Chapter 2
Expression Platforms for Functional Metagenomics: Emerging Technology Options Beyond *Escherichia coli*

Anna Lewin, Rahmi Lale, and Alexander Wentzel

Abstract *Escherichia coli* is the prime workhorse for various metagenomic applications due to the multitude of efficient tools available for genetic manipulation and controlled heterologous gene expression. However, metagenome-based bioprospecting efforts continuously target a wider spectrum of ecological niches in order to harvest new enzymes and bioactive compounds for industrial and medical applications from the enormous pool of natural microbial diversity. Consequently, the development of robust and flexible screening platforms that allow functional evaluation of an expanded fraction of the highly diverse metagenomic information is widely addressed in Functional Metagenomics research. The heterologous recognition of transcriptional regulators and promotors, diverse codon usages among environmental microorganisms, and sufficient supply of precursors for secondary metabolite formation are major challenges that are addressed by an increasing spectrum of alternative expression and host systems. This includes optimized broad host-range transfer and expression vectors, screening hosts for improved gene expression and metabolite formation, as well as cell-free expression systems to cover proteins that due to toxicity are inaccessible by in vivo screening methods. In this chapter, we provide a current overview of the state of the art of selected expression systems and host organisms useful for functional metagenome screening for new enzymes and bioactive metabolites, as emerging options beyond what is currently available in and for *E. coli*.

A. Lewin • A. Wentzel (✉)
Department of Biotechnology and Nanomedicine, SINTEF Materials and Chemistry, Trondheim, Norway
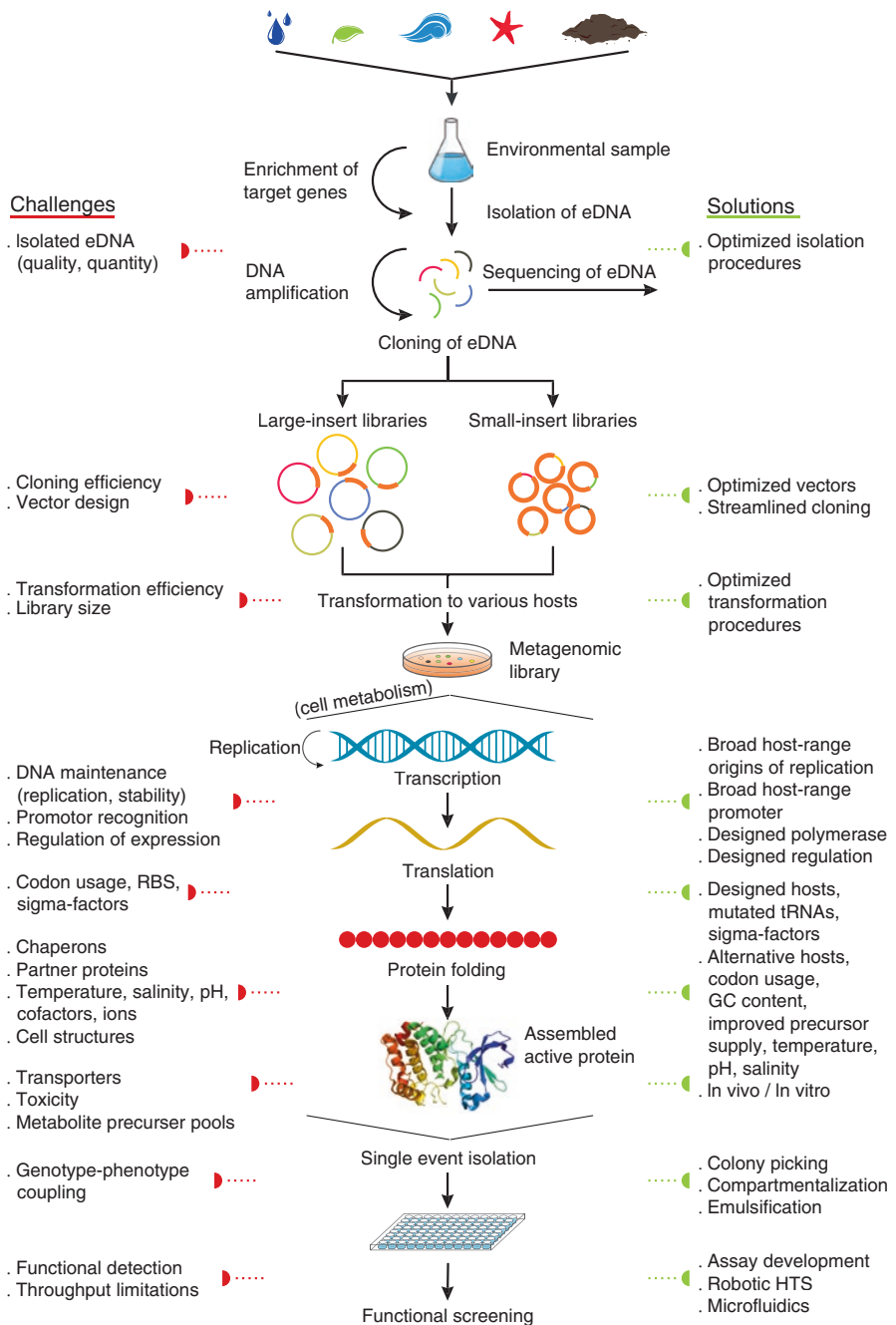e-mail: Alexander.Wentzel@sintef.no

R. Lale
Department of Biotechnology and Food Science, PhotoSynLab, Norwegian University of Science and Technology, Trondheim, Norway

2.1   Introduction

Metagenomics has since its introduction in the late 1990s (Handelsman et al. 1998) proven to be a powerful tool for describing microbial communities and their metabolic potentials irrespectively of cultivability. Over the years, both sequence- and function-based screening approaches have led to the discovery of numerous new enzymes and metabolites fulfilling various academic and industrial needs (Ferrer et al. 2015; Fernandez-Arrojo et al. 2010; Novakova and Farkasovsky 2013). The pipeline for Functional Metagenomics spans from sampling, isolation of high-quality environmental DNA (eDNA), and its cloning (including vector design) to metagenomic library construction (including host transformation and transfer), heterologous gene expression, and production of functional molecules in amounts sufficient for detection in high throughput screening (Fig. 2.1). The function-based screening route of metagenome-based bioprospecting therewith complements the sequence-based route, in which eDNA is sequenced using next-generation sequencing methods and resulting sequence datasets mined bioinformatically for genes of interest (Lewin et al. 2013).

Irrespective of the chosen screening route, successful bioprospecting of a metagenomic library starts with the isolation of the eDNA. Its quality and quantity are of major importance for the achievable number of clones of the constructed library and consequently the representation of biodiversity in an environmental sample (Zhou et al. 1996). In order to capture as much of the biodiversity as possible, the applied DNA isolation procedures need to be highly effective in sampling from the diverse microorganisms inhabiting the selected environment (Kakirde et al. 2010). In addition, isolated DNA needs to have a high degree of purity and be free of contaminating substances, such as humic acids that are often present in soil and hamper efficient library construction (Tebbe and Vahjen 1993). Several studies document eDNA isolation procedures that resulted in contamination-free high molecular weight (HMW) DNA (Zhou et al. 1996; Brady 2007; Liles et al. 2008; Pel et al. 2009; Cheng et al. 2014). Contaminating compounds co-isolated with the eDNA can also be successfully removed by gel electrophoretic methods, including conventional (Craig et al. 2010), pulse-field (Cheng et al. 2014), or nonlinear electrophoresis (Pel et al. 2009), followed by size selection of the random fragmented DNA, prior to cloning.

New and improved enzyme discovery is currently the largest field of application for Functional Metagenomics tools. Aside from the catalytic function itself, beneficial properties like robustness under harsh conditions or high activity at low temperatures are often required in industrial applications. Consequently, dependent on the aims of a bioprospecting approach, different environments might serve as eDNA sources (Taupp et al. 2011). The microbial habitat to be sampled usually reflects the desired properties, i.e., subjecting a metagenomic library originating from a thermal vent or a hot deep subsurface oil reservoir to thermostable enzyme screening is likely to have a higher success rate compared to subjecting a

Challenges

. Isolated eDNA
  (quality, quantity)

. Cloning efficiency
. Vector design

. Transformation efficiency
. Library size

. DNA maintenance
  (replication, stability)
. Promotor recognition
. Regulation of expression

. Codon usage, RBS,
  sigma-factors

. Chaperons
. Partner proteins
. Temperature, salinity, pH,
  cofactors, ions
. Cell structures

. Transporters
. Toxicity
. Metabolite precurser pools

. Genotype-phenotype
  coupling

. Functional detection
. Throughput limitations

Solutions

. Optimized isolation
  procedures

. Optimized vectors
. Streamlined cloning

. Optimized
  transformation
  procedures

. Broad host-range
  origins of replication
. Broad host-range
  promoter
. Designed polymerase
. Designed regulation

. Designed hosts,
  mutated tRNAs,
  sigma-factors

. Alternative hosts,
  codon usage,
  GC content,
  improved precursor
  supply, temperature,
  pH, salinity
. In vivo / In vitro

. Colony picking
. Compartmentalization
. Emulsification

. Assay development
. Robotic HTS
. Microfluidics

Environmental sample

Enrichment of
target genes

Isolation of eDNA

DNA
amplification

Sequencing of eDNA

Cloning of eDNA

Large-insert libraries   Small-insert libraries

Transformation to various hosts

Metagenomic
library

(cell metabolism)

Replication

Transcription

Translation

Protein folding

Assembled
active protein

Single event isolation

Functional screening

Fig. 2.1 Graphical representation of the Functional Metagenomics biodiscovery pipeline with its key challenges and potential solutions

glacier or permafrost soil-originating library to the same screening. In many examples such directed metagenomics sampling strategies aiming to increase probability of finding the desired properties have proven successful (Vester et al. 2015; Taupp et al. 2011). Selected examples are, among many others, a cold-adapted esterase enzyme from Antarctic desert soil (Hu et al. 2012), hydrolytic enzymes from cow rumen metagenome (Ferrer et al. 2007), and thermostable lipolytic enzymes from water, sediment, and biofilm samples from the Azores, Portugal (Leis et al. 2015a). However, due to often lower microbial density within some, particularly extreme environments, sufficient DNA yields may not be readily obtainable (Kennedy et al. 2008; Vester et al. 2015; Kotlar et al. 2011). In such cases, isolated metagenomic DNA can be subjected to isothermal amplification (like Phi29 whole genome amplification, WGA) in order to increase DNA yields prior to cloning (Rodrigue et al. 2009; Zhang et al. 2006). However, the challenge of this technology with respect to the formation of amplification artifacts, like chimeras, duplications, and inversions, needs to be considered. Therefore it is well suited for small-insert libraries for the purpose of enzyme discovery, but less suitable for large-insert library cloning where intact biosynthetic gene clusters are targeted.

Following sampling and successful isolation, the eDNA is usually either sequenced directly or cloned in suitable vectors for functional screening approaches (Sect. 2.2). The choice of the vector usually depends on the envisioned eDNA insert sizes, as well as the screening targets and methodology. However, for successful expression of genetic information contained in metagenomic DNA libraries, several additional factors need to be taken into account (Fig. 2.1). Suitable vector systems need to carry host-compatible selection markers, replicate stably and autonomously (ideally in combination with the possibility to control the copy number), may contain functional gene regulatory elements like inducible promotors for high level expression, and preferably enable vector transfer to other host organisms. Suitable host organisms in turn need to provide functionality of the vector elements involved in the production of functional products and allow efficient transcription and translation (Sect. 2.3). In addition, proper folding, possible cofactor supply, sufficient precursor availability for metabolite product formation, as well as means for nontoxic product localization, like secretion mechanisms, are needed. In order to meet the different demands for functional expression, such as codon usage, different assay temperatures, precursor requirements, etc. (Lam and Charles 2015; Uchiyama and Miyazaki 2009), different approaches can be applied in order to maximize the probability of successful expression (Fig. 2.1). *E. coli* systems designed and optimized for this purpose have so far been most widely used and extensively covered elsewhere (Guazzaroni et al. 2015). The scope of this chapter is therefore to summarize developments of various hosts and heterologous expression systems for functional metagenome screening beyond the common systems available for *E. coli* only.

2.2 Cloning and Expression Vectors for Environmental DNA

Selection of a suitable vector system for random metagenomic library construction will largely be guided by (1) the expected DNA size encoding the targeted compound of interest, (2) the envisioned subsequent screening approach involving one or more expression hosts, (3) the desired design of the library to be established, as well as in some occasions (4) the quantity of DNA available (Sect. 2.2.1). For approaches to identify new enzymes, small-insert libraries with eDNA sizes of 5–10 kb will in most cases be sufficient to obtain a sufficiently large number of complete gene sequences. Isolation of DNA for small-insert libraries is normally straightforward, since DNA shearing is not a major concern. However, it needs to be considered that a library with an average insert size of 10 kb will require 3–20 times more clones compared to a library with inserts of 30–40 kb to cover the same amount of genetic potential (Sabree et al. 2009). Hence comparably larger amounts of DNA are needed. To identify encoded functions that rely on single genes or small gene loci (e.g., enzyme function or genetic determinants of antibiotic resistance (Riesenfeld et al. 2004)), small-insert libraries are normally sufficient (Kakirde et al. 2010; Sabree et al. 2009). However, in cases where a desired function depends on multiple gene products, libraries harboring larger inserts are needed. These are normally constructed as cosmid, fosmid (30–40 kb), or bacterial artificial chromosome (BAC) libraries (up to ≥100 kb). The construction of comprehensive large-insert libraries can be very laborious, both with respect to the isolation of HMW DNA and successful cloning and transformation of the host. In addition, the lower stability of large inserts in the generated library needs to be considered. Also, the aspect of a higher degree of degradation of low guanine + cytosine (G + C) content DNA and some DNA modifications, which can impair cloning of HMW DNA, can result in a bias within large-insert libraries (Danhorn et al. 2012).

