysis of fosmid-encoded proteins by 2-DE	[54•]
Pseudomonas putida	γ-proteobacteria	negative	conjugation	pigmentation		[35••]
Xanthomonas campestris	γ-proteobacteria	negative	conjugation			[54•]

be expressed owing to differential transcriptional, translational or posttranslational controls including, but not limited to, promoter recognition and initiation factors, codon usage, ribosomal entry and protein folding (reviewed in [3]). Current estimates based on *in silico* detection of compatible expression signals suggest that *E. coli* can support expression of approximately 40% of genes within a subset of 32 taxonomically diverse genomes with wide-ranging variation in expression potential between genomes (7–73%) [30••]. A more recent study using microarrays and reverse-transcription polymerase chain reaction (RT-PCR) to assess incidental transcription in *E. coli* observed widespread transcription of heterologous DNA driven by native promoter elements with concomitant changes in the transcriptional profile of the *E. coli* host [31]. For effective functional screening, widespread expression potential is considered essential for efficient recovery of active clones. However, gene toxicity has the potential to limit screening efficiency. Indeed, the toxic side effects of cloned DNA in *E. coli* have been a limiting factor in whole genome sequencing projects resulting in gapped assemblies [32]. Adopting alternative technologies can circumvent gaps imposed by gene toxicity. However, for functional metagenomic screens, genetic modifications to *E. coli* or the use of alternative expression systems are required to overcome gene toxicity.

Just as using multiple sequencing technologies promotes successful genome assembly, using multiple expression systems has potential to increase recovery of active clones in functional metagenomic screens. For example, a codon usage calculation for *E. coli* K12 revealed a difference of 25.6% relative to *Streptomyces coelicolor* and related actinobacterial species suggesting limited or less efficient gene expression. Utilization of an alternative host such as *Rhodococcus jostii* RHA1 with more compatible initiation factors and a codon usage difference of 6.4% has the potential to increase expression of actinobacterial clones within a metagenomic library [33]. Indeed, phenotypic traits of interest and the taxonomic composition of a metagenome should inform host selection. For example, Martinez *et al.* developed multiple screening hosts and compatible shuttle vectors to screen for antibiotic production from BACs derived from soil DNA [34]. More recently, Craig *et al.* utilized broad host range vectors in a parallel screen for bioactive small molecules in six different proteobacterial hosts [35••]. The value-added capacity to shuttle clone libraries between alternative hosts with different expression capabilities should not be underestimated and highlights the need for continued development of broad host range vectors and library retrofitting methods.

Developing effective screens

Screens relying on genetic selection and survival of the host are the most powerful owing to the digital nature of the output, growth or no growth. Phenotypic traits amenable to this form of selection include antibiotic or heavy metal resistance or the complementation of genes mutated or missing in the host (for more information see [36•]). Classical screens without selection have relied on visible changes in colony color, morphology or clearing (halo) using agar plating methods (for more information see [36•]). Such visible screens are labour-intensive but require no special equipment beyond careful observation. However, signal-to-noise associated with colony detection is low, and weakly expressing positive clones are often missed. Moreover, visible changes are difficult to quantify, even compared against plates assayed in the same manner. Screens conducted in 96-well or 384-well plate format offer possibilities to apply high throughput methodology, including colony picking robots, liquid handlers and microplate readers. The application of automation for readout not only decreases labour costs, but increases reproducibility and comparability between samples.

