



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2009

Genomic islands: tools of bacterial horizontal gene transfer and evolution

Juhas, M ; van der Meer, J R ; Gaillard, M ; Harding, R M ; Hood, D W ; Crook, D W

Abstract: Bacterial genomes evolve through mutations, rearrangements or horizontal gene transfer. Besides the core genes encoding essential metabolic functions, bacterial genomes also harbour a number of accessory genes acquired by horizontal gene transfer that might be beneficial under certain environmental conditions. The horizontal gene transfer contributes to the diversification and adaptation of microorganisms, thus having an impact on the genome plasticity. A significant part of the horizontal gene transfer is or has been facilitated by genomic islands (GEIs). GEIs are discrete DNA segments, some of which are mobile and others which are not, or are no longer mobile, which differ among closely related strains. A number of GEIs are capable of integration into the chromosome of the host, excision, and transfer to a new host by transformation, conjugation or transduction. GEIs play a crucial role in the evolution of a broad spectrum of bacteria as they are involved in the dissemination of variable genes, including antibiotic resistance and virulence genes leading to generation of hospital 'superbugs', as well as catabolic genes leading to formation of new metabolic pathways. Depending on the composition of gene modules, the same type of GEIs can promote survival of pathogenic as well as environmental bacteria.

DOI: <https://doi.org/10.1111/j.1574-6976.2008.00136.x>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-24443>

Journal Article

Accepted Version

Originally published at:

Juhas, M; van der Meer, J R; Gaillard, M; Harding, R M; Hood, D W; Crook, D W (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiology Reviews*, 33(2):376-393.

DOI: <https://doi.org/10.1111/j.1574-6976.2008.00136.x>

FEMS Microbiology Reviews, 2009, 33(2), 376-393

<http://www3.interscience.wiley.com/journal/121495628/abstract?CRETRY=1&SRETRY=0>

Editor: Eva Top

The definitive version is available at www.blackwell-synergy.com

Genomic islands: tools of bacterial horizontal gene transfer and evolution

Mario Juhas¹, Jan Roelof van der Meer², Muriel Gaillard², Rosalind M. Harding³, Derek W. Hood⁴ and Derrick W. Crook¹

¹Clinical Microbiology and Infectious Diseases, Nuffield Department of Clinical Laboratory Sciences, University of Oxford, OX3 9DU Oxford, UK. ²Department of Fundamental Microbiology, University of Lausanne, CH 1015, Lausanne, Switzerland.

³Departments of Zoology and Statistics, University of Oxford, OX1 3TG Oxford, UK.

⁴Molecular Infectious Diseases Group, The Weatherall Institute of Molecular Medicine, University of Oxford, OX3 9DS Oxford, UK.

*Corresponding author: Mario Juhas, PhD, Clinical Microbiology and Infectious Diseases, Nuffield Department of Clinical Laboratory Sciences, University of Oxford, Headington, OX3 9DU Oxford, UK, Tel: (0044) 01865 222194, Fax: (0044) 01865 222195, Email: mario.juhas@ndcls.ox.ac.uk, Current email: mario.juhas@access.uzh.ch

Running title: Genomic islands and horizontal gene transfer

Key words: horizontal gene transfer/ genomic island/ evolution/ pathogenicity/
biodegradation

ABSTRACT

Bacterial genomes evolve through mutations, rearrangements or horizontal gene transfer. Besides the core genes encoding essential metabolic functions, bacterial genomes also harbor a number of accessory genes acquired by horizontal gene transfer that usually encode adaptive traits that might be beneficial under certain environmental conditions. The horizontal gene transfer contributes to the diversification and adaptation of microorganisms, thus having a significant impact on the genome plasticity. A significant part of the horizontal gene transfer is facilitated by genomic islands. Genomic islands are horizontally acquired segments of DNA capable of integration into the chromosome of the host, excision, and transfer to a new host by transformation, conjugation or transduction. Genomic islands play a crucial role in the evolution of a broad spectrum of pathogenic and environmental bacteria as they are involved in the dissemination of a number of genes, including antibiotic resistance and virulence genes leading to generation of hospital “superbugs”, as well as dissemination of catabolic gene clusters and the formation of new metabolic pathways. Work over the past few years showed that depending on the composition of gene modules, the same type of genomic islands can promote survival of pathogenic as well as environmental bacteria.

INTRODUCTION

Genomes of bacterial species can evolve through a variety of processes including mutations, rearrangements or horizontal gene transfer. Information gathered over the past few years from a rapidly increasing number of sequenced genomes has shown that besides the core genes encoding essential metabolic functions, bacterial genomes also harbor a variable number of accessory genes acquired by horizontal gene transfer that encode adaptative traits that might be beneficial for bacteria under certain growth or environmental conditions (Schmidt & Hensel, 2004).

Many of the accessory genes acquired by horizontal transfer form syntenic blocks recognized as genomic islands (GEIs). The formation of GEIs contributes to the diversification and adaptation of microorganisms, thus having a significant impact on the genome plasticity and evolution, the dissemination of antibiotic resistance and virulence genes, and formation of catabolic pathways. As more and more DNA elements are being detected by bacterial genome sequencing projects, and, at the same time, more information has become available on their life style, this review will summarize recent advances in our understanding of their distribution, evolution and mechanistic modes of behavior. One big class of highly conserved integrative and conjugative elements (ICEs) occurring in β - and γ -*Proteobacteria* will be particularly highlighted.

GENERAL FEATURES OF GENOMIC ISLANDS

GEIs are probably the latest addition to the types of elements comprising the horizontal gene pool. The concept of pathogenicity islands (PAIs) was originally coined in the late

1980s by J. Hacker and colleagues who investigated the genetic basis of virulence of uropathogenic isolates of *Escherichia coli* (UPEC) (Hacker, *et al.*, 1990). PAIs found in their study were unstable chromosomal regions with variable virulence-associated characteristics and phenotypes (Groisman & Ochman, 1996). Nowadays it is appreciated that GEIs represent a much broader and more diverse group of DNA elements than PAIs only, with a large variety of sizes (20-500 kb) and abundance in bacterial genomes (Dobrindt, *et al.*, 2004). Different GEI families, some probably evolutionary very ancient, have been recognized on the basis of predicted sequence and functional homologies (Burrus, *et al.*, 2002, Juhas, *et al.*, 2007b, Vernikos & Parkhill, 2008). The coding capacity of GEIs is not limited to pathogenicity functions, but can be very diverse, including such traits as symbiosis (Sullivan, *et al.*, 2002), sucrose and aromatic compound metabolism (Gaillard, *et al.*, 2006), mercury resistance or siderophore synthesis (Larbig, *et al.*, 2002b). Bioinformatics studies have shown that GEIs tend to carry more “novel” genes (i.e., those that do not have orthologues in other species) than the rest of the genome (Hsiao, *et al.*, 2005). This suggests that GEIs have become strongly selected for adaptive and auxiliary functions. The fact that GEIs come in a large spectrum of varieties in terms of genetic organization and functionality makes it more difficult to provide an exact definition of a GEI. Here we propose that the term GEI should be used for the overarching family of discrete “DNA elements”, which are part of a cell’s chromosome and can drive or have driven strain differentiation.

Most GEIs known to date share following features (Fig. 1):

(1) GEIs are relatively large segments of DNA, usually in the range between 10 and 200 kb detected by comparisons among closely related strains. Discrete DNA regions

detected by comparative genome sequencing with sizes smaller than 10 kb have been named genomic islets (Hacker & Kaper, 2000).

(2) GEIs may be recognized by nucleotide statistics (e.g., GC content, cumulative GC skew, tetranucleotide frequencies or codon usage) that usually differ from the rest of the chromosome.

(3) GEIs are often inserted at tRNA genes.

(4) GEIs are often flanked by 16 to 20 bp perfect or almost perfect direct repeats (DR). DRs usually arise by the site-specific integration of the GEIs into the target site and can act as recognition sequences for their enzymatic excision (Schmidt & Hensel, 2004).

(5) GEIs often harbor functional or cryptic genes encoding integrases or factors related to plasmid conjugation systems or phages involved in GEI transfer.

(6) GEIs often carry insertion elements or transposons, which may have been implicated in mobilizing genetic material onto or deleting DNA from the element. (Buchrieser, *et al.*, 1998, Gal-Mor & Finlay, 2006).

(7) GEIs often carry genes offering a selective advantage for host bacteria. According to their gene content, GEIs are often described as pathogenicity, symbiosis, metabolic, fitness or resistance islands (Dobrindt, *et al.*, 2004, Schmidt & Hensel, 2004).

EVOLUTIONARY ORIGINS OF GENOMIC ISLANDS

Although most of the GEIs known so far fit within the above described definition, a significant number of elements lack one or more of the hallmark indications. Since many GEIs have only been put forward by data from sequencing projects, but not by experimental evidence on their mode of action, many elements may actually be in an evolutionary state of regression, as has been suggested earlier by Dobrindt *et al.*

(Dobrindt, *et al.*, 2004). Sequence and phylogenetic comparisons show that GEIs tend to fall within structurally distinct genus-specific families, but looking across different genus boundaries leads to emergence of universally distributed structural GEI components (Vernikos & Parkhill, 2008). This suggests that GEIs may have arisen multiple times independently during evolution, and can thus only be seen as a superfamily of mobile elements on the basis of analogous core and structural features, rather than being phylogenetically related (Juhas, *et al.*, 2007b, Vernikos & Parkhill, 2008). As the original definition of GEIs was put forward 10 years ago when only 12 complete bacterial genomes were available, it is plausible to assume that other GEIs with novel and unusual features will be discovered with the increasing number of bacterial genomes being sequenced. In the more overarching view we propose here, GEI would encompass other categories of elements, such as ICE/conjugative transposons (which were proposed to be one functionally similar group of integrative and conjugative elements (Burrus, *et al.*, 2002, Burrus & Waldor, 2004)), integrated plasmids, non-replicative but excisable elements (NBU from *Bacteroides* (Shoemaker, *et al.*, 2000)), and perhaps even cryptic or damaged prophages (Fig. 2).

Work over the last few years has suggested that GEIs have had multiple parallel evolutionary origins, as some of them clearly contain phage and others conjugative plasmid-related genes. In addition, it has become clear that mobile elements form by combination of functional modules, which makes it more difficult to categorize them (Toussaint & Merlin, 2002, Frost, *et al.*, 2005, Leplae, *et al.*, 2006). For example, as it is common to find multiple prophages in various stages of functionality in a bacterial chromosome, it is plausible to assume that some have cointegrated with the

chromosome and slowly acquired other genes of non-phage origin (Dobrindt, *et al.*, 2004, Klockgether, *et al.*, 2004). Indeed, for a number of GEIs, the similarities between GEI-encoded and known plasmid transfer systems are so striking that one can rightfully assume that the plasmid transfer genes formed the origin for the progenitor hybrid GEI element. This is the case, for example, for the *Mesorhizobium loti* R7A symbiosis island (Sullivan, *et al.*, 2002, Ramsay, *et al.*, 2006), for the CTn4371 biphenyl transposon of *Ralstonia oxalatica* (Toussaint, *et al.*, 2003), and the SXT element of *Vibrio cholerae* (Beaber, *et al.*, 2002). This need not necessarily be the case for all GEIs, however. Recent experiments with a family of ICE/GEIs with deep evolutionary relationships suggested that some GEIs may have evolutionary very ancient self-transfer modes, for which no current plasmid relatives are known. This conclusion was drawn from observations for three members of this family, namely ICEHin1056 from *Haemophilus influenzae*, pKLC102 from *Pseudomonas aeruginosa* and ICEclc from *Pseudomonas* sp. strain B13. These three GEIs are examples of what is known as ICE, fully functional elements, capable of integration into the chromosome of the host, excision, and self-transfer by conjugation to a new host and reintegration (Dimopoulou, *et al.*, 2002, van der Meer & Sentschilo, 2003, Gaillard, *et al.*, 2006, Klockgether, *et al.*, 2007, Juhas, *et al.*, 2007a). As will be outlined further below, their self-transfer system forms a distinct type IV secretion system (T4SS), which is only very distantly related to other known plasmid-encoded T4SSs.

