Genes without frontiers?

Douda Bensasson, Jeffrey L. Boore and Kaare M. Nielsen[*†]

Evolutionary Genomics Department, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA; * Department of Pharmacy, University of Tromso, N9037 Tromso, Norway; [†] The Norwegian Institute for Gene Ecology, Tromso, Norway.

Corresponding author: Douda Bensasson
Dept of Biological Sciences
Imperial College London
Silwood Park
Buckhurst Rd
Ascot, Berks SL5 7PY, UK
email. d.bensasson@imperial.ac.uk

Word count for abstract & main text : 3021

Number of references: 48

Running title : Horizontal gene transfer from recombinant vectors

Keywords : horizontal gene transfer, biosafety, natural transformation

Abstract

For bacteria, the primary genetic barrier against the genetic exchange of DNA that is not self-transmissible is dissimilarity in the bacterial DNA sequences concerned. Genetic exchange by homologous recombination is frequent among close bacterial relatives and recent experiments have shown that it can enable the uptake of closely linked non-homologous foreign DNA. Artificial vectors are mosaics of mobile DNA elements from free-living bacterial isolates and so bear a residual similarity to their ubiquitous natural progenitors. This homology is tightly linked to the multitude of different DNA sequences that are inserted into synthetic vectors. Can homology between vector and bacterial DNA enable the uptake of these foreign DNA inserts? In this review we investigate pUC18 as an example of an artificial vector and consider whether its homology to broad host-range antibiotic resistance transposons and plasmid origins of replication could enable the uptake of insert DNA in the light of studies of homology-facilitated foreign DNA uptake. We also discuss the disposal of recombinant DNA, its persistence in the environment and whether homologies to pUC18 may exist in naturally competent bacteria. Most DNA that is inserted into the cloning site of artificial vectors would be of little use to a bacterium, but perhaps not all. 202 words

Gene flow across distantly related bacterial groups (horizontal gene transfer) is a major feature of bacterial evolution (Maynard Smith *et al*, 1991; Campbell, 2000; Ochman *et al*, 2000; Gogarten *et al*, 2002). This evolution need not be slow. The intense selection pressure imposed on microbial communities by worldwide antibiotic use reveals that new multi-resistance plasmids can arise from diverse origins and spread in less than five decades (Hartl and Dykhuizen, 1984; Davies, 1994). In this case, the antibiotic resistance genes that spread so rapidly were mostly carried by mobile genetic elements.

Genetic exchange of DNA that is not self-transmissible can occur within a similar timeframe (Maynard Smith *et al*, 1991; Bowler *et al*, 1994), but such genetic exchanges occur by homologous recombination and so are usually restricted to very closely related bacteria. Thus the primary genetic barrier against genetic exchange of DNA that is not self-transmissible is dissimilarity in the bacterial DNA sequences concerned (DuBose *et al*, 1988; Rayssiguier *et al*, 1989; Maynard Smith *et al*, 1991; Vulic *et al*, 1997; Feil *et al*, 2001). The rate of homologous recombination is high enough to obliterate the phylogenetic signal among some close bacterial relatives (Feil *et al*, 2001), and for the nucleotide sites of some species, homologous recombination is a far more likely prospect than a point substitution

(Guttman and Dykhuizen, 1994; Feil *et al*, 1999; Feil *et al*, 2000; Feil *et al*, 2001). Recent experiments show that close DNA sequence similarity can also facilitate the stable acquisition of linked non-homologous foreign DNA sequences (Nielsen *et al*, 2000a; Nielsen *et al*, 2000b; de Vries and Wackernagel, 2002; Kay *et al*, 2002; Prudhomme *et al*, 2002).

The synthetic vectors that are heavily used in industry and research are mosaics of mobile DNA elements from natural bacterial isolates (Pouwels *et al*, 1988). Care was taken to render these vectors non-transmissible by removal of the genes necessary for their mobility (Berg *et al*, 1975; NIH guidelines for research involving recombinant DNA molecules: Appendix I, April 2002). Because of their natural origins, artificial vectors still bear a residual similarity to DNA of naturally occurring bacteria. If this residual similarity were to enable recombination between bacterial DNA and only the parts of the artificial vectors that are homologous, there would most likely be no discernable effect on the bacterium involved. However, a plethora of different DNA sequences is inserted into artificial vectors and this raises the question whether homology between vector and bacterial DNA can facilitate the uptake of these foreign DNA inserts. In this review, we address this question by considering the extent of vector similarity to natural bacterial sequences, the requirements of

homology-mediated foreign DNA uptake, the persistence of plasmid DNA and its bacterial uptake in the environment, and the possible effects of natural selection.