The choice of suitable vector systems is usually also related to the available expression host organism for subsequent screening experiments (Sect. 2.3). Moreover, for some targets, screening in multiple hosts can increase the hit rates (Mullany 2014). Hence library transfer and broad host-range capabilities of an expression vector (Sect. 2.2.2) can be desired characteristics (Craig et al. 2010; Aakvik et al. 2009; Kakirde et al. 2010).

2.2.1 Small- and Large-Insert Random Cloning Vectors

Cloning vectors useful for small-insert metagenomic library construction usually contain a defined promoter for transcription of the inserted DNA sequence. In some cases they are even equipped with two promoters (dual promotor vectors), flanking both sides of the cloning site in order to achieve gene expression regardless of insert orientation (Lammle et al. 2007). The promoters can have different, independent

induction mechanisms in order to achieve expression in only one direction at a time to prevent potential mRNA duplex formation that may result in lower protein production (Lale et al. unpublished). For cloning and construction of small-insert metagenomic libraries in *Escherichia coli* as primary host organisms, standard cloning vectors, such as pUC derivatives, pBluescript SK(+), and pTOPO, or their derivatives (Mullany 2014; Sabree et al. 2009) are frequently used.

In order to allow metagenomic library clones to cover entire biosynthetic pathways, like secondary metabolite clusters, large-insert libraries are required. Such libraries can be generated as cosmids or fosmids based on phage packaging of the eDNA ligated to a respective vector fragment or for very large inserts (up to 100 kb or more) as BACs (Kakirde et al. 2010; Danhorn et al. 2012). Fosmid and cosmid cloning vectors carry inserts of 30–40 kb, and both approaches utilize phage-based transfer of the cloned DNA into the host, usually *E. coli*. Consequently, the resulting library clones carry inserts within a narrow size range, determined by the packing capacity of the phage particle, and generally rely on gene expression from promoters included in the cloned insert. Cosmids are hybrid plasmids containing *cos* sequences from the λ phage, whereas fosmids are based on the F-factor replicon from *E. coli*. Compared to cosmids, fosmids are more tightly regulated with respect to copy number and are hence more stable (Kim et al. 1992; Kakirde et al. 2010). Both cosmids and fosmids are designed to carry antibiotic resistance markers and have broad host-range capabilities (Craig et al. 2010; Cheng et al. 2014; Aakvik et al. 2009; Wexler et al. 2005). Due to the frequent use of both cosmid and fosmid systems for metagenomic library construction, several variants (including commercial ones) are available (Lam et al. 2015; Mullany 2014; Kim et al. 1992; Parks and Graham 1997; Li et al. 2011; Terron-Gonzalez et al. 2013).

For random cloning of very large inserts, 40–100 kb and above, BACs are normally used, relying on the F-factor replicon (Danhorn et al. 2012; Shizuya et al. 1992). BAC vectors have been used in several metagenomic studies (Brady 2007) using, e.g., soil samples (Rondon et al. 2000) and murine bowel microbiota (Yoon et al. 2013). Similar to fosmids and cosmids, there are different BAC systems available, with some of them allowing inducible high copy numbers (Mullany 2014; Warburton et al. 2009; Wild et al. 2002) and/or having broad host-range capability (Mullany 2014; Aakvik et al. 2009; Kakirde et al. 2010). The US-based company Lucigen Corp. (Madison, WI; www.lucigen.com) has developed dedicated broad host-range vector systems for use in Functional Metagenomics. The pBAC-SBO and pSMART-BAC-S vectors both attribute efficient library construction in *E. coli* and are transferable to both Gram-positive and Gram-negative hosts. They have features allowing selection in several host organisms and gene expression from both insert-flanking regions, and are inducible in copy number (see Chap. 1). pSMART-BAC-S vector provides integration in the host genome only, whereas the pBAC-SBO vector allows both chromosomal integration, as well as extrachromosomal propagation in the recipient.

For DNA experiencing superhelical stress due to, e.g., regions dense in tandem and/or inverted repeats, cloning into circular plasmids can be challenging. In such cases, linear plasmids, such as the pJAZZ vector series (Lucigen), have been

designed which can carry large DNA inserts and contain features like transcriptional terminators flanking the cloning site to hinder vector-insert transcriptional interference (Godiska et al. 2010).

2.2.2 Broad Host-Range Expression Vectors

Depending on the desired activity, functional screening in different (or several) hosts can be of high value. As mentioned, *E. coli* is the most commonly used host both for library construction and functional screening. However, for certain screening activities, such as thermostable enzymes, or for bioactive secondary metabolite production, hosts like *Thermus thermophilus* (Angelov et al. 2009) and *Streptomyces* (or other *Actinobacteria*), respectively, might be beneficial due to their inherent features (Kakirde et al. 2010; Martinez et al. 2004) (see Sects. 2.3.1 and 2.3.2). Metagenomic libraries can be constructed directly in the host where they will be screened. However, the number of transformants obtained is often much lower in such hosts compared to the number of clones that can be obtained in *E. coli*. Thus, the common method is to utilize shuttle and/or broad host-range vectors for library construction in *E. coli*, which allows library transfer and screening in the host organism of choice. There are various such vectors available, both for small and large inserts. *E. coli–Bacillus subtilis* shuttle systems (plasmid and BAC) have been used for screening soil metagenomes for antimicrobial activities (Biver et al. 2013), and the pMDB14 vector (McMahon et al. 2012) can be shuttled between *E. coli*, *Pseudomonas putida*, and *Streptomyces lividans*, allowing gene expression in different hosts, similar to other systems reported (Sosio et al. 2000; Martinez et al. 2004). For development of psychrophilic expression systems, *E. coli* shuttle vectors such as a pGEM derivative and a pJRD215 derivative have been constructed, allowing the transfer of constructed libraries from *E. coli* to, e.g., *Psychrobacter* sp. and *Shewanella livingstonensis* (Cavicchioli et al. 2011; Miyake et al. 2007; Tutino et al. 2001). Also, *E. coli–T. thermophilus* shuttle systems have been designed (Angelov et al. 2009; Leis et al. 2015b). Apart from these, several other broad host-range systems have been developed. The pUvBBAC system supports replication in both Gram-positive and Gram-negative bacteria and allows functional screening in *Listeria* hosts (Hain et al. 2008). pGNS-BAC-1 presents opportunities for a copy induction in *E. coli*, as well as replication and functional screening in a broad spectrum of Gram-negative species (Kakirde et al. 2010). The pRS44 plasmid system (Aakvik et al. 2009) has been constructed both as fosmid and BAC system, which enables induction based on control on the vector copy number in *E. coli* and conjugative transfer into other hosts. In addition to the transferable BAC systems, several broad host-range cosmid vectors have also been reported (Craig et al. 2010; Cheng et al. 2014; Wexler et al. 2005).

In order to exploit the benefits of metagenomic library screening in several hosts with complementary features (Martinez et al. 2004; Leis et al. 2015a, b), efficient library transfer between host strains is of high importance. Though library vector

isolation by simple plasmid DNA extraction followed by re-transformation into the alternative host is possible, conjugation is in most cases the transformation method of choice. This is generally regardless of whether the library originally was constructed as a cosmid (Wexler et al. 2005; Craig et al. 2010; Cheng et al. 2014), fosmid (Aakvik et al. 2009), or BAC (Kakirde et al. 2010). The conjugative transfer of abovementioned vectors requires the full set of *tra* genes to be present in the donor (F positive) strain. Libraries to be transferred are often large, and therefore library transfer is preferably done in a high throughput fashion, similar to the high throughput conjugation procedure described by Martinez and co-workers (Martinez et al. 2004).

2.3   Expression Host Organisms

As previously mentioned, *E. coli* is presently the most commonly used expression host in metagenomic functional screening efforts (Ekkers et al. 2012; Kennedy et al. 2008; Aakvik et al. 2009; Rondon et al. 2000; Parachin and Gorwa-Grauslund 2011). Several dedicated tools for metagenomic library screening have also been developed for *E. coli*, such as engineered strains suitable for stable replication and copy control of large vectors. These remain at one single copy prior to screening to minimize the potential toxic effects of insert-encoded proteins or other produced metabolites (Taupp et al. 2011). *E. coli* strains have been modified for optimized heterologous expression, e.g., by expression of heterologous sigma factors that allow recognition of a wider range of promoter structures than *E. coli* wild-type strains (Gaida et al. 2015), and for heterologous expression of polyketide synthase (PKS) encoding secondary metabolite gene clusters and production of derivable natural products (Zhang et al. 2015). However, even engineered *E. coli* strains are not in all cases the best-suited hosts with respect to expressing metagenome-encoded functions. This accounts particularly for screening of metagenomic libraries harboring eDNA from extreme environments at conditions not compatible with *E. coli*'s natural lifestyle as a mesophilic human commensal, like very high or low temperatures. In addition, functional expression of genes from species that are phylogenetically distant from *E. coli* can be challenging (Warren et al. 2008). This can be due to, e.g., the differences in codon usage, improper promoter recognition, lack of transcription and/or translation factors, hampered protein folding, absence of cofactors, gene product toxicity, and absence of precursor metabolites. It has been shown that only approximately 40% of all genes can be heterologously expressed in *E. coli* (Gabor et al. 2004). Therefore, the use of multiple, complementary screening hosts has been proposed to express more of the diversity within a metagenomic library (Liebl et al. 2014). Table 2.1 summarizes the most commonly used as well as high potential future host systems for functional metagenome screening.

Table 2.1 Key features of frequently used as well as high potential future hosts for functional expression and screening of metagenomic libraries

Host organism	Key features	Key references
Escherichia coli	+ Well established as a heterologous expression and screening host	Ekkers et al. (2012), Kennedy et al. (2008), Aakvik et al. (2009), and Rondon et al. (2000), Parachin and Gorwa-Grauslund (2011)
	+ Fully developed toolbox for genetic manipulation, easy manageable	Taupp et al. (2011) and Gaida et al. (2015)
	+ Genetic transfer systems well established	
	+ Designed/optimized strains for different screening purposes available	Gabor et al. (2004) and Zhang et al. (2015)
	− *Only distantly related to many environmental microbes, codon usage*	Warren et al. (2008)
	− *Restricted to mesophilic cultivation an screening*	Liebl et al. (2014)
	− *Limited precursor availability for secondary metabolite formation*	
	− *Restrictions with respect to cofactor availability*	
	− *Potential toxicity effects due to limits in protein secretion, Gram-negative*	
Thermus thermophilus	+ Allows cultivation and in vivo screening high temperatures	Tabata et al. (1993) and Cava et al. (2009)
	+ Natural competence for DNA uptake	Hidaka et al. (1994)
	+ Efficient transformation protocols available	Schwarzenlander and Averhoff (2006)
	+ Thermostable resistance markers and other genetic tools (promoters, origin of replication, etc.) available	Matsumura and Aiba (1985), Liao et al. (1986), Nakamura et al. (2005), and Tamakoshi et al. (1997)
	+ *T. thermophilus*–*E. coli* shuttle vectors available	Lasa et al. (1992) and Wayne and Xu (1997)
	+ Chromosomal integration well established	de Grado et al. (1999)
	− *Only a limited set of selection markers available*	
	− *Only a few promoter alternatives developed*	
Streptomyces spp.	+ Full set of genetic tools available	Gust et al. (2004), Kieser et al. (2000), and Jones et al. (2013)

Table 2.1 (continued)

Host organism	Key features	Key references
Streptomyces coelicolor	+ Optimized strains for heterologous gene cluster expression available	Gomez-Escribano and Bibb (2011, 2012, 2014)
	+ High G + C content, thus complementary to other expression hosts	Gomez-Escribano and Bibb (2014)
	+ Assembly platform for secondary metabolite production machinery	Shima et al. (1996), Okamoto-Hosoya et al. (2000), and Hu et al. (2002)
	+ Natural provision of precursors for sec. metabolite formation	
	+ Gram-positive, efficient protein/ enzyme secretion	
	− *Mycelial growth phenotype, advanced cultivation systems necessary*	Wentzel et al. (2012a)
Rhodobacter capsulatus	+ Suitable for expression of membrane proteins	Liebl et al. (2014)
Gluconobacter oxydans	+ Acid tolerant	Liebl et al. (2014)
Burkholderia graminis	+ *E. coli* alternative	Craig et al. (2010)
Caulobacter vibrioides	+ *E. coli* alternative	Craig et al. (2010)
Pseudomonas putida	+ Stress tolerant	Craig et al. (2010) and Troeschel et al. (2010)
	+ Capable of producing secondary metabolites	Loeschcke and Thies (2015)
Ralstonia metallidurans	+ Robustness at extreme conditions, broad screening spectrum	Craig et al. (2010) and Mergeay et al. (2003)
Pseudomonas fluorescens	+ *E. coli* alternative	Aakvik et al. (2009)
	+ Secretion pathway	Retallack et al. (2006)
Xanthomonas campestris	+ *E. coli* alternative	Aakvik et al. (2009)
	+ Increased stabilities of proteins produced	Leza et al. (1996)
Sinorhizobium meliloti	+ *E. coli* alternative	Cheng et al. (2017)
	+ Chromosomal integration established	Heil et al. (2012)
Agrobacterium tumefaciens	+ *E. coli* alternative, plant symbiont	Craig et al. (2010), Troeschel et al. (2010), and Murai (2013)
	+ Chromosomal integration established	Heil et al. (2012)