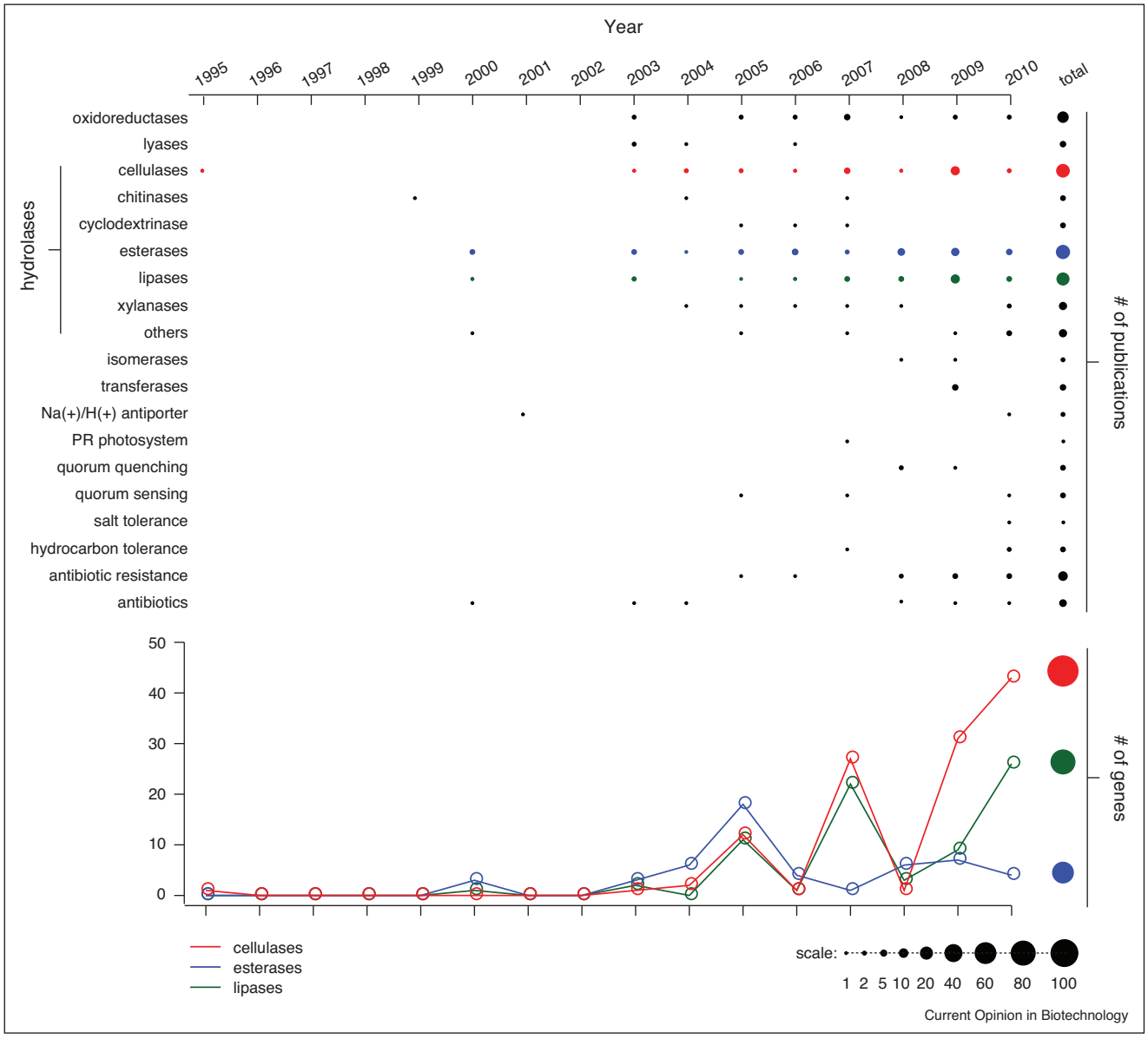
Please cite this article in press as: Taupp M, et al. The art and design of functional metagenomic screens, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.02.010

To overcome potential problems associated with intracellular accumulation of enzyme activity, cell lysis can be applied in microtiter plates to increase assay sensitivity [11,37]. Common cell lysis methods use enzymes or detergents (for visualized experiment see [38•]). While SDS is commonly used for denaturing protein studies, non-ionic detergents such as Triton X or Tween are ideal for functional screening. Both detergents effectively disrupt lipid-lipid and lipid–protein interactions, allowing for enzyme escape from cells, while keeping expressed proteins in native or active conformations. A novel alternative to enzymatic or chemical treatment in 96-well or 384-well plate formats uses a lytic gene cassette from bacteriophage Lambda under the control of a UV-inducible promoter [39]. The expression of cloned gene products is accompanied by host cell lysis eliminating potential cellular toxicity and enabling more efficient substrate utilization in the absence of additional liquid handling steps.