The natural homologies of artificial vectors

How similar are the sequences of artificial vector DNA to those of naturally occurring bacterial DNA? We use the relatively small cloning vector pUC18 (a pBR322 derivative) to address this question, because it is one of the simplest synthetic vectors, is commonly used, and is similar in sequence and organisation to many other artificial vectors.

The 2,686 bp plasmid vector pUC18 is a mosaic of three different naturally occurring DNA sequences: an origin of replication and its flanking DNA from the *Escherichia coli* ColE1 plasmid, an antibiotic resistance gene and its flanking DNA from the transposon Tn*3* of a naturally occurring resistance plasmid, and part of a β-d-galactosidase gene (LacZ) and flanking DNA from *E. coli* chromosomal DNA that may have originated from transposon Tn*951* (Cornelis *et al*, 1978). The remaining 55 bp originate from the mostly synthetic multiple cloning site (MCS, or polylinker region originally developed in the bacteriophage M13) into which recombinant DNA is inserted (Fig. 1) (Pouwels *et al*, 1988). The “M13” universal primers

that are generally used for PCR of recombinant DNA bind to sites that are actually homologous only to the β-d-galactosidase gene.

The largest component of pUC18 is the Tn*3*-like region containing the 860 bp TEM-1 β-lactamase gene (*bla*$_{TEM-1}$), which confers resistance to the antibiotic ampicillin, and other penicillins (β-lactam antibiotics). The Tn*3*-like region of pUC18 (1209 bp) is almost identical (97.5-99.8% similarity) to all reported natural TEM-type DNA sequences (Table 1). These include the TEM-type β-lactamases that are resistant to the new generation of synthetic β-lactams and those that are resistant to inhibitors manufactured to suppress β-lactamase activity. These highly related TEM-type sequences have spread, on Tn*3*-like transposons, to an extremely broad range of bacterial taxa. The TEM-type sequences deposited in GenBank alone (Table 1) represent species of commensal and pathogenic bacteria of the Enterobacteriaceae, other Proteobacteria such as *Pseudomonas aeruginosa*, *Stenotrophomonas maltophilia*, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, and the phylogenetically distant gum-dwelling *Capnocytophaga ochracea* of the Flavobacteria (Rosenau *et al*, 2000). TEM-type β-lactamases are carried by plasmid-borne transposons, to which pUC18 has at least 99% similarity, including Tn*1*, Tn*2*, Tn*3*, Tn*801*, Tn*841*, Tn*4353* and Tn*SF1* (Table 1).

The broad distribution of these sequences is less surprising considering that β-lactams are the most commonly prescribed antibiotics worldwide, and that the single most common form of genetic resistance to these antibiotics is TEM-type β-lactamase production (Therrien and Levesque, 2000). Early work on antibiotic resistance has shown that TEM-type β-lactamases occur in a broad range of bacterial species, in many different countries, and on many different conjugative resistance plasmids, whose incompatibility groups include FI, FII, Iα, N, A-C complex, H, L , M, P, S, T, W, X and Y (Matthew and Hedges, 1976). The Tn*3*-like region of pUC18, and most ampicillin resistant vectors, is derived from one of the first antibiotic resistance factors to be isolated in 1963 (Datta and Kontomichalou, 1965; Sutcliffe, 1978), and it is not surprising that one of the first natural isolates of an antibiotic resistance gene represents one of the most ubiquitous types.

The pUC18 ColE1-like replication region has at least 83% sequence identity to the replication origins of many naturally occurring, small, transmissible plasmids of the Enterobacteriaceae (Table 1). In some cases similarities flank the pUC18 cloning site, such that they also flank any DNA that is inserted into the vector. This is also the case for the β-d-galactosidase (LacZ) region (Table 1 and Fig. 1).

The cloning vector pUC18 is not unusual among synthetic vectors. Most of the artificial vector sequences in the GenBank vector database share pUC18's ColE1-like origin of replication (671 of 911 sequences), and its Tn*3*-like region conferring ampicillin resistance (648 of 911). Other genes commonly used as selective markers for vectors are also derived from mobile, naturally occurring bacterial DNA sequences. For example, neomycin resistance (*npt*II) is conferred by a region of transposon Tn*5*, chloramphenicol resistance is from transposon Tn*9*, and kanamycin resistance (*npt*I) is from transposon Tn*903* (Pouwels *et al*, 1988).