Table 2.1   (continued)

Host organism	Key features	Key references
Bacillus subtilis	+ Gram-positive model organism	Biver et al. (2013)
	+ Fully developed toolbox for genetic manipulation, easy manageable	
	+ Low GC, thus complementary to other expression hosts	
	+ Secretion pathway (enzyme production)	Wong (1995) and Zobel et al. (2015)
R. eutropha H16	+ *E. coli* alternative	Gruber et al. (2015)
Saccharomyces cerevisiae	+ Eukaryotic (yeast)	Damon et al. (2011)
	+ Fully developed toolbox for genetic manipulation	
	+ Secretion pathway (enzyme production)	Strausberg and Strausberg (2001)
	+ Protein posttranslational modifications, glycosylation	Holz et al. (2003)

2.3.1   Extremophiles as Expression Hosts for Metagenome Screening

Similar to sampling and cloning of metagenomic DNA from an environment that matches the desired properties of an enzyme, it appears reasonable to use a heterologous expression host for metagenomic library screening that functions optimally under respective conditions, such as in vivo screening for thermostable enzymes at elevated temperature in a thermophilic host. In terms of thermophilic hosts for heterologous gene expression, the hyperthermophilic, Gram-negative bacterium *T. thermophilus* (*Deinococcus-Thermus* phylum), growing optimally at temperatures as high as 85 °C, is the most well-studied species (Tabata et al. 1993; Cava et al. 2009). Several *T. thermophilus* strains have been genome sequenced, and their natural competence (Hidaka et al. 1994; Koyama et al. 1986) renders them very efficient in taking up external DNA without source discrimination (Schwarzenlander and Averhoff 2006). In addition, *T. thermophilus* has been shown to acquire DNA by conjugation, however, not as effectively as by utilizing its natural competence (Ramirez-Arcos et al. 1998; Cava et al. 2007).

A large number of genetic tools have been developed to genetically amend *T. thermophilus* (Tamakoshi et al. 1997; de Grado et al. 1999). In the 1980s a selection marker in the form of a thermostable version of a kanamycin resistance was developed using mutagenesis (Matsumura and Aiba 1985; Liao et al. 1986), allowing antibiotic-based selection for transformed *T. thermophilus* cells. Since then other selectable markers stable at high temperature have been used such as the bleomycin-binding protein conferring bleomycin resistance (Brouns et al. 2005),

and a hygromycin B phosphotransferase evolved to thermostability (Nakamura et al. 2005). Several plasmids and vectors are available, like the cryptic pTT8 plasmid (Koyama et al. 1990) used to transfer genes into *T. thermophilus*. pTT8 has also been supplemented with the gene providing thermostable kanamycin resistance described above, resulting in the selectable cloning vector pMKM001 (Mather and Fee 1992) and variants thereof. The *Thermus*-compatible plasmids have also been engineered further into *E. coli–Thermus* shuttle vectors by integration of the cryptic *Thermus* vectors with commonly used *E. coli* plasmids from the pUC series, resulting in several variants, e.g., pMY1-3 and pLU1-4 (Lasa et al. 1992; Wayne and Xu 1997). In addition, there are other plasmids available for *Thermus*, like plasmid pTA103 (Chu et al. 2006), the pS4C, and pL4C plasmids harboring both integrase and transposase (Ruan and Xu 2007) as well as the widely used pMK18 vector carrying the multiple cloning site from pUC18 (de Grado et al. 1999).

As a consequence of the available genetic tools for *T. thermophilus*, these strains have been used as thermophilic cell factories to complement production of certain proteins in, e.g., *E. coli*. *T. thermophilus* has been used to homologously produce *Tth* DNA polymerase more efficiently than in *E. coli* (Moreno et al. 2005), as well as for production of an active thermostable Mn-dependent catalase which failed to express in *E. coli* (Hidalgo et al. 2004). *T. thermophilus* has also been successfully used in metagenomic approaches, e.g., by Angelov and co-workers (2009). In their work, large-insert fosmid libraries were constructed in *E. coli* and transferred to a *T. thermophilus* host. Screening was performed in both species, resulting in different hit spectra. This clearly illustrates the benefits of high-temperature screening for thermostable enzymes. The same authors also constructed a pCC1fos derivative (denoted pCT3FK) which carries *T. thermophilus* HB27 chromosomal DNA sequences which allow integration in the host chromosome by homologous recombination (Angelov et al. 2009). This vector has been used in the screening of a metagenomic library for thermostable esterases in both *E. coli* and *T. thermophilus* hosts, resulting in a higher number of thermostable enzyme candidates in the *T. thermophilus* than in the *E. coli* screening (Leis et al. 2015a, b).

On the opposite end of the temperature range, cold environments provide a large understudied biodiversity. Particularly psychrophilic enzymes from such environments are sought due to their unique characteristics, i.e., high activity at low and moderate temperatures, necessitating lower enzyme concentrations to achieve a similar performance compared to higher temperature homologues. Psychrophilic enzymes are considered to be less stable compared to their mesophilic homologues, as their structural flexibility enables them to function at low temperatures and imparts a decreased thermal stability (Feller 2013). However, the biodiscovery of relevant gene functions from these environments is limited to their expression and function in mesophilic hosts. For instance, the utilization of *E. coli* as a host for the expression of psychrophilic enzymes limits the growth temperature to around 15 °C, which presents a significant barrier to their exploitation in biotechnology (Struvay and Feller 2012).

There are several examples where *E. coli* has been successfully used in the production of cold-adapted enzymes (Cavicchioli et al. 2011; Wang et al. 2010; Zhang

and Zeng 2008). However, the total number of such reports is comparably low, reflecting significant challenges. Two strategies to overcome these challenges are (1) low-temperature adaptation of existing mesophilic expression systems and (2) the development of new psychrophilic expression hosts. The former approach includes engineering the mesophilic expression host for sufficient growth at low temperatures to promote correct folding of recombinant proteins. The co-expression of Cpn60 and Cpn10 from *Oleispira antarctica*, cold-adapted homologues of the *E. coli* GroELS chaperonins, provided *E. coli* with an operational folding system at 4–12 °C (Ferrer et al. 2003). This led to improved growth at low temperatures and enhanced solubility of the recombinant proteins produced. Another example is the utilization of cold-shock promoter systems together with solubility partners for psychrophilic genes in *E. coli.* Bjerga and Williamson showed that cspA-driven expression of maltose-binding protein (MBP), thioredoxin (TRX), small ubiquitin-like modifier (SUMO), and trigger factor (TF) encoding gene fusion enabled high level production of soluble protein (Bjerga and Williamson 2015).

Dedicated host-expression systems for the production of cold-adapted products have been developed, such as the pTAUp and pTADw vectors for *Psychrobacter,* found to replicate by rolling circle mechanisms (Tutino et al. 2000). Also, the cryptic replicon plasmid pMtBL from *Pseudoalteromonas* sp. has been used as a psychrophilic expression vector, shown to have a broad host-range profile compatible to not only psychrophiles but also mesophilic species after fusion with a pGEM derivate (Tutino et al. 2001). Other broad host-range vectors for cold-adapted expression include a variant of pJRD215 carrying a regulatory promoter from *Shewanella* and a β-lactamase reporter from *Desulfotalea* (Miyake et al. 2007) and a shuttle vector based on the p54 plasmid originating from a psychrophilic *Arthrobacter* sp. isolated from a Greenland glacier and pUC18. The latter example resulted in a low-temperature expression system transferrable to not only *E. coli* but also some high G + C Gram-positive bacteria (Miteva et al. 2008).

2.3.2   *Actinobacteria as Hosts for Heterologous Natural Product Formation*

The phylum *Actinobacteria* comprises a comprehensive and diverse group of Gram-positive bacteria predominantly with a mycelial lifestyle. They are potent producers of a plethora of natural products with a wide spectrum of medical applications, including antibacterial, antifungal, anthelmintic, and immunosuppressant compounds (Barka et al. 2016). Among them, members of the *Streptomyces* taxon are particularly prolific in this respect, accounting for the majority of antibiotics in medical use today (Hopwood 2007). Actinomycete genomes contain a multitude of secondary metabolite gene clusters (Bentley et al. 2002; Ohnishi et al. 2008; Oliynyk et al. 2007; Udwary et al. 2007) of which, however, only a subset is expressed and the respective compounds produced under laboratory conditions. Hence, the majority of gene clusters remains silent, rendering them cryptic, with functions yet to be

discovered. Also among *Actinobacteria*, cultivable strains represent only a minute fraction of the entire diversity (Maldonado et al. 2005), leaving a vast resource of new potential drug candidates untapped, unless new methods to enable cultivation (Zengler et al. 2002), or efficiently allow the heterologous realization of their genetic potential, become available. In that respect, well-described members of the *Actinobacteria* themselves, like the model species *Streptomyces coelicolor*, have been proposed as hosts for the heterologous expression of natural product gene clusters (Gomez-Escribano and Bibb 2011, 2012). Their versatility with respect to expressing complex biosynthetic gene clusters, their high G + C codon usage, and the provision of important precursors necessary to simultaneously form natural products of different compound classes (like polyketides, non-ribosomal peptides, lantibiotics, etc.) are excellent rationales to select such strains for metagenome screening for new bioactive compounds. In addition, these *Streptomyces* spp. strains might prove useful in accessing the potential of cryptic gene clusters of cultivable strains by heterologous expression. In-depth understanding of gene regulation and precursor supply will be instrumental in optimizing model *Actinobacteria* as functional metagenome screening hosts.

S. coelicolor has been extensively studied with respect to the regulation of secondary metabolite production, and all necessary genetic tools for genetic manipulation, like plasmids and inducible promoters, and large-insert library tools for chromosomal integration (Gust et al. 2004; Kieser et al. 2000; Jones et al. 2013), are fully developed. Also, new tools for fast and efficient genome editing, like the CRISPR/Cas system (Garneau et al. 2010), have been optimized and applied to *Actinobacteria* to make deletions and directed genomic mutations (Tong et al. 2015; Huang et al. 2015; Cobb et al. 2015). Though its applicability to introduce larger gene clusters into the *Streptomyces* genome is currently limited, it can be expected that this technology will develop into a powerful tool for reprogramming *Actinobacteria* for the production of new bioactive compounds. Wild-type *S. coelicolor* produces several antibiotic compounds of different classes, including the polyketides actinorhodin (Act, Rudd and Hopwood 1979) and coelimycin (Cpk, Gomez-Escribano et al. 2012), the prodiginine undecylprodigiosin (Red, Feitelson et al. 1985), the lipopeptide calcium-dependent antibiotic (CDA, Hopwood and Wright 1983), and the plasmid-encoded cyclopentanoid methylenomycin (Mmy, Wright and Hopwood 1976). However, its genome sequence revealed a much larger potential of bioactive compounds, represented by more than 20 different, mostly non-expressed gene clusters for secondary metabolites (Bentley et al. 2002). Extensive research has been performed to detect and study cryptic gene clusters (Medema et al. 2011; Nett et al. 2009; Zerikly and Challis 2009; Baltz 2008) and ultimately activate them for product formation (Ochi et al. 2014; Rutledge and Challis 2015; Yoon and Nodwell 2014; Zhu et al. 2014). However, regulation of antibiotic production by *S. coelicolor* is complex and needs to be understood in depth when considering it as a generic cell factory for heterologous natural product formation.

Several factors are involved in triggering antibiotic production in *Streptomyces* in correlation with the species' life cycle (Bibb 2005; van Wezel and McDowall 2011).

Nutrient depletion and cessation of growth induce morphological differentiation and antibiotic production via the stringent response and guanosine tetra- and pentaphosphate (p)ppGpp (Potrykus and Cashel 2008). Programmed cell death and the release of N-acetyl glucosamine (GlcNAc) trigger the onset of development and antibiotic production via the global regulator DasR (Rigali et al. 2006, 2008). Also, induced mycelial fragmentation by overexpression of cell division activator protein SsgA affects antibiotic production in *S. coelicolor* (van Wezel et al. 2009). From responses of the global regulatory network, information is passed on to pathway-specific activators encoded within biosynthetic gene clusters, usually controlled in a growth phase-dependent manner (Wietzorrek and Bibb 1997). Once produced in sufficient amount, these are solely responsible for all further downstream regulation of the biosynthetic gene cluster expression. Removal of pathway-specific regulators (Smanski et al. 2012) or exchange of native promotors (Du et al. 2013) as well as overexpression of export proteins (Huo et al. 2012) have led to improved production yields of platencin, gougerotin, and bottromycin, respectively.

Taking all the different layers of regulation into account will be the key for developing *Streptomyces* into potent heterologous production platforms for natural product discovery, from both silent gene cluster in cultivable microorganisms and realizing the biosynthetic potential in environmental metagenomes. *S. coelicolor* has been extensively used as heterologous expression platform for antibiotic gene clusters as recently reviewed by Gomez-Escribano and Bibb (2014). By successively deleting the biosynthetic gene clusters for Act, Red, CDA, and Cpk in the plasmid-free (thus Mmy negative) wild-type M145 of *S. coelicolor*, a strain (M1146) was obtained with largely reduced background of bioactive compounds produced and secreted to the medium (Gomez-Escribano and Bibb 2011). In the same work, additional introduction of point mutations in the genes rpoB and rpsL, encoding the RNA polymerase β-subunit and the ribosomal protein S12, respectively, (strain M1154) led to a pleiotropic increase in the level of secondary metabolite production. Each of these mutations had previously been shown to enhance antibiotic production levels in *Streptomyces* without negative effects on growth (Shima et al. 1996; Okamoto-Hosoya et al. 2000; Hu et al. 2002) and has been proposed as a new strategy to activate silent gene clusters for new drug discovery (Ochi and Hosaka 2013). M1146 and M1154 have been successfully applied for the heterologous production of numerous antibiotics of diverse classes (Gomez-Escribano and Bibb 2014).