Once and future screens

Functional metagenomic screens for hydrolytic activity currently dominate the literature and this trend is pro-

Figure 2



A dot plot showing the number of publications reporting functional identification of enzyme activity in metagenomic libraries in each year from 1995 to 2010 (top). Below, line graph showing the number of reported genes for the three most commonly identified enzyme classes over the same time period. The number of genes and publications of hydrolases has outpaced the identification of other enzymes in recent years, partly owing to increased demand for more efficient biomass conversion of renewable bioenergy sources.

Please cite this article in press as: Taupp M, et al. The art and design of functional metagenomic screens, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.02.010

jected to increase with the renewed interest in the development of improved biomass conversion processes (Figure 2). Indeed, robust and scalable assays for cellulase activity were developed more than three decades ago [40]. These methods can be applied and extended in both solid-phase and liquid-phase screens on agar plates or in 96-well or 384-well formats [41] (for visualized protocol see [38•]). The development of novel substrates and detection methods along with increased screening throughput remain the primary activation barriers to more efficient and variable screens. The recent application of promoter trap vectors to functional metagenomic screens show great promise for high throughput discovery. Uchiyama *et al.* used a promoterless green fluorescent protein (GFP) expression vector for use in shotgun cloning of environmental DNA. Substrate-induced gene expression of the resulting clones enabled the recovery of GFP positive clones grown in liquid culture using fluorescence-activated cell sorting (FACS) [42•,43,44]. Similarly, the use of colorimetric or fluorometric substrates for the detection of enzyme activity allows for optical detection of positive clones, amenable to high throughput methods including automation. Optical detection approaches can be combined effectively with flow cytometry and cell sorting to further increase screening throughout [45].

Recently, microfluidic approaches have been demonstrated for a number of techniques including protein detection [46•,47], microRNA expression profiling [48], sequencing [49] and cell culture [50•] (reviewed in [51]). Droplet-based microfluidic approaches provide an ideal reaction vessel for functional metagenomic screening, analogous to wells in 96-well or 384-well plates in nanoliter volumes. Assay sensitivity is increased because fewer molecules are required to bring the concentration of product above the threshold of detection. Correspondingly, fewer cells are required to produce a sufficient quantity of enzyme, allowing the growth phase of a screen to be significantly shortened (reviewed in [52]). Future screens based on microfluidic detection will not only reduce experimental cycles but provide enormous cost savings on fluorogenic substrates used to detect many enzyme activities with high sensitivity.

Future prospects

A combination of *E. coli* and broad host range expression vectors provides a powerful parallel screening platform for the development of efficient activity-based screens. High throughput functional screens will become increasing important as sensitive and effective validation and discovery tools. Indeed, by building on established high throughput technologies such as FACS and microfluidic systems, functional metagenomic screens will experience orders of magnitude increases in screening throughput. All that is required for this method is a suitable and creatively deployed substrate that provides a detectable readout upon specific gene activation or product presentation.

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Riesenfeld CS, Schloss PD, Handelsman J: **Metagenomics: genomic analysis of microbial communities**. *Annu Rev Genet* 2004, **38**:525-552.

2. Shuman HA, Silhavy TJ: **The art and design of genetic screens: Escherichia coli**. *Nat Rev Genet* 2003, **4**:419-431.

3. Uchiyama T, Miyazaki K: **Functional metagenomics for enzyme discovery: challenges to efficient screening**. *Curr Opin Biotechnol* 2009, **20**:616-622.

4. Pope PB, Denman SE, Jones M, Tringe SG, Barry K, Malfatti SA, McHardy AC, Cheng JF, Hugenholtz P, McSweeney CS *et al.*: **Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores**. *Proc Natl Acad Sci USA* 2010, **107**:14793-14798.

5. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B: **The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics**. *Nucleic Acids Res* 2009, **37**:D233-238.

6. •• Warnecke F, Luginbuhl P, Ivanova N, 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.
Metagenomic shotgun sequencing of an environment naturally enriched for glycosyl hydrolases yielded more than 200 gene modules correlating to CAZy families, but only showed functional activity in 33 cases, demonstrating a gap between *in silico* and functional screening results.