A hitchhiker's guide to homology

How much homologous DNA sequence is required for the bacterial acceptance of how much foreign DNA? In a recent experiment, de Vries and Wackernagel (2002) grew Gram-negative naturally competent bacteria (*Acinetobacter* sp. BD413) in the presence of linear naked DNA. As little as 183 bp of homology between the linear naked DNA and a plasmid in the bacteria was enough to facilitate detectable levels of foreign DNA uptake. Further experiments, which used 1 kb of homology, showed that this homology enabled the frequent bacterial acquisition of a working bleomycin resistance gene that itself had no homology to the plasmid. This 1 kb of homology frequently anchored the stable

integration of as much as 2.6 kb of foreign DNA and about 1 kb on average (de Vries and Wackernagel, 2002).

Prudhomme *et al* (2002) present very similar findings for *Streptococcus pneumoniae,* a naturally competent Gram-positive bacterium, where less than 1 kb of one-sided homology is sufficient for the chromosomal uptake of over 2 kb of foreign DNA. In the absence of DNA homology no stable integration of the naked DNA with bacterial DNA could be detected, but when homology to bacterial DNA flanks both sides of the foreign DNA the rate of uptake is higher than in the case of one-sided homology (de Vries and Wackernagel, 2002; Prudhomme *et al*, 2002) and close to the rate at which naked DNA is able to enter the bacterium (Nielsen *et al*, 2000a; Nielsen *et al*, 2000b; de Vries *et al*, 2001). Prudhomme *et al* also show that a single homologous region (942 bp) facilitates the integration of whole plasmids into bacterial chromosomes at a 100-fold greater rate than foreign DNA uptake from the same plasmid when linearised.

Several studies estimate the degree of similarity required for homologous recombination in different types of bacteria. The frequency of recombination of DNA molecules in bacteria, such as the enterobacteria (Vulic *et al*, 1997; Vulic *et al*, 1999), *Bacillus* (Majewski and Cohan, 1998), and *Streptococcus* (Majewski *et al*,

2000), increases exponentially as the nucleotide similarity of the participating DNA molecules increases. Most wild-type bacteria abort effective recombination with DNA that is over 1-2% diverged from their own, however, the constraints on sequence similarity for homologous recombination can be relaxed to ~20% divergence in certain “mutator” individuals (Rayssiguier *et al*, 1989; Vulic *et al*, 1997; Vulic *et al*, 1999; Majewski *et al*, 2000), and others predict recombination between molecules that are up to 30% diverged (Townsend *et al,* 2003). Mutator individuals make up 1-2% of natural *E. coli* and *Salmonella* populations (LeClerc *et al*, 1996; Matic *et al*, 1997), though one study of patients with cystic fibrosis, found that on average 20% of the pathogenic isolates of *Pseudomonas aeruginosa* examined showed mutator phenotypes (Oliver *et al*, 2000). Homologous recombination in mutators could explain the finding that *Streptococcus pneumoniae* strains became resistant to antibiotics in the last 50 years through changes to their penicillin-binding protein genes by at least two independent recombination events, involving DNA sequences that were up to 20% divergent (Maynard Smith *et al*, 1991).

Most of the homologies described in Table 1 are between pUC18 and Tn*3*-like transposons. Most of these are less than 1% diverged from naturally occurring bacterial DNA sequences and span

more than 1 kb. If artificial vectors and their inserts were available to bacteria, such sequence similarity would seem sufficient to facilitate the bacterial uptake of insert DNA. The sequences with which pUC18 has Tn*3*-like homology are on broad host-range conjugative plasmids and are associated with broad host-range mobile elements (Table 1). If a new DNA sequences could be associated with such mobile elements this may increase the chance of it spreading to other bacteria. The homology of pUC18 to plasmids with ColE1-like origins of replication is also sufficiently long for the uptake of inserted DNA sequences, but natural and artificial sequences are up to 17% diverged so homology-facilitated uptake of foreign DNA would seem likely only in mutators. Homologies of ColE1-like origins of replication are generally associated with small non-conjugative plasmids (Table 1) so even if recombination did occur with these it probably would not significantly increase the chances of spread to other bacterial species.

Bacterial competence and naked DNA

Are artificial vectors and their inserts available to free-living bacteria? If DNA is associated with the appropriate mobile elements in living bacteria, genetic exchange can proceed by conjugation or transduction and thus reach a broad range of bacteria at a potentially high rate. Artificial cloning vectors are not mobile, however, and the

laboratory microorganisms that propagate them are killed prior to their environmental release. Recombinant DNA will not be propagated in the environment unless free-living bacteria pick it up by natural transformation. Given that most of the mobile elements with which artificial vectors have homology are the very ones that carried antibiotic resistance genes across bacterial phyla, the product of an initial natural transformation event could then spread in a relatively unimpeded way.