A further optimization of the existing heterologous host strains of *S. coelicolor* as an optimized Superhost for new antibiotics discovery from environmental metagenomes may be guided by the comprehensive knowledge of physiology and gene regulation of antibiotic production, as well as systems biology understanding of this species. A dedicated fermentation strategy for system scale studies of metabolic switching in *S. coelicolor* has been established (Wentzel et al. 2012a), allowing reproducible cultivations of *S. coelicolor* and high-resolution time-scale sampling for full 'omics analysis (Battke et al. 2011). The dynamic architecture of the metabolic switch in *S. coelicolor* was studied at the gene expression (Nieselt et al. 2010), the proteome (Thomas et al. 2012) and the metabolome level (Wentzel

et al. 2012b). By studying the effect of different mutations, the complex regulatory interplay of nitrogen and phosphate metabolism was elucidated (Martin et al. 2012; Waldvogel et al. 2011). A genome scale model for *S. coelicolor* is available (Alam et al. 2010), and detailed insight in the structure of the transcription factor mediated regulatory network has been gained (Iqbal et al. 2012).

In addition to *S. coelicolor*, other *Actinobacteria* species have been considered as heterologous expression hosts. *S. avermitilis*, for example, has been engineered as an expression host for heterologous gene clusters (Komatsu et al. 2013), and also *S. lividans* and *S. albus* as well as *Saccharopolyspora* (Baltz 2010) have been used for that purpose. *S. lividans* was used as host organism in successful screening for anti-mycobacterial compounds (Wang et al. 2000), and both *S. lividans* and *S. albus* have been shown to be able to produce products from an introduced Type II PKS pathway (King et al. 2009). *Nonomuraea* sp. ATCC 39727 heterologously produced microbisporicin and planosporicin (Marcone et al. 2010) more efficiently than as *Streptomyces* hosts (Foulston and Bibb 2010; Sherwood and Bibb 2013), indicating potential benefits of using several actinobacterial expression hosts for bioactive compound screening of metagenome libraries. *Streptomyces* spp. have proven to be useful in heterologous gene cluster expression and functional screening for associated bioactivity (Kakirde et al. 2010; Martinez et al. 2005). Screening of a BAC library from soil DNA produced in *E. coli* and transferred to *Pseudomonas putida* (low G + C) and *Streptomyces lividans* (high G + C) resulted in different expression patterns (Martinez et al. 2004), indicating usefulness of the high G + C *Streptomyces* hosts as complement to other metagenome screening platforms for bioactivity, like polyketide production-optimized *E. coli* BTRA (Zhang et al. 2015).

Recently, the "*Tectomicrobia*" candidate phylum including the "*Entotheonella*" candidate genus has been discovered by a combined single cell- and metagenomics-based approach to describe microbial consortia producing bioactive polyketides and peptides in association with the marine sponge species *Theonella swinhoei* (Wilson et al. 2014). This study exemplifies the huge potential of marine environments to identify new compounds produced by non-cultivable microbial strains. The genetic optimization of different actinobacterial model strains for natural product formation will help in establishing a platform of different optimized host strains that in combination can potentially be useful in functional screening also for new natural products from such biodiversity with an increased success rate.

2.3.3 *Other Expression Hosts for Metagenome Screening*

There are several other species apart from *E. coli* and those discussed above (Sects. 2.3.1 and 2.3.2) that have been considered as hosts for metagenome expression and screening, all with their respective benefits and drawbacks. These species can contribute to building a flexible platform for multi-host expression and screening of microbial metagenomes as suggested before (Liebl et al. 2014).

Mesophilic hosts applied for metagenomic screening, apart from *E. coli* and the *Actinobacteria* covered in detail above (Sect. 2.3.2), include species like *Agrobacterium tumefaciens* (alphaproteobacteria), *Burkholderia graminis* (betaproteobacteria), *Caulobacter vibrioides* (alphaproteobacteria), *Pseudomonas putida* (gammaproteobacteria), and *Ralstonia metallidurans* (betaproteobacteria) that have been used to screen a soil metagenome (Craig et al. 2010). Also, the alphaproteobacterium *Rhizobium leguminosarum* has been used in metagenome screening for alcohol/aldehyde dehydrogenases (Wexler et al. 2005). Other mesophilic host bacteria utilized in metagenomic screening include *Rhodobacter capsulatus* and *Gluconobacter oxydans* (Liebl et al. 2014), where *R. capsulatus* has been shown to be suitable for expression of membrane proteins, and *G. oxydans* to be tolerant to screening at acidic conditions. Also, the low G + C, Gram-positive bacterium *Bacillus subtilis*, widely used for recombinant enzyme production due to its capability to secrete protein in the medium, has been used in metagenome screening (Biver et al. 2013). Similarly, species of *Burkholderia*, *Sphingomonas*, and *Pseudomonas* (Ekkers et al. 2012; Martinez et al. 2004) have been used, and, by using the bacterial symbiont *Sinorhizobium meliloti* as expression host, a greater diversity of clones was found compared to screening in *E. coli* (Lam et al. 2015). In addition, the gammaproteobacteria *Pseudomonas fluorescens* and *Xanthomonas campestris* (Aakvik et al. 2009) as well as integrase-mediated recombination of libraries in hosts *S. meliloti* and *Agrobacterium tumefaciens* (Heil et al. 2012) have been shown to be applicable for functional metagenome screening.

Even though prokaryotic hosts have been applied successfully in screening of metagenomic DNA libraries with content including eukaryotic DNA (Geng et al. 2012), eukaryotic host systems may be an important area for further development of metagenomic tools and expression hosts. Even though much more prokaryotic vector-host systems have been developed and used through history, there are genetic tools available for yeasts such as *Saccharomyces cerevisiae* (e.g., Drew and Kim 2012) and *Pichia pastoris* (e.g., Daly and Hearn 2005), as well as filamentous fungi, for example, *Aspergillus* (Nevalainen et al. 2005). A mutant strain of *S. cerevisiae*, defective in di-/tripeptide uptake, has been used in a functional screening of a soil metagenome library for the identification of novel oligopeptide transporters (Damon et al. 2011), demonstrating the potential of eukaryotic hosts in functional screening of environmental metagenomes.

2.4 In Vitro Expression Systems for Functional Metagenomics

Cell-free protein synthesis (CFPS) covers the in vitro transcription of coding DNA to mRNA and its subsequent translation into polypeptide and functional protein by using cell extracts. CFPS is a field in rapid development with the potential to make large impact in both protein production and screening for new enzyme functions in

the future. The first CFPS system of *E. coli* was already introduced in 1961, with the main purpose of studying the process of translation (Matthaei and Nirenberg 1961). Since then, a multitude of advanced CFPS systems using extracts of organisms from all three domains of life, including from Bacteria, Archaea, fungi, plants, insects, and mammals (Zemella et al. 2015), has been developed. With their open nature, CFPS systems bypass a number of limitations existing in cellular, in vivo expression systems, as they are highly flexible with respect to the physicochemical environment, the reaction conditions, and the reaction format for gene expression to take place. In addition, they allow incorporation of nonnatural amino acids/cofactors, avoid biological background, and are not constraint by cell viability in response to toxic proteins being produced. In the absence of membranes to be bypassed, almost unlimited use of substrates for screening of gene libraries is enabled, and library sizes that are not restricted by transformation efficiency of expression host cells. This renders CFPS an increasingly recognized alternative option to cell-based expression systems for both protein screening and production (Catherine et al. 2013).

Several key challenges associated with CFPS have recently been successfully addressed and mitigated, such as low productivities, quality and quantity constraints of DNA templates, posttranslational modifications, and clonal separation for genotype-phenotype coupling. Low productivity has been a major issue due to the rapid depletion of the chemical energy carrier ATP and stoichiometric accumulation of phosphate, binding vital magnesium ions. The development of ATP regeneration methods, in particular utilization of the intact glycolytic pathway to produce ATP from glucose by oxidative phosphorylation (Jewett and Swartz 2004; Calhoun and Swartz 2007; Kim and Kim 2009), represented a major breakthrough in achieving larger protein amounts. Moreover, in situ supply of glucose by hydrolysis of polymeric carbohydrates like maltodextrin or starch could be implemented to control the ATP delivery rate (Wang and Zhang 2009). Other metabolic functions in crude cell extracts for CFPS were used to be beneficial, for example, for the provision of cofactors for produced target enzymes (Kwon et al. 2013).

Several studies have suggested solutions to the challenges connected to high template amounts required, as well as high exonucleolytic degradation of linear DNA template in crude cell extracts. In addition to sufficient template preparation by PCR-based methods (Sawasaki et al. 2002; Endo and Sawasaki 2004), the use of isothermal DNA amplification in connection with CFPS (Kumar and Chernaya 2009) was shown to enable high throughput protein synthesis based on very small amounts of template DNA. mRNA stabilization by inclusion of the terminal stem-loop structures and depletion of extracts from RNase E led to greatly improved protein production (Ahn et al. 2005). More relevant for expression library screening, the protection of linear DNA templates and improved protein production was shown by inhibiting the RecBCD nuclease in *E. coli* extracts by addition of bacteriophage Lambda Gam (Sitaraman et al. 2004). This was also shown to be achieved by using extracts of *E. coli* in which the endonuclease I gene *endA* was removed and the *recBCD* operon was replaced by the Lambda recombination system (Michel-Reydellet et al. 2005). Also the tethering of linear DNA ends to microbeads in an agarose matrix led to improved DNA template stability (Lee et al. 2012).

For posttranslational modifications during in vitro synthesis of eukaryotic proteins, for example, for pharmaceutical applications, several eukaryotic CFPS systems have been developed as recently reviewed by Zemella and co-workers ((Zemella et al. 2015) and references therein). This includes systems based on *S. cerevisiae*, the fall armyworm *Spodoptera frugiperda*, rabbit reticulocytes, CHO cells, and different human cell lines. The set of well-documented eukaryotic CFPS systems also includes plant systems from tobacco BY-2 and the widely used cell-free expression system based on wheat germ embryos which represents a high yield system with correct folding of many protein types, including disulfide-rich proteins (Takai et al. 2010).

In vitro compartmentalization (IVC) represents one possible solution to the demand for clonal separation and genotype–phenotype coupling in cell-free screening systems. Being early addressed by the SIMPLEX approach (Rungpragayphan et al. 2003) using diluted single-molecule templates for PCR and subsequent CFPS in a microtiter format, emulsion-based approaches bear the possibility of substantial library sizes. Small aqueous droplets are prepared in a continuous oil phase to isolate templates in individual micro-reactors for isothermal or PCR-based amplification (Courtois et al. 2008) and CFPS. This represents a promising platform for enzyme activity screening against a wide array of substrates using either FACS- or microfluidics-based screening and sorting methods (Kintses et al. 2010).

The insight in biodiversity and the huge metabolic potential in nature provided by the recent revolutions in next-generation sequencing have renewed attention in the potential of CFPS. Consequently, key improvements have been triggered, greatly expanding the applicability of cell-free systems to HT gene expression and even large-scale protein production (Zemella et al. 2015). CFPS and suitable screening systems may form an ideal platform for the functional screening of enzymes using genomic and metagenomic DNA, independent of the limitations of cell-based systems. In a recent example, a cow rumen metagenomic library was screened for glycoside hydrolases using cell-free expression and utilizing the energy-providing effect of glucose in CFPS extracts (Kim et al. 2011). Energy generation in this case started with the polysaccharides cellulose, xylan, amylose, as well as a small amount of glucose. Enzymatic substrate degradation in a feedback loop then led to increased glucose amounts, ultimately leading to an indicator-detectable pH drop due to acid by-products (Kim et al. 2011).

This example shows that optimized CFPS systems in combination with smart assay design represent a powerful option for expression screening for microbial enzymes with high versatility, in particular when combined with platforms for ultra-high throughput analysis and sorting. Further developments in this field will likely include expansion of CFPS systems to additional microbial species, including from extreme environments, as eDNA from extreme environments may fail to be transcribed or translated by *E. coli* extracts (Angelov et al. 2009). Hence, "unconventional" microbial systems for functional expression are demanded (Liebl et al. 2014). Pure component systems and extracts have already been described for extremophiles from both Bacteria and Archaea (Endoh et al. 2007; Zhou et al. 2012), including *Thermus*, *Pyrococcus*, *Sulfolobus*, and *Thermococcus* (Hethke et al. 1996; Tachibana et al. 1996; Ruggero et al. 2006), which might be a valuable resource for future systems.

2.5   Outlook

Metagenomics has proven to be a powerful tool to describe environmental microbial biodiversity and exploit it for metabolic functions of relevance for commercial applications. With the ever-advancing throughput of next-generation sequencing technologies, (meta)genomic DNA sequence databases are filling rapidly, and based on that, our insight into the huge and diverse metabolic potential existing in nature has never been deeper. However, identification of useful functions is ultimately still dependent on experimental proof. Though in silico predictions are constantly improving, the field of Functional Metagenomics will continue to develop as it directly and efficiently links desired function to its determining source code, the eDNA.

