

Isolation of Metagenomic High Molecular Weight DNA from Environmental Water Samples for the Construction of Metagenomic Fosmid Libraries

Dilara Begum, James Murray, and Fred Hyde — EPICENTRE, Madison, WI

Abstract

The construction of metagenomic libraries using a fosmid cloning system is rapidly becoming the method of choice for exploring environmental microbial communities that are unculturable or difficult to culture. However, the challenge remains primarily in isolation of genomic DNA from environmental samples. Current protocols for isolation and processing of high molecular weight (HMW) DNA suitable for fosmid cloning are long and tedious. We introduce a simple method to aid in direct isolation of HMW metagenomic DNA from environmental water samples. Our method is rapid and ensures maximum recovery of DNA, offering significant advantages over current methods: it eliminates the need for shearing and size selection of DNA using pulsed-field gel electrophoresis, steps in construction of fosmid libraries, and thus minimizes sample loss. This greatly increases the chances of recovering DNA from low-abundance organisms present in the environmental sample. In this report, we demonstrate the successful production of metagenomic libraries using DNA isolated from water samples.

Methods

Overview

A schematic overview of the protocol for DNA isolation and fosmid cloning is presented in Fig. 1. The protocol offers several advantages over current methods:

- No bead-beating is required.
- No agarose plugs, mechanical shearing, or size selection are needed, thereby minimizing sample loss.
- No columns or phenol/chloroform extractions are required.
- The procedure yields random-sheared, fosmid cloning-ready DNA in 90 minutes.
- Improved yields increase the probability of recovering DNA from low-abundant organisms present in the collective genomes in the environmental sample.