Some bacteria, for example *Streptococcus pneumoniae*, *Acinetobacter* sp. Strain BD13, *Pseudomonas stutzeri*, *Neisseria gonorrhoeae* (reviewed in Lorenz and Wackernagel, 1994), are naturally competent, that is they are able to accept naked DNA from the environment into their cells. The various experiments that show how homology facilitates natural transformation, also illustrate how readily these bacterial species can absorb naked DNA (Nielsen *et al*, 2000a; Nielsen *et al*, 2000b; de Vries *et al*, 2001; de Vries and Wackernagel, 2002). It is likely that some naturally competent bacteria carry elements homologous to artificial vectors because of the broad host-range of the elements to which pUC18 is homologous (Table 1; Matthew and Hedges, 1976).

There is tight homology between pUC18 and mobile elements in *Neisseria* species (Table 1), which are naturally competent.

However, *Neisseria* (and *Haemophilus*) only take up DNA that is connected to specific uptake sequences (Solomon and Grossman, 1996). In *Neisseria* the specific uptake sequence required is 10 nucleotides long and there is no exact match to this in the pUC18 DNA sequence. In other species, such as *Acinetobacter* and *Pseudomonas stutzeri*, DNA uptake is not sequence specific.

Several genes resembling those that encode the cellular machinery required for natural transformation exist in bacterial species that are not known for their natural competence, for example in *E. coli*, *Salmonella typhi*, *Klebsiella pneumoniae* and *Pseudomonas aeruginosa* (Finkel and Kolter, 2001). Some species are also known to be naturally competent in particular situations (Lorenz and Wackernagel, 1994; Demaneche *et al*, 2001b) and so could potentially absorb artificial DNA if it is available in their environment. For example, *E. coli* may become competent in mineral water, when hungry, when in human food, and when in soil that is struck by lightning (Bauer *et al*, 1999; Finkel and Kolter, 2001; Woegerbauer *et al*, 2002; Demaneche *et al*, 2001a), and *Pseudomonas fluorescens* is competent in soil but not *in vitro* (Demaneche *et al*, 2001b).

Could free-living bacteria be exposed to recombinant DNA in the environment? DNA can persist in the environment for thousands

of years (Austin *et al*, 1997), though probably not in a form where the genes it may encode are readily utilizable. Experiments so far have demonstrated the transformation and unharmed activity of genes after a few days in soil (Nielsen *et al*, 1997; Sikorski *et al*, 1998).

There are no recommendations or guidelines on the disposal of recombinant DNA, only on the disposal of living genetically modified organisms (e.g. NIH guidelines for research involving recombinant DNA molecules (2002): Section I-B, Section III-F-1, Appendix G and Appendix K). Recombinant DNA is most likely poured down the sink in large and diverse quantities and genes can stay intact even when heated or chemically treated along with the killed microbial biomass from which they came (Andersen *et al*, 2001).

Natural Selection

Most of the DNA that is inserted into artificial vectors would probably not benefit free-living bacteria. Without a phenotypic benefit to recipient bacteria it is unlikely that non-homologous DNA picked up from artificial vectors would spread or persist (Campbell, 2000; Berg and Kurland, 2002), and it is unclear why such acquisitions would be of any concern. Given the large-scale genome sequencing of deadly microorganisms or the scope of research into virulence, however, there are almost certainly DNA inserts released in artificial

vectors that would have a chance of persistence and would be of concern if acquired by some bacteria.

Even genetic variants that do not confer a beneficial trait sometimes spread among bacteria if linked to an advantageous locus (Guttman and Dykhuizen, 1994). The plasmids and transposons to which artificial vectors have homologies generally confer a whole suite of potentially beneficial traits, for example, multiple drug resistance or pathogenic characteristics (Table 1). Close linkage to such traits increases the chances of natural recombinants persisting in the environment even if newly acquired foreign DNA does not confer an immediate selective advantage on recipient bacteria.

Conclusions

Many factors influence the risk that the natural homology of artificial vectors could facilitate the bacterial uptake of recombinant DNA inserts. Because several of these factors are notoriously difficult to quantify in an open environment, the degree to which this risk is realistic remains unknown. The processes involved occur at low frequencies, but could act on a large number of different molecules, bacteria and potential environmental situations. Given the mobility of the elements to which artificial vectors have homology, the

acquisition of an undesirable trait that is beneficial to bacteria need only happen once for potentially far-reaching consequences. Most DNA that is inserted into artificial vectors would most likely pose no risk to human health or the environment even if acquired by free-living bacteria.

Acknowledgements

Many thanks for helpful comments and criticisms on various incarnations of the manuscript to Daniel L. Hartl, M. Pilar Francino, Richard H. Thomas, Casey M. Bergman, and anonymous reviewers. K.M.N was supported by The Research Council of Norway (MU: biodiversity program) and the Norwegian Institute for Gene Ecology; D.B. was partly supported by a Royal Society – Fulbright fellowship. Part of this work was performed under the auspices of the U.S. Department of Energy, Office of Biological and Environmental Research, by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC0376SF00098.