LIFE-CYCLES OF GENOMIC ISLANDS

Since not all GEIs have the same components, it is difficult to speak of one unifying GEI life cycle. Interestingly, a large number of GEIs, for which self-mobility has been

demonstrated, can excise from their chromosomal location, encode the full capacity for horizontal self-transfer to another cell, and can reintegrate into the target site in the new host's chromosome. GEIs that exhibit simultaneously all these features and self-transfer by *conjugation* are part of an increasingly well defined group of elements that have been named Integrative Conjugative Elements (ICEs) (Burrus & Waldor, 2004). ICEs also include conjugative transposons, a terminology that had mostly been used for elements that originated in Gram-positive bacteria, and, according to some authors, should be reserved for elements which can target multiple different integration sites (Burrus, *et al.*, 2002). Since some GEIs do not self-transfer by conjugation but by phage packaging, release and infection, they can not be called ICEs.

The schematic life cycle for mobile GEI would thus consist of the following steps (Fig. 3):

- (1) Acquisition of the GEI by a host through horizontal gene transfer.
- (2) Integration of the GEI into the host chromosome by site-specific recombination.
- (3) Development of the GEI by genetic rearrangements, gene loss or acquisition of other mobile genetic elements.
- (4) Excision of the GEI from the chromosome.
- (5) Transfer of the GEI to another recipient.

A number of active GEIs have been studied to various levels of their life-style, coding capacity or regulation of transfer, and thus contributed to the general mechanistic modes of GEIs. For reasons only partially understood, GEIs are often inserted in the 3' end of tRNA genes (Ravatn, *et al.*, 1998b, Williams, 2002, Larbig, *et al.*, 2002b). Insertion is catalyzed by site-specific phage-like recombinases called integrases, which are usually

encoded by the GEI itself. Integrases are not strictly conserved among all GEIs, and various different tRNA genes can be targeted (Williams, 2002). Several GEI-encoded integrases relate to the lambda, P4 or XerD families (van der Meer, *et al.*, 2001, Burrus, *et al.*, 2002, Mohd-Zain, *et al.*, 2004). The integrases are also implicated in the excision of GEIs, a process that may be assisted by an excisionase (Burrus & Waldor, 2003, Lesic, *et al.*, 2004, Ramsay, *et al.*, 2006). The *int* gene encoding integrase is often situated at one extremity of the island and adjacent to the tRNA gene in the integrated GEI form. Following excision, both GEI ends close up to form a single copy of the recombination site (*attP*) (Ravatn, *et al.*, 1998a, Doublet, *et al.*, 2005, Qiu, *et al.*, 2006, Ramsay, *et al.*, 2006). Excision is mostly non-replicative and a single copy of the recombination site is again formed on the chromosome (*attB*) (Burrus & Waldor, 2003, Ubeda, *et al.*, 2007). GEI can reintegrate from the excised form back into the *attB*-site of the same host, or, after transfer to a new cell, in a suitable *attB*-site in that one. Integration is the result of a site-specific recombination between a 15 to 20 bp motif within the 3' extremity of the tRNA gene (Ravatn, *et al.*, 1998b, Williams, 2002). Except in a few cases (Klockgether, *et al.*, 2004, Ubeda, *et al.*, 2007), excised GEIs don't seem to replicate independently from the host's chromosome but rely on reintegration or horizontal transfer to a new host and reintegration to proliferate. Intermediate forms may also exist, as was demonstrated in a recent study on SaPIs (staphylococcal pathogenicity islands) which are phage-type GEIs. SaPIs, that have intimate relationships with certain temperate phages involving phage-induced excision, replication and packaging, were also proven to be capable of an existence in a self-replicating plasmid-like state (Ubeda, *et al.*, 2007, Ubeda, *et al.*, 2008). A large body of

knowledge exists on integration and excision processes of conjugative transposons in e.g., *Bacteroides*, the details of which, however, lay outside the scope of this review.

TRANSFER OF GENOMIC ISLANDS BETWEEN BACTERIA

As suggested above: a wide variety of GEIs are intimately connected to phages and conjugative plasmids through evolutionary origins. As a consequence, besides transformation, their transfer proceeds often via conjugation and transduction (Jain, *et al.*, 2002, Chen, *et al.*, 2005). GEIs do not necessarily encode the whole process of self-transfer, several cases are known in which GEIs can be packaged by another co-residing lysogenic phage or mobilized by a plasmid or ICE conjugative system (Shoemaker, *et al.*, 2000).

Transformation has been observed in a number of Gram-positive or Gram-negative, pathogenic or environmental bacteria. Naturally transformable bacteria like *Acinetobacter* sp. ADP1, *Bacillus subtilis*, *Streptococcus pneumoniae*, *Neisseria gonorrhoeae*, *Pseudomonas stutzeri*, *Ralstonia solanacearum* or *Haemophilus influenzae* take up free DNA from their surrounding environment (Smith, *et al.*, 1995, Hamoen, *et al.*, 2001, Barbe, *et al.*, 2004, Meier & Wackernagel, 2005, Fall, *et al.*, 2007). Part of the foreign DNA acquired by natural transformation is degraded, but part could be incorporated into the host's genome thus contributing to evolution of bacterial species. Naturally transformable bacteria need to acquire a physiological state called "competence" through expression of a number of genes encoding DNA uptake and processing systems prior to transformation. Uptake systems of naturally transformable bacteria are often composed of components resembling subunits of the type IV pili and

type II secretion systems (Chen, *et al.*, 2005). Several bacteria, including *N. gonorrhoeae* and *H. influenzae* have a strong bias to take up species-specific DNA for transformation. Efficient species-specific DNA uptake in *N. gonorrhoeae* and *H. influenzae* requires the presence of an approximately 10 nucleotide long DNA uptake sequence, that is found at a high frequency (approximately 1400 copies) in the respective genomes (Goodman & Scocca, 1988, Smith, *et al.*, 1995). Recently a novel type IV secretion system (T4SS) has been identified in *N. gonorrhoeae* that secretes chromosomal DNA in the surrounding environment in a non-contact dependent manner (Hamilton, *et al.*, 2005). This T4SS is localized in the large, horizontally acquired gonococcal genetic island (GGI), present in the chromosome of *N. gonorrhoeae*; thus by facilitating chromosomal DNA secretion this GEI also encodes the mechanism of its own dissemination. Interestingly, the SOS response to antibiotic stress-induced DNA damage has been shown to induce genetic transformability of bacteria and hence to promote horizontal dissemination of antibiotic resistance genes (Beaber, *et al.*, 2004, Prudhomme, *et al.*, 2006).

Conjugation is the process of DNA transfer from donor to recipient through a specialized apparatus that consists of a cell-envelope spanning translocation channel joined to a tube-like structure known as a pilus in Gram-negative bacteria or to the surface-associated adhesins in Gram-positive bacteria (Chen, *et al.*, 2005). Conjugation systems constitute part of a large and versatile family of type IV secretion (T4SS)-dependent transport systems (Chen, *et al.*, 2005, Christie, *et al.*, 2005). T4SSs are usually encoded by multiple genes organized into a single operon. Based on the organization of genetic determinants, shared homologies and evolutionary relationships,

T4SSs have been classified into several types (Juhas, *et al.*, 2008). Type F and P T4SSs, often referred to as type IVA systems, resemble the archetypal VirB/VirD4 system of *Agrobacterium tumefaciens* and are considered to be the paradigm of type IV secretion. Type I T4SSs, also referred to as type IVB systems, resembling the archetypal Dot/Icm system are found in two intracellular bacterial pathogens, *Legionella pneumophila* and *Coxiella burnetii* (Segal, *et al.*, 2005). The most recently described T4SSs that are evolutionarily distant from all previously described, GI T4SSs, play a key role in the horizontal transfer of a wide variety of GEIs derived from a broad spectrum of bacteria, including *Haemophilus* spp., *Pseudomonas* spp., *Erwinia carotovora*, *Salmonella enterica* serovar Typhi, *L. pneumophila* and others (Juhas, *et al.*, 2007a, Juhas, *et al.*, 2007b, Juhas, *et al.*, 2008). This lineage of T4SSs has been named GI-type in order to emphasize the fact that it was found to be associated only with certain GEIs (Juhas, *et al.*, 2007a).

Transduction is the process of DNA transfer from one bacterium to another via bacterial viruses, bacteriophages. Many bacteriophages are able to transfer bacterial genes, including GEIs, as passengers in their genomes. One good example is the SaPI family of *Staphylococcus aureus* islands (Maiques, *et al.*, 2007). Members of this family were shown to be induced to excise and replicate by certain resident temperate phages that also play a role in their packaging into a small phage-like particles (Ubeda, *et al.*, 2003, Maiques, *et al.*, 2007, Ubeda, *et al.*, 2007) that are transferred from donor to recipient cells at frequencies commensurate with the plaque-forming titer of the phage (Ruzin, *et al.*, 2001). Interestingly, SaPI_{bov2}, a member of the SaPI family of GEIs, has been shown to be induced to replicate by different staphylococcal phages, encapsidated and

transferred to a variety of recipient bacteria, including different *Staphylococcus* strains (Maiques, *et al.*, 2007). Certain genomic regions of staphylococci rather represent slightly deteriorated prophages that could be mobilized by other phages. GEI transfer may mechanistically resemble certain aspects indicative of an existence of phage- and plasmid-like ancestors: the transfer of such islands may include self replicating plasmid-like states (Ubeda, *et al.*, 2007). Several other examples of GEIs from a wide variety of bacterial species have been reported recently to be transferred by bacteriophages, including HPI (Yersinia high-pathogenicity island) of *Yersinia pseudotuberculosis* (Lesic, *et al.*, 2004) or GEIs of the marine cyanobacterium *Prochlorococcus* (Coleman, *et al.*, 2006).

Active GEI transfer has been described in a number of cases, although for the great majority of potential GEIs detected by genome sequencing projects, no information is as yet available on their transfer capacities. For example, conjugative transfer of ICEHin1056, calculated as a number of transconjugants divided by a number of recipients, proceeds at a frequency of around 10^{-1} to 10^{-2} between two *H. influenzae* strains (Juhas, *et al.*, 2007b). ICEclc of *Pseudomonas* sp. strain B13, distant member of the same ICEHin1056 subfamily, is self-transferable at similar frequencies to *P. putida*, *Cupriavidus necator* or *P. aeruginosa* (Gaillard, *et al.*, 2006). Recently another member of this ICEHin1056 subfamily of GEIs, PAPI-1 (*Pseudomonas aeruginosa* pathogenicity island-1) was shown to transfer from donor strain into *P. aeruginosa* recipient strains that do not harbor this island naturally. Furthermore, PAPI-1 has been demonstrated to exist in an extrachromosomal circular form and to reintegrate into the host genome following excision from the chromosome (Qiu, *et al.*, 2006). Finally, also

the GEI pKLC102 of *P. aeruginosa* was shown to be highly mobile and capable of transfer at the frequencies similar to *clc* or ICE*Hin1056* (Klockgether, *et al.*, 2007). Interestingly; however, no excision or transfer have been demonstrated for other members of this subfamily of GEIs, like PAGI-2 and PAGI-3 (Klockgether, *et al.*, 2007).

Recent research indicates that the host background has a strong influence on the transferability of GEIs. The T4SS transfer module was found to be one of the most conserved parts of the ICE*Hin1056* subfamily of GEIs (Juhas, *et al.*, 2007b). However, the analysis of the conjugation transfer frequencies showed that these GEIs were transferred between closely related haemophili strains with significantly different frequencies. To test whether these differences were either due to the respective host strain background or to the gene content of the island, haemophilis' GEIs were transferred from the same donor into the same recipient strain. In this setting, the conjugal transfer frequencies were almost constant, which is indicative of the different host strains introducing variations in the conjugal efficiency in the initial experiment (Juhas, *et al.*, 2007b). Interestingly, also transfer frequencies of ICE*clc* were largely dependent on the type of donor cell, even within closely related strains of *P. aeruginosa* (Gaillard, *et al.*, 2008). These studies have clearly shown that host background has a tremendous impact on the transferability of GEIs.

