



Extensive Range of Plant Genomic Libraries Made with the CopyControl™ Fosmid Library Production Kit

Ken Manning, Warwick Plant Genomic Libraries Limited

Warwick Plant Genomic Libraries Limited is a new company dedicated to supplying high quality plant genomic libraries to the international research community.¹ I have worked with the co-founder Professor Graham Seymour for many years, and have a clear understanding of the benefits of having the highest quality of genomic resources. Our own research has led to the successful cloning of a gene for a transcription factor regulating the ripening process in the tomato.² The expertise and technical know-how gained in this work now forms the foundation of a growing range of libraries made with the state-of-the-art CopyControl™ Fosmid Library Production Kit from EPICENTRE Biotechnologies.

The CopyControl Advantage

To use the CopyControl Fosmid Library Production Kit, genomic DNA is first sheared into approximately 40 kb fragments, end-repaired to form blunt ends, and when necessary, size selected using agarose gel electrophoresis. Finally, the size selected DNA is ligated into the cloning-ready pCC1FOS™ Vector (EPICENTRE) and packaged using the high efficiency MaxPlax™ Lambda Packaging Extracts (>10⁹ pfu/μg for phage lambda). Lambda phage are only viable if the amount of DNA packaged into the phage head is between 38 and 51 kb.³ Thus, the average insert size for the Foxglove library

(FIG 1) is 38 kb, and the number of clones without an insert is less than one percent.

The CopyControl Fosmid Library Production Kit offers several advantages over BAC and conventional cosmid cloning. The isolation of suitable 40 kb DNA fragments for cloning into the pCC1FOS Vector is much simpler and far less time-consuming because there is no need to isolate high molecular weight DNA, or perform partial restriction enzyme digestions, which often require extensive optimization. Moreover, a restriction endonuclease digestion is not a random unbiased process due to the uneven distribution of restriction sites in genomic DNA. The CopyControl Fosmid Library Production Kit uses the inherent randomness of mechanical shearing of the DNA to increase the probability that ALL sequences will be represented in the library.

Fosmids contain the origin of replication from the F' episome of *E. coli*, which maintains the clones as single copy, thereby enhancing insert stability. Cosmids, in contrast, contain a ColE1 origin of replication, which maintains clones at relatively high copy number. Continuous propagation of some clones at high copy number may result in deletions, inversions, and rearrangements of the DNA.⁴ Furthermore, the CopyControl pCC1FOS Vector contains an inducible, high copy origin, *oriV*. A short incubation in the presence of an inducer (the CopyControl™ Induction Solution) increases copy number for higher yields and higher purity DNA without compromising insert stability.⁵

High cloning efficiencies (>10⁶ fosmids from 100 ng of DNA) also make it easy to achieve full genome coverage in about 2 days. As shown in FIG 1, the 1-C value or genome size for Foxglove is approximately 1.2 X 10⁹ bp, and the library contains 1.5 X 10⁶ clones with an average insert size of 38,000 bp. Thus, the genome coverage is: (1.5 X 10⁶ clones) (38,000 bp)/1.2 X 10⁹ bp, or 47 genomes worth of Foxglove DNA. A 5-fold coverage indicates the chance of finding a particular genomic sequence in a library is approximately 99%. Most of the CopyControl Fosmid libraries have a

genome coverage greater than 20-fold. Even the huge genome of *Aloe vera*, five times larger than that of the human genome, has more than a 4-fold sequence representation in a typical fosmid library.

Of particular concern for the construction of plant genomic libraries is minimizing the contamination of nuclear DNA with chloroplast DNA. Chloroplast DNA can comprise up to 20% of the total DNA extracted from leaves, and homology between it and some nuclear sequences can confound genome walking and hybridization. To estimate the extent of contaminating chloroplast sequences, our libraries are screened for the presence of the chloroplast-specific large subunit of ribulose-bisphosphate carboxylase/oxygenase gene. As demonstrated for the Foxglove library, the frequency of chloroplast-specific fosmid clones was less than 0.00051%. Moreover, the abundance of nuclear encoded 18S ribosomal genes (1.8%) is as expected.

Conclusion

Warwick Plant Genomic Libraries makes the only commercially available plant genomic libraries created with large and unbiased inserts to give the best possible representation of the DNA of each species. By using EPICENTRE's CopyControl Fosmid Library Production Kit, we are confident that our genome libraries are of the highest possible quality.

References

1. <http://www.wpgl.co.uk/home.htm>
2. *Nature Genetics*, article accepted.
3. Murialdo, H. (1991) *Annu. Rev. Biochem.* **60**, 125.
4. Kim, U.J. et al., (1992) *Nucleic Acids Res.* **20**(5), 1083.
5. *EPICENTRE Forum* (2002), **9**(1), 3.

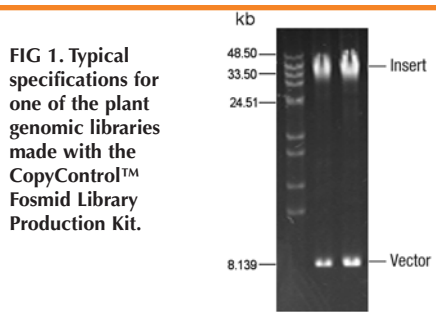


FIG 1. Typical specifications for one of the plant genomic libraries made with the CopyControl™ Fosmid Library Production Kit.

Species: Foxglove (<i>Digitalis purpurea</i>)	
Genome size (1-C value ¹):	1201 Mbp
Number of clones in primary library:	1.50 x 10 ⁶
Average insert size ² (see gel):	38 kb
Fold genome coverage:	47
Proportion of empty clones: ³	0.26%
Chloroplast contamination: ^{3,4}	0.00051%
Abundance of 18S rRNA: ³	1.8%

¹Data from Royal Botanic Gardens, Kew

²NotI digestion of 500 clones

³Determined by real time PCR

⁴Ribulose bis-phosphate carboxylase large subunit

www.EpiBio.com/ccfos.asp

CopyControl™ Fosmid Library Production Kit	
CCFOS110	1 Kit
CopyControl™ HTP Fosmid Library Production Kit	
CCFOS059	1 Kit
Both kits are for producing up to 10 complete and unbiased fosmid libraries.	