Table 1 Examples of homology between naturally occurring bacterial DNA and pUC18.

DNA isolated from	*Homology length (bp)*	*% S*[a]	*Description*	*Natural mobile element*	*Comments*	*Accession*
Proteobacteria						
Betaproteobacteria						
Neisseriales						
Neisseria meningitidis strains MC9690-129, and MC9690-130	1209	99.4%	*tnp*R, promoter, bla_{TEM-1}, repeat	Transposon Tn*3* on plasmid pAB6	5.6kb plasmid has a mobilization and a replication protein	AF126482
Neisseria gonorrhoeae strain GC-182	1209	99.4%	*tnp*R, promoter, *bla*, repeat	Transposon Tn*2* on plasmid pJD4	7.4kb plasmid has a mobilization and a replication protein	NC_002098
Gammaproteobacteria						
Pseudomonales						
Pseudomonas aeruginosa unknown strain	1209	99.3%	*tnp*R, promoter, *bla*, repeat	Transposon Tn*1* on plasmid RK2, and RP1	60kb conjugative plasmid, broad host-range IncP-alpha	NC_001621
unknown strain	1209	99.2%	*tnp*R, promoter, bla_{TEM-21}, repeat	Transposon Tn*801*	Chromosomal, linked to IS6100 element	AF466526
unknown strain	>1062	99.8%	*tnp*R, promoter, *bla*	Transposon Tn*3*	Broad host-range transposon	X50604

	>1062	99.4%	*tnp*R, promoter, *bla*	Transposon Tn*2*	Broad host-range transposon	X50607
	>1062	99.2%	*tnp*R, promoter, *bla*	Transposon Tn*1*	Broad host-range transposon	X50606
Xanthomondales						
Stenotrophomonas maltophilia strain J675Ia, UK	>1071	98.8%	*tnp*R, promoter, bla_{TEM-2}	Tn1-like transposon	Mobilizes to broad host-range conjugative plasmids	AJ251946
Enterobacteriales						
Escherichia coli many different strains	861 - 1209	98.7-99.1%	*tnp*R, promoter, *bla*, repeat	Tn*1*, *2* or *3*-like transposons	> 16 different TEM-type β-lactamases.	see below[b]
strain SE53	672[c]	86%[d]	ColE1-like replication origin	Plasmid pUD2380	8.5kb mobilizable resistance plasmid	AJ008006
strain K12 P678-54	629[c]	87%[d]	ColE1-like replication origin	Plasmid pCloDF13	10kb colicinogenic plasmid	NC_002119
Klebsiella pneumoniae many different strains	855 - 1209	99%	*tnp*R, promoter, *bla*, repeat	Tn*1*, *2* or *3*-like transposons	9 different TEM-type β-lactamases. TEM-3 is on multiresistance conjugative plasmid pCFF04 Inc7-M	See below[e]
Klebsiella oxytoca strain ROM	> 1082	99.3%	*tnp*R, promoter, bla_{TEM-12}, repeat	Tn*841* transposon on plasmid pOZ201.	100kb conjugative plasmid.	M88143

strain Kox443	> 862	99.1%	*tnp*R, promoter, bla_{TEM-59}, repeat	Conjugative plasmid	50kb conjugative plasmid.	AF062386
strain 26W	> 1005	99.1%	*tnp*R, promoter, part of bla_{TEM-10}			U09188
strain KH11 subgroup 2	401[c]	83%	ColE1-like replication origin	Plasmid pTKH11	8kb cryptic resistance plasmid	Y17716
strain NBL63 subgroup 3	625[c]	88%	ColE1-like replication origin	Plasmid pNBL63	7kb cryptic resistance plasmid	Y17846
Shigella flexneri strains K24, K25, Hong Kong	>1067	99.3%	*tnp*R, promoter, *bla*	Tn*2* on conjugative plasmid	120kb broad host-range multiresistance plasmid	U48775
strain SH595	1209	99.5%	*tnp*R, promoter, *bla*, repeat	Transposon Tn*SF1*	Tn-21-like transposon	AF188331
Salmonella enteritidis strain YMC4199 serovar Saintpaul, Korea	>836	99%	bla_{TEM-52}	Plasmid pYMC4199		AF126444
Salmonella typhimurium unknown strain	1209 & 545	99% & 95%	*tnp*R, promoter, *bla*, repeat, ColE1-like replication origin	Plasmid pNTP16	8.3kb mobilizable multi-resistance plasmid	NC_002090
Salmonella enterica serovar Borreze	469[c] & 137	87% & 92%	ColE1-like replication origin	Plasmid pWQ799	6.9kb mobilizable plasmid carrying genes for O:54 antigen expression	L39794
Salmonella panama unknown strain	684[c]	94%	ColE1-like replication origin	Plasmid RSF1030	8.3kb resistance plasmid carrying Tn2	J01784