REGULATION OF GENOMIC ISLANDS AND ADAPTIVE BEHAVIOUR

Still only little information is available on the regulation of or environmental conditions influencing GEI transfer. In fact, for most GEIs it was initially assumed that transfer

would be “spontaneous” or “constitutive”. However, it has been suggested that in a number of cases, tightly regulated events underlay the onset of GEI self-mobilization (Jain, *et al.*, 2003). This is not so surprising, especially given the high level of control on the life-style of temperate bacteriophages. A better understanding of the conditions for self-transfer, the factors augmenting or decreasing transfer rates, and the (self-) regulation of the onset of the transfer process is of great importance, both for our appreciation of the impact of horizontal gene transfer on the evolution of microorganisms and for the practical purpose of judging the potential distribution of transgenes or antibiotic resistance genes in natural microbial populations (Nielsen & Townsend, 2004). Evidence from the few GEI models studied so far suggests indeed quite the opposite as “constitutive” behavior: a variety of regulatory modes and signals which determine GEI self-transfer.

One of the most striking regulatory modes for GEI-behavior was revealed from studies on the tetracycline-determinants in *Bacteroides*. Conjugative transfer of the ICEs CTnDOT and CTnERL from *Bacteroides* is stimulated up to 10000-fold when cells are grown in the presence of tetracycline. This effect was found to be the result of an induction of two regulatory genes, *rteA* and *rteB*, both of which stimulate transcription of a third factor *rteC* that influences excision of the element (Cheng, *et al.*, 2001, Whittle, *et al.*, 2002). From the work on ICE*clc* of *Pseudomonas* sp. strain B13, we know that *clc* transfer is strongly enhanced in stationary phase in a bistable fashion (e.g., only $\approx 5\%$ of all cells engage in transfer). Transfer of *clc* correlates to an increase of expression from the *intB13* integrase gene, which is stimulated by the product of the gene *inrR* (Sentchilo, *et al.*, 2003b). Upon excision and formation of a circular

intermediate, a strong promoter - otherwise located at the other end of ICE_{Ecl} facing outwards, is placed in front of the *intB13* gene favouring the re-integration process (Sentchilo, *et al.*, 2003b). Transfer of the 108-kb GEI called PAPI-1, which was discovered in *P. aeruginosa* strain PA14 and is similar to the element pKLC102, proceeds via excision, formation of an intermediate circular form and reintegration into either of the two *tRNA^{Lys}* genes in *P. aeruginosa* (PA4541 and PA0976) (Qiu, *et al.*, 2006). The authors of this work could demonstrate that a *soj* gene encoded by PAPI-1 itself was required for the maintenance of the element, both in integrated and in circular form (Qiu, *et al.*, 2006). Their hypothesis was that Soj protects the circular form of PAPI-1 either directly from degradation or indirectly by promoting the integration of the circular form back into the chromosome. The major evidence for this was the finding that *soj* is expressed by the circular form at early stationary phase. The Soj protein is related to the ParA family of proteins, which are responsible for correct segregation of low-copy plasmids during cell division.

Transfer regulation proceeds differently in the SXT element of *V. cholerae*. Excision of SXT is favored by an excisionase Xis (Burrus & Waldor, 2003), which; however, also inhibits its integration. In SXT, *xis* and *int* are convergent genes that do not appear to be co-regulated. SXT transfer is strongly enhanced under stress conditions and is dependent on the SOS response (Beaber, *et al.*, 2004). Interestingly, the *V. cholerae* SOS response is eluded in particular by two antibiotics, ciprofloxacin and trimethoprim, for which the SXT element encodes resistance determinants. The mechanism is thought to proceed as follows: in the presence of an SOS stimulus, the SXT-encoded repressor SetR is cleaved, resulting in the expression of two SXT-encoded genes *setC* and *setD*, which are activators for the *int* and *tra* genes of the element (Burrus & Waldor, 2003).

Also excision of *ICEBsI*, a mobile element found in the genome of *Bacillus subtilis*, is stimulated by global DNA damage in addition to an intercellular peptide signaling. This behavior was found to be dependent on the factor ImmR, which regulates expression of a number of *ICEBsI* genes and is responsible for immunity to superinfection (Auchtung, *et al.*, 2007).

Excision of the *ICEMISymR7A* symbiosis island of *Mesorhizobium loti* strain R7A is also stimulated by a novel recombination directionality factor called RdfS, which is encoded by the gene *msi109* (Ramsay, *et al.*, 2006). Transfer of the *ICEMISymR7A* also requires a putative relaxase, RlxS. The genes *rdfS* and *rlxS* are part of the same cluster of which two other genes are homologues to the conjugative protein TraF (Ramsay, *et al.*, 2006). Similar to the *clc* element, also the excised form of *ICEMISymR7A* was more abundant in stationary than exponential phase of *M. loti* and experimental evidence suggested that this excision was under quorum-sensing control (Ramsay, *et al.*, 2006).

Another example for excisionase requirement is the integrase of the *Yersinia* high-pathogenicity island (HPI), which cannot alone promote efficiently the excision of HPI. In this case, it was demonstrated that a factor called Hef RDF (recombination directionality factor), which is encoded by HPI, is required for excision (Lesic, *et al.*, 2004). Although the level of *hef* expression severely affected the rate of HPI excision, it had little or no effect on *int* transcription, and the authors concluded that Hef could not act as transcriptional regulator. Another RDF called Rox (for Regulator of excision) was shown to stimulate excision of the *Shigella* resistance locus pathogenicity island in *Shigella flexneri* (Luck, *et al.*, 2004). Similarly to Hef, this Rox protein is not an activator of *int* transcription, although it showed 66% sequence similarity to AlpA from the phage CP4-57. AlpA is a transcription factor for the *int* gene of the phage CP4-57 in

E. coli K-12, regulating the excision of the prophage from the bacterial chromosome (Trempey, *et al.*, 1994). The examples described above suggest that GEI transfer is a highly regulated process with a variety of developed regulatory modes.

As discussed below, GEIs play an important role in bacterial genome evolution in general and in adaptation to changing conditions, in either clinical, industrial or natural environments. The outcome of these adaptations is obvious from the development of antibiotic resistancies, pathogenicity or catabolic functions. Thus it would be extremely interesting to find specific features which make GEIs so successful, either in self-transfer, host-entry or establishment in a host. Are GEIs considered as parasites by a new host? Do GEIs have means to trick a host and avoid its defense systems? Are all GEIs alike in this respect or do exceptions exist? Relatively little information is available concerning these questions and the nature of GEI-host interactions in general.

Most information on the host-DNA invader interactions comes from conjugative plasmids and phages. In general, the frequency of successful DNA exchange between bacteria belonging to different genera will depend on many factors: the degree of homology between the transferred DNA and the bacterial host, the metabolic compatibility, adaptations to their abiotic environment, gene expression systems, gene-transfer mechanisms, the mismatch repair and restriction endonuclease systems. For example, the transfer efficiency of the broad-host range IncP-1 plasmid RK2, as measured by the number of transfer events per donor present, was dramatically influenced by the nature of the donor-recipient combination: between *E. coli* strains or from *E. coli* to *P. putida* RK2 transfer was much less frequent than between *P. putida*

strains. This was attributed to species-specific differences in RK2 gene expression (Bingle & Thomas, 2001). The recipient cell can also limit the entry or establishment of the incoming DNA by surface exclusion, a process by which a barrier seems to be created by cells that already carry the genes for a closely related transfer apparatus (Frost, *et al.*, 1994). Horizontally acquired DNA that confers a selective advantage to the host obviously has the potential to spread further among suitable recipients within a bacterial population under the appropriate selective conditions (Thomas & Nielsen, 2005).

Upon successful transfer, the newly incoming DNA must still be maintained in order to ensure its long-term survival in the new host. If there is a clear selective advantage conferred by the acquired DNA for the host, and no major fitness cost under non-selective conditions, it is less likely that the horizontally acquired DNA will be lost. Incoming DNA on a plasmid must be able to replicate independently but synchronously to the host's chromosomal replication and cell division. This process is usually guaranteed by a plasmid-specific system, such as killing daughter cells without partitioned plasmids. In the absence of appropriate replication, plasmids may still 'survive' by recombining into the host's chromosome, which is promoted by the presence of suitable sequences for homologous recombination. On the other hand, GEIs depend on site-specific reintegration into an appropriate chromosomal target site, in the absence of which the element is unlikely to be maintained. However, once integrated, a GEI will be automatically maintained by chromosomal replication. Loss can occur when the GEI excises as there seems to be a strong advantage for growth of cells without GEIs. Such a scenario was recently demonstrated in a study on the UPEC *E. coli* isolate 536 which contains five PAIs, some of which can become deleted during

chronic infection. The authors could show that the integrase of one of the PAIs could actually mediate the excision of another PAI. This suggests an existence of a unidirectional cross-talk between integrases of different PAIs (Hochhut, *et al.*, 2006). Similar cross-talk phenomena had previously been detailed by work on the *Bacteroides* element NBU1 (non-replicating B*acteroides* unit), which can be co-excised and mobilized by the conjugative transposon CTnERL (Shoemaker, *et al.*, 2000).

Although it is generally stated that GEIs give a selective advantage to the host cell (Dobrindt, *et al.*, 2004), this has not been extensively experimentally tested yet. In fact, results from conjugative plasmids would suggest the opposite, namely that cells which acquire a conjugative plasmid go through a period of fitness loss (Dahlberg & Chao, 2003). This fitness loss can be caused by particular functions carried on the plasmid which are detrimental upon expression in the new host (e.g., regulatory proteins interfering with the global gene expression network), or by proteins expressed from plasmid which have a direct phenotypic effect (e.g. antibiotic resistance) (Nguyen, *et al.*, 1989). Bacterial populations having received a newly incoming plasmid were shown to adapt to their previous fitness by spontaneous mutations which reduce or repress the plasmid specific effects (Dahlberg & Chao, 2003). Plasmids with both, very low to unmeasurable, and very large cost to the host were detected and their ability to persist in a bacterial population was shown to be influenced by the host strain background (De Gelder, *et al.*, 2007, Schluter, *et al.*, 2007). Very recently it was demonstrated that ICE_{ec} had only a very minor effect on fitness of host *P. aeruginosa* strains (Gaillard, *et al.*, 2008). Although this was a single case so far, it might point to this type of ICE/GEI having specific mechanisms to reduce fitness cost in the host (Gaillard, *et al.*, 2008).

There is increasing evidence that GEI expression is globally influenced by the host in which it resides. Thus GEIs usually constitute part of global regulatory networks and genes on the particular GEI can be regulated by regulators present on the same GEI, by regulators harbored by other GEIs or by regulators encoded by the host bacterium. Similarly, GEI-borne regulators often play a role in the regulation of genes on the bacterial chromosome.

Regulators frequently contributing to regulation of GEIs comprise the two-component response regulator family, AraC family, alternative sigma factors and histone-like proteins. Paradigmatic regulatory networks involving intensive cross-talk between GEI-borne regulators and gene components of the host genome include *Salmonella* pathogenicity islands (SPI-1 and SPI-2) of *Salmonella enterica*, *Vibrio* pathogenicity island (VPI) of *Vibrio cholerae* and the LEE of *Escherichia coli* and have been reviewed elsewhere (Hacker & Kaper, 2000, Schmidt & Hensel, 2004).

Recent studies into the regulatory role of the histone-like nucleoid structuring protein (H-NS) have shed a new light into the regulation of some bacterial GEIs. H-NS is a pleiotropic regulator that modulates gene expression of Gram-negative bacteria in response to environmental stimuli, such as temperature and osmolarity (Hommais, *et al.*, 2001). H-NS represents the bacterial functional equivalent of histones, plays an important role in a local supercoiling of DNA and has higher affinity for curved DNA (Navarre, *et al.*, 2006). Work from several laboratories exploiting recent DNA microarray technology has shown that H-NS plays a key role in the selective silencing of horizontally acquired genes (Lucchini, *et al.*, 2006, Navarre, *et al.*, 2006). Expression of more than 400 genes was shown to be upregulated in the *Salmonella hns* mutant, out of which more than 90 % were acquired by the horizontal gene transfer (Navarre, *et al.*,

2006). The GC content of most of the H-NS repressed genes was lower than the average GC content of the host *Salmonella* genome, thus leading to the conclusion that H-NS can selectively silence horizontally acquired genes by targeting sequences with proportionally high AT content (Navarre, *et al.*, 2006). Recently conserved sequence motifs have been identified that represent the high-affinity DNA binding sites for H-NS (Lang, *et al.*, 2007). In accordance with previously published studies, these motifs occur in AT rich regions of DNA both within operons and in genes harboured by the pathogenicity associated GEIs (Lang, *et al.*, 2007). As described above, the differences in the nucleotide statistics are among the characteristic features of GEIs, thus targeting sequences with different GC content by H-NS represent an elegant and efficient mechanism of regulation of newly acquired GEIs.

