Investigation of a metagenomic library for aerobic toluene degradation genes

Emna Bouhajja[1], Isabelle George[2], Mark R. Liles[3] and Spiros Agathos[1]

[1] Laboratoire de Génie Biologique, Earth and Life, Université Catholique de Louvain, Place Croix de Sud 2/ L7.05.19, 1348 Louvain-la-Neuve, Belgium, [2] Université Libre de Bruxelles, Laboratoire d'Ecologie des Systèmes Aquatiques, Campus de la Plaine CP 221, Boulevard du Triomphe, 1050 Bruxelles, Belgium, and Laboratoire de Biologie Marine, Campus du Solbosch, CP 160/15, 50 Av. F.D. Roosevelt, 1050 Bruxelles, Belgium, [3] Department of Biological Sciences, Auburn University, Room 101, Rouse Life Sciences Building, 120 W. Samford Avenue, Auburn, Alabama 36849, USA.

Correspondence to: emna.bouhajja@uclouvain.be

A metagenomic study on sediment from a former gasworks site located in Düsseldorf-Flingern was conducted to assess the tar oil hydrocarbon degradation potential at that site. A sediment sample was collected from 6.70 m to 7.10 m below surface in the anoxic contaminant plume. This plume consisted mainly of benzene, toluene, ethylene, xylene and polyaromatic hydrocarbons. The first step was to extract enough high quality, high molecular weight DNA despite low cell density within the sediment. The extraction method used direct lysis of bacterial cells within the sediment followed by a phenol-chloroform extraction. The DNA was then purified using a centrifugal concentrator coupled with Q-Sepharose treatment. This protocol resulted in pure DNA of appropriate size range to construct a metagenomic library in the pCC1FOS vector with the CopyControl^{TM} fosmid library system (Epicentre). The recombinant *E.coli* clones were robotically picked in duplicate onto nylon membranes at Libragen, France. The macroarrays contained 53,760 fosmid clones with an average insert size of 40 kb representing 537 *E.coli* genome equivalents. They were screened by colony blot hybridization for toluene mono-oxygenase. We used the primer set TMOA-F/TMOA-R to amplify the *tmoA* gene from the same DNA that was used for library construction to generate a digoxygenin-labeled-probe for downstream library screening. A total of 200 clones exhibited chemiluminescence after hybridization with the digoxigenin-labeled *tmoA* probe (in duplicate), indicating the presence of a toluene mono-oxygenase cluster and therefore a

possible toluene degradation pathway. Assuming an equal abundance of genomes within the library, about 40 % of the screened genomes contained a *tmoA* gene. Only 34 clones out of these 200 positive *tmoA* clones identified by hybridization gave a *tmoA* PCR product. This result suggests that the remaining 166 clones contain mismatches with the TMOA-F/TMOA-R primer set. These clones may contain a *tmoA* gene that is divergent from previously described sequences. Sequencing of these clones is expected to considerably enrich the catalogue of known genes involved in toluene degradation.