Republic of Iraq
Ministry of Higher Education
& Scientific Research
University of Al-Qadisiyah
College of Veterinary Medicine

Pathogenicity islands

A Graduation Project Submitted to the Department Council of the Internal and Preventive Medicine-College of Veterinary Medicine/ University of Al-Qadisiyah in a partial fulfillment of the requirements for the Degree of Bachelor of Science in Veterinary Medicine and Surgery.

By
هنادي محمود كاظم

Supervised by
الدكتور احمد جاسم نعمة

2020 A.D. 1441 A.H.

فَتَعَٰلَى ٱللَّهُ ٱلْمَلِكُ ٱلْحَقُّ ۗ وَلَا تَعْجَلْ بِٱلْقُرْءَانِ مِن قَبْلِ أَن يُقْضَىٰٓ
إِلَيْكَ وَحْيُهُۥ ۖ وَقُل رَّبِّ زِدْنِى عِلْمًا ﴿١١٤﴾

صَدَقَ ٱللَّهُ ٱلْعَظِيمُ

من سورة طه

Certificate of Supervisor

I certify that the project entitled (----------) was prepared by ---------- under my supervision at the College of Veterinary Medicine / University of Al-Qadisiyah.

Supervisor

Dept. of ----------

Coll. Of Vet.Med./ Univ. of Al-Qadisiyah.

-- / -- / 2020

Certificate of Department

We certify that ---------- has finished his/her Graduation Project entitled (----------) and candidate it for debating.

Instructor

Dr. Muthanna H. Hussain

-- / -- / 2020

Head of Dept of Int. and Prev. Med.

-- / -- / 2020

Acknowledgment

subject **Page number**

First of all we thank God for the most mercy for enabling us to present this project in the best form that we wanted to be, we would like to thank our supervisor of this project, *Dr. Ahmed Jassim* for his valuable help and advice to come out with this project. We thank our faculty and doctors that provided us with all the knowledge. Most of all we are all thankful for our families for their endless love, assistance, support and encouragement. And for our friends for their understanding and support for us to complete this project

CHAPTER 1

INTRODUCTION

Pathogenicity islands (PAIs) are gene clusters incorporated in the genome, chromosomally or extra chromosomally, of pathogenic organisms, but are usually absent from those nonpathogenic organisms of the same or closely related species.[2][4][5] They may be located on a bacterial chromosome or may be transferred within a plasmid or can be found in bacteriophage genomes.[2] The GC-content and codon usage of pathogenicity islands often differs from that of the rest of the genome, potentially aiding in their detection within a given DNA sequence, unless the donor and recipient of the PAI have similar GC-content[2]

PAIs are discrete genetic units flanked by direct repeats, insertion sequences or tRNA genes, which act as sites for recombination into the DNA. Cryptic mobility genes may also be present, indicating the provenance as transduction.[3] PAIs are flanked by direct repeats; the sequence of bases at two ends of the inserted sequence are the same. They carry functional genes, such as integrases, transposases, phagocytosis, or part of insertion sequences, to enable insertion into host DNA.[2] PAIs are often associated with tRNA genes, which target sites for this integration event.[2] They can be transferred as a single unit to new bacterial cells, thus conferring virulence to formerly benign strains.[4]

PAIs, a type of mobile genetic element, may range from 10-200 kb and encode genes which contribute to the virulence of the respective pathogen.[2] Pathogenicity islands carry genes encoding one or more virulence factors, including, but not limited to, adhesions, secretion systems (like type III secretion system), toxins, invasions, modulins, effectors, super antigens, iron uptake systems, o-antigen synthesis, serum resistance, immunoglobulin A proteases, apoptosis, capsule synthesis, and plant tumorigenesis via A. tumefactions.[2]

There are various combinations of regulation involving pathogenicity islands. The first combination is that the pathogenicity island contains the genes to regulate the virulence genes encoded on the PAI.[2] The second combination is that the pathogenicity island contains the genes to regulate genes located outside of the pathogenicity island.[2] Additionally, regulatory genes outside of the PAI may regulate virulence genes in the pathogenicity island.[2] Regulation genes typically encoded on PAIs include AraC-like proteins and two-component response regulators.[2]

PAIs can be considered unstable DNA regions as they are susceptible to deletions or mobilization.[2] This may be due to the structure of PAIs, with direct repeats, insertion sequences and association with tRNA that

enables deletion and mobilization at higher frequencies.[3] Additionally, deletions of pathogenicity islands inserted in the genome can result in disrupting tRNA and subsequently affect the metabolism of the cell.[4]

The P fimbriae island contains virulence factors such as hemolysis, pili, cytotoxic enclosing factor, and uropathogenic specific protein (USP).[6] Yersinia pastis high pathogenicity island I has genes regulating iron uptake and storage. Salmonella SP1 and SP2 sites.[4] Rhodococcus equi virulence plasmid pathogenicity island encodes virulence factors for proliferation in macrophages.