CONTRIBUTION OF GENOMIC ISLANDS TO HORIZONTAL GENE TRANSFER AND BACTERIAL EVOLUTION

It is widely recognized that horizontal gene transfer facilitated by GEIs has played a crucial role in the evolution of bacterial species. This is attributed not only to simple acquisition and loss of GEI-borne genes, but also to the possibility of GEIs transferring parts of a host's chromosomal DNA. Upon excision from the host genome, GEIs can play a role in the transfer of parts of the host chromosome into the recipient bacteria (Hochhut, *et al.*, 2000). Furthermore, the presence of a wide variety of secretion systems on many GEIs suggests that these can be used not only for the transfer of GEIs and the GEI-encoded products, but also for the transfer of the host's chromosomal DNA. One good example is the secretion of chromosomal DNA in the gonococcus via the GGI- encoded T4SS (Hamilton, *et al.*, 2005). Chromosomal DNA secreted via the

GGI-encoded type T4SS may be subsequently taken up by natural transformation, thus facilitating recombination that contributes to antigenic variation and the spread of antibiotic resistance (Hamilton, *et al.*, 2005). GEIs could undergo a recombination with the host's chromosome having a significant impact on the evolution of the host bacterium as GEI-borne genes often encode important clinical or fitness traits. Whether encoding genes involved in pathogenicity or biodegradation, GEIs can facilitate evolution by “quantum leaps” as their acquisition or loss can rapidly and dramatically alter the lifestyle of a bacterium (Groisman & Ochman, 1996, Ochman, *et al.*, 2000, van der Meer & Sentschilo, 2003, Vernikos, *et al.*, 2007). It should be noted, that many of the genes found in GEIs are novel of unknown function, with no detectable homologues in other species, but nevertheless might have a role and confer selective advantage to the host organism (Hsiao, *et al.*, 2005).

Contribution of genomic islands to evolution of pathogenic bacteria

Antibiotic resistance represents one of the most frequent and well studied traits associated with GEIs. The emergence and dissemination of antibiotic resistance is a serious threat to public health as it puts the successful treatment of infectious diseases in increasing doubt. This change in bacterial populations from near universal susceptibility to increasing antibiotic resistance worldwide in a few decades is illustrative of the remarkable capacity for bacterial adaptation, albeit a man made threat. Many aspects of the accumulation of antibiotic resistance are poorly understood, one prominent aspect is how such resistance has disseminated globally so rapidly.

Investigations into the emergence of antibiotic resistance in *H. influenzae* suggest that it is an illustrative model for how resistance genes have become associated with a GEI.

Sequence analysis of a *H. influenzae* antibiotic resistant GEI, ICEHin1056, revealed that this island belonged within a much larger family (Mohd-Zain, *et al.*, 2004, Dimopoulou, *et al.*, 2007). Diverged GEIs of this family co-evolve independently with a wide range of Proteobacteria, including *Haemophilus* spp., *Pseudomonas* spp., *Yersinia enterocolitica*, *S. enterica* serovar Typhi, *Ralstonia metallidurans* and others (Mohd-Zain, *et al.*, 2004). A recent study investigating GEIs of haemophili has shown that the ICEHin1056 sub-family of ICE/GEIs is diverse and has not recently emerged in haemophili (Juhas, *et al.*, 2007b). However, the low sequence diversity and distinctive GC content of the Tn3s found in the haemophili's GEIs suggests more recent acquisition of the transposon-borne antibiotic resistance genes. Distribution between core and accessory genes (transposons) of the haemophili's GEIs is remarkably similar to the host *H. influenzae* "supragenome" and conforms to the Distributed Genome Hypothesis (Juhas, *et al.*, 2007b). This hypothesis proposes that the full complement of genes available to a pathogenic bacterial species exists in a supragenome pool, which is not contained by any particular strain (Hogg, *et al.*, 2007). The modular structure of the core genes of the haemophili's GEIs suggests that they are functionally constrained and, acting in concert, they play a major role in the successful propagation and survival of these GEIs within their hosts' "supragenomes". Their importance is overtly manifested by the rapid dissemination of antibiotic resistance worldwide among *H. influenzae* and *H. parainfluenzae* strains (Juhas, *et al.*, 2007b). The recent acquisition of antibiotic resistance genes and their rapid global spread over the past 30 to 40 years is an exemplar of how GEIs contribute to bacterial diversification and adaptation.

Staphylococcus aureus is a potentially pathogenic bacterium that plays a role in a broad spectrum of diseases and is a major cause of hospital-acquired infections worldwide. Methicillin-resistant *S. aureus* (MRSA) often referred to as the hospital “superbug” in the press represents one of the most serious threats to public health due to its resistance to a wide variety of antibiotics. This antibiotic resistance is facilitated mostly by genes located on a GEI designated staphylococcal cassette chromosome *mec* (SCC*mec*) (Katayama, *et al.*, 2000). Five SCC*mec* subtypes identified so far, range in size from 20-70 kb and confer resistance to methicillin, kanamycin, tobramycin, bleomycin, penicillins, heavy metals, tetracycline, macrolide, lincosamide and streptogramin (Ito, *et al.*, 2001, Deurenberg, *et al.*, 2007). GEI SCC*mec* is interesting in a sense that it does not contain phage-related and *tra* genes or transposases and is transferred between bacteria with the help of two site-specific recombinases that catalyze its chromosomal excision and reintegration (Ito, *et al.*, 2004, Noto & Archer, 2006). The origin of GEI SCC*mec* remains to be elucidated; however, it is hypothesized that it could originate from other staphylococci, namely *Staphylococcus sciuri* or *Staphylococcus epidermidis*. This is suggested by the high amino-acid sequence similarities of the methicillin resistance gene *mecA* products of *S. sciuri* and *S. aureus* as well as by an experiment where the *S. sciuri mecA* gene induced methicillin resistance in a formerly methicillin-sensitive *S. aureus* strain (Wu, *et al.*, 2001). Furthermore, *S. aureus mecA* was identical to that identified in an *S. epidermidis* isolate from the same individual, thus suggesting that MRSA strain has arisen *in vivo* by horizontal transfer of *mecA* between two staphylococcal species (Wielders, *et al.*, 2001, Deurenberg, *et al.*, 2007). While one theory hypothesizes that all MRSA clones have a common ancestor (Kreiwirth, *et al.*, 1993), another theory suggests that SCC*mec* was introduced several times into different

S. aureus lineages (Musser & Kapur, 1992, Fitzgerald, *et al.*, 2001, Enright, *et al.*, 2002, Qi, *et al.*, 2005).

Enterococcus faecalis is also one of the leading agents of nosocomial infections of surgical sites, the urinary tract and bloodstream (Richards, *et al.*, 2000). Most of the virulent strains of *E. faecalis* harbour a 150 kb GEI consisting of 129 open reading frames, including those encoding a wide variety of toxins, cytolysin and the surface proteins Esp and aggregation substance (Shankar, *et al.*, 2002). Recently, the horizontal transfer of the *E. faecalis* GEI has been demonstrated (Coburn, *et al.*, 2007). This has many implications for the evolution and diversity of *E. faecalis* as it suggests a mechanism for the conversion of commensal *E. faecalis* strains to virulent ones by acquisition of the GEI encoded virulence traits (Coburn, *et al.*, 2007).

Results from recent studies have shown that GEIs also played a key role in the evolution of pathogenic human and mammalian *Mycobacterium* spp. (Gutierrez, *et al.*, 2005, Becq, *et al.*, 2007). Speciation events in ancestral *Mycobacterium* spp. that were initially environmental bacteria occurred only relatively recently about a million years ago and are linked to the invasion of the genome by foreign DNA (Gutierrez, *et al.*, 2005, Rosas-Magallanes, *et al.*, 2006, Becq, *et al.*, 2007). Several *Mycobacterium tuberculosis* GEIs harbor genes that have been previously identified as virulence genes in other bacteria (Pethe, *et al.*, 2004, Stewart, *et al.*, 2005, Becq, *et al.*, 2007). A good example, illustrating the role played by the horizontally acquired GEIs in the evolution of *Mycobacterium* spp. pathogenic to mammals, is GEI Rv0986-8, as the Rv0986-8-

borne genes are required for the binding of *M. tuberculosis* to eukaryotic cells and *in vitro* trafficking (Pethe, *et al.*, 2004, Becq, *et al.*, 2007, Rosas-Magallanes, *et al.*, 2007).

Siderophores-mediated iron uptake is important for pathogenic as well as environmental bacteria. HPI is a 36- to 43-kb GEI that encodes the siderophore yersiniabactin-mediated iron uptake system (Lesic, *et al.*, 2004). This GEI was first found in *Yersinia* spp. but has been subsequently identified in a broad spectrum of *Enterobacteriaceae*. The presence of highly homologous HPI-borne ORFs among different species suggests recent acquisition of the HPI by these bacteria; however, the exact mechanism mediating the horizontal transfer of these islands awaits further investigation (Lesic & Carniel, 2005).

Many bacterial species exploit specialized secretion systems to transfer macromolecules across bacterial membranes. GEIs of a wide variety of bacterial pathogens encode type III secretion systems (T3SS) and type IV secretion systems (T4SS), which by transfer of proteins or nucleoprotein complexes directly mediate pathogenicity and horizontal gene transfer. *S. enterica* serovar Typhi is a Gram-negative facultative intracellular pathogen that is a causative agent of gastroenteritis and typhoid fever (Shea, *et al.*, 1996). The divergence of *Salmonella* and *E. coli* from their common ancestor occurred approximately 100-140 million years ago and these bacterial species have each acquired and lost more than 3 Mb of novel DNA since their divergence (Vernikos, *et al.*, 2007). First stage of *Salmonella* infection characterized by the colonization and invasion of intestinal epithelial cells is usually followed by the replication within host's macrophages. Two T3SSs harbored by two different GEIs, SPI-1 and SPI-2 play a

crucial role in the pathogenesis of *S. enterica*. T3SS of SPI-1 is important for the penetration of intestinal epithelium, while T3SS of SPI-2 has been hypothesized to be required solely after bacterium has gained access to host's macrophages. Recent findings with a mouse model of typhoid indicate that this is even more complicated, as SPI-2 was shown to be expressed also during early stages of pathogenesis prior to penetrating the intestine. This suggests that in addition to other functions, SPI-2 may be also involved in preparing *Salmonella* to successfully resist the harsh antimicrobial environment within macrophages (Brown, *et al.*, 2005). T3SSs harbored by *Salmonella* islands SPI-1 and SPI-2 are only two of many examples of GEIs-borne T3SSs enhancing pathogenicity of their host bacteria. Besides T3SSs, GEIs can also encode T4SSs. Various T4SSs associated with GEIs include for instance the GI-like family of T4SSs described above. The T4SSs are unique among other bacterial secretion systems due to their ability to transfer both proteins and nucleoprotein complexes. They can deliver bacterial effector proteins to host cells, thus contributing directly to pathogenicity. Furthermore, they can mediate horizontal gene transfer, thus facilitating the evolution of pathogens through dissemination of virulence genes (Juhás, *et al.*, 2008).

Adhesion, either interbacterial or to specific receptors of host cells represents another important pathogenicity trait that is often associated with GEIs. *Vibrio* pathogenicity island (VPI) encodes the toxin-coregulated pilus (TCP) that represents the major intestinal adherence factor of *V. cholerae*, the causative agent of cholera (Karaolis, *et al.*, 1999). TCP is a type IV pilus that mediates interbacterial adherence resulting in formation of microcolonies, as well as secretion of a soluble colonization factor, TcpF,

crucial for successful colonization of host (Kirn, *et al.*, 2003). Importance of the VPI-encoded TCP is overtly manifested by the inability of the TCP-deficient mutants of *V. cholerae* to colonize and cause disease in both humans and mice (Tripathi & Taylor, 2007).