unknown strain			replication origin		plasmid carrying Tn*2*	
Yersinia pestis strain CO-92 Biovar Orientalis	972[c]	95-98%	ColE1-like replication origin	Plasmid pPCP1	9.6kb virulence plasmid essential for plague transmission by flea bight	AL109969
Yersinia enterolitica strain 29807, biogroup 1A	447 [c]	90%	ColE1-like replication origin	Plasmid p29807	2.7kb cryptic plasmid	AJ32618
strain 842 (O:8,19)	>118 & >160 & 38	96.6%, 96.3%, 100%	LacI, flank & β-d-galactosidase & Tn*3* inverted repeat	Transposon Tn*951* on 50kb conjugative plasmid pGC1	16.6kb lac transposon	M82979, M25019
Proteus mirabilis unknown strain	>1084	99.4%	*tnp*R, promoter, bla_{IRT-18}			AJ012256
strain H223b, Tunisia	>885	99.3%	bla_{TEM-1}	Plasmid pCMY-4		Y18200
strain FI-14, Italy	>1008	98.8%	bla_{TEM-72}	Plasmid pPM14	40kb plasmid	AF157413
Morganella morganii strain FFLM15	>1095	98.8%	*tnp*R, promoter, bla_{TEM-10}	Conjugative plasmid	50kb conjugative plasmid	AF093512
strain Mm126	>919	98.9%	bla_{TEM-21}	Plasmid pMm126	40kb conjugative multiresistance plasmid	AF052748
strain FI-13, Italy	>1008	98.8%	bla_{TEM-72}	Plasmid pMN13	40kb plasmid	AF157553
Serratia marcescens strain S5 (epidemic)	>1016	97.5%	bla_{TEM-AQ}	Plasmid pSM5	54kb plasmid	X97254

unknown strain	220	100%	*tnp*R, promoter, part of *bla*	part of Tn*3* on plasmid pAZ007	Multidrug resistance plasmid	M23634
Pantoea agglomerans from a kitchen sink	556[c]	89%[d]	ColE1-like replication origin	Plasmid pPIGDM1	2.5kb plasmid	AF014880
Providencia stuartii strain VR1, Italy	997	98.2%	bla_{TEM-60}	Plasmid pVR-1	20kb plasmid	AF047171
Edwardsiella ictaluri unknown strain	467[c]	84%[d]	ColE1-like replication origin	Plasmid pEI1	4.8kb plasmid	NC_002497

[a] % similarity to pUC18. [b] V00613, AF188200, X57972, U37195, Y17584, AF190692-5, AB049569. [c] BLAST alignment was gapped. [d] weighted mean of sections aligned by *BLA*ST.

pUC18 (Accession: L08752) was used to BLAST completed microbial genomes and all DNA sequences with the organism name “plasmids”. NCBI sequence entries containing keywords such as “*bla* NOT vector” were also compared to pUC18 using pairwise BLAST. To avoid contaminating results from accidental inclusions of vector sequences in GenBank sequence entries, results were excluded if they fulfilled several criteria: (i) they were not diverged from pUC18, (ii) if homology was at the start or the end of a sequence entry, (iii) if this homology was not a subset of a longer region of homology to the sequence of the naturally occurring Tn*3* transposon (Accession: V00613), or ColE1 (Accession: NC_001371), (iv) if reading of publications associated with the DNA sequence revealed that a synthetic vector with homology to pUC18 had been used to generate the DNA sequence.

Figure 1 The natural origins of plasmid vector pUC18. These were summarised using the details given in the GenBank entries of pUC18 (Accession: L08752), cloning vector pBR322 (J01749), cloning vector M13mp18 (X02513), the *E. coli* lactose operon (J01636), transposon Tn*3* (V00613), the ColE1 plasmid (NC_001371) and references therein, and by aligning these sequences with pUC18, using the NCBI "*BLA*ST 2 sequences" facility at <www.ncbi.nlm.nih.gov/*BLA*ST/>.

References

Andersen J, Schafer T, Jorgensen P, Moller S (2001). Using inactivated microbial biomass as fertilizer: the fate of antibiotic resistance genes in the environment. *Res. Microbiol.* **152**: 823-833.