E. coli and genetic tools developed for this species have been the first choice in Functional Metagenomics research, both with respect to library construction, recombinant expression, and functional screening. However, it is presently obvious that *E. coli* has some shortcomings, especially in the light of the growing spectrum of ecological niches and greater microbial diversity being accessed and a broader spectrum of metabolic functions and properties aimed to be exploited. Therefore, along with *E. coli*, which itself is still being improved further as a screening host for specific target classes, other microbial model systems, potentially more suitable for screenings for particular functions of interest, have emerged in recent years. These include, for example, thermophilic and psychrophilic systems for respective enzyme discovery and actinobacterial systems for secondary metabolite gene cluster expression and bioactive compound formation.

New and better tools are demanded and continuously developed to increase efficiency at the different steps of the Functional Metagenomics biodiscovery pipeline (Fig. 2.1). In addition to dedicated sampling and efficient DNA extraction procedures from diverse natural environments and developments within metagenomic (small- and large-insert) library cloning technology, several other aspects are in focus. Vector development for heterologous expression in and transfer between multiple host species (broad host-range) as well as optimization of different host species to heterologously express genes for bioactive functions will likely continue to converge. In particular, a higher efficiency in large-insert cloning of eDNA and its shuffling between different expression hosts allowing screening in in different organisms with complementary features and capabilities has proven to generate complementary hits (Liebl et al. 2014). It is therefore still highly desired to improve the functional metagenomic pipeline for metagenome-based bioactive compound discovery by means of new expression and screening platforms. Several different host organisms may be included, and shuffling of metagenomic libraries between these, connected to multiple host screening, is of potentially high value. It can be expected that newly developed expression systems aim to be optimal within screening for specific targeted applications (specific enzyme functions or bioactive compounds) or product properties. The concept of specifically accessing environments providing desired properties (e.g., of an enzyme of choice) and subsequently using

screening hosts that perform optimally at similar conditions can be expected to produce further valuable output in the future. In addition, within this concept, the metabolic optimization of the host species from different phyla or even domains (including Archaea and Eukaryotes) may be pursued. The integration of new host species of phyla other than the *Actinobacteria* may expand the options to access biodiversity for medical compound discovery and thus mitigate the threat of antibiotic resistance, as well as help fighting deadly diseases, including cancer.

System biology understanding, the application of new genome editing tools, and synthetic biology principles will guide new approaches to optimize host strains for heterologous expression of metagenomic genes and formation of new natural products. Optimized Superhosts for bioactivity screening based on different model *Actinobacteria* will enable heterologous expression of biosynthetic gene clusters and compound formation from uncultured bacteria. Well-established thermophilic and psychrophilic host species will be good candidates for further optimization with respect to high- and low-temperature screening. Thereby, optimal hosts should attribute, among others, stable cloning vector maintenance, sensitivity toward relevant antibiotics for selection purposes, and suitable transcription and translation machinery. In addition, they should ensure correct folding, cofactor provision and insertion, relevant precursor supply, as well as counteract toxic effects from product formation (e.g., by product export mechanisms).

In vivo systems for Functional Metagenomics come with their inherent set of challenges, like limitations in achievable library sizes and the spectrum of usable substrates for screening. Consequently, cell-free (in vitro) expression systems have lately emerged as a potential alternative in functional metagenome screening for enzymatic functions (Sect. 2.4). In vitro expression systems still have their own limitations, in particular regarding large-scale production, which, however, is not very relevant for screening and biodiscovery, requiring only small amounts of product. Key challenges are constantly being addressed with new research, and solutions to key bottlenecks have already been found. An expanded spectrum of CFPS species, similar to the diversification of in vivo expression systems, as well as hybrid systems combining beneficial components of different species, can be expected to become available soon. Thus, in combination with ongoing developments of compartmentalization and miniaturization of screening technology, as achievable by, e.g., using advanced microfluidics devices, in vitro systems may become a potential future alternative to in vivo systems in Functional Metagenomics.

References

Aakvik T, Degnes KF, Dahlsrud R, Schmidt F, Dam R, Yu L, Volker U, Ellingsen TE, Valla S (2009) A plasmid RK2-based broad-host-range cloning vector useful for transfer of metagenomic libraries to a variety of bacterial species. FEMS Microbiol Lett 296(2):149–158. doi:10.1111/j.1574-6968.2009.01639.x

Ahn JH, Chu HS, Kim TW, Oh IS, Choi CY, Hahn GH, Park CG, Kim DM (2005) Cell-free synthesis of recombinant proteins from PCR-amplified genes at a comparable productivity to that

of plasmid-based reactions. Biochem Biophys Res Commun 338(3):1346–1352. doi:10.1016/j. bbrc.2005.10.094

Alam MT, Merlo ME, Consortium S, Hodgson DA, Wellington EM, Takano E, Breitling R (2010) Metabolic modeling and analysis of the metabolic switch in Streptomyces coelicolor. BMC Genomics 11:202. doi:10.1186/1471-2164-11-202

Angelov A, Mientus M, Liebl S, Liebl W (2009) A two-host fosmid system for functional screening of (meta)genomic libraries from extreme thermophiles. Syst Appl Microbiol 32(3):177–185. doi:10.1016/j.syapm.2008.01.003

Baltz RH (2008) Renaissance in antibacterial discovery from actinomycetes. Curr Opin Pharmacol 8(5):557–563. doi:10.1016/j.coph.2008.04.008

Baltz RH (2010) Streptomyces and Saccharopolyspora hosts for heterologous expression of secondary metabolite gene clusters. J Ind Microbiol Biotechnol 37(8):759–772. doi:10.1007/s10295-010-0730-9

Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Klenk HP, Clement C, Ouhdouch Y, van Wezel GP (2016) Taxonomy, physiology, and natural products of actinobacteria. Microbiol Mol Biol Rev 80(1):1–43. doi:10.1128/MMBR.00019-15

Battke F, Herbig A, Wentzel A, Jakobsen ØM, Bonin M, Hodgson DA, Wohlleben W, Ellingsen TE, Consortium S, Nieselt K (2011) A technical platform for generating reproducible expression data from Streptomyces coelicolor batch cultivations. In: Arabnia HR, Tran QN (eds) Software tools and algorithms for biological systems. Springer, New York, pp 3–15. doi:10.1007/978-1-4419-7046-6_1

Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, Hopwood DA (2002) Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature 417(6885):141–147. doi:10.1038/417141a

Bibb MJ (2005) Regulation of secondary metabolism in streptomycetes. Curr Opin Microbiol 8(2):208–215. doi:10.1016/j.mib.2005.02.016

Biver S, Steels S, Portetelle D, Vandenbol M (2013) Bacillus subtilis as a tool for screening soil metagenomic libraries for antimicrobial activities. J Microbiol Biotechnol 23(6):850–855. doi:10.4014/jmb.1212.12008

Bjerga GEK, Williamson AK (2015) Cold shock induction of recombinant Arctic environmental genes. BMC Biotechnol 15(1):1–12. doi:10.1186/s12896-015-0185-1

Brady SF (2007) Construction of soil environmental DNA cosmid libraries and screening for clones that produce biologically active small molecules. Nat Protoc 2(5):1297–1305. doi:10.1038/nprot.2007.195

Brouns SJ, Wu H, Akerboom J, Turnbull AP, de Vos WM, van der Oost J (2005) Engineering a selectable marker for hyperthermophiles. J Biol Chem 280(12):11422–11431. doi:10.1074/jbc.M413623200

Calhoun KA, Swartz JR (2007) Energy systems for ATP regeneration in cell-free protein synthesis reactions. Methods Mol Biol 375:3–17. doi:10.1007/978-1-59745-388-2_1

Catherine C, Lee KH, Oh SJ, Kim DM (2013) Cell-free platforms for flexible expression and screening of enzymes. Biotechnol Adv 31(6):797–803. doi:10.1016/j.biotechadv.2013.04.009

Cava F, Laptenko O, Borukhov S, Chahlafi Z, Blas-Galindo E, Gomez-Puertas P, Berenguer J (2007) Control of the respiratory metabolism of Thermus thermophilus by the nitrate respiration conjugative element NCE. Mol Microbiol 64(3):630–646. doi:10.1111/j.1365-2958.2007.05687.x

Cava F, Hidalgo A, Berenguer J (2009) Thermus thermophilus as biological model. Extremophiles 13(2):213–231. doi:10.1007/s00792-009-0226-6

Cavicchioli R, Charlton T, Ertan H, Mohd Omar S, Siddiqui KS, Williams TJ (2011) Biotechnological uses of enzymes from psychrophiles. Microb Biotechnol 4(4):449–460. doi:10.1111/j.1751-7915.2011.00258.x

Cheng J, Pinnell L, Engel K, Neufeld JD, Charles TC (2014) Versatile broad-host-range cosmids for construction of high quality metagenomic libraries. J Microbiol Methods 99:27–34. doi:10.1016/j.mimet.2014.01.015

Cheng J, Romantsov T, Engel K, Doxey AC, Rose DR, Neufeld JD, Charles TC (2017) Functional metagenomics reveals novel beta-galactosidases not predictable from gene sequences. PLOS ONE 12(3):e0172545. doi:10.1371/journal.pone.0172545

Chu SF, Shu HY, Lin LC, Chen MY, Tsay SS, Lin GH (2006) Characterization of a rolling-circle replication plasmid from Thermus aquaticus NTU103. Plasmid 56(1):46–52. doi:10.1016/j.plasmid.2006.01.005

Cobb RE, Wang YJ, Zhao HM (2015) High-efficiency multiplex genome editing of streptomyces species using an engineered crispr/cas system. ACS Synth Biol 4(6):723–728. doi:10.1021/sb500351f

Courtois F, Olguin LF, Whyte G, Bratton D, Huck WT, Abell C, Hollfelder F (2008) An integrated device for monitoring time-dependent in vitro expression from single genes in picolitre droplets. Chembiochem 9(3):439–446. doi:10.1002/cbic.200700536

Craig JW, Chang FY, Kim JH, Obiajulu SC, Brady SF (2010) Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental dna libraries in diverse proteobacteria. Appl Environ Microbiol 76(5):1633–1641. doi:10.1128/Aem.02169-09

Daly R, Hearn MT (2005) Expression of heterologous proteins in Pichia pastoris: a useful experimental tool in protein engineering and production. J Mol Recognit 18(2):119–138. doi:10.1002/jmr.687

Damon C, Vallon L, Zimmermann S, Haider MZ, Galeote V, Dequin S, Luis P, Fraissinet-Tachet L, Marmeisse R (2011) A novel fungal family of oligopeptide transporters identified by functional metatranscriptomics of soil eukaryotes. ISME J 5(12):1871–1880. doi:10.1038/ismej.2011.67

Danhorn T, Young CR, DeLong EF (2012) Comparison of large-insert, small-insert and pyrosequencing libraries for metagenomic analysis. ISME J 6:2056–2066. doi:10.1038/ismej.2012.35

de Grado M, Castan P, Berenguer J (1999) A high-transformation-efficiency cloning vector for Thermus thermophilus. Plasmid 42(3):241–245. doi:10.1006/plas.1999.1427

Drew D, Kim H (2012) Preparation of saccharomyces cerevisiae expression plasmids. Methods Mol Biol 866:41–46. doi:10.1007/978-1-61779-770-5_4

Du D, Zhu Y, Wei JH, Tian YQ, Niu G, Tan HR (2013) Improvement of gougerotin and nikkomycin production by engineering their biosynthetic gene clusters. Appl Microbiol Biotechnol 97(14):6383–6396. doi:10.1007/s00253-013-4836-7

Ekkers DM, Cretoiu MS, Kielak AM, van Elsas JD (2012) The great screen anomaly-a new frontier in product discovery through functional metagenomics. Appl Microbiol Biotechnol 93(3):1005–1020. doi:10.1007/s00253-011-3804-3

Endo Y, Sawasaki T (2004) High-throughput, genome-scale protein production method based on the wheat germ cell-free expression system. J Struct Funct Genom 5(1–2):45–57. doi:10.1023/B:JSFG.0000029208.83739.49

Endoh T, Kanai T, Imanaka T (2007) A highly productive system for cell-free protein synthesis using a lysate of the hyperthermophilic archaeon, Thermococcus kodakaraensis. Appl Microbiol Biotechnol 74(5):1153–1161. doi:10.1007/s00253-006-0753-3

Feitelson JS, Malpartida F, Hopwood DA (1985) Genetic and biochemical characterization of the red gene cluster of Streptomyces coelicolor A3(2). J Gen Microbiol 131(9):2431–2441. doi:10.1099/00221287-131-9-2431

Feller G (2013) Psychrophilic enzymes: from folding to function and biotechnology. Scientifica (Cairo) 2013:28. doi:10.1155/2013/512840

Fernandez-Arrojo L, Guazzaroni ME, Lopez-Cortes N, Beloqui A, Ferrer M (2010) Metagenomic era for biocatalyst identification. Curr Opin Biotechnol 21(6):725–733. doi:10.1016/j.copbio.2010.09.006

Ferrer M, Chernikova TN, Yakimov MM, Golyshin PN, Timmis KN (2003) Chaperonins govern growth of Escherichia coli at low temperatures. Nat Biotechnol 21(11):1266–1267. doi:10.1038/nbt1103-1266b