Other GEI-encoded adherence factors include for instance P fimbriae and S fimbriae of uropathogenic *E. coli* (UPEC) and intimin of enteropathogenic and enterohemorrhagic *E. coli* (EPEC and EHEC, respectively) (reviewed elsewhere (Hacker & Kaper, 2000)). Both EPEC and EHEC infections are the leading cause of infantile diarrhea. Multiple separate acquisitions of the locus of enterocyte effacement (LEE) GEI and of the EPEC adherence factor plasmid (EAF) seems to be responsible for the evolution of the EPEC pathotype within *E. coli* (Lacher, *et al.*, 2007). Furthermore, microarray and whole genome PCR scanning analyses have shown that LEE GEI constitutes one of the most conserved parts of different EHEC strains. This study suggests that independent infections of similar but distinct bacteriophages carrying virulence determinants, including LEE GEI are deeply involved in the evolution of EHEC strains belonging to different *E. coli* lineages (Ogura, *et al.*, 2007).

Horizontal transfer of GEIs was also shown to play a major role in the evolution of a number of other bacterial pathogens, including LIPI of *L. monocytogenes* (Vazquez-Boland, *et al.*, 2001), SHI-0, SHI-1, SHI-2 of *Shigella flexneri* (Nie, *et al.*, 2006), BfPAI of *Bacteroides fragilis* (Franco, 2004) or *cag* of *Helicobacter pylori* (Gressmann, *et al.*, 2005).

Contribution of genomic islands to evolution of environmental bacteria and making the link to pathogenicity

Interestingly, while GEIs have been increasingly associated with the arisal and distribution of virulence functions or antibiotic resistance genes, at the same time they appear to be implicated in a multitude of other adaptive traits, which collectively could be called of “environmental relevance”. For example, a number of ICE/GEIs have been identified in β - and γ -*Proteobacteria* which carry gene clusters for the degradation of chlorinated and nitroaromatic compounds (Gaillard, *et al.*, 2006) or biphenyls (Toussaint, *et al.*, 2003). Even dehalogenases in *Dehalococcoides ethenogenes* have been associated with integrated DNA elements, whether these are mobile or not remains to be determined (Seshadri, *et al.*, 2005).

One of these ICE/GEIs, the highly mobile 103-kb ICE*clc* of *Pseudomonas* sp. strain B13, often used as a model for the behavior of catabolic GEIs (Sentchilo, *et al.*, 2003a, Sentchilo, *et al.*, 2003b) is a good example of how GEIs can contribute to the adaptation of environmental strains to use polluting compounds as new carbon sources. Various classical experiments have used *Pseudomonas* sp. strain B13 to develop new metabolic pathways in a single step, such as for chlorobiphenyl or chlorobenzene degradation (Reineke, 1998). The recently revealed molecular basis for this metabolic “complementation” of recipient bacteria also showed that ICE*clc* is in a continuing state of further adaptation. First of all, ICE*clc* carries a 9.5-kb *amn* gene cluster for aminophenol degradation with an aberrant nucleotide usage suggesting its past acquisition by a precursor element. Then, the genome of the environmental bacterium *Burkholderia xenovorans* LB400 contains an element almost 100 % identical to ICE*clc*,

except for additional genes allowing degradation of *o*-halobenzoates (Chain, *et al.*, 2006, Gaillard, *et al.*, 2006). Finally, in the groundwater isolate *Ralstonia* sp. strain JS705 an element similar to ICE*clc* was detected with a 10-kb insertion of another catabolic pathway gene cassette, necessary for chlorobenzene degradation (Muller, *et al.*, 2003). This development model for ICE*clc* of acquisition (and loss) of gene modules became even more striking when it was discovered that about 40% of catabolic *clc* elements are highly similar to a number of other ICE/GEIs (Larbig, *et al.*, 2002b, Mohd-Zain, *et al.*, 2004, Gaillard, *et al.*, 2006) (Fig.4). This region was therefore suspected to contain the information for self-transfer, which was more recently confirmed by the analysis of the T4SS genes of the ICE*Hin1056* from *H. influenzae* that showed significant homology to counterpart ICE*clc* genes (Juhas, *et al.*, 2007a). The core region of this group of elements was found in a variety of other bacteria (Fig.4), thus showing that this evolutionarily ancient ICE-type has been very successful in transferring into a large host-range (Mohd-Zain, *et al.*, 2004, Juhas, *et al.*, 2007a). More recently similar GEI fragments were also detected in *Xylella fastidiosa*, *Xanthomonas campestris*, *Rubrivivax gelatinosus*, *Azoarcus*, *Cupriavidus* and arsenic oxidizing bacterium *Herminiimonas arsenicoxydans* (Gaillard, *et al.*, 2006, Muller, *et al.*, 2007).

Particularly interesting was the close similarity between ICE*clc* and a large widely distributed group of GEIs of *P. aeruginosa*, including PAGI-2 and PAGI-3 (Larbig, *et al.*, 2002a, Larbig, *et al.*, 2002b, Klockgether, *et al.*, 2007). PAGI-2 and PAGI-3 were initially characterized from one clinical and one environmental *P. aeruginosa* strain (Larbig, *et al.*, 2002b), but after a large screening of clinical and environmental isolates of *P. aeruginosa* were found to be very common among those, in addition to the even

more promiscuous element pKLC102/PAPI-1 (Wolfgang, *et al.*, 2003, He, *et al.*, 2004, Qiu, *et al.*, 2006, Klockgether, *et al.*, 2007). The genetic load, in addition to the core region, was very different between *clc*, PAPI-2 or PAPI-3, demonstrating that these GEIs can easily accommodate and distribute regions of 60-70 kb of highly dissimilar DNA. The fact that this genetic load is found mostly between the integrase gene (near one end of the element) and the core region suggests that insertions here are selectively neutral and do not compromise the GEI's life cycle. A majority of accessory genes were found to be located in this position also in the closely related *Haemophilus* spp. ICEHin1056 subfamily of GEIs (Juhas, *et al.*, 2007b). This GEI-type is therefore an extreme example of the near continuum between characteristics judged as pathogenicity (ICEHin1056 of *H. influenzae*, or PAPI-2 and PAPI-3 of *P. aeruginosa*) and environmentally (ICE*clc* of *Pseudomonas* sp. strain B13) related (Fig.4).

Various other recent examples of whole genome comparisons show that GEIs have been implicated in adaptation to a wide range of different life styles and carry a large variety of functions. For example, the 49.6-kb GEI AGI-3 of *E. coli* EXPEC strain BEN2908, which is integrated in the *selC* tRNA locus, encodes various sugar transporters and sugar metabolic functions (Chouikha, *et al.*, 2006). *Burkholderia cenocepacia* isolates often harbor an instable 31.7 kb GEI which provides its host bacterium with the ability to synthesize N-acyl homoserine lactones and fatty acids and also encodes further amino acid transporters and metabolic genes (Baldwin, *et al.*, 2004). Furthermore, *Corynebacterium efficiens* carries four GEIs, of 70, 55, 40 and 32 kb, one of which (CEGI4) has the most characteristics of a fully functional GEI. Apart from an integrase, all other open reading frames encode unknown proteins (Zhang & Zhang, 2005). Recent

comparison of several *P. aeruginosa* isolates has revealed presence of multiple novel GEIs. One of those, the GEI RGP29 in *P. aeruginosa* PA2192, is a 224-kb element containing within its sequence another discernable GEI, called the *dit* island. The *dit* island encodes a full metabolic pathway for abietane diterpenoids (Mathee, *et al.*, 2008). Also the RGP5 element of *P. aeruginosa* PA14 encodes various transporters and genes known to be involved in iron metabolism (Mathee, *et al.*, 2008). *Magnetospirillum gryphiswaldense* contains a 130 kb unstable region of probable ancient GEI origin, but now full with insertion elements (42 copies) that cause frequent recombinations and deletions of parts of this genomic region. Interestingly, this region encodes many genes involved in magnetosome biomineralization (Ullrich, *et al.*, 2005). Suspected ancient GEIs with deteriorated integrases for which no transfer or excision could be demonstrated were detected in the anaerobic bacterium *Geobacter sulfurreducens* (Butler, *et al.*, 2007). One such region has a size of 300 kb and contained many genes implicated in anaerobic metabolism of benzoate, phenol, *p*-cresol and 4-hydroxybenzoate. Free-living water and soil-borne bacteria with no specific pathogenicity characteristics but which are thought to provide reservoirs for more pathogenic bacteria have now also been found to harbor GEIs. One of those is *Arcobacter butzleri*, an *Epsilonproteobacterium* related to *H. pylori* (Miller, *et al.*, 2007). Three potential GEIs have been found in this bacterium, the largest of which (ABG11) had a size of 26.9 kb and was integrated in *tRNA^{Leu}* with an integrase gene nearby. The island encodes 29 genes, none of which; however, have known orthologues in other bacterial genomes. More recent genome comparisons of multiple *Xylella fastidiosa* and *Xanthomonas campestris* strains have again reinforced the importance of GEIs in strain differentiation and provision of potential virulence functions (da Silva, *et*

al., 2007, He, *et al.*, 2007). For example, *Xylella* strains from citrus fruits carried one or more copies of the GI1 island, which encodes fimbrillin synthesis, haemolysin production and LPS synthesis (da Silva, *et al.*, 2007). Various rearrangements were observed in the largest (100 kb) island of *X. campestris* named XVR13 integrated in *tRNA^{Gly}* (He, *et al.*, 2007).

CONCLUDING REMARKS

Work over a number of years has shown the importance of GEIs in promoting horizontal gene transfer and distributing a wide range of adaptive functions for the host bacterium. Hence, compared to many self-transferable plasmids, GEIs seem to do the same job. Conceptually, GEIs may have a number of advantages over a plasmid, one of the most notable being that GEIs are integrated at a precise location in the host's chromosome, which is not damaged by the insertion and is restored following excision. Thus, unlike replicating plasmid molecules, GEIs do not need to continuously ensure coordinated replication, partitioning or specific maintenance, and because there is usually only a single copy of the GEI present per genome, its replication "cost" may not be as strong burden to the host cell (Gaillard, *et al.*, 2008). However, under certain conditions, GEIs can undergo excision, self-transfer and reintegration into a new host. More interestingly, several findings, including tetracycline induced transfer of the conjugative transposon CTnDot of *Bacteroides* (Cheng, *et al.*, 2000), DNA damage and SOS stress response induced transfer of the SXT element of *V. cholerae* (Beaber, *et al.*, 2004), and 3-chlorobenzoate augmented transfer of the *clc* element of *Pseudomonas* sp. strain B13 (Sentchilo, *et al.*, 2003a), indicate a much more regulated train of events for the transfer of GEIs that is fine-tuned in response to environmental signals. Thus the question arises, whether unwillingly by changing the conditions for bacteria in hospitals (antibiotic stress) and environment (pollution), we are generating selective conditions which promote the success of self-transferable and responsive GEIs. As shown on multiple examples in this review, GEIs play a crucial role in the evolution of a broad spectrum of pathogenic or environmental bacteria. Furthermore, several lines of evidence suggest existence of an evolutionary ancient GEIs spread in a versatile groups

of otherwise unrelated bacterial species. Besides conserved set of core genes required for maintenance, these evolutionary ancient GEIs also harbor a variable number of other gene modules whose composition is strongly dependent on the life-style of the particular host bacterial species, which can be either pathogenic or environmental (Fig.4). The important contribution of many research groups around the world, has led to GEI-facilitated horizontal gene transfer being one of the most rapidly evolving fields of microbiology research.

ACKNOWLEDGEMENTS

We would like to cordially thank all scientists who contributed with their results and ideas to the rapidly advancing field of the horizontal gene transfer and apologize to those whose work was not cited due to space limitations. MJ was supported from the grant from the Medical Research Council, UK; and JM and MG were supported from the Swiss National Science Foundation (grant nr. 3100-67229).

