Identification of an antibacterial protein by functional screening of a human oral metagenomic library

Preeti Arivaradarajan[1], Gunasekaran Paramasamy[1], Sean P. Nair[2], Elaine Allan[2], Peter Mullany[2#]

[1]Department of Genetics, Centre for Excellence in Genomic Sciences, School of Biological Sciences, Madurai Kamaraj University, Madurai, India.

[2]Department of Microbial Diseases, UCL Eastman Dental Institute, London, United Kingdom.

#Corresponding author

Mailing address: Department of Microbial Diseases, UCL Eastman Dental Institute, 256 Gray's Inn Road, University College London, WC1X 8LD, United Kingdom.

Phone: 44(0)203 456 1223. Fax: 44 (0)203 456 1127.

E-mail: p.mullany@ucl.ac.uk.

ABSTRACT

Screening of a bacterial artificial chromosome (BAC) library containing metagenomic DNA from human plaque and saliva allowed the isolation of four clones producing antimicrobial activity. Three of these were pigmented and encoded homologues of glutamyl-tRNA reductase (GluTR), an enzyme involved in the C5 pathway leading to tetrapyrole synthesis, and one clone had antibacterial activity with no pigmentation. The latter contained a BAC with an insert of 15.6 kb. Initial attempts to localise the gene(s) responsible for antimicrobial activity by sub-cloning into pUC-based vectors failed. A new plasmid for toxic gene expression (pTGEX) was designed enabling localisation of the antibacterial activity to a 4.7 kb HindIII fragment. Transposon mutagenesis localised the gene to an open reading frame of 483 bp designated antibacterial protein1 (*abp1*). Abp1 was 94% identical to a hypothetical protein of *Neisseria subflava* (accession number WP_004519448.1). An *E. coli* clone expressing Abp1 exhibited antibacterial activity against *Bacillus subtilis* BS78H, *Staphylococcus epidermidis* NCTC 11964 and B4268 , and *Staphylococcus aureus* NCTC 12493 ATCC 35696 and NCTC 11561 . However no antibacterial activity was observed against *Pseudomonas aeruginosa* ATCC 9027, *Neisseria subflava* ATCC A1078, *Escherichia coli* K12 JM109, and BL21(DE3) *Fusobacterium*

Deleted:

Deleted: (methicillin resistant, MRSA) *nucleatum* ATCC 25586 and NCTC 11326, *Prevotella intermedia* ATCC 25611, *Veillonella parvula* ATCC 10790 or *Lactobacillus casei* NCTC 6375.

INTRODUCTION

After the industrialization of penicillin production a plethora of antibiotics came into the market (Fernandes, 2006). However overuse of antibiotics acts as a selection pressure resulting in the emergence of resistant bacteria. This is exacerbated by the fact that antibiotic resistance genes are often borne on mobile genetic elements which may promote their rapid dissemination. Transfer of resistance genes to common pathogens makes otherwise treatable diseases untreatable; for a recent review see Juhas, 2015. Thus, there is an urgent need for new antimicrobials.

Microorganisms are known producers of antimicrobials (Tawiah *et al*., 2012) however access to all antimicrobial producers is hindered by the fact that not all are cultivable in the laboratory. Metagenomics allows access to the genetic potential of whole microbial communities in an environmental sample (Mullany, 2014) and has allowed the isolation of novel antimicrobial products. For example, investigation of the microbial communities associated with the marine sponge, *Cymbastela concentric,* and the green alga, *Ulva australis,* led to the identification of three novel hydrolytic enzymes with antibacterial activity (Yung *et al*., 2011). The pigments, indirubin and indigo, were discovered to have antibacterial activity by screening a fosmid library of a forest soil metagenome (Lim *et al*., 2005). The broad spectrum antibiotics, turbomycin A and B, from a soil metagenomic library were reported by Gillespie *et al*. (2002). The antibiotics, violacein and palmitoylputrescine, were isolated from soil and bromeliad tank water metagenomic libraries, respectively (Brady *et al*., 2001; Brady & Clardy, 2004). Palmitoylputrescine was not found in a previous screen for antibacterials amongst cultivable bacteria, indicating that metagenomic screening allows access to natural products that cannot be found using other methods. Recently a novel method for growing previously uncultivable soil bacteria has been developed which enabled the isolation of a new antibiotic, teixobactin, further emphasising the potential of non-cultivable bacteria as a rich source of useful products (Ling *et al*., 2015).