7. Ferrer M, Beloqui A, Golyshina OV, Plou FJ, Neef A, Chernikova TN, Fernandez-Arrojo L, Ghazi I, Ballesteros A, Elborough K *et al.*: **Biochemical and structural features of a novel cyclodextrinase from cow rumen metagenome**. *Biotechnol J* 2007, **2**:207-213.

8. Wexler M, Bond PL, Richardson DJ, Johnston AW: **A wide host-range metagenomic library from a waste water treatment plant yields a novel alcohol/aldehyde dehydrogenase**. *Environ Microbiol* 2005, **7**:1917-1926.

9. Brennerova MV, Josefiova J, Brenner V, Pieper DH, Junca H: **Metagenomics reveals diversity and abundance of meta-cleavage pathways in microbial communities from soil highly contaminated with jet fuel under air-sparging bioremediation**. *Environ Microbiol* 2009, **11**:2216-2227.

10. Ono A, Miyazaki R, Sota M, Ohtsubo Y, Nagata Y, Tsuda M: **Isolation and characterization of naphthalene-catabolic genes and plasmids from oil-contaminated soil by using two cultivation-independent approaches**. *Appl Microbiol Biotechnol* 2007, **74**:501-510.

11. Suenaga H, Ohnuki T, Miyazaki K: **Functional screening of a metagenomic library for genes involved in microbial degradation of aromatic compounds**. *Environ Microbiol* 2007, **9**:2289-2297.

12. Hardeman F, Sjoling S: **Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment**. *FEMS Microbiol Ecol* 2007, **59**:524-534.

13. Rhee JK, Ahn DG, Kim YG, Oh JW: **New thermophilic and thermostable esterase with sequence similarity to the hormone-sensitive lipase family, cloned from a metagenomic library**. *Appl Environ Microbiol* 2005, **71**:817-825.

14. Ferrer M, Golyshina OV, Chernikova TN, Khachane AN, dos Santos VAPM, Yakimov MM, Timmis KN, Golyshin PN: **Microbial enzymes mined from the Urania deep-sea hypersaline anoxic basin**. *Chem Biol* 2005, **12**:895-904.

Please cite this article in press as: Taupp M, et al. The art and design of functional metagenomic screens, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.02.010

15. Mo X-c, Chen C-l, Pang H, Feng Y, Feng J-x: **Identification and characterization of a novel xylanase derived from a rice straw degrading enrichment culture**. *Appl Environ Microbiol* 2010.

16. Neufeld JD, Wagner M, Murrell JC: **Who eats what, where and when? Isotope-labelling experiments are coming of age**. *ISME J* 2007, **1**:103-110.

17. • Dunford EA, Neufeld JD: **DNA stable-isotope probing (DNA-SIP)**. *J Vis Exp* 2010.
A useful visual experiment demonstrating stable-isotope probing methods with emphasis on controlled experiments.

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

19. •• 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.
The amplification of DNA from a single microbial cell is discussed, allowing for the genomic sequencing of organisms resistant to both culture and culture-independent assay methods.

20. Lasken RS: **Genomic DNA, amplification by the multiple displacement amplification (MDA) method**. *Biochem Soc Trans* 2009, **037**:450-453.

21. • Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall AM, Mead D, Azam F, Rohwer F: **Genomic analysis of uncultured marine viral communities**. *Proc Natl Acad Sci USA* 2002, **99**:14250-14255.
The use of LASLs allows the metagenomic analysis of a viral community in *E. coli* despite the presence of potential toxic genes and non-native nucelotide modifications.

22. Schmitz JE, Daniel A, Collin M, Schuch R, Fischetti VA: **Rapid DNA library construction for functional genomic and metagenomic screening**. *Appl Environ Microbiol* 2008, **74**:1649-1652.

23. Neufeld JD, Chen Y, Dumont MG, Murrell JC: **Marine methylotrophs revealed by stable-isotope probing, multiple displacement amplification and metagenomics**. *Environ Microbiol* 2008, **10**:1526-1535.

24. • Taupp M, Lee S, Hawley A, Yang J, Hallam SJ: **Large insert environmental genomic library production**. *J Vis Exp* 2009.
A useful visual experiment detailing the production of fosmid libraries from environmental genomic DNA.