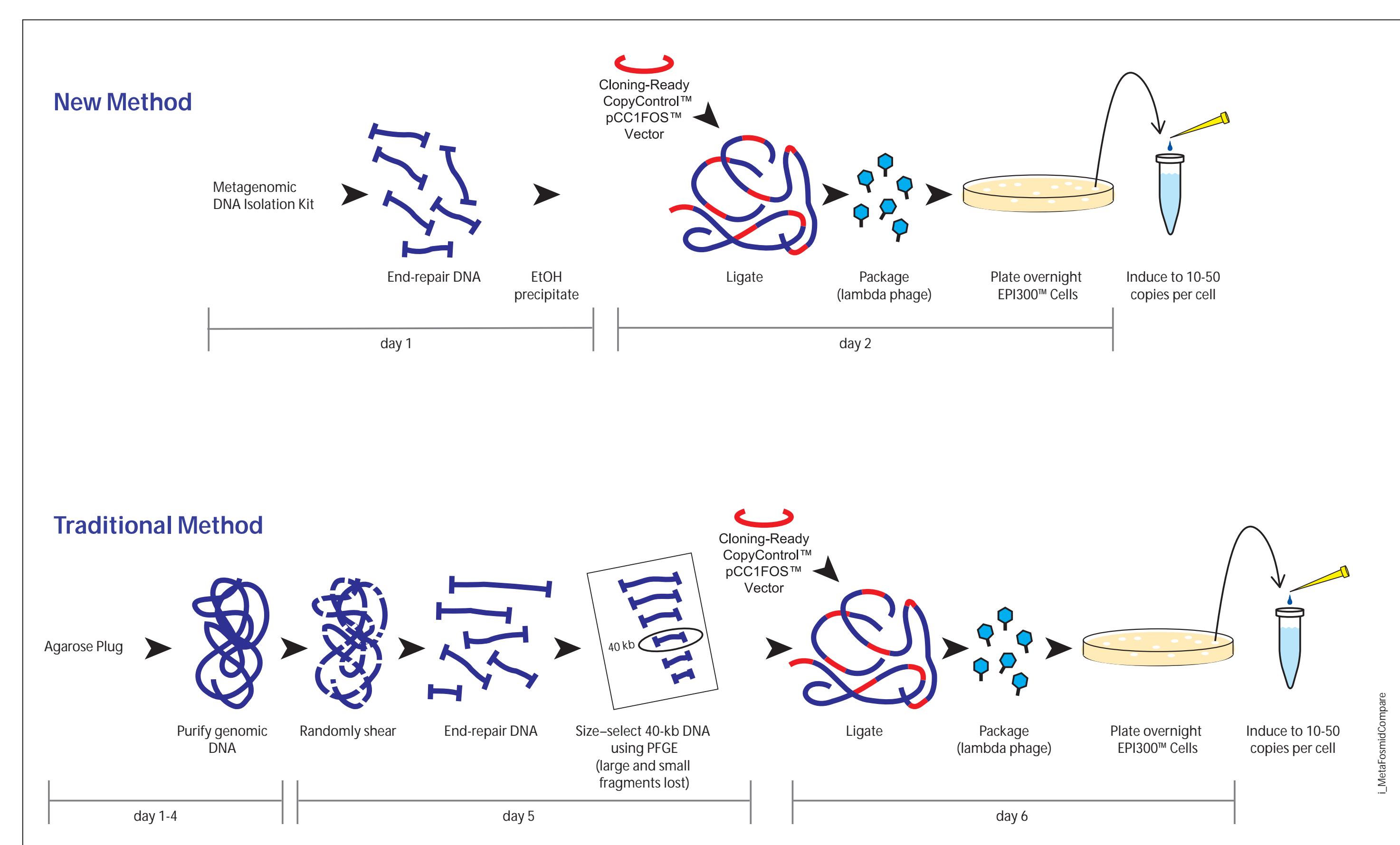


Fig. 1. Overview of the DNA isolation and library construction procedure.

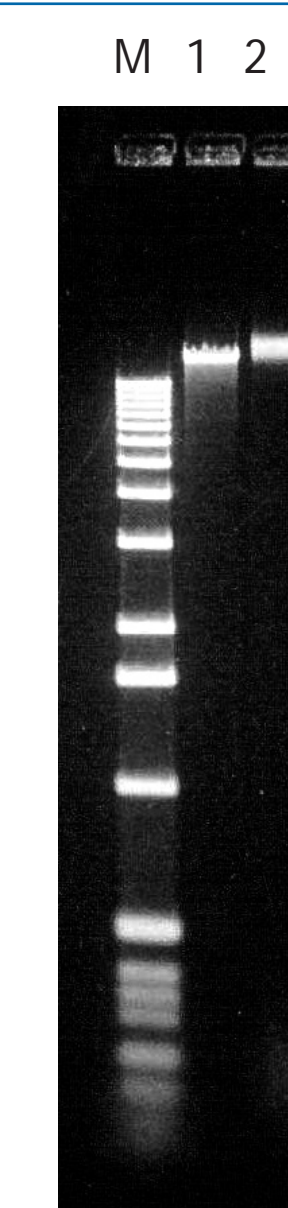
Construction of Metagenomic Fosmid Libraries

A 250-ml water sample was collected from Lake Mendota in Madison, WI and filtered, first through Miracloth (Calbiochem, EMD Biosciences) to remove large particles, and then through a 0.45-micron membrane filter to trap bacteria. Using the kit, the bacteria were washed off the filter and then lysed with lysis solution that contains ReadyLyse™ Lysozyme. The sheared DNA was purified by isopropanol precipitation and resuspended in 50 µl of TE buffer. Aliquots of the DNA were analyzed by agarose gel electrophoresis, confirming the size range of the sheared DNA was approximately 40 kb (Fig. 2). Using the CopyControl™ Fosmid Library Construction Kit (EPICENTRE), the sheared DNA was end-repaired to generate blunt, 5'-phosphorylated ends and then ligated directly into the cloning-ready CopyControl pCC1FOS Vector. The ligated DNA was then packaged using an ultra-high efficiency MaxPlax™ Lambda Packaging Extract (EPICENTRE) and plated on phage T1-resistant *E. coli* cells to produce the library. DNA from several randomly chosen clones was purified using the Direct Lysis Fosmid96 Kit (EPICENTRE). *Not I* digestion of the fosmid DNA produced inserts that were approximately 40 kb in size (Fig. 3).

Results

Metagenomic DNA from Lake Water

Fig. 2. Fosmid-ready genomic DNA was isolated directly from lake water. Lane M, kilobase ladder; lane 1, 40-kb control DNA; lane 2, DNA isolated from lake water sample.