Austin JJ, Smith AB, Thomas RH (1997). Palaeontology in a molecular world: the search for authentic ancient DNA. *Trends Ecol. Evol.* **12**: 303-306.

Bauer F, Hertel C, Hammes WP (1999). Transformation of *Escherichia coli* in foodstuffs. *Syst. Appl. Microbiol.* **22**: 161-168.

Berg OG, Kurland CG (2002). Evolution of microbial genomes: sequence acquisition and loss. *Mol. Biol. Evol.* **19**: 2265-2276.

Berg P, Baltimore D, Brenner S, Roblin RO, Singer MF (1975). Summary statement of the Asilomar conference on recombinant DNA molecules. *Proc. Nat. Acad. Sci. USA* **72**: 1981-1984.

Bowler LD, Zhang Q-Y, Riou J-Y, Spratt BG (1994). Interspecies recombination between the *penA* genes of *Neisseria meningitidis* and commensal *Neisseria* species during the emergence of penicillin resistance in *N. meningitidis*: natural events and laboratory simulation. *J. Bacteriol.* **176**: 333-337.

Campbell AM (2000). Lateral gene transfer in prokaryotes. *Theor. Pop. Biol.* **57**: 71-77.

Cornelis G, Ghosal D, Saedler H (1978). Tn951: a new transposon carrying a lactose operon. *Mol. Gen. Genet.* **160**: 215-224.

Datta N, Kontomichalou P (1965). Penicillinase synthesis controlled by infectious *R* factors in Enterobacteriaceae. *Nature* **208**: 239-241.

Davies J (1994). Inactivation of antibiotics and the dissemination of resistance genes. *Science* **264**: 375-382.

de Vries J, Meier P, Wackernagel W (2001). The natural transformation of the soil bacteria *Pseudomonas stutzeri* and *Acinetobacter* sp. by transgenic plant DNA strictly depends on homologous sequences in the recipient cells. *FEMS Microbiol. Lett.* **195**: 211-215.

de Vries J, Wackernagel W (2002). Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. *Proc. Nat. Acad. Sci. USA* **99**: 2094-2099.

Demaneche S, Bertolla F, Buret F, Nalin R, Sailland A, Auriol P, Vogel TM, Simonet P (2001a). Laboratory-scale evidence for lightning-mediated gene transfer in soil. *Appl. Environ. Microbiol.* **67**: 3440-3444.

Demaneche S, Kay E, Gourbiere F, Simonet P (2001b). Natural transformation of *Pseudomonas fluorescens* and *Agrobacterium tumefaciens* in soil. *Appl. Environ. Microbiol.* **67**: 2617-2621.

DuBose RF, Dykhuizen DE, Hartl DL (1988). Genetic exchange among natural isolates of bacteria: recombination within the *phoA* gene of *Escherichia coli*. *Proc. Nat. Acad. Sci. USA* **85**: 7036-7040.

Feil EJ, Holmes EC, Bessen DE, Chan M-S, Day NPJ, Enright MC, Goldstein R, Hood DW, Kalia A, Moore CE, Zhou J, Spratt BG (2001). Recombination

within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. *Proc. Nat. Acad. Sci. USA* **98**: 182-187.

Feil EJ, Maiden CJ, Achtman M, Spratt BG (1999). The relative contributions of recombination and mutation to the divergence of clones of *Neisseria meningitidis*. *Mol. Biol. Evol.* **16**: 1496-1502.

Feil EJ, Maynard Smith J, Enright MC, Spratt BG (2000). Estimating recombinational parameters in *Streptococcus pneumoniae* from multilocus sequence typing data. *Genetics* **154**: 1439-1450.

Finkel SE, Kolter R (2001). DNA as a nutrient: novel role for bacterial competence gene homologs. *J. Bacteriol.* **183**: 6288-6293.

Gogarten JP, Doolittle WF, Lawrence JG (2002). Prokaryotic evolution in the light of gene transfer. *Mol. Biol. Evol.* **19**: 2226-2238.

Guttman DS, Dykhuizen DE (1994). Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science* **266**: 1380-1383.

Hartl DL, Dykhuizen DE (1984). The population genetics of *Escherichia coli*. *Annu. Rev. Genet.* **18**: 31-68.

Kay E, Vogel TM, Bertolla F, Nalin R, Simonet P (2002). In situ transfer of antibiotic resistance genes from transgenic (transplastomic) tobacco plants to bacteria. *Appl. Environ. Microbiol.* **68**: 3345-3351.

LeClerc JE, Li B, Payne WL, Cebula TA (1996). High mutation frequencies among *Escerichia coli* and *Salmonella* pathogens. *Science* **274**: 1208-1211.

Lorenz MG, Wackernagel W (1994). Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol. Rev.* **58**: 563-602.