Deleted: palmitoylputrescine

Deleted: Palmitoylputrerescine

In this study we investigated the human oral metagenome by functional screening for antimicrobial production. A novel antibacterial protein was isolated from a BAC library of human plaque and saliva metagenomic DNA. Additionally, a vector for efficient cloning and stable maintenance of toxic genes was designed and used.

MATERIALS AND METHODS

Bacterial cultures and vectors used.

Escherichia coli strains and plasmids used in this study are listed in Table 1. The indicator bacteria tested were: *Bacillus subtilis* BS78H (a chloramphenicol resistant derivative of CU2189), *Staphylococcus epidermidis* NCTC 11964 and B4268, *Staphylococcus aureus* NCTC 12493 NCTC 11561 and ATCC 35696 , *Pseudomonas aeruginosa* ATCC 9027, *Neisseria subflava* ATCC A1078, *Escherichia coli* K12 JM109 and *E. coli* B BL21(DE3) , *Fusobacterium nucleatum* ATCC 25586 and NCTC 11326 , *Prevotella intermedia* ATCC 25611, *Veillonella parvula* ATCC 10790 and *Lactobacillus casei* NCTC 6375.

Deleted: (MRSA)

Deleted: ,

In order to mediate stable cloning of genes toxic to *E. coli,* the vector, pTGEX (plasmid for toxic gene expression) was designed (Fig. 1). This was constructed by DNA2.0, USA to our specifications: plasmid pJ421 containing a p15a (Chang & Cohen, 1978) origin of replication was modified by insertion of the pUC19 *lacZ* fragment containing a multiple cloning site allowing cloning flexibility and blue white selection. DNA cloned into pTGEX is tightly controlled by an isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible phage T5 promoter. The vector has transcription terminators, designed and synthesised by DNA 2.0 and placed upstream and downstream of the multiple cloning sites.

Functional screening.

Clones from a previously described human oral metagenomic library (Warburton *et al*., 2009) were grown on Luria Bertani (LB) agar containing chloramphenicol (12.5 µg/ml) and IPTG (0.5 mM) for 24 h at 37°C. Thereafter the plates were incubated at 22-25°C for 5 days. For the antibacterial assay, a 16 h *B. subtilis* culture (in LB broth) was diluted to obtain an optical density at 600 nm (OD_{600}) of 0.3. This suspension was diluted 1 in 100 in to 0.5% LB agar and overlaid on to the LB agar plate containing the metagenomic clones. The plates were incubated for 24 h at 37°C followed by 3-4 days incubation at 22-25°C and were scored for a zone of inhibited *B. subtilis* growth.

Any clones showing antibacterial activity against *B. subtilis* were also tested against *S. epidermidis, S. aureus, P. aeruginosa, N. subflava, E. coli* and *L. casei* which were grown aerobically at 37 °C in brain-heart infusion (BHI) broth. Cultures of *F. nucleatum*, *P. intermedia*, and *V. parvula* were grown in BHI broth in an anaerobic atmosphere (80% N_2, 10% CO_2, 10% H_2).

Subcloning and DNA sequencing.

The pTGEX plasmid was used for sub-cloning in *E. coli* DH5α. Sub-clones from each experiment were grown on LB agar containing kanamycin (30 µg/ml) and IPTG (1 mM) for 12 h at 37°C followed by 12 h of incubation at 22-25°C and thereafter scored for antibacterial activity as described above. The insert DNA from the sub-clone showing antibacterial activity was sequenced using pTGEX sequencing primers, pTGEXFP (5'-TTACGAGCTTCATGCACAG-3') and pTGEXRP (5'-AGGGTTATTGTCTCATGAGC-3') which flank the cloning site.