Size Analysis of the Metagenomic Library

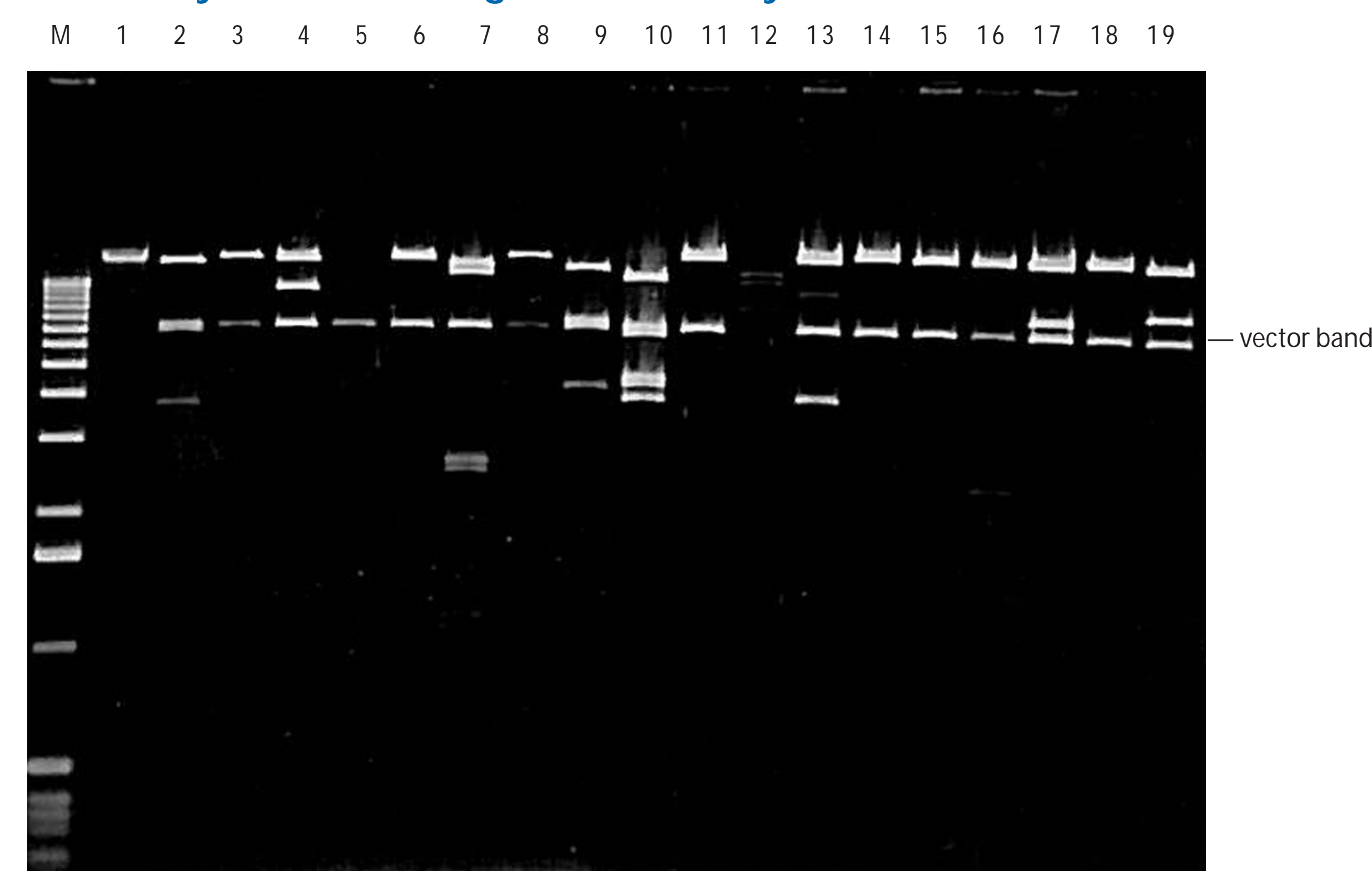


Fig. 3. *Not I* digestion of fosmid DNA isolated from randomly chosen clones verifies a 40-kb insert size. Lane M, kilobase ladder; lane 1, 40-kb control insert; lanes 2-19, *Not I*-digested fosmid DNA.