Ferrer M, Beloqui A, Golyshina OV, Plou FJ, Neef A, Chernikova TN, Fernandez-Arrojo L, Ghazi I, Ballesteros A, Elborough K, Timmis KN, Golyshin PN (2007) Biochemical and structural features of a novel cyclodextrinase from cow rumen metagenome. Biotechnol J 2(2):207–213. doi:10.1002/biot.200600183

Ferrer M, Martinez-Martinez M, Bargiela R, Streit WR, Golyshina OV, Golyshin PN (2016) Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. Microb Biotechnol 9(1):22–34. doi:10.1111/1751-7915.12309

Foulston LC, Bibb MJ (2010) Microbisporicin gene cluster reveals unusual features of lantibiotic biosynthesis in actinomycetes. Proc Natl Acad Sci U S A 107(30):13461–13466. doi:10.1073/pnas.1008285107

Gabor EM, Alkema WB, Janssen DB (2004) Quantifying the accessibility of the metagenome by random expression cloning techniques. Environ Microbiol 6(9):879–886. doi:10.1111/j.1462-2920.2004.00640.x

Gaida SM, Sandoval NR, Nicolaou SA, Chen Y, Venkataramanan KP, Papoutsakis ET (2015) Expression of heterologous sigma factors enables functional screening of metagenomic and heterologous genomic libraries. Nat Commun 6:7045. doi:10.1038/ncomms8045

Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadan AH, Moineau S (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468(7320):67–71. doi:10.1038/nature09523

Geng A, Zou G, Yan X, Wang Q, Zhang J, Liu F, Zhu B, Zhou Z (2012) Expression and characterization of a novel metagenome-derived cellulase Exo2b and its application to improve cellulase activity in Trichoderma reesei. Appl Microbiol Biotechnol 96(4):951–962. doi:10.1007/s00253-012-3873-y

Godiska R, Mead D, Dhodda V, Wu C, Hochstein R, Karsi A, Usdin K, Entezam A, Ravin N (2010) Linear plasmid vector for cloning of repetitive or unstable sequences in Escherichia coli. Nucleic Acids Res 38(6):e88. doi:10.1093/nar/gkp1181

Gomez-Escribano JP, Bibb MJ (2011) Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters. Microb Biotechnol 4(2):207–215. doi:10.1111/j.1751-7915.2010.00219.x

Gomez-Escribano JP, Bibb MJ (2012) Streptomyces coelicolor as an expression host for heterologous gene clusters. Methods Enzymol 517:279–300. doi:10.1016/B978-0-12-404634-4.00014-0

Gomez-Escribano JP, Bibb MJ (2014) Heterologous expression of natural product biosynthetic gene clusters in Streptomyces coelicolor: from genome mining to manipulation of biosynthetic pathways. J Ind Microbiol Biotechnol 41(2):425–431. doi:10.1007/s10295-013-1348-5

Gomez-Escribano JP, Song LJ, Fox DJ, Yeo V, Bibb MJ, Challis GL (2012) Structure and biosynthesis of the unusual polyketide alkaloid coelimycin P1, a metabolic product of the cpk gene cluster of Streptomyces coelicolor M145. Chem Sci 3(9):2716–2720. doi:10.1039/c2sc20410j

Gruber S, Schwab H, Koefinger P (2015) Versatile plasmid-based expression systems for gram-negative bacteria—general essentials exemplified with the bacterium ralstonia eutropha H16. New Biotechnol 32(6):552–558. doi:10.1016/j.nbt.2015.03.015

Guazzaroni M-E, Silva-Rocha R, Ward RJ (2015) Synthetic biology approaches to improve biocatalyst identification in metagenomic library screening. Microb Biotechnol 8(1):52–64. doi:10.1111/1751-7915.12146

Gust B, Chandra G, Jakimowicz D, Tian YQ, Bruton CJ, Chater KF (2004) Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces. Adv Appl Microbiol 54:107–128. doi:10.1016/S0065-2164(04)54004-2

Hain T, Otten S, von Both U, Chatterjee SS, Technow U, Billion A, Ghai R, Mohamed W, Domann E, Chakraborty T (2008) Novel bacterial artificial chromosome vector pUvBBAC for use in studies of the functional genomics of Listeria spp. Appl Environ Microbiol 74(6):1892–1901. doi:10.1128/AEM.00415-07

Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol 5(10):R245–R249. doi:10.1016/S1074-5521(98)90108-9

Heil JR, Cheng J, Charles TC (2012) Site-specific bacterial chromosome engineering: PhiC31 integrase mediated cassette exchange (IMCE). J Vis Exp 61. doi:10.3791/3698

Hethke C, Geerling AC, Hausner W, de Vos WM, Thomm M (1996) A cell-free transcription system for the hyperthermophilic archaeon Pyrococcus furiosus. Nucleic Acids Res 24(12):2369–2376. doi:10.1093/nar/24.12.2369

Hidaka Y, Hasegawa M, Nakahara T, Hoshino T (1994) The entire population of Thermus thermophilus cells is always competent at any growth phase. Biosci Biotechnol Biochem 58(7):1338–1339. doi:10.1271/bbb.58.1338

Hidalgo A, Betancor L, Moreno R, Zafra O, Cava F, Fernandez-Lafuente R, Guisan JM, Berenguer J (2004) Thermus thermophilus as a cell factory for the production of a thermophilic Mn-dependent catalase which fails to be synthesized in an active form in Escherichia coli. Appl Environ Microbiol 70(7):3839–3844. doi:10.1128/AEM.70.7.3839-3844.2004

Holz C, Prinz B, Bolotina N, Sievert V, Büssow K, Simon B, Stahl U, Lang C (2003) Establishing the yeast Saccharomyces cerevisiae as a system for expression of human proteins on a proteome-scale. J Struct Funct Genom 4(2–3):97–108. doi:10.1023/A:1026226429429

Hopwood DA (2007) Streptomyces in nature and medicine: the antibiotic makers. Oxford University Press, New York

Hopwood DA, Wright HM (1983) CDA is a new chromosomally determined antibiotic from Streptomyces coelicolor A3(2). J Gen Microbiol 129:3575–3579. doi:10.1099/00221287-129-12-3575

Hu H, Zhang Q, Ochi K (2002) Activation of antibiotic biosynthesis by specified mutations in the rpoB gene (encoding the RNA polymerase beta subunit) of Streptomyces lividans. J Bacteriol 184(14):3984–3991. doi:10.1128/JB.184.14.3984-3991.2002

Hu XP, Heath C, Taylor MP, Tuffin M, Cowan DA (2012) A novel extremely alkaliphilic and cold-active esterase from Antarctic desert soil. Extremophiles 16:79–86. doi:10.1007/s00792-011-0407-y

Huang H, Zheng GS, Jiang WH, Hu H, Lu YH (2015) One-step high-efficiency CRISPR/Cas9-mediated genome editing in Streptomyces. Acta Biochim Biophys Sin 47(4):231–243. doi:10.1093/abbs/gmv007

Huo L, Rachid S, Stadler M, Wenzel SC, Muller R (2012) Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. Chem Biol 19(10):1278–1287. doi:10.1016/j.chembiol.2012.08.013

Iqbal M, Mast Y, Amin R, Hodgson DA, Consortium S, Wohlleben W, Burroughs NJ (2012) Extracting regulator activity profiles by integration of de novo motifs and expression data: characterizing key regulators of nutrient depletion responses in Streptomyces coelicolor. Nucleic Acids Res 40(12):5227–5239. doi:10.1093/nar/gks205

Jewett MC, Swartz JR (2004) Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. Biotechnol Bioeng 86(1):19–26. doi:10.1002/bit.20026

Jones AC, Gust B, Kulik A, Heide L, Buttner MJ, Bibb MJ (2013) Phage p1-derived artificial chromosomes facilitate heterologous expression of the FK506 gene cluster. PLoS One 8(7):e69319. doi:10.1371/journal.pone.0069319

Kakirde KS, Parsley LC, Liles MR (2010) Size does matter: application-driven approaches for soil metagenomics. Soil Biol Biochem 42(11):1911–1923. doi:10.1016/j.soilbio.2010.07.021

Kennedy J, Marchesi JR, Dobson AD (2008) Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments. Microb Cell Factories 7:27. doi:10.1186/1475-2859-7-27

Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical Streptomyces genetics. John Innes Foundation, Norwich

Kim HC, Kim DM (2009) Methods for energizing cell-free protein synthesis. J Biosci Bioeng 108(1):1–4. doi:10.1016/j.jbiosc.2009.02.007

Kim UJ, Shizuya H, de Jong PJ, Birren B, Simon MI (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. Nucleic Acids Res 20(5):1083–1085. doi:10.1093/nar/20.5.1083

Kim HC, Kim TW, Kim DM (2011) Prolonged production of proteins in a cell-free protein synthesis system using polymeric carbohydrates as an energy source. Process Biochem 2011(46):1366–1369. doi:10.1016/j.procbio.2011.03.008

King RW, Bauer JD, Brady SF (2009) An environmental DNA-derived type II polyketide biosynthetic pathway encodes the biosynthesis of the pentacyclic polyketide erdacin. Angew Chem 48:6257–6261. doi:10.1002/anie.200901209

Kintses B, van Vliet LD, Devenish SR, Hollfelder F (2010) Microfluidic droplets: new integrated workflows for biological experiments. Curr Opin Chem Biol 14(5):548–555. doi:10.1016/j.cbpa.2010.08.013

Komatsu M, Komatsu K, Koiwai H, Yamada Y, Kozone I, Izumikawa M, Hashimoto J, Takagi M, Omura S, Shin-ya K, Cane DE, Ikeda H (2013) Engineered Streptomyces avermitilis host for heterologous expression of biosynthetic gene cluster for secondary metabolites. ACS Synth Biol 2(7):384–396. doi:10.1021/sb3001003

Kotlar HK, Lewin A, Johansen J, Throne-Holst M, Haverkamp T, Markussen S, Winnberg A, Ringrose P, Aakvik T, Ryeng E, Jakobsen K, Drabløs F, Valla S (2011) High coverage sequencing of DNA from microorganisms living in an oil reservoir 2.5 kilometres subsurface. Environ Microbiol Rep 3(6):674–681. doi:10.1111/j.1758-2229.2011.00279.x

Koyama Y, Hoshino T, Tomizuka N, Furukawa K (1986) Genetic transformation of the extreme thermophile Thermus thermophilus and of other Thermus spp. J Bacteriol 166(1):338–340. doi: 0021-9193/86/040338-03$02.00/0

Koyama Y, Okamoto S, Furukawa K (1990) Cloning of alpha- and beta-galactosidase genes from an extreme thermophile, Thermus strain T2, and their expression in Thermus thermophilus HB27. Appl Environ Microbiol 56(7):2251–2254. doi: 0099-2240/90/072251-04$02.00/0

Kumar G, Chernaya G (2009) Cell-free protein synthesis using multiply-primed rolling circle amplification products. BioTechniques 47:637–639. doi:10.2144/000113171

Kwon YC, Oh IS, Lee N, Lee KH, Yoon YJ, Lee EY, Kim BG, Kim DM (2013) Integrating cell-free biosyntheses of heme prosthetic group and apoenzyme for the synthesis of functional P450 monooxygenase. Biotechnol Bioeng 110(4):1193–1200. doi:10.1002/bit.24785

Lam KN, Charles TC (2015) Strong spurious transcription likely a cause of DNA insert bias in typical metagenomic clone libraries. Microbiome 3:22. doi:10.1101/013763

Lam KN, Cheng J, Engel K, Neufeld JD, Charles TC (2015) Current and future resources for functional metagenomics. Front Microbiol 6:1196. doi:10.3389/fmicb.2015.01196

Lammle K, Zipper H, Breuer M, Hauer B, Buta C, Brunner H, Rupp S (2007) Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. J Biotechnol 127(4):575–592. doi:10.1016/j.jbiotec.2006.07.036

Lasa I, de Grado M, de Pedro MA, Berenguer J (1992) Development of Thermus-Escherichia shuttle vectors and their use for expression of the Clostridium thermocellum celA gene in Thermus thermophilus. J Bacteriol 174(20):6424–6431. doi: 0021-9193/92/206424-08$02.00/0

Lee KH, Lee KY, Byun JY, Kim BG, Kim DM (2012) On-bead expression of recombinant proteins in an agarose gel matrix coated on a glass slide. Lab Chip 12(9):1605–1610. doi:10.1039/c2lc21239k

Leis B, Angelov A, Mientus M, Li HJ, Pham VTT, Lauinger B, Bongen P, Pietruszka J, Goncalves LG, Santos H, Liebl W (2015a) Identification of novel esterase-active enzymes from hot environments by use of the host bacterium Thermus thermophilus. Front Microbiol 6:275. doi:10.3389/frricb.2015.00275

Leis B, Heinze S, Angelov A, Pham VT, Thürmer A, Jebbar M, Golyshin PN, Streit WR, Daniel R, Liebl W (2015b) Functional screening of hydrolytic activities reveals an extremely thermostable cellulase from a deep-sea archaeon. Front Bioeng Biotechnol 3:95. doi:10.3389/fbioe.2015.00095

Lewin A, Wentzel A, Valla S (2013) Metagenomics of microbial life in extreme temperature environments. Curr Opin Biotechnol 24(3):516–525. doi:10.1016/j.copbio.2012.10.012

Leza A, Palmeros B, García JO, Galindo E, Soberón-Chávez G (1996) Xanthomonas campestris as a host for the production of recombinant Pseudomonas aeruginosa lipase. J Ind Microbiol 16(1):22–28. doi:10.1007/BF01569917