Bioinformatics analysis.

DNA sequences were analysed using tools available at the National Centre for Biotechnology Information (NCBI), ORF finder (http://www.ncbi.nlm.nih.gov/) and BLAST (Altschul *et al*., 1990). A putative promoter was assigned using a promoter prediction program, BPROM (http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfindb). The transcription terminator was recognized using the web server, ARNold (Naville *et al*., 2011). The putative signal peptide and cleavage site was identified by the online tool, Prediction of Protein Localization Sites (PSORT) (Gardy *et al*., 2005).

Transposon mutagenesis.

For transposon mutagenesis, the Entranceposon (an artificial Mu transposon) containing a kanamycin resistance gene was used. Transposon mutant libraries were constructed following the manufacturer's instructions (Template Generation System II kit; TGS, F-702; Finnzyme, Finland). The transposition reaction mixture was transformed into the strains of *E. coli* EPI300-T1[R] carrying the plasmids encoding antibacterials, and transformants were selected on LB agar containing chloramphenicol (12.5 µg/ml) and kanamycin (20 µg/ml). The transformants were grown for 12 h at 37°C, with shaking at 200 rpm in 96 well flat bottom plates with 110 µl LB broth containing kanamycin. 5 µl of the 12 h culture was spotted on to LB agar supplemented with chloramphenicol (12.5 µg/ml) and IPTG (0.5 mM) and the plates were incubated overnight at 37°C. After overnight incubation, an overlay assay using *B. subtilis* was performed as described above and mutants not showing a zone of inhibition were selected. To localise the transposon insertion site, plasmids isolated from the mutants were digested with HindIII and analyzed by gel electrophoresis. Mutant plasmids with single transposon insertions were sequenced using transposon specific primers (SeqW and SeqE provided by the manufacturer) and the transposon insertion point identified.

Nucleotide sequence submission.

The nucleotide sequence of the insert from p112C was deposited in GenBank under the accession number KF955286.

RESULTS AND DISCUSSION

Screening of a human plaque and saliva metagenomic library allowed the isolation of four clones producing antibacterial products.

19,200 metagenomic clones containing DNA from human pooled plaque and saliva were screened and three brownish-red pigmented clones, BAC28G, BAC108J and BAC110F, containing BACs with insert sizes of 17.7 kb, 5.4 kb and 21.3 kb, respectively, and one non-pigmented clone, HOAb112C, which contained a BAC with an insert of 15.6 kb, gave zones of growth inhibition of *B. subtilis*.

In vitro transposon mutagenesis of plasmids from the pigmented clones showed that the loss of function mutants invariably had a transposon inserted in an ORF (of identical sequence in all three pigmented clones) homologous to *hemA* which encodes glutamyl-tRNA reductase (GluTR), the first enzyme in the tetrapyrrole C5 biosynthetic pathway (Srivastava & Beale, 2005). The predicted amino acid sequence from the metagenomic clones was 99% identical to GluTR of a number of *Neisseria* species (e.g. *N. mucosa* WP_003748579), 96% identical to GluTR of *V. parvula* (YP_003312383) and 91% identical to GluTR of *Neisseria flavescens* (WP_003679373). Pigmented metagenomic clones with antibacterial activity have been found in gene libraries constructed from rice paddy soil DNA and from metagenomic DNA from the marine sponge, *Discodermia calyx* (Kim *et al*., 2009; He *et al*., 2012). Biochemical analysis of the pigments from the soil metagenome showed that they were coproporphyrin and minor porphyin intermediates and the gene encoding GluTR (*hemA*) was shown to be responsible for pigment production. As *E. coli* also contains a *hemA* gene it is likely that the pigmentation is due to increased levels of GluTR as a result of either the increase in copy number of the *hemA* gene or increased expression in the metagenomic clone. Therefore it is likely that the GluTR-encoding clones from the oral metagenome were also producing antibacterial porphyrins.