REFERENCES

- Auchtung JM, Lee CA, Garrison KL & Grossman AD (2007) Identification and characterization of the immunity repressor (ImmR) that controls the mobile genetic element ICEBs1 of *Bacillus subtilis*. *Mol Microbiol* 64: 1515-1528.
- Baldwin A, Sokol PA, Parkhill J & Mahenthiralingam E (2004) The *Burkholderia cepacia* epidemic strain marker is part of a novel genomic island encoding both virulence and metabolism-associated genes in *Burkholderia cenocepacia*. *Infect Immun* 72: 1537-1547.
- Barbe V, Vallenet D, Fonknechten N, *et al.* (2004) Unique features revealed by the genome sequence of *Acinetobacter* sp. ADP1, a versatile and naturally transformation competent bacterium. *Nucleic Acids Res* 32: 5766-5779.
- Beaber JW, Hochhut B & Waldor MK (2002) Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from *Vibrio cholerae*. *J Bacteriol* 184: 4259-4269.
- Beaber JW, Hochhut B & Waldor MK (2004) SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* 427: 72-74.
- Becq J, Gutierrez MC, Rosas-Magallanes V, Rauzier J, Gicquel B, Neyrolles O & Deschavanne P (2007) Contribution of horizontally acquired genomic islands to the evolution of the tubercle bacilli. *Mol Biol Evol* 24: 1861-1871.
- Bingle LE & Thomas CM (2001) Regulatory circuits for plasmid survival. *Curr Opin Microbiol* 4: 194-200.
- Brown NF, Vallance BA, Coombes BK, Valdez Y, Coburn BA & Finlay BB (2005) *Salmonella* pathogenicity island 2 is expressed prior to penetrating the intestine. *PLoS Pathog* 1: e32.

- Buchrieser C, Prentice M & Carniel E (1998) The 102-kilobase unstable region of *Yersinia pestis* comprises a high-pathogenicity island linked to a pigmentation segment which undergoes internal rearrangement. *J Bacteriol* 180: 2321-2329.
- Burrus V & Waldor MK (2003) Control of SXT integration and excision. *J Bacteriol* 185: 5045-5054.
- Burrus V & Waldor MK (2004) Shaping bacterial genomes with integrative and conjugative elements. *Res Microbiol* 155: 376-386.
- Burrus V, Pavlovic G, Decaris B & Guedon G (2002) Conjugative transposons: the tip of the iceberg. *Mol Microbiol* 46: 601-610.
- Butler JE, He Q, Nevin KP, He Z, Zhou J & Lovley DR (2007) Genomic and microarray analysis of aromatics degradation in *Geobacter metallireducens* and comparison to a *Geobacter* isolate from a contaminated field site. *BMC Genomics* 8: 180.
- Chain PS, Denev VJ, Konstantinidis KT, *et al.* (2006) *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. *Proc Natl Acad Sci U S A* 103: 15280-15287.
- Chen I, Christie PJ & Dubnau D (2005) The ins and outs of DNA transfer in bacteria. *Science* 310: 1456-1460.
- Cheng Q, Paszkiet BJ, Shoemaker NB, Gardner JF & Salyers AA (2000) Integration and excision of a *Bacteroides* conjugative transposon, CTnDOT. *J Bacteriol* 182: 4035-4043.
- Cheng Q, Sutanto Y, Shoemaker NB, Gardner JF & Salyers AA (2001) Identification of genes required for excision of CTnDOT, a *Bacteroides* conjugative transposon. *Mol Microbiol* 41: 625-632.
- Chouikha I, Germon P, Bree A, Gilot P, Moulin-Schouleur M & Schouler C (2006) A selC-associated genomic island of the extraintestinal avian pathogenic *Escherichia*

- coli strain BEN2908 is involved in carbohydrate uptake and virulence. *J Bacteriol* 188: 977-987.**
- Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S & Cascales E (2005) Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu Rev Microbiol* 59: 451-485.**
- Coburn PS, Baghdayan AS, Dolan GT & Shankar N (2007) Horizontal transfer of virulence genes encoded on the *Enterococcus faecalis* pathogenicity island. *Mol Microbiol* 63: 530-544.**
- Coleman ML, Sullivan MB, Martiny AC, Steglich C, Barry K, Delong EF & Chisholm SW (2006) Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* 311: 1768-1770.**
- da Silva VS, Shida CS, Rodrigues FB, *et al.* (2007) Comparative genomic characterization of citrus-associated *Xylella fastidiosa* strains. *BMC Genomics* 8: 474.**
- Dahlberg C & Chao L (2003) Amelioration of the cost of conjugative plasmid carriage in *Escherichia coli* K12. *Genetics* 165: 1641-1649.**
- De Gelder L, Ponciano JM, Joyce P & Top EM (2007) Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology* 153: 452-463.**
- Deurenberg RH, Vink C, Kalenic S, Friedrich AW, Bruggeman CA & Stobberingh EE (2007) The molecular evolution of methicillin-resistant *Staphylococcus aureus*. *Clin Microbiol Infect* 13: 222-235.**
- Dimopoulou ID, Russell JE, Mohd-Zain Z, Herbert R & Crook DW (2002) Site-specific recombination with the chromosomal *tRNA(Leu)* gene by the large conjugative *Haemophilus* resistance plasmid. *Antimicrob Agents Chemother* 46: 1602-1603.**

- Dimopoulou ID, Kartali SI, Harding RM, Peto TE & Crook DW (2007) Diversity of antibiotic resistance integrative and conjugative elements among haemophili. *J Med Microbiol* 56: 838-846.
- Dobrindt U, Hochhut B, Hentschel U & Hacker J (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* 2: 414-424.
- Doublet B, Boyd D, Mulvey MR & Cloeckaert A (2005) The Salmonella genomic island 1 is an integrative mobilizable element. *Mol Microbiol* 55: 1911-1924.
- Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H & Spratt BG (2002) The evolutionary history of methicillin-resistant Staphylococcus aureus (MRSA). *Proc Natl Acad Sci U S A* 99: 7687-7692.
- Fall S, Mercier A, Bertolla F, *et al.* (2007) Horizontal gene transfer regulation in bacteria as a "spandrel" of DNA repair mechanisms. *PLoS ONE* 2: e1055.
- Fitzgerald JR, Sturdevant DE, Mackie SM, Gill SR & Musser JM (2001) Evolutionary genomics of Staphylococcus aureus: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. *Proc Natl Acad Sci U S A* 98: 8821-8826.
- Franco AA (2004) The Bacteroides fragilis pathogenicity island is contained in a putative novel conjugative transposon. *J Bacteriol* 186: 6077-6092.
- Frost LS, Ippen-Ihler K & Skurray RA (1994) Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol Rev* 58: 162-210.
- Frost LS, Leplae R, Summers AO & Toussaint A (2005) Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* 3: 722-732.
- Gaillard M, Pernet N, Vogne C, Hagenbuchle O & van der Meer JR (2008) Host and invader impact of transfer of the clc genomic island into Pseudomonas aeruginosa PAO1. *Proc Natl Acad Sci U S A* 105: 7058-7063.

- Gaillard M, Vallaeyts T, Vorholter FJ, *et al.* (2006) The *clc* element of *Pseudomonas* sp. strain B13, a genomic island with various catabolic properties. *J Bacteriol* 188: 1999-2013.
- Gal-Mor O & Finlay BB (2006) Pathogenicity islands: a molecular toolbox for bacterial virulence. *Cell Microbiol* 8: 1707-1719.
- Goodman SD & Scocca JJ (1988) Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. *Proc Natl Acad Sci U S A* 85: 6982-6986.
- Gressmann H, Linz B, Ghai R, *et al.* (2005) Gain and loss of multiple genes during the evolution of *Helicobacter pylori*. *PLoS Genet* 1: e43.
- Groisman EA & Ochman H (1996) Pathogenicity islands: bacterial evolution in quantum leaps. *Cell* 87: 791-794.
- Gutierrez MC, Brisse S, Brosch R, *et al.* (2005) Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathog* 1: e5.
- Hacker J & Kaper JB (2000) Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* 54: 641-679.
- Hacker J, Bender L, Ott M, Wingender J, Lund B, Marre R & Goebel W (1990) Deletions of chromosomal regions coding for fimbriae and hemolysins occur in vitro and in vivo in various extraintestinal *Escherichia coli* isolates. *Microb Pathog* 8: 213-225.
- Hamilton HL, Dominguez NM, Schwartz KJ, Hackett KT & Dillard JP (2005) *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel type IV secretion system. *Mol Microbiol* 55: 1704-1721.
- Hamoen LW, Haijema B, Bijlsma JJ, Venema G & Lovett CM (2001) The *Bacillus subtilis* competence transcription factor, ComK, overrides LexA-imposed

- transcriptional inhibition without physically displacing LexA. *J Biol Chem* 276: 42901-42907.
- He J, Baldini RL, Deziel E, *et al.* (2004) The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proc Natl Acad Sci U S A* 101: 2530-2535.
- He YQ, Zhang L, Jiang BL, *et al.* (2007) Comparative and functional genomics reveals genetic diversity and determinants of host specificity among reference strains and a large collection of Chinese isolates of the phytopathogen *Xanthomonas campestris* pv. *campestris*. *Genome Biol* 8: R218.
- Hochhut B, Marrero J & Waldor MK (2000) Mobilization of plasmids and chromosomal DNA mediated by the SXT element, a constin found in *Vibrio cholerae* O139. *J Bacteriol* 182: 2043-2047.
- Hochhut B, Wilde C, Balling G, *et al.* (2006) Role of pathogenicity island-associated integrases in the genome plasticity of uropathogenic *Escherichia coli* strain 536. *Mol Microbiol* 61: 584-595.
- Hogg JS, Hu FZ, Janto B, *et al.* (2007) Characterization and modeling of the *Haemophilus influenzae* core and supragenomes based on the complete genomic sequences of Rd and 12 clinical nontypeable strains. *Genome Biol* 8: R103.
- Hommais F, Krin E, Laurent-Winter C, *et al.* (2001) Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol* 40: 20-36.
- Hsiao WW, Ung K, Aeschliman D, Bryan J, Finlay BB & Brinkman FS (2005) Evidence of a large novel gene pool associated with prokaryotic genomic islands. *PLoS Genet* 1: e62.

- Ito T, Ma XX, Takeuchi F, Okuma K, Yuzawa H & Hiramatsu K (2004) Novel type V staphylococcal cassette chromosome mec driven by a novel cassette chromosome recombinase, *ccrC*. *Antimicrob Agents Chemother* 48: 2637-2651.
- Ito T, Katayama Y, Asada K, Mori N, Tsutsumimoto K, Tiensasitorn C & Hiramatsu K (2001) Structural comparison of three types of staphylococcal cassette chromosome mec integrated in the chromosome in methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 45: 1323-1336.
- Jain R, Rivera MC, Moore JE & Lake JA (2002) Horizontal gene transfer in microbial genome evolution. *Theor Popul Biol* 61: 489-495.
- Jain R, Rivera MC, Moore JE & Lake JA (2003) Horizontal gene transfer accelerates genome innovation and evolution. *Mol Biol Evol* 20: 1598-1602.
- Juhas M, Crook DW & Hood DW (2008) Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cell Microbiol*.
- Juhas M, Crook DW, Dimopoulou ID, Lunter G, Harding RM, Ferguson DJ & Hood DW (2007a) Novel type IV secretion system involved in propagation of genomic islands. *J Bacteriol* 189: 761-771.
- Juhas M, Power PM, Harding RM, *et al.* (2007b) Sequence and functional analyses of *Haemophilus* spp. genomic islands. *Genome Biol* 8: R237.
- Karaolis DK, Somara S, Maneval DR, Jr., Johnson JA & Kaper JB (1999) A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* 399: 375-379.
- Katayama Y, Ito T & Hiramatsu K (2000) A new class of genetic element, staphylococcus cassette chromosome mec, encodes methicillin resistance in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 44: 1549-1555.
- Kirn TJ, Bose N & Taylor RK (2003) Secretion of a soluble colonization factor by the TCP type 4 pilus biogenesis pathway in *Vibrio cholerae*. *Mol Microbiol* 49: 81-92.