The SaPI family of Staphylococcus aureus pathogenicity islands, mobile genetic elements, encode super antigens, including the gene for toxic shock syndrome toxin, and are mobilized at high frequencies by specific bacteriophages.[7]

Phage encoded Cholera toxin of Vibrio cholerae, Diphtheria toxin of Corynebacterium diphtheriae, Neurotoxins of Clostridium botulinum and Cytotoxin of Pseudomonas aeruginosa.[3]

H. pylori has two strains, one being more virulent than the other due to the presence of the Cag pathogenicity island.[3]

CHAPTER 2

(REVIEW AND LITERATURES)

PAIs-Related Features

When comparing the genomic region of PAIs and the remaining parts of the host genome, we can usually find that PAIs have their own genomic characteristics such as containing mobility genes, containing virulence genes, and having their own sequence signature. Figure 1 illustrates a schematic view of a PAI. The PAI associated features and corresponding measurement methods summarized

in Table 1. **Figure 1.** A schematic view of a pathogenicity island with associated features.

The PAI region has biased sequence composition. The PAI regions are associated with virulence genes (vir1, vir2, vir3, and vir4), phage-related genes (phag1 and phag2), mobile genes (int and trans), hypothetic protein genes (hypo1, hypo2, and hypo3), insertion sequence elements, direct repeats, and tRNA gene.

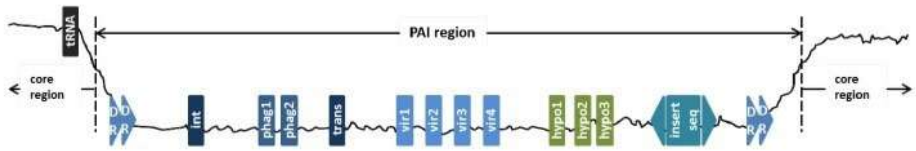

Table 1. A list of Pathogenicity Islands (PAI)-associated features and measurement methods.

PAI-Associated Features	Feature Measurement Methods
Different genomic sequence signature	Compute G+C content, GC-skew, codon usage, or other sequence signature tools
Presence of virulence factors	Search through virulence factor database such as VFDB
Presence of mobility genes (integrases, transposes)	Search through NCBI-nr/nt, UniprotKB, Pfam or COG database
High percentage of phage-related genes	Search through NCBI-nr/nt, UniprotKB, Pfam or COG database
Presence of tRNA genes	Use tRNA gene search tool of tRNAscan-SE
High percentage of hypothetic protein genes	Search through NCBI-nr/nt, UniprotKB, Pfam or COG database
Presence of direct repeats	Use repeat finder software REPuter
Presence of insertion sequences	Search through ISfinder database

Genomic Sequence Signature

In general, each genome has its unique genome signature, which can be measured by G+C content, dinucleotide frequencies (or other *k*-mer frequencies), and codon usage. This is because PAIs were originally transferred from other pathogenic bacteria, plasmids, or phages, and, thus,

the genomic sequence structure of PAIs is different from that of the rest of host genome.

G+C Content and GC-Skew

The G+C (%) contents (*i.e.*, the percentage of guanine and cytosine bases) in PAIs are often different from that of the host organisms. For instance, the G+C content of the Uropathogenic *E. coli* core genome was 51%, while the G+C content was 41% in PAI I, II, IV, and V [6–8]. In the genome of Enteropathogenic *E. coli*, the G+C content of PAI of LEE was only 39% [9]. G+C content differences between PAIs and the core genomes have also been found in other genomes, such as *D. noddosus* [10], *H. pylori* [11], and *Y. pestis* [12].

A slightly different measure can also be used for measuring the difference between PAI and core region. For instance, a large scale comparative genomic analyses of 1,088 bacterial and Archaeal genomes showed that PAIs were anchored around switch sites of GC-skew (sGCS), which was measured by [G−C]/[G+C] [13].

k-Mer Frequency

The measurements of dinucleotides or high-order oligonucleotide frequencies have been increasingly used [14].

Theoretically, the higher-order measurement used, the more accurate to differentiate two genomes, given the assumption that the genomic region for measurement is long enough to evaluate all combinations of oligonucleotide patterns (or words). For instance, if 6-mer frequency is used, then there will be 46 = 4,096 words, and, thus, a genomic region with at least several *kb* is required if 6-mer frequency is used. We have seen several approaches such as Alien Hunter [15] and Centroid [16] that used *k*-mer frequencies to predict island regions.