Li C, Zhang F, Kelly WL (2011) Heterologous production of thiostrepton A and biosynthetic engineering of thiostrepton analogs. Mol Biosyst 7(1):82–90. doi:10.1039/c0mb00129e

Liao H, McKenzie T, Hageman R (1986) Isolation of a thermostable enzyme variant by cloning and selection in a thermophile. Proc Natl Acad Sci U S A 83(3):576–580. doi: pnas00307-0057

Liebl W, Angelov A, Juergensen J, Chow J, Loeschcke A, Drepper T, Classen T, Pietruszka J, Ehrenreich A, Streit WR, Jaeger KE (2014) Alternative hosts for functional (meta) genome analysis. Appl Microbiol Biotechnol 98(19):8099–8109. doi:10.1007/s00253-014-5961-7

Liles MR, Williamson LL, Rodbumrer J, Torsvik V, Goodman RM, Handelsman J (2008) Recovery, purification, and cloning of high-molecular-weight DNA from soil microorganisms. Appl Environ Microbiol 74(10):3302–3305. doi:10.1128/AEM.02630-07

Loeschcke A, Thies S (2015) Pseudomonas putida—a versatile host for the production of natural products. Appl Microbiol Biotechnol 99(15):6197–6214. doi:10.1007/s00253-015-6745-4

Maldonado LA, Stach JE, Pathom-aree W, Ward AC, Bull AT, Goodfellow M (2005) Diversity of cultivable actinobacteria in geographically widespread marine sediments. Antonie Van Leeuwenhoek 87(1):11–18. doi:10.1007/s10482-004-6525-0

Marcone GL, Foulston L, Binda E, Marinelli F, Bibb M, Beltrametti F (2010) Methods for the genetic manipulation of Nonomuraea sp. ATCC 39727. J Ind Microbiol Biotechnol 37(10):1097–1103. doi:10.1007/s10295-010-0807-5

Martin JF, Santos-Beneit F, Rodriguez-Garcia A, Sola-Landa A, Smith MC, Ellingsen TE, Nieselt K, Burroughs NJ, Wellington EM (2012) Transcriptomic studies of phosphate control of primary and secondary metabolism in Streptomyces coelicolor. Appl Microbiol Biotechnol 95(1):61–75. doi:10.1007/s00253-012-4129-6

Martinez A, Kolvek SJ, Yip CL, Hopke J, Brown KA, MacNeil IA, Osburne MS (2004) 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 70(4):2452–2463. doi:10.1128/AEM.70.4.2452-2463.2004

Martinez A, Kolvek SJ, Hopke J, Yip CL, Osburne MS (2005) Environmental DNA fragment conferring early and increased sporulation and antibiotic production in Streptomyces species. Appl Environ Microbiol 71(3):1638–1641. doi:10.1128/AEM.71.3.1638-1641.2005

Mather MW, Fee JA (1992) Development of plasmid cloning vectors for Thermus thermophilus HB8: expression of a heterologous, plasmid-borne kanamycin nucleotidyltransferase gene. Appl Environ Microbiol 58(1):421–425. doi: 0099-2240/92/010421-05$02.0O/O

Matsumura M, Aiba S (1985) Screening for thermostable mutant of kanamycin nucleotidyltransferase by the use of a transformation system for a thermophile, Bacillus stearothermophilus. J Biol Chem 260(28):15298–15303. doi: 260/28/15298

Matthaei JH, Nirenberg MW (1961) The dependence of cell-free protein synthesis in *E. coli* upon RNA prepared from ribosomes. Biochem Biophys Res Commun 28(4):404–408. doi: pnas00214-0066

McMahon MD, Guan C, Handelsman J, Thomas MG (2012) Metagenomic analysis of Streptomyces lividans reveals host-dependent functional expression. Appl Environ Microbiol 78(10):3622–3629. doi:10.1128/AEM.00044-12

Medema MH, Breitling R, Bovenberg R, Takano E (2011) Exploiting plug-and-play synthetic biology for drug discovery and production in microorganisms. Nat Rev Microbiol 9(2):131–137. doi:10.1038/nrmicro2478

Mergeay M, Monchy S, Vallaeys T, Auquier V, Benotmane A, Bertin P, Taghavi S, Dunn J, van der Lelie D, Wattiez R (2003) Ralstonia metallidurans, a bacterium specifically adapted to toxic metals: towards a catalogue of metal-responsive genes. FEMS Microbiol Rev 27(2–3):385–410. doi:10.1016/s0168-6445(03)00045-7

Michel-Reydellet N, Woodrow K, Swartz J (2005) Increasing PCR fragment stability and protein yields in a cell-free system with genetically modified Escherichia coli extracts. J Mol Microbiol Biotechnol 9(1):26–34. doi:10.1159/000088143

Miteva V, Lantz S, Brenchley J (2008) Characterization of a cryptic plasmid from a Greenland ice core Arthrobacter isolate and construction of a shuttle vector that replicates in psychrophilic high G+C Gram-positive recipients. Extremophiles 12(3):441–449. doi:10.1007/s00792-008-0149-7

Miyake R, Kawamoto J, Wei YL, Kitagawa M, Kato I, Kurihara T, Esaki N (2007) Construction of a low-temperature protein expression system using a cold-adapted bacterium, Shewanella sp. strain Ac10, as the host. Appl Environ Microbiol 73(15):4849–4856. doi:10.1128/AEM.00824-07

Moreno R, Haro A, Castellanos A, Berenguer J (2005) High-level overproduction of His-tagged Tth DNA polymerase in Thermus thermophilus. Appl Environ Microbiol 71(1):591–593. doi:10.1128/AEM.71.1.591-593.2005

Mullany P (2014) Functional metagenomics for the investigation of antibiotic resistance. Virulence 5(3):443–447. doi:10.4161/viru.28196

Murai N (2013) Review: plant binary vectors of Ti plasmid in agrobacterium tumefaciens with a broad host-range replicon of pRK2, pRi, pSa or pVS1. AJPS 4:932–939. doi:10.4236/ajps.2013.44115

Nakamura A, Takakura Y, Kobayashi H, Hoshino T (2005) In vivo directed evolution for thermo-stabilization of Escherichia coli hygromycin B phosphotransferase and the use of the gene as a selection marker in the host-vector system of Thermus thermophilus. J Biosci Bioeng 100(2):158–163. doi:10.1263/jbb.100.158

Nett M, Ikeda H, Moore BS (2009) Genomic basis for natural product biosynthetic diversity in the actinomycetes. Nat Prod Rep 26(11):1362–1384. doi:10.1039/b817069j

Nevalainen KM, Te'o VS, Bergquist PL (2005) Heterologous protein expression in filamentous fungi. Trends Biotechnol 23(9):468–474. doi:10.1016/j.tibtech.2005.06.002

Nieselt K, Battke F, Herbig A, Bruheim P, Wentzel A, Jakobsen OM, Sletta H, Alam MT, Merlo ME, Moore J, Omara WA, Morrissey ER, Juarez-Hermosillo MA, Rodriguez-Garcia A, Nentwich M, Thomas L, Iqbal M, Legaie R, Gaze WH, Challis GL, Jansen RC, Dijkhuizen L, Rand DA, Wild DL, Bonin M, Reuther J, Wohlleben W, Smith MC, Burroughs NJ, Martin JF, Hodgson DA, Takano E, Breitling R, Ellingsen TE, Wellington EM (2010) The dynamic architecture of the metabolic switch in Streptomyces coelicolor. BMC Genomics 11:10. doi:10.1186/1471-2164-11-10

Novakova J, Farkasovsky M (2013) Bioprospecting microbial metagenome for natural products. Biologia 68(6):1079–1086. doi:10.2478/s11756-013-0246-7

Ochi K, Hosaka T (2013) New strategies for drug discovery: activation of silent or weakly expressed microbial gene clusters. Appl Microbiol Biotechnol 97(1):87–98. doi:10.1007/s00253-012-4551-9

Ochi K, Tanaka Y, Tojo S (2014) Activating the expression of bacterial cryptic genes by rpoB mutations in RNA polymerase or by rare earth elements. J Ind Microbiol Biotechnol 41(2):403–414. doi:10.1007/s10295-013-1349-4

Ohnishi Y, Ishikawa J, Hara H, Suzuki H, Ikenoya M, Ikeda H, Yamashita A, Hattori M, Horinouchi S (2008) Genome sequence of the streptomycin-producing microorganism Streptomyces griseus IFO 13350. J Bacteriol 190(11):4050–4060. doi:10.1128/JB.00204-08

Okamoto-Hosoya Y, Sato TA, Ochi K (2000) Resistance to paromomycin is conferred by rpsL mutations, accompanied by an enhanced antibiotic production in Streptomyces coelicolor A3(2). J Antibiot 53(12):1424–1427. doi:10.7164/antibiotics.53.1424

Oliynyk M, Samborskyy M, Lester JB, Mironenko T, Scott N, Dickens S, Haydock SF, Leadlay PF (2007) Complete genome sequence of the erythromycin-producing bacterium Saccharopolyspora erythraea NRRL23338. Nat Biotechnol 25(4):447–453. doi:10.1038/nbt1297

Parachin NS, Gorwa-Grauslund MF (2011) Isolation of xylose isomerases by sequence- and function-based screening from a soil metagenomic library. Biotechnol Biofuels 4:9. doi:10.1186/1754-6834-4-9

Parks RJ, Graham FL (1997) A helper-dependent system for adenovirus vector production helps define a lower limit for efficient dna packaging. J Virol 71(4):3293–3298. doi: 0022-538X/97/$04.0010

Pel J, Broemeling D, Mai L, Poon HL, Tropini G, Warren RL, Holt RA, Marziali A (2009) Nonlinear electrophoretic response yields a unique parameter for separation of biomolecules. Proc Natl Acad Sci U S A 106(35):14796–14801. doi:10.1073/pnas.0907402106

Potrykus K, Cashel M (2008) (p)ppGpp: still magical? Annu Rev Microbiol 62:35–51. doi:10.1146/annurev.micro.62.081307.162903

Ramirez-Arcos S, Fernandez-Herrero LA, Marin I, Berenguer J (1998) Anaerobic growth, a property horizontally transferred by an Hfr-like mechanism among extreme thermophiles. J Bacteriol 180(12):3137–3143. doi: 0021-9193/98/$04.0010

Retallack D, Schneider JC, Chew L, Ramseier T, Allen J, Patkar A, Squires C, Talbot H, Mitchell J (2006) Pseudomonas fluorescens—a robust expression platform for pharmaceutical protein production. Microb Cell Factories 5(1):1–1. doi:10.1186/1475-2859-5-s1-s28

Riesenfeld CS, Goodman RM, Handelsman J (2004) Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. Environ Microbiol 6(9):981–989. doi:10.1111/j.1462-2920.2004.00664.x

Rigali S, Nothaft H, Noens EE, Schlicht M, Colson S, Muller M, Joris B, Koerten HK, Hopwood DA, Titgemeyer F, van Wezel GP (2006) The sugar phosphotransferase system of Streptomyces coelicolor is regulated by the GntR-family regulator DasR and links N-acetylglucosamine metabolism to the control of development. Mol Microbiol 61(5):1237–1251. doi:10.1111/j.1365-2958.2006.05319.x

Rigali S, Titgemeyer F, Barends S, Mulder S, Thomae AW, Hopwood DA, van Wezel GP (2008) Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by Streptomyces. EMBO Rep 9(7):670–675. doi:10.1038/embor.2008.83

Rodrigue S, Malmstrom RR, Berlin AM, Birren BW, Henn MR, Chisholm SW (2009) Whole genome amplification and de novo assembly of single bacterial cells. PLoS One 4(9):e6864. doi:10.1371/journal.pone.0006864

Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne MS, Clardy J, Handelsman J, Goodman RM (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl Environ Microbiol 66(6):2541–2547. doi:10.1128/AEM.66.6.2541-2547.2000

Ruan L, Xu X (2007) Sequence analysis and characterizations of two novel plasmids isolated from Thermus sp. 4C. Plasmid 58(1):84–87. doi:10.1016/j.plasmid.2007.04.001

Rudd BA, Hopwood DA (1979) Genetics of actinorhodin biosynthesis by Streptomyces coelicolor A3(2). J Gen Microbiol 114(1):35–43. doi:10.1099/00221287-114-1-35

Ruggero D, Creti R, Londei P (2006) In vitro translation of archaeal natural mRNAs at high temperature. FEMS Micobiol Lett 107(1):89–94. doi:10.1111/j.1574-6968.1993.tb06009.x

Rungpragayphan S, Nakano H, Yamane T (2003) PCR-linked in vitro expression: a novel system for high-throughput construction and screening of protein libraries. FEBS Lett 540(1–3):147–150. doi:10.1016/S0014-5793(03)00251-5

Rutledge PJ, Challis GL (2015) Discovery of microbial natural products by activation of silent biosynthetic gene clusters. Nat Rev Microbiol 13(8):509–523. doi:10.1038/nrmicro3496

Sabree ZL, Rondon MR, Handelsman J (2009) Metagenomics. In: Schaechter M (ed) Encyclopedia of microbiology. Elsevier, Amsterdam, pp. 622–632. doi:10.1016/B978-012373944-5.00034-1

Sawasaki T, Hasegawa Y, Tsuchimochi M, Kamura N, Ogasawara T, Kuroita T, Endo Y (2002) A bilayer cell-free protein synthesis system for high-throughput screening of gene products. FEBS Lett 514(1):102–105. doi:10.1016/S0014-5793(02)02329-3