Sub-cloning of the antibacterial encoding genes required the construction of a new vector

The non-pigmented clone, HOAb112C, contained a plasmid designated p112C with an insert of 15.6 kb. In order to localize the region responsible for antibacterial activity, sub-cloning of the three

HindIII fragments (7 kb, 4.7 kb & 3.9 kb) was attempted. Despite multiple attempts, the gene encoding antimicrobial activity could not be cloned into pUC based vectors. However, the three HindIII fragments were successfully cloned in pTGEX (see materials and methods for vector details) to generate pTGEX7, pTGEX4.7 and pTGEX3.9 and antibacterial activity was only observed in *E. coli* containing pTGEX4.7 (Fig. 2).

The utility of pTGEX was further demonstrated by sub-cloning the 5.4 kb insert from pBAC108J, containing the gene involved in production of a pigmented antibacterial metabolite (see above) into pTGEX to generate pTGEX5.4. The radius of the zone of inhibition in pTGEX5.4 was 0.4 cm while that in the original BAC clone was 0.1 cm (data not shown). Presumably this is because in pTGEX5.4 the gene in the insert is under the control of a strong ITPG inducible promoter, whereas in pCC1BAC its likely to be under the control of its own promoter. Furthermore, un induced pCC1BAC has a copy number of 1 whereas pTGEX5 has a copy number of 10-12. The larger zone of inhibition produced by subcloning in pTGEX, may prove useful for the easier identification of clones producing antibacterial products.

DNA sequence analysis of a clone encoding a novel antibacterial protein.

DNA sequence analysis of pTGEX4.7 showed that it contained six *orf*s (Fig. 3) with homology to genes of *Neisseria* species: starting at the the 5' end, the first partial *orf* showed 82% identity to the gene *fadD* coding for long-chain-fatty -acid--CoA ligase (NCBI accession number CBN86964), the second orf showed 86% identity to the gene *trmU* coding for tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase (accession number ADO32009), the third showed 83% identity to a gene coding for a hypothetical protein (accession number ACF28991) which we subsequently designated antibacterial protein1 (*abp1*). The fourth *orf* showed 73% identity to *dgk* coding for diacylglycerol kinase (accession number CBN86961), the fifth orf displayed 84% identity to *gshB* coding for glutathione synthetase (accession number ADZ01159) and the sixth partial *orf* showed 80% identity to *recN* coding for DNA repair protein RecN (accession number CBN87706).

To precisely identify the gene(s) responsible for antibacterial activity, p112C was subjected to *in vitro* transposon mutagenesis. This clone was chosen as a target for mutagenesis as it was thought it may contain regulatory genes involved in the modulation of antimicrobial production. Two mutants, 112CTM3D10 and 112CTM1H7, which had lost antibacterial activity and which each had a single transposon insertion were obtained. In 112CTM3D10, the transposon was inserted in the 483 bp *orf* (*abp1*) whose translation product showed 96% identity to a hypothetical protein (WP_003681960) of *N. flavescens*. Homology searches in the conserved domain database showed that Abp1 belonged to a family of proteins (putative membrane or periplasmic proteins) with unknown function. However, the mutagenesis results clearly showed that the gene encoded a product with antibacterial activity. Interestingly, in 112CTM1H7, the *abp1* coding sequence was intact and the transposon was inserted between the -10 and -35 boxes of a predicted upstream promoter suggesting that transposon insertion has prevented transcription. Downstream of the gene, a Rho-independent transcription terminator was predicted (Fig. 4).

At the N-terminus of Abp1, a 19 residue signal peptide was recognized by PSORT (Fig. 4) which is predicted to direct the protein to the periplasmic space (certainty=0.93).

HOAb112 has antimicrobial activity against a range of bacteria

HOAb112 had antibacterial activity against *S. epidermidis* NCTC 11964 and B4268 , *S. aureus* NCTC 12493, ATCC 35696 and NCTC 11561 and *B. subtilis* CU2189. However no antibacterial activity against *P. aeruginosa* ATCC 9027, *N. subflava* ATCC A1078, *E. coli* K12 JM109 and BL21(DE3), *F. nucleatum* ATCC 25586 and NCTC 11326 , *P. intermedia* ATCC 25611, *V. parvula* ATCC 10790 or *L. casei* NCTC 6375 was observed.