Codon Usage

Codon usage is another useful feature to tell the differences between two genomes. Generally speaking, each genome has its own preferred codon usage, and thus the codon usage in a genome region will be significantly different than the rest of host genome if this region was transferred from outsider. SIGI-HMM software uses codon usage bias to predict GIs [17].

Caveat

While PAIs have skewed sequence composition, highly expressed genes (HEGs) (including ribosomal related genes, chaperonin genes, transcription and termination factor genes, energy metabolism genes, recombination and repair genes, and electron transport genes) may also have codon usage bias and dinucleotide bias [18]. In this scenario, using sequence composition information only to detect islands in pathogenic bacteria will lead to the problem of *false positives* (*i.e.*, predicted PAIs might actually be HEGs).

On the other hand, it is possible that the donor and recipient organisms have similar sequence composition, thus, making it difficult to dig out those real PAIs sporadically distributed in the core genome. Furthermore, even the donor and recipient organisms have different sequence compositions, it is very likely that the PAI region will be eventually ameliorated, a process that makes the sequence composition (or codon usage) of the alien genomic region (*i.e.*, PAIs) be similar to that of the core genome, so that the integrated region can be adapted to enhance expression [19]. A recent large scale genomic study of 1,088 bacterial and Archaeal genomes has shown the newer acquired PAIs were closer to sGCS than the older ones, implying that the older PAIs are in the process of amelioration [13]. In this scenario, using sequence composition

information only will lead to the problem of *false negatives* (*i.e.*, the actual PAIs may not be discovered easily).

EVOLUTION AND TRANSFER OF PAI

The observation that important virulence factors are present in very similar forms in different bacteria may be explained by horizontal gene transfer. Different scenarios can be considered to explain the transfer between bacterial strains and species.[14]

Natural Transformation

Certain bacteria are capable of natural transformation. During certain phases of growth, transport systems are expressed that allow the uptake of free DNA from the environment.

Although the majority of this foreign DNA will be degraded, some fragments that harbor "useful" genes are integrated into the genome of the recipient and maintained. It appears possible that this mechanism allows uptake of DNA from distantly related species that will be maintained as the selective pressure selects for the newly acquired features.[14]

PAI and Plasmids

Similar clusters of virulence genes are present in PAI and on virulence plasmids, indicating that episomal and chromosomal locations are possible for the same gene cluster. It was observed that certain clusters of virulence genes are present in

PAI of some pathogens but also on virulence plasmids in other bacteria. The T3SS required for invasion of epithelial cells by *Shigella* spp. is encoded by the *mxi* and *spa* genes located on a virulence plasmid, and a related gene cluster that is required for the invasiveness of *Salmonella enterica* is located in SPI-1 in a chromosomal location.

Conjugation can allow the transfer of plasmids between bacteria. These plasmids can then replicate a utonomously from the bacterial chromosome, but under certain conditions plasmids may also integrate into the chromosome. Conversely, the formation of episomal elements has been reported for certain PAI of *Staphylococcus aureus*. Thus, plasmids could be another means of transfer of PAI between bacteria.[15]

Transduction

Bacteriophages have been isolated from virtually all bacterial species; even obligate intracellular pathogens such as *Chlamydia* spp. contain specific phages. Bacteriophages are able to transfer bacterial virulence genes as passengers in their genomes.

The occasional transfer of virulence genes by phages allows the recipient bacteria to colonize new habitats, such as new host organisms or specific anatomic sites. This extension also allows a more efficient spread of the bacteriophages. Thus, the transfer of bacterial virulence genes as passengers in the viral genome can also be an evolutionary benefit for the bacteriophage.

A well-characterized example of the contribution of bacteriophages to the evolution of bacterial virulence is found in *V. cholerae* (see "*Vibrio cholerae*" below).

Many PAI are too large to be transferred as passengers in bacteriophage genomes. For example, gene clusters on PAI encoding T3SS or T4SS comprise 25 to 40 kb DNA, which is almost equivalent to the total genome size of a bacteriophage.[16]

In these cases, other mechanisms are conceivable. Certain bacteriophages are capable of generalized transduction. Normally, for the replication of the phage within the host bacterium, copies of the phage genome are packaged into phage heads.

During replication, the host DNA is fragmented. Occasionally, the enzymes involved in packaging the phage genome erroneously pack a fragment of the host genome into the phage head.

Since the resulting particles are still able to infect a new bacterial host, a fragment of the bacterial DNA can be transduced.[17]

Given sufficient sequence similarity, recombination may occur and the transduced fragment is integrated into the genome of the new host.