Majewski J, Cohan FM (1998). The effect of mismatch repair and heteroduplex formation on sexual isolation in *Bacillus*. *Genetics* **148**: 13-18.

Majewski J, Zawadzki P, Pickerill P, Cohan FM, Dowson CG (2000). Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. *J. Bacteriol.* **182**: 1016-1023.

Matic I, Radman M, Taddei F, Picard F, Doit C, Bingen E, Denamur E, Elion J (1997). Highly variable mutation rates in commensal and pathogenic *Escherichia coli*. *Science* **277**: 1833-1834.

Matthew M, Hedges RW (1976). Analytical isoelectric focusing of R Factor-determined beta-lactamases: correlation with plasmid compatibility. *J. Bacteriol.* **125**: 713-718.

Maynard Smith J, Dowson CG, Spratt BG (1991). Localized sex in bacteria. *Nature* **349**: 29-31.

Nielsen KM, Smalla K, van Elsas J (2000). Natural transformation of *Acinetobacter* sp. strain BD413 with cell lysates of *Acinetobacter* sp., *Pseudomonas fluorescens*, and *Burkholderia cepacia* in soil microcosms. *Appl. Environ. Microbiol.* **66**: 206-212.

Nielsen KM, van Elsas JD, Smalla K (2000). Transformation of *Acinetobacter* sp. BD413(pFG∆*ntpII*) with Transgenic Plant DNA in Soil Microcosms and Effects

of Kanamycin on Selection of Transformants. *Appl. Environ. Microbiol.* **66**: 1237-1242.

Nielsen KM, van Weerelt MDM, Berg TN, Bones AM, Hagler AN, van Elsas JD (1997). Natural transformation and availability of transforming DNA to *Acinetobacter calcoaceticus* in soil microcosms. *Appl. Environ. Microbiol.* **63**: 1945-1952.

NIH guidelines for research involving recombinant DNA molecules (April 2002): http://www4.od.nih.gov/oba/rac/guidelines/guidelines.html

Ochman H, Lawrence JG, Groisman EA (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299-304.

Oliver A, Canton R, Campo P, Baquero F, Blasquez J (2000). High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science* **288**: 1251-1253.

Pouwels PH, Enger-Valk BE, Brammar WJ (1988) *Cloning vectors*. Elsevier

Prudhomme M, Libante V, Claverys J-P (2002). Homologous recombination at the border: Insertion-deletions and the trapping of foreign DNA in *Streptococcus pneumoniae*. *Proc. Nat. Acad. Sci. USA* **99**: 2100-2105.

Rayssiguier C, Thaler DS, Radman M (1989). The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch repair mutants. *Nature* **342**: 396-401.

Rosenau A, Cattier B, Gousset N, Harriau P, Philippon A, Quentin R (2000). *Capnocytophaga ochracea*: characterization of a plasmid-encoded extended-

spectrum TEM-17 beta-lactamase in the phylum *Flavobacter-Bacteroides*. *Antimicrob. Agents Chemother.* **44**: 760-762.

Sikorski J, Graupner S, Lorenz MG, Wackernagel W (1998). Natural genetic transformation of *Pseudomonas stutzeri* in a non-sterile soil. *Microbiol.* **144**: 569-576.

Solomon JM, Grossman AD (1996). Who's competent and when: regulation of natural genetic competence in bacteria. *Trends Genet.* **12**: 150-155.

Sutcliffe JG (1978). Nucleotide sequence of the ampicillin resistance gene of *Escerichia coli* plasmid pBR322. *Proc. Nat. Acad. Sci. USA* **75**: 3737-3741.

Therrien C, Levesque RC (2000). Molecular basis of antibiotic resistance and beta-lactamase inhibition by mechanism-based inactivators: perspectives and future directions. *FEMS Microbiol. Rev.* **24**: 251-262.

Townsend JP, Nielsen KM, Fisher DS, Hartl DL (2003). Horizontal acquisition of divergent chromosomal DNA in bacteria: effects of mutator phenotypes. Genetics **164**:13-21

Vulic M, Dionisio F, Taddel F, Radman M (1997). Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. *Proc. Nat. Acad. Sci. USA* **94**: 9763-9767.

Vulic M, Lenski RE, Radman M (1999). Mutation, recombination, and incipient speciation of bacteria in the laboratory. *Proc. Nat. Acad. Sci. USA* **96**: 7348-7351.

Woegerbauer M, Jenni B, Thalhammer F, Graninger W, Burgmann H (2002). Natural genetic transformation of clinical isolates of *Escherichia coli* in urine and water. *Appl. Environ. Microbiol.* **68**: 440-443.