- Klockgether J, Reva O, Larbig K & Tummler B (2004) Sequence analysis of the mobile genome island pKLC102 of *Pseudomonas aeruginosa* C. *J Bacteriol* 186: 518-534.
- Klockgether J, Wurdemann D, Reva O, Wiehlmann L & Tummler B (2007) Diversity of the abundant pKLC102/PAGI-2 family of genomic islands in *Pseudomonas aeruginosa*. *J Bacteriol* 189: 2443-2459.
- Kreiswirth B, Kornblum J, Arbeit RD, *et al.* (1993) Evidence for a clonal origin of methicillin resistance in *Staphylococcus aureus*. *Science* 259: 227-230.
- Lacher DW, Steinsland H, Blank TE, Donnenberg MS & Whittam TS (2007) Molecular evolution of typical enteropathogenic *Escherichia coli*: clonal analysis by multilocus sequence typing and virulence gene allelic profiling. *J Bacteriol* 189: 342-350.
- Lang B, Blot N, Bouffartigues E, *et al.* (2007) High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res* 35: 6330-6337.
- Larbig K, Kiewitz C & Tummler B (2002a) Pathogenicity islands and PAI-like structures in *Pseudomonas* species. *Curr Top Microbiol Immunol* 264: 201-211.
- Larbig KD, Christmann A, Johann A, *et al.* (2002b) Gene islands integrated into tRNA(Gly) genes confer genome diversity on a *Pseudomonas aeruginosa* clone. *J Bacteriol* 184: 6665-6680.
- Leplae R, Lima-Mendez G & Toussaint A (2006) A first global analysis of plasmid encoded proteins in the ACLAME database. *FEMS Microbiol Rev* 30: 980-994.
- Lesic B & Carniel E (2005) Horizontal transfer of the high-pathogenicity island of *Yersinia pseudotuberculosis*. *J Bacteriol* 187: 3352-3358.
- Lesic B, Bach S, Ghigo JM, Dobrindt U, Hacker J & Carniel E (2004) Excision of the high-pathogenicity island of *Yersinia pseudotuberculosis* requires the combined actions of its cognate integrase and Hef, a new recombination directionality factor. *Mol Microbiol* 52: 1337-1348.

- Lucchini S, Rowley G, Goldberg MD, Hurd D, Harrison M & Hinton JC (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog* 2: e81.
- Luck SN, Turner SA, Rajakumar K, Adler B & Sakellaris H (2004) Excision of the Shigella resistance locus pathogenicity island in Shigella flexneri is stimulated by a member of a new subgroup of recombination directionality factors. *J Bacteriol* 186: 5551-5554.
- Maiques E, Ubeda C, Tormo MA, Ferrer MD, Lasa I, Novick RP & Penades JR (2007) Role of staphylococcal phage and SaPI integrase in intra- and interspecies SaPI transfer. *J Bacteriol* 189: 5608-5616.
- Mathee K, Narasimhan G, Valdes C, *et al.* (2008) Dynamics of Pseudomonas aeruginosa genome evolution. *Proc Natl Acad Sci U S A* 105: 3100-3105.
- Meier P & Wackernagel W (2005) Impact of mutS inactivation on foreign DNA acquisition by natural transformation in Pseudomonas stutzeri. *J Bacteriol* 187: 143-154.
- Miller WG, Parker CT, Rubenfield M, *et al.* (2007) The Complete Genome Sequence and Analysis of the Epsilonproteobacterium Arcobacter butzleri. *PLoS ONE* 2: e1358.
- Mohd-Zain Z, Turner SL, Cerdeno-Tarraga AM, *et al.* (2004) Transferable antibiotic resistance elements in Haemophilus influenzae share a common evolutionary origin with a diverse family of syntenic genomic islands. *J Bacteriol* 186: 8114-8122.
- Muller D, Medigue C, Koechler S, *et al.* (2007) A tale of two oxidation states: bacterial colonization of arsenic-rich environments. *PLoS Genet* 3: e53.
- Muller TA, Werlen C, Spain J & Van Der Meer JR (2003) Evolution of a chlorobenzene degradative pathway among bacteria in a contaminated groundwater mediated by a genomic island in Ralstonia. *Environ Microbiol* 5: 163-173.

- Musser JM & Kapur V (1992) Clonal analysis of methicillin-resistant *Staphylococcus aureus* strains from intercontinental sources: association of the *mec* gene with divergent phylogenetic lineages implies dissemination by horizontal transfer and recombination. *J Clin Microbiol* 30: 2058-2063.
- Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ & Fang FC (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. *Science* 313: 236-238.
- Nguyen TN, Phan QG, Duong LP, Bertrand KP & Lenski RE (1989) Effects of carriage and expression of the Tn10 tetracycline-resistance operon on the fitness of *Escherichia coli* K12. *Mol Biol Evol* 6: 213-225.
- Nie H, Yang F, Zhang X, *et al.* (2006) Complete genome sequence of *Shigella flexneri* 5b and comparison with *Shigella flexneri* 2a. *BMC Genomics* 7: 173.
- Nielsen KM & Townsend JP (2004) Monitoring and modeling horizontal gene transfer. *Nat Biotechnol* 22: 1110-1114.
- Noto MJ & Archer GL (2006) A subset of *Staphylococcus aureus* strains harboring staphylococcal cassette chromosome *mec* (SCC*mec*) type IV is deficient in CcrAB-mediated SCC*mec* excision. *Antimicrob Agents Chemother* 50: 2782-2788.
- Ochman H, Lawrence JG & Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* 405: 299-304.
- Ogura Y, Ooka T, Asadulghani, *et al.* (2007) Extensive genomic diversity and selective conservation of virulence-determinants in enterohemorrhagic *Escherichia coli* strains of O157 and non-O157 serotypes. *Genome Biol* 8: R138.
- Pethe K, Swenson DL, Alonso S, Anderson J, Wang C & Russell DG (2004) Isolation of *Mycobacterium tuberculosis* mutants defective in the arrest of phagosome maturation. *Proc Natl Acad Sci U S A* 101: 13642-13647.

- Prudhomme M, Attaiech L, Sanchez G, Martin B & Claverys JP (2006) Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneumoniae*. *Science* 313: 89-92.**
- Qi W, Ender M, O'Brien F, Imhof A, Ruef C, McCallum N & Berger-Bachi B (2005) Molecular epidemiology of methicillin-resistant *Staphylococcus aureus* in Zurich, Switzerland (2003): prevalence of type IV SCCmec and a new SCCmec element associated with isolates from intravenous drug users. *J Clin Microbiol* 43: 5164-5170.**
- Qiu X, Gurkar AU & Lory S (2006) Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 103: 19830-19835.**
- Ramsay JP, Sullivan JT, Stuart GS, Lamont IL & Ronson CW (2006) Excision and transfer of the *Mesorhizobium loti* R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS. *Mol Microbiol* 62: 723-734.**
- Ravatn R, Studer S, Zehnder AJ & van der Meer JR (1998a) Int-B13, an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase *clc* element of *Pseudomonas* sp. Strain B13. *J Bacteriol* 180: 5505-5514.**
- Ravatn R, Studer S, Springael D, Zehnder AJ & van der Meer JR (1998b) Chromosomal integration, tandem amplification, and deamplification in *Pseudomonas putida* F1 of a 105-kilobase genetic element containing the chlorocatechol degradative genes from *Pseudomonas* sp. Strain B13. *J Bacteriol* 180: 4360-4369.**
- Reineke W (1998) Development of hybrid strains for the mineralization of chloroaromatics by patchwork assembly. *Annu Rev Microbiol* 52: 287-331.**

- Richards MJ, Edwards JR, Culver DH & Gaynes RP (2000) Nosocomial infections in combined medical-surgical intensive care units in the United States. *Infect Control Hosp Epidemiol* 21: 510-515.
- Rosas-Magallanes V, Deschavanne P, Quintana-Murci L, Brosch R, Gicquel B & Neyrolles O (2006) Horizontal transfer of a virulence operon to the ancestor of *Mycobacterium tuberculosis*. *Mol Biol Evol* 23: 1129-1135.
- Rosas-Magallanes V, Stadthagen-Gomez G, Rauzier J, *et al.* (2007) Signature-tagged transposon mutagenesis identifies novel *Mycobacterium tuberculosis* genes involved in the parasitism of human macrophages. *Infect Immun* 75: 504-507.
- Ruzin A, Lindsay J & Novick RP (2001) Molecular genetics of SaPI1--a mobile pathogenicity island in *Staphylococcus aureus*. *Mol Microbiol* 41: 365-377.
- Schluter A, Szczepanowski R, Puhler A & Top EM (2007) Genomics of IncP-1 antibiotic resistance plasmids isolated from wastewater treatment plants provides evidence for a widely accessible drug resistance gene pool. *FEMS Microbiol Rev* 31: 449-477.
- Schmidt H & Hensel M (2004) Pathogenicity islands in bacterial pathogenesis. *Clin Microbiol Rev* 17: 14-56.
- Segal G, Feldman M & Zusman T (2005) The Icm/Dot type-IV secretion systems of *Legionella pneumophila* and *Coxiella burnetii*. *FEMS Microbiol Rev* 29: 65-81.
- Sentchilo V, Zehnder AJ & van der Meer JR (2003a) Characterization of two alternative promoters for integrase expression in the *clc* genomic island of *Pseudomonas* sp. strain B13. *Mol Microbiol* 49: 93-104.
- Sentchilo V, Ravatn R, Werlen C, Zehnder AJ & van der Meer JR (2003b) Unusual integrase gene expression on the *clc* genomic island in *Pseudomonas* sp. strain B13. *J Bacteriol* 185: 4530-4538.
- Seshadri R, Adrian L, Fouts DE, *et al.* (2005) Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes*. *Science* 307: 105-108.

- Shankar N, Baghdayan AS & Gilmore MS (2002) Modulation of virulence within a pathogenicity island in vancomycin-resistant *Enterococcus faecalis*. *Nature* 417: 746-750.
- Shea JE, Hensel M, Gleeson C & Holden DW (1996) Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. *Proc Natl Acad Sci U S A* 93: 2593-2597.
- Shoemaker NB, Wang GR & Salyers AA (2000) Multiple gene products and sequences required for excision of the mobilizable integrated *Bacteroides* element NBU1. *J Bacteriol* 182: 928-936.
- Smith HO, Tomb JF, Dougherty BA, Fleischmann RD & Venter JC (1995) Frequency and distribution of DNA uptake signal sequences in the *Haemophilus influenzae* Rd genome. *Science* 269: 538-540.
- Stewart GR, Patel J, Robertson BD, Rae A & Young DB (2005) Mycobacterial mutants with defective control of phagosomal acidification. *PLoS Pathog* 1: 269-278.
- Sullivan JT, Trzebiatowski JR, Cruickshank RW, *et al.* (2002) Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A. *J Bacteriol* 184: 3086-3095.
- Thomas CM & Nielsen KM (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev Microbiol* 3: 711-721.
- Toussaint A & Merlin C (2002) Mobile elements as a combination of functional modules. *Plasmid* 47: 26-35.
- Toussaint A, Merlin C, Monchy S, Benotmane MA, Leplae R, Mergeay M & Springael D (2003) The biphenyl- and 4-chlorobiphenyl-catabolic transposon Tn4371, a member of a new family of genomic islands related to IncP and Ti plasmids. *Appl Environ Microbiol* 69: 4837-4845.

- Trempey JE, Kirby JE & Gottesman S (1994) Alp suppression of Lon: dependence on the *slpA* gene. *J Bacteriol* 176: 2061-2067.
- Tripathi SA & Taylor RK (2007) Membrane association and multimerization of TcpT, the cognate ATPase ortholog of the *Vibrio cholerae* toxin-coregulated-pilus biogenesis apparatus. *J Bacteriol* 189: 4401-4409.
- Ubeda C, Barry P, Penades JR & Novick RP (2007) A pathogenicity island replicon in *Staphylococcus aureus* replicates as an unstable plasmid. *Proc Natl Acad Sci U S A* 104: 14182-14188.
- Ubeda C, Tormo MA, Cucarella C, Trotonda P, Foster TJ, Lasa I & Penades JR (2003) Sip, an integrase protein with excision, circularization and integration activities, defines a new family of mobile *Staphylococcus aureus* pathogenicity islands. *Mol Microbiol* 49: 193-210.
- Ubeda C, Maiques E, Barry P, *et al.* (2008) SaPI mutations affecting replication and transfer and enabling autonomous replication in the absence of helper phage. *Mol Microbiol* 67: 493-503.
- Ullrich S, Kube M, Schubbe S, Reinhardt R & Schuler D (2005) A hypervariable 130-kilobase genomic region of *Magnetospirillum gryphiswaldense* comprises a magnetosome island which undergoes frequent rearrangements during stationary growth. *J Bacteriol* 187: 7176-7184.
- van der Meer JR & Sentschilo V (2003) Genomic islands and the evolution of catabolic pathways in bacteria. *Curr Opin Biotechnol* 14: 248-254.
- van der Meer JR, Ravatn R & Sentschilo V (2001) The *clc* element of *Pseudomonas* sp. strain B13 and other mobile degradative elements employing phage-like integrases. *Arch Microbiol* 175: 79-85.