Characterization of the Library

Total Number of Fosmid Clones: 100,000

Clone Size: ~40 kb

Sequence Coverage: 100,000 X 40 kb = 4 Gb

Equivalent to 5X sequence coverage of 100 5-megabase microbial genomes

Sequencing of the Metagenomic Library

DNA isolated from 96 randomly picked fosmid clones from the water metagenomic fosmid library was end-sequenced using the pCC1Fosmid primers and was analyzed on an ABI 3730 Genetic Analyzer (Table 1). The BLAST sequence homology searches were conducted against three databases: GenBank nr, Genome Survey Sequence (GSS), and environmental sample (env-nt). A representative sample of the matches against env-nt is shown in Table 2. The results indicate that a majority of the sequence data obtained from the metagenomic analysis potentially represents novel species.

Table 1. Summary of sequencing analysis of the metagenomic fosmid library.

Number of sequences analyzed	70
Number of clones that showed significant match	8
Number of clones that did not show any significant match	62

Table 2. Representative sample of BLAST results.

Metagenomic Clones	Accession/ Description	Score	E Value	Match Length (bp)
MGLw-1C	gi 129568090 gb AACY024122579.1 Marine metagenome 1098101634...	127	8e-27	123
	gi 129571067 gb AACY024119602.1 Marine metagenome 1098213017...	105	3e-20	93
	gi 131591210 gb AACY022124186.1 Marine metagenome 1092351330...	71.9	4e-10	60
MGLw-2C	gi 132602310 gb AACY021110761.1 Marine metagenome 1093022077...	155	4e-35	166
	gi 59967137 gb AAF01067830.1 Metagenome sequence 2662324_fa...	145	4e-32	161
	gi 60088957 gb AAF01079838.1 Metagenome sequence XZS98949.x...	143	1e-31	132
MGLw-3G	gi 133248732 gb AACY020471240.1 Marine metagenome 1096626629...	117	2e-23	203
	gi 133530342 gb AACY020180644.1 Marine metagenome 1096626804...	95.6	7e-17	504
	gi 133234124 gb AACY020485848.1 Marine metagenome 1096626806...	65.9	6e-08	416
MGLw-5A	gi 133136967 gb AACY020583005.1 Marine metagenome 1092953001...	178	2e-42	214
	gi 179991701 gb ABO01014378.1 Saltern metagenome 41560987, ...	67.9	5e-09	45
	gi 132010008 gb AACY021702053.1 Marine metagenome 1095390199...	65.9	2e-08	45
MGLw-7E	gi 133417768 gb AACY020302451.1 Marine metagenome 1096626379...	444	6e-122	463
	gi 132322652 gb AACY021387333.1 Marine metagenome 1095522132...	240	2e-60	421
	gi 133413235 gb AACY020306984.1 Marine metagenome 1096626384...	95.6	6e-17	124
MGLw-8E	gi 130207674 gb AACY023484243.1 Marine metagenome ctg_110166...	69.9	4e-09	59
	gi 60089463 gb AAF01061869.1 Metagenome sequence XZS99643.g...	67.9	2e-08	36
	gi 130951462 gb AACY022745825.1 Marine metagenome ctg_110166...	65.9	7e-08	26
MGLw-8B	gi 133402111 gb AACY020318108.1 Marine metagenome 1096626396...	127	2e-26	96
	gi 133102597 gb AACY020617375.1 Marine metagenome 1092347005...	127	2e-26	185
	gi 133178502 gb AACY020541470.1 Marine metagenome 1096626799...	119	5e-24	96
MGL-w-12D	gi 152165121 gb ABEF01039910.1 Marine metagenome HOTS_Contig...	117	2e-23	200
	gi 180992368 gb ABON01004509.1 Freshwater metagenome 4120665...	111	1e-21	80
	gi 132469708 gb AACY021241325.1 Marine metagenome 1093017104...	101	1e-18	198

Conclusions

The Metagenomic DNA Isolation Kit for Water was developed for rapid isolation of 40-kb genomic DNA directly from microbes present in water samples. The kit maximizes DNA yields for fosmid library construction by making traditional shearing and size-fractionation steps obsolete. The kit facilitates construction of a high-coverage fosmid genomic library of the collective species present in any water sample.