25. •• Yung PY, Burke C, Lewis M, Egan S, Kjelleberg S, Thomas T: **Phylogenetic screening of a bacterial, metagenomic library using homing endonuclease restriction and marker insertion**. *Nucl Acids Res* 2009, **37**: e144-.
The authors present a powerful selection-based method for recovering large insert clones haboring phylogenetic anchors from a metagenomic library.

26. Sorensen HP, Mortensen KK: **Advanced genetic strategies for recombinant protein expression in Escherichia coli**. *J Biotechnol* 2005, **115**:113-128.

27. Wild J, Hradecna Z, Szybalski W: **Conditionally amplifiable BACs: switching from single-copy to high-copy vectors and genomic clones**. *Genome Res* 2002, **12**:1434-1444.

28. •• Martinez A, Bradley AS, Waldbauer JR, Summons RE, DeLong EF: **Proteorhodopsin photosystem gene expression enables photophosphorylation in a heterologous host**. *Proc Natl Acad Sci USA* 2007, **104**:5590-5595.
An elegant functional metagenomic screen that provides a useful workflow for the characterization of gene pathways. The authors visually screened for PR-systems expressed in the absence of exogenous retinal and identified PR-containing fosmids capable of retinal prodution. Transposon mutagenesis coupled with liquid chromoatography identified a complete retinal biosynthetic operon.

29. Martinez A, Tyson GW, DeLong EF: **Widespread known and novel phosphonate utilization pathways in marine bacteria revealed by functional screening and metagenomic analyses**. *Environ Microbiol* 2010, **12**:222-238.

30. •• Gabor EM, Alkema WBL, Janssen DB: **Quantifying the accessibility of the metagenome by random expression cloning techniques**. *Environ Microbiol* 2004, **6**:879-886.
The authors calculate gene expression potential of 32 complete genomes in *E. coli* revealing significant differences in modes of gene expression between different taxonomic groups and suggesting that approximately 40% of genes in all genomes may be readily recovered by random cloning in *E. coli.*

31. Warren RL, Freeman JD, Levesque RC, Smailus DE, Flibotte S, Holt RA: **Transcription of foreign DNA in Escherichia coli**. *Genome Res* 2008, **18**:1798-1805.

32. Sorek R, Zhu Y, Creevey CJ, Francino MP, Bork P, Rubin EM: **Genome-wide experimental determination of barriers to horizontal gene transfer**. *Science* 2007, **318**:1449-1452.

33. McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D *et al.*: **The complete genome of Rhodococcus sp. RHA1 provides insights into a catabolic powerhouse**. *Proc Natl Acad Sci USA* 2006, **103**:15582-15587.

34. Martinez A, Kolvek SJ, Yip CLT, Hopke J, Brown KA, MacNeil IA, Osburne MS: **Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts**. *Appl Environ Microbiol* 2004, **70**:2452-2463.

35. •• Craig JW, Chang FY, Kim JH, Obiajulu SC, Brady SF: **Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria**. *Appl Environ Microbiol* 2010, **76**:1633-1641.
The authors report on the development of broad host range vectors and phenotypic screening of environmental DNA libraries in six different proteobacteria. They describe clone-specific small molecule production, and identify the genes responsible for the production of these small molecules.

36. • Streit WR, Daniel R (Eds): *Metagenomics: Methods and Protocols*, vol 668. New York: Humana Press; 2010.
The editors present a useful compendium of metagenomic library production and functional screening methods.

37. Gloux K, Leclerc M, Iliozer H, L'Haridon R, Manichanh C, Corthier G, Nalin R, Blottiere HM, Dore J: **Development of high-throughput phenotyping of metagenomic clones from the human gut microbiome for modulation of eukaryotic cell growth**. *Appl Environ Microbiol* 2007, **73**:3734-3737.

38. • Mewis K, Taupp M, Hallam SJ: **A high throughput screen for biomining cellulase activity from metagenomic libraries**. *J Vis Exp* 2011.
A useful visual experiment demonstrating a robust, liquid based protocol for high throughput screening of carbohydrate active enzymes.

39. Li S, Xu L, Hua H, Ren C, Lin Z: **A set of UV-inducible autolytic vectors for high throughput screening**. *J Biotechnol* 2007, **127**:647-652.