- Vazquez-Boland JA, Dominguez-Bernal G, Gonzalez-Zorn B, Kreft J & Goebel W (2001) Pathogenicity islands and virulence evolution in *Listeria*. *Microbes Infect* 3: 571-584.
- Vernikos GS & Parkhill J (2008) Resolving the structural features of genomic islands: A machine learning approach. *Genome Res* 18: 331-342.
- Vernikos GS, Thomson NR & Parkhill J (2007) Genetic flux over time in the *Salmonella* lineage. *Genome Biol* 8: R100.
- Whittle G, Shoemaker NB & Salyers AA (2002) The role of *Bacteroides* conjugative transposons in the dissemination of antibiotic resistance genes. *Cell Mol Life Sci* 59: 2044-2054.
- Wielders CL, Vriens MR, Brisse S, *et al.* (2001) In-vivo transfer of *mecA* DNA to *Staphylococcus aureus* [corrected]. *Lancet* 357: 1674-1675.
- Williams KP (2002) Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. *Nucleic Acids Res* 30: 866-875.
- Wolfgang MC, Kulasekara BR, Liang X, *et al.* (2003) Conservation of genome content and virulence determinants among clinical and environmental isolates of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* 100: 8484-8489.
- Wu SW, de Lencastre H & Tomasz A (2001) Recruitment of the *mecA* gene homologue of *Staphylococcus sciuri* into a resistance determinant and expression of the resistant phenotype in *Staphylococcus aureus*. *J Bacteriol* 183: 2417-2424.
- Zhang R & Zhang CT (2005) Genomic islands in the *Corynebacterium efficiens* genome. *Appl Environ Microbiol* 71: 3126-3130.

FIGURE LEGENDS

Figure 1. General features of genomic islands. GEIs are relatively large segments of mobile DNA whose nucleotide characteristics often differ from the rest of the chromosome. GEIs are often inserted at tRNA genes and flanked by direct repeats (DR). GEIs typically harbor genes encoding factors involved in genetic mobility, such as integrases, transposases and insertion sequences (IS). According to their gene content GEIs can be described as pathogenicity, symbiosis, metabolic, fitness or resistance islands.

Figure 2. Variable types of genomic islands. GEIs come in a large spectrum of varieties and encompass other categories of elements, such as conjugative transposons/ICEs, integrated plasmids, non-replicative but excisable elements and cryptic or damaged prophages. Grey-shaded areas point to self-mobile GEIs.

Figure 3. “Life-cycle” of genomic islands. The “life-cycle” of GEI consists of the following steps: (1) Acquisition by horizontal gene transfer, (2) Integration into the host’s chromosome by site-specific recombination, (3) Development of the GEI by genetic rearrangements, gene loss (3a) or gene acquisition (3b), (4) Excision from the chromosome, (5) Transfer to another recipient.

Figure 4. Various functions encoded by genomic islands of the same family. Modified Artemis Comparison Tool view of four GEIs of the same family: *ICEHin1056*, *clc*, *ICEHpaT3T1* and PAPI. Homologous sequences (minimum cut-off = 50) are indicated by grey lines joining regions of the four GEIs. Gene modules homologous in all shown GEIs are represented by white boxes, while modules specific for individual GEIs are indicated by black boxes. The figure shows that depending on the composition of gene modules, members of the

same family of GEIs can promote survival of pathogenic (*ICEHin1056*, *ICEHpaT3T1*, PAPI of *H. influenzae*, *H. parainfluenzae* and *P. aeruginosa*, respectively) as well as environmental (*clc* of *Pseudomonas* sp B13) bacteria. ATB resist.=antibiotics resistance, T4SS=type IV secretion, Integr.=integration.

Figure 1

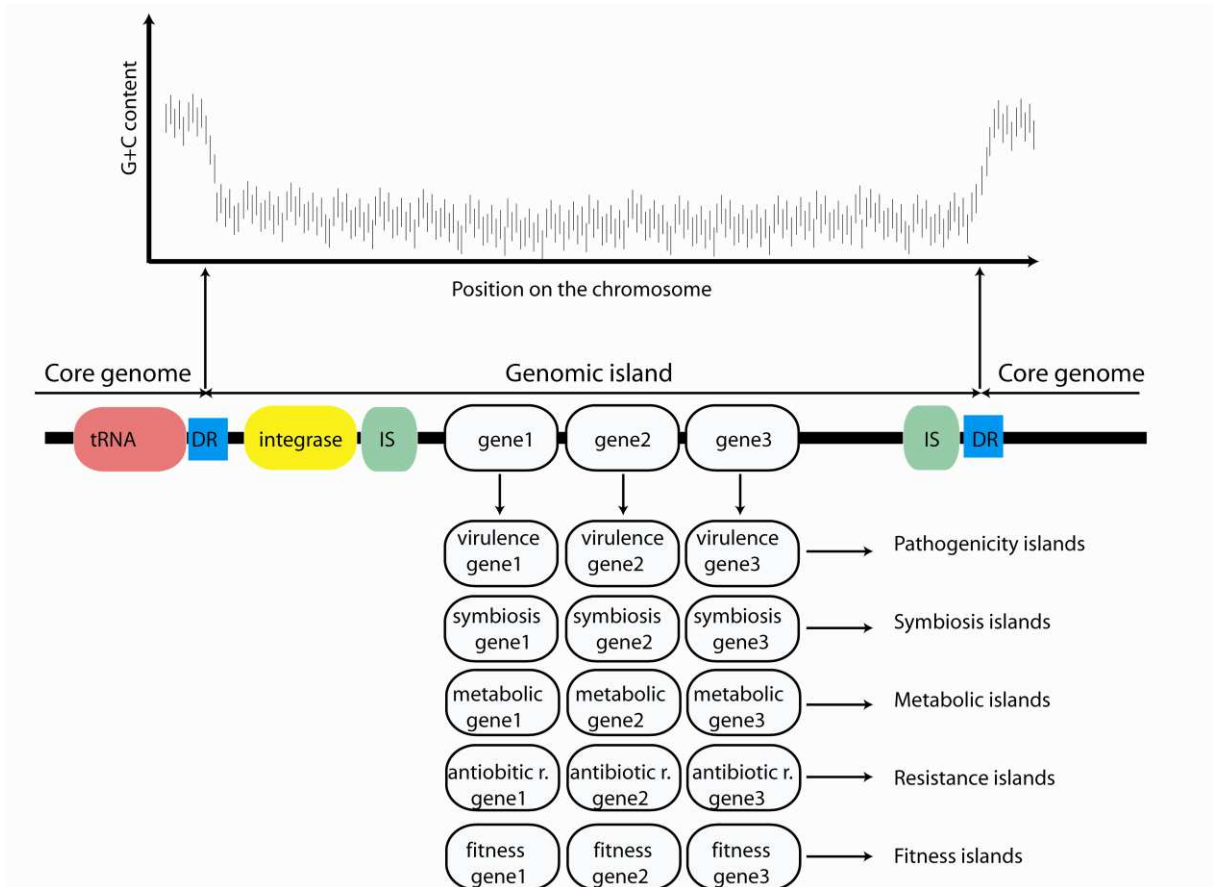


Figure 2

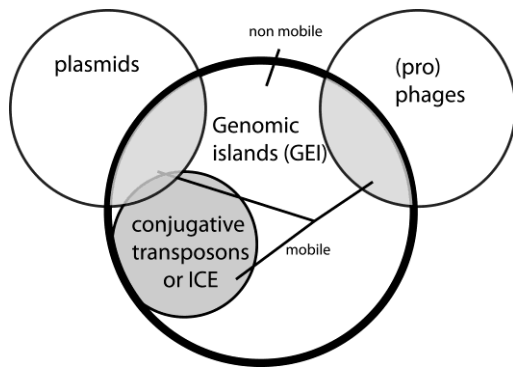


Figure 3

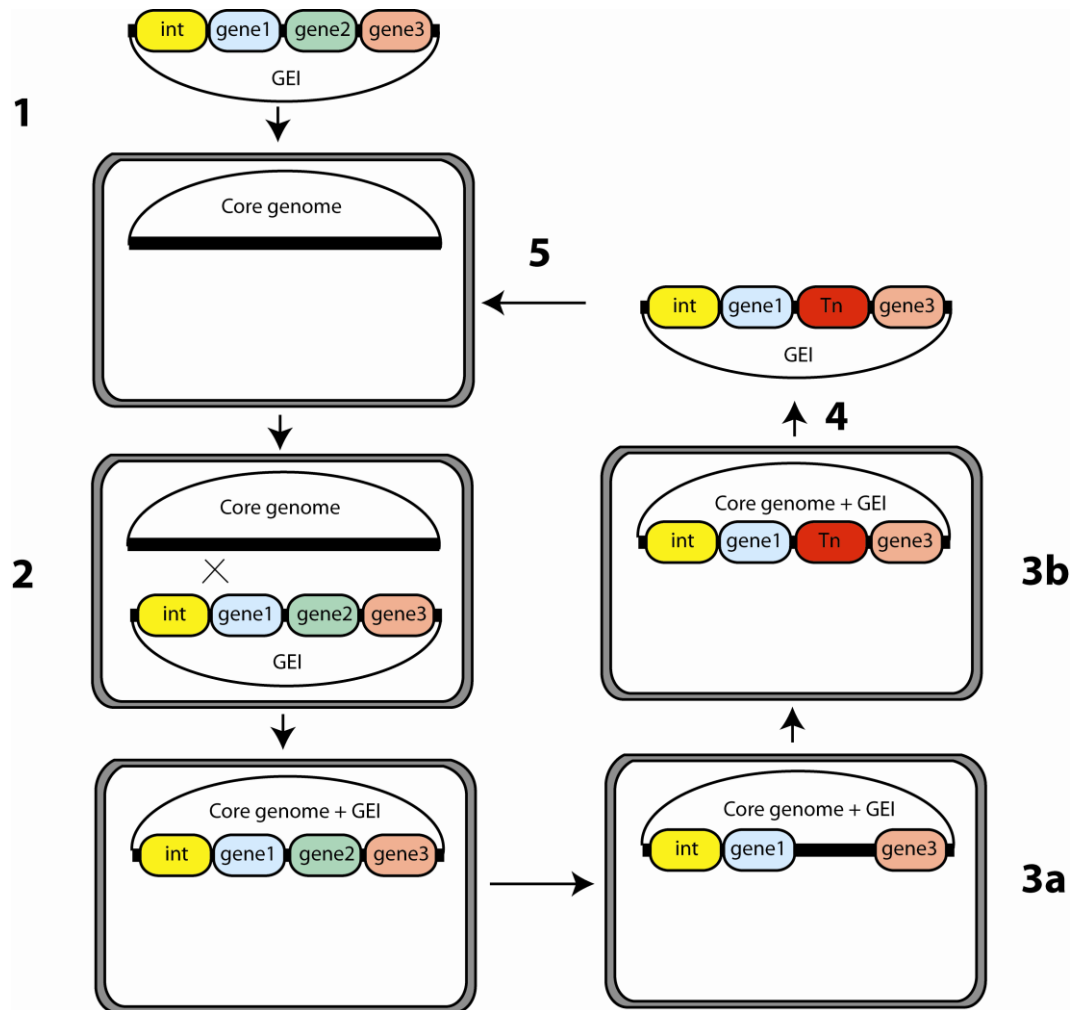


Figure 4