Deleted:

The narrow antibacterial spectrum of Abp1 may be useful therapeutically.

In summary, this work has shown that the human oral metagenome is a source of antimicrobial agents. A potential problem in accessing these in functional metagenomic studies is their toxicity to the host. To overcome this we have designed the vector pTGEX which allows the cloning of toxic genes. In future work we plan to use this vector to construct gene libraries and screen for antimicrobial production.

ACKNOWLEDGEMENTS

The financial support by Commonwealth Scholarship Commission in the United Kingdom through Commonwealth Split-site Doctoral Scholarship is acknowledged. The authors are grateful to Dr. Haitham Hussain from Department of Microbial Diseases, UCL Eastman Dental Institute, London for providing *Bacillus subtilis* BS78H.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215:** 403-410.

Brady SF & Clardy J (2004) Palmitoylputrescine, an antibiotic isolated from the heterologous expression of DNA extracted from bromeliad tank water. *J Nat Prod* **67:** 1283-1286.

Brady SF, Chao CJ, Handelsman J & Clardy J (2001) Cloning and heterologous expression of a natural product biosynthetic gene cluster from eDNA. *Org Lett* **3:** 1981-1984.

Chang AC & Cohen SN (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134:** 1141-56.

Fernandes P (2006) Antibacterial discovery and development--the failure of success? *Nat Biotechnol* **24:** 1497-1503.

Gardy JL, Laird MR, Chen F, Rey S, Walsh CJ, Ester M & Brinkman FS (2005) PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. Bioinformatics **21:** 617-623.

Gillespie DE, Brady SF, Bettermann AD, Cianciotto NP, Liles MR, Rondon MR, Clardy J, Goodman RM & Handelsman J (2002) Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. *Appl Environ Microbiol* **68:** 4301-4306.

He R, Wakimoto T, Takeshige Y, Egami Y, Kenmoku H, Ito T, Wang B, Asakawa Y & Abe I (2012) Porphyrins from a metagenomic library of the marine sponge *Discodermia calyx*. *Mol Biosyst* **8:** 2334-2338.

Juhas M (2015) Horizontal gene transfer in human pathogens. *Crit Rev Microbiol* **41:** 101-108.

Kim JS, Lim HK, Lee MH, Park JH, Hwang EC, Moon BJ & Lee SW (2009) Production of porphyrin intermediates in *Escherichia coli* carrying soil metagenomic genes. *FEMS Microbiol Lett* **295:** 42-49.

Lim HK, Chung EJ, Kim JC, Choi GJ, Jang KS, Chung YR, Cho KY & Lee SW (2005) Characterization of a forest soil metagenome clone that confers indirubin and indigo production on *Escherichia coli*. *Appl Environ Microbiol* **71:** 7768-7777.

Ling LL, Schneider T, Peoples AJ *et al*. (2015) A new antibiotic kills pathogens without detectable resistance. *Nature*, **517:** 455-459.

Mullany P (2014) Functional metagenomics for the investigation of antibiotic resistance. *Virulence* **5:** 443-447.

Naville M, Ghuillot-Gaudeffroy A, Marchais A & Gautheret D (2011) ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol* **8:** 11-13.

Srivastava A & Beale SI (2005) Glutamyl-tRNA reductase of *Chlorobium vibrioforme* is a dissociable homodimer that contains one tightly bound heme per subunit. *Journal of Bacteriology* **187:** 4444-4450.

Tawiah AA, Gbedema SY, Adu F, Boamah VE & Annan K (2012) Antibiotic producing microorganisms from River Wiwi, Lake Bosomtwe and the Gulf of Guinea at Doakor Sea Beach, Ghana. *BMC Microbiol* **12:** 234-241.

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:** 273-276.

Yung PY, Burke C, Lewis M, Kjelleberg S & Thomas T (2011) Novel antibacterial proteins from the microbial communities associated with the sponge *Cymbastela concentrica* and the green alga *Ulva australis*. *Appl Environ Microbiol* **77:** 1512-1515.