40. Teather RM, Wood PJ: **Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen**. *Appl Environ Microbiol* 1982, **43**:777-780.

41. Feng Y, Duan CJ, Pang H, Mo XC, Wu CF, Yu Y, Hu YL, Wei J, Tang JL, Feng JX: **Cloning and identification of novel cellulase genes from uncultured microorganisms in rabbit cecum and characterization of the expressed cellulases**. *Appl Microbiol Biotechnol* 2007, **75**:319-328.

42. • Uchiyama T, Abe T, Ikemura T, Watanabe K: **Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes**. *Nat Biotechnol* 2005, **23**:88-93.
The authors describe a high throughput method for fluorescent detection of gene expression in metagenomic libraries allowing for a rapid alternative to enzyme activity-based screening.

43. Uchiyama T, Watanabe K: **The SIGEX scheme: high throughput screening of environmental metagenomes for the isolation of novel catabolic genes**. *Biotechnol Genet Eng Rev* 2007, **24**:107-116.

44. Uchiyama T, Watanabe K: **Substrate-induced gene expression (SIGEX) screening of metagenome libraries**. *Nat Protoc* 2008, **3**:1202-1212.

Please cite this article in press as: Taupp M, et al. The art and design of functional metagenomic screens, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.02.010

45. Tracy BP, Gaida SM, Papoutsakis ET: **Flow cytometry for bacteria: enabling metabolic engineering, synthetic biology and the elucidation of complex phenotypes**. *Curr Opin Biotechnol* 2010, **21**:85-99.

46. • Aharoni A, Thieme K, Chiu CP, Buchini S, Lairson LL, Chen H, Strynadka NC, Wakarchuk WW, Withers SG: **High-throughput screening methodology for the directed evolution of glycosyltransferases**. *Nat Methods* 2006, **3**:609-614.
The authors take advantage of the ultra-high throughput of FACS to screen single cells and identifiy phenotypic variants from directed evolution studies.

47. Diercks AH, Ozinsky A, Hansen CL, Spotts JM, Rodriguez DJ, Aderem A: **A microfluidic device for multiplexed protein detection in nano-liter volumes**. *Anal Biochem* 2009, **386**:30-35.

48. Petriv OI, Kuchenbauer F, Delaney AD, Lecault V, White A, Kent D, Marmolejo L, Heuser M, Berg T, Copley M *et al.*: **Comprehensive microRNA expression profiling of the hematopoietic hierarchy**. *Proc Natl Acad Sci USA* 2010, **107**:15443-15448.

49. Pushkarev D, Neff NF, Quake SR: **Single-molecule sequencing of an individual human genome**. *Nat Biotechnol* 2009, **27**:847-850.

50. • Zengler K, Toledo G, Rappe M, Elkins J, Mathur EJ, Short JM, Keller M: **Cultivating the uncultured**. *Proc Natl Acad Sci* 2002, **99**:15681-15686.
The authors present microfluidic technology facilitating the growth of pure cultures from uncultured single cells in proximity, suggesting the ability of cultivating co-dependent organisms.

51. Link AJ, Jeong KJ, Georgiou G: **Beyond toothpicks: new methods for isolating mutant bacteria**. *Nat Rev Microbiol* 2007, **5**:680-688.

52. Cardenas E, Tiedje JM: **New tools for discovering and characterizing microbial diversity**. *Curr Opin Biotechnol* 2008, **19**:544-549.

53. Ferrer M, Beloqui A, Timmis KN, Golyshin PN: **Metagenomics for mining new genetic resources of microbial communities**. *J Mol Microbiol Biotechnol* 2009, **16**:109-123.

54. • Aakvik T, Degnes KF, Dahlsrud R, Schmidt F, Dam R, Yu L, Volker U, Ellingsen TE, Valla S: **A plasmid RK2-based broad-host-range cloning vector useful for transfer of metagenomic libraries to a variety of bacterial species**. *FEMS Microbiol Lett* 2009, **296**:149-158.
The authors describe a vector system that provides a stable and diverse platform for metagenomic analysis across bacterial species.

Please cite this article in press as: Taupp M, et al. The art and design of functional metagenomic screens, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.02.010