Schwarzenlander C, Averhoff B (2006) Characterization of DNA transport in the thermophilic bacterium Thermus thermophilus HB27. FEBS J 273(18):4210–4218. doi:10.1111/j.1742-4658.2006.05416.x

Sherwood EJ, Bibb MJ (2013) The antibiotic planosporicin coordinates its own production in the actinomycete Planomonospora alba. Proc Natl Acad Sci U S A 110(27):E2500–E2509. doi:10.1073/pnas.1305392110

Shima J, Hesketh A, Okamoto S, Kawamoto S, Ochi K (1996) Induction of actinorhodin production by rpsL (encoding ribosomal protein S12) mutations that confer streptomycin resistance in Streptomyces lividans and Streptomyces coelicolor A3(2). J Bacteriol 178(24):7276–7284. doi: 0021-9193/96/$04.0010

Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. Proc Natl Acad Sci U S A 89(18):8794–8797. doi: pnas01092-0395

Sitaraman K, Esposito D, Klarmann G, Le Grice SF, Hartley JL, Chatterjee DK (2004) A novel cell-free protein synthesis system. J Biotechnol 110(3):257–263. doi:10.1016/j.jbiotec.2004.02.014

Smanski MJ, Casper J, Peterson RM, Yu Z, Rajski SR, Shen B (2012) Expression of the platencin biosynthetic gene cluster in heterologous hosts yielding new platencin congeners. J Nat Prod 75(12):2158–2167. doi:10.1021/np3005985

Sosio M, Giusino F, Cappellano C, Bossi E, Puglia AM, Donadio S (2000) Artificial chromosomes for antibiotic-producing actinomycetes. Nat Biotechnol 18(3):343–345. doi:10.1038/73810

Strausberg RL, Strausberg SL (2001) Overview of protein expression in saccharomyces cerevisiae. In: Current protocols in protein science. Wiley, Hoboken. doi:10.1002/0471140864.ps0506s02

Struvay C, Feller G (2012) Optimization to low temperature activity in psychrophilic enzymes. Int J Mol Sci 13(9):11643–11665. doi:10.3390/ijms130911643

Tabata K, Kosuge T, Nakahara T, Hoshino T (1993) Physical map of the extremely thermophilic bacterium Thermus thermophilus HB27 chromosome. FEBS Lett 331(1–2):81–85. doi:10.1016/0014-5793(93)80301-A

Tachibana A, Tanaka T, Taniguchi M, Oi S (1996) Evidence for farnesol-mediated isoprenoid synthesis regulation in a halophilic archaeon, Haloferax volcanii. FEBS Lett 379(1):43–46. doi:10.1016/0014-5793(95)01479-9

Takai K, Sawasaki T, Endo Y (2010) The wheat-germ cell-free expression system. Curr Pharm Biotechnol 11(3):272–278. doi:10.1016/j.febslet.2014.05.061

Tamakoshi M, Uchida M, Tanabe K, Fukuyama S, Yamagishi A, Oshima T (1997) A new Thermus-Escherichia coli shuttle integration vector system. J Bacteriol 179(15):4811–4814. doi: 0021-9193/97/$04.0010

Taupp M, Mewis K, Hallam SJ (2011) The art and design of functional metagenomic screens. Curr Opin Biotechnol 22(3):465–472. doi:10.1016/j.copbio.2011.02.010

Tebbe CC, Vahjen W (1993) Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant-DNA from bacteria and a yeast. Appl Environ Microbiol 59(8):2657–2665. doi: aem00037-0325

Terron-Gonzalez L, Medina C, Limon-Mortes MC, Santero E (2013) Heterologous viral expression systems in fosmid vectors increase the functional analysis potential of metagenomic libraries. Sci Rep 3:1107. doi:10.1038/srep01107

Thomas L, Hodgson DA, Wentzel A, Nieselt K, Ellingsen TE, Moore J, Morrissey ER, Legaie R, Consortium S, Wohlleben W, Rodriguez-Garcia A, Martin JF, Burroughs NJ, Wellington EM, Smith MC (2012) Metabolic switches and adaptations deduced from the proteomes of Streptomyces coelicolor wild type and phoP mutant grown in batch culture. Mol Cell Proteomics 11(2):M111.013797. doi:10.1074/mcp.M111.013797

Tong YJ, Charusanti P, Zhang LX, Weber T, Lee SY (2015) CRISPR-Cas9 based engineering of actinomycetal genomes. ACS Synth Biol 4(9):1020–1029. doi:10.1021/acssynbio.5b00038

Troeschel S, Drepper T, Leggewie C, Streit W, Jaeger K-E (2010) Novel tools for the functional expression of metagenomic DNA. In: Streit WR, Daniel R (eds) Metagenomics, Methods in molecular biology, vol 668. Humana Press, New York, pp 117–139. doi:10.1007/978-1-60761-823-2_8

Tutino ML, Duilio A, Moretti MA, Sannia G, Marino G (2000) A rolling-circle plasmid from Psychrobacter sp. TA144: evidence for a novel rep subfamily. Biochem Biophys Res Commun 274(2):488–495. doi:10.1006/bbrc.2000.3148

Tutino ML, Duilio A, Parrilli E, Remaut E, Sannia G, Marino G (2001) A novel replication element from an Antarctic plasmid as a tool for the expression of proteins at low temperature. Extremophiles 5(4):257–264. doi:10.1007/s007920100203

Uchiyama T, Miyazaki K (2009) Functional metagenomics for enzyme discovery: challenges to efficient screening. Curr Opin Biotechnol 20(6):616–622. doi:10.1016/j.copbio.2009.09.010

Udwary DW, Zeigler L, Asolkar RN, Singan V, Lapidus A, Fenical W, Jensen PR, Moore BS (2007) Genome sequencing reveals complex secondary metabolome in the marine actinomycete Salinispora tropica. Proc Natl Acad Sci U S A 104(25):10376–10381. doi:10.1073/pnas.0700962104

van Wezel GP, McDowall KJ (2011) The regulation of the secondary metabolism of Streptomyces: new links and experimental advances. Nat Prod Rep 28(7):1311–1333. doi:10.1039/c1np00003a

van Wezel GP, McKenzie NL, Nodwell JR (2009) Chapter 5. Applying the genetics of secondary metabolism in model actinomycetes to the discovery of new antibiotics. Methods Enzymol 458:117–141. doi:10.1016/S0076-6879(09)04805-8

Vester JK, Glaring MA, Stougaard P (2015) Improved cultivation and metagenomics as new tools for bioprospecting in cold environments. Extremophiles 19(1):17–29. doi:10.1007/s00792-014-0704-3

Waldvogel E, Herbig A, Battke F, Amin R, Nentwich M, Nieselt K, Ellingsen TE, Wentzel A, Hodgson DA, Wohlleben W, Mast Y (2011) The PII protein GlnK is a pleiotropic regulator for morphological differentiation and secondary metabolism in Streptomyces coelicolor. Appl Microbiol Biotechnol 92(6):1219–1236. doi:10.1007/s00253

Wang F, Hao J, Yang C, Sun M (2010) Cloning, expression, and identification of a novel extracellular cold-adapted alkaline protease gene of the marine bacterium strain YS-80-122. Appl Biochem Biotechnol 162(5):1497–1505. doi:10.1007/s12010-010-8927-y

Wang GYS, Graziani E, Waters B, Pan W, Li X, McDermott J, Meurer G, Saxena G, Andersen RJ, Davies J (2000) Novel natural products from soil DNA libraries in a Streptomycete host. Organic Letters 2(16):2401–2404. doi: 10.1021/ol005860z

Wang Y, Zhang YH (2009) Cell-free protein synthesis energized by slowly-metabolized maltodextrin. BMC Biotechnol 9:58. doi:10.1186/1472-6750-9-58

Warburton P, Roberts AP, Allan E, Seville L, Lancaster H, Mullany P (2009) Characterization of tet(32) genes from the oral metagenome. Antimicrob Agents Chemother 53(1):273–276. doi:10.1128/AAC.00788-08

Warren RL, Freeman JD, Levesque RC, Smailus DE, Flibotte S, Holt RA (2008) Transcription of foreign DNA in Escherichia coli. Genome Res 18(11):1798–1805. doi:10.1101/gr.080358.108

Wayne J, Xu SY (1997) Identification of a thermophilic plasmid origin and its cloning within a new Thermus-E. coli shuttle vector. Gene 195(2):321–328. doi:10.1016/S0378-1119(97)00191-1

Wentzel A, Bruheim P, Overby A, Jakobsen OM, Sletta H, Omara WA, Hodgson DA, Ellingsen TE (2012a) Optimized submerged batch fermentation strategy for systems scale studies of metabolic switching in Streptomyces coelicolor A3(2). BMC Syst Biol 6:59. doi:10.1186/1752-0509-6-59

Wentzel A, Sletta H, Consortium S, Ellingsen TE, Bruheim P (2012b) Intracellular metabolite pool changes in response to nutrient depletion induced metabolic switching in streptomyces coelicolor. Metabolites 2(1):178–194. doi:10.3390/metabo2010178

Wexler M, Bond PL, Richardson DJ, Johnston AW (2005) A wide host-range metagenomic library from a waste water treatment plant yields a novel alcohol/aldehyde dehydrogenase. Environ Microbiol 7(12):1917–1926. doi:10.1111/j.1462-2920.2005.00854.x

Wietzorrek A, Bibb M (1997) A novel family of proteins that regulates antibiotic production in streptomycetes appears to contain an OmpR-like DNA-binding fold. Mol Microbiol 25(6):1181–1184. doi:10.1046/j.1365-2958.1997.5421903.x

Wild J, Hradecna Z, Szybalski W (2002) Conditionally amplifiable BACs: switching from single-copy to high-copy vectors and genomic clones. Genome Res 12:1434–1444. doi:10.1101/gr.130502

Wilson MC, Mori T, Ruckert C, Uria AR, Helf MJ, Takada K, Gernert C, Steffens UA, Heycke N, Schmitt S, Rinke C, Helfrich EJ, Brachmann AO, Gurgui C, Wakimoto T, Kracht M, Crusemann M, Hentschel U, Abe I, Matsunaga S, Kalinowski J, Takeyama H, Piel J (2014) An environmental bacterial taxon with a large and distinct metabolic repertoire. Nature 506(7486):58–62. doi:10.1038/nature12959

Wong S-L (1995) Advances in the use of Bacillus subtilis for the expression and secretion of heterologous proteins. Curr Opin Biotechnol 6(5):517–522. doi:10.1016/0958-1669(95)80085-9

Wright LF, Hopwood DA (1976) Identification of the antibiotic determined by the SCP1 plasmid of Streptomyces coelicolor A3(2). J Gen Microbiol 95(1):96–106. doi:10.1099/00221287-95-1-96

Yoon V, Nodwell JR (2014) Activating secondary metabolism with stress and chemicals. J Ind Microbiol Biotechnol 41(2):415–424. doi:10.1007/s10295-013-1387-y

Yoon MY, Lee KM, Yoon Y, Go J, Park Y, Cho YJ, Tannock GW, Yoon SS (2013) Functional screening of a metagenomic library reveals operons responsible for enhanced intestinal colonization by gut commensal microbes. Appl Environ Microbiol 79(12):3829–3838. doi:10.1128/AEM.00581-13

Zemella A, Thoring L, Hoffmeister C, Kubick S (2015) Cell-free protein synthesis: pros and cons of prokaryotic and eukaryotic systems. Chembiochem 16(17):2420–2431. doi:10.1002/cbic.201500340

Zengler K, Toledo G, Rappe M, Elkins J, Mathur EJ, Short JM, Keller M (2002) Cultivating the uncultured. Proc Natl Acad Sci U S A 99(24):15681–15686. doi:10.1073/pnas.252630999

Zerikly M, Challis GL (2009) Strategies for the discovery of new natural products by genome mining. Chembiochem 10(4):625–633. doi:10.1002/cbic.200800389

Zhang JW, Zeng RY (2008) Molecular cloning and expression of a cold-adapted lipase gene from an Antarctic deep sea psychrotrophic bacterium Pseudomonas sp 7323. Mar Biotechnol 10(5):612–621. doi:10.1007/s10126-008-9099-4

Zhang K, Martiny AC, Reppas NB, Barry KW, Malek J, Chisholm SW, Church GM (2006) Sequencing genomes from single cells by polymerase cloning. Nat Biotechnol 24(6):680–686. doi:10.1038/nbt1214

Zhang G, Li Y, Fang L, Pfeifer BA (2015) Tailoring pathway modularity in the biosynthesis of erythromycin analogs heterologously engineered in *E. coli*. Sci Adv 1(4):e1500077. doi:10.1126/sciadv.1500077

Zhou JZ, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62(2):316–322. doi: 0099-2240/96/$04.0010

Zhou Y, Asahara H, Gaucher EA, Chong S (2012) Reconstitution of translation from Thermus thermophilus reveals a minimal set of components sufficient for protein synthesis at high temperatures and functional conservation of modern and ancient translation components. Nucleic Acids Res 40(16):7932–7945. doi:10.1093/nar/gks568

Zhu H, Sandiford SK, van Wezel GP (2014) Triggers and cues that activate antibiotic production by actinomycetes. J Ind Microbiol Biotechnol 41(2):371–386. doi:10.1007/s10295-013-1309-z

Zobel S, Kumpfmüller J, Süssmuth R, Schweder T (2015) Bacillus subtilis as heterologous host for the secretory production of the non-ribosomal cyclodepsipeptide enniatin. Appl Microbiol Biotechnol 99(2):681–691. doi:10.1007/s00253-014-6199-0