PAI do not occur only in human pathogens; they have also been found in animal and plant pathogens. Examples are the *hrp* islands of *Pseudomonas syringae* and *Xanthomonas campestris*, and islands in animal pathogenic salmonellae and staphylococci. They are distributed throughout the bacterial world, and horizontal transfer may be facilitated by plasmids and phages or by bacteria, which are competent for the uptake of free DNA by natural transformation.[17]

INTEGRATION SITES OF PAI

Integration of PAI into the bacterial chromosome is a sitespecific event. Most PAI currently known have inserted at the 3_ end of tRNA loci. Also, phage attachment sites frequently are located in this region. However, certain genes, and infrequently intergenic regions in operons are used by PAI. In members of the *Enterobacteriaceae*, the *selC* locus is an insertion site frequently used by functionally different PAI in *E. coli*, *Shigella* spp., and *S. enterica* (Fig. 4). The overlapping sequences of tRNA loci and PAI are within the 3_ end of the tRNA genes, are usually 15 to 20 nucleotides long and encode the 3_ side of the acceptor-T_C stem-loop region of a tRNA up to the conserved CCA end (156). The molecular basis of the use of tRNA genes as integration sites is not fully understood, but three hypotheses are plausible.[18]

Specific tRNAs are associated with a PAI, so that the encoded tRNA may be used to read codons of the associated PAI. This has been shown for the

leuX tRNA gene encoding the rare tRNALeuX . Expression of *leuX* is necessary for the synthesis of virulence factors encoded on PAI536. Since basic cellular genes that are not involved in pathogenicity also are modulated by *leuX* and since the association of PAI with specific tRNA genes is not found in other islands, this thesis is not favored.[18]

A second hypothesis would include the presence of multiple copies of tRNA genes, providing multiple insertion sites and amplification of pathogenicity factors. This is, however, not true for *selC* and *leuX*, which occur in single copies. The third, and most plausible, hypothesis suggests that the conserved structure in tRNA genes provides structural motifs that facilitate the integration and excision of PAI and also phages .[19]

This emphasizes that integration and excision are catalyzed by integrases.

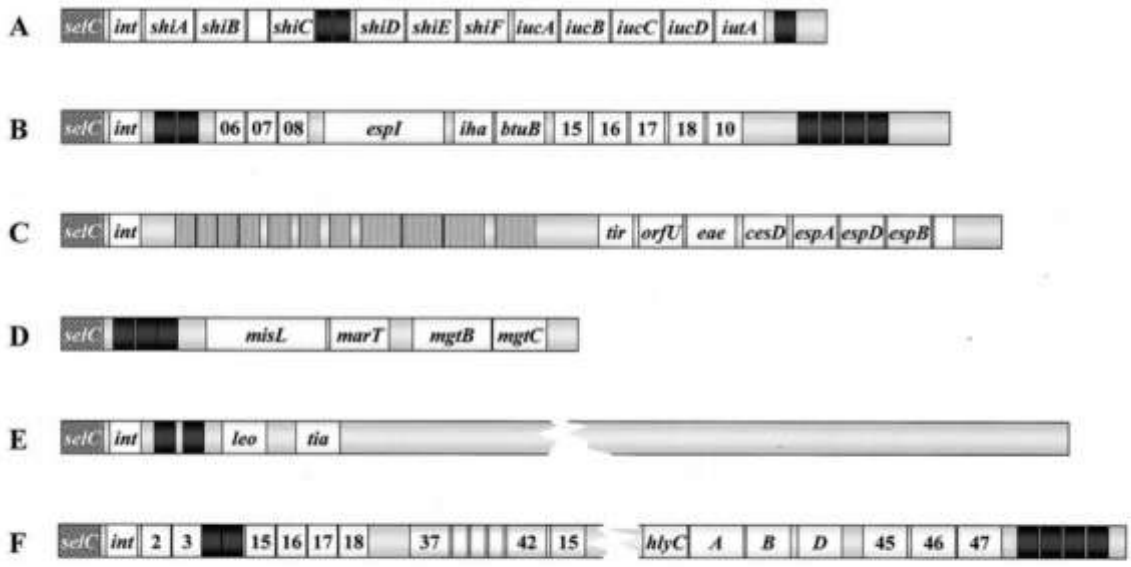

FIG. 4. Comparison of various PAI integrated at the *selC* locus. This schematic drawing of PAI demonstrates that the *selC* tRNA locus may have served as an integration site of PAI with different functions in different organisms either by means of a phage integrase or by other unknown events. (A) SHI-1 of *S. flexneri*; (B) LPA of STEC; (C) LEE of EPEC; (D) SPI-3 of *S. enterica*; (E) Tia-PAI of ETEC; (F) PAI I_{536} of UPEC. Numbers and gene designations are adapted from the original papers (20, 68, 82, 94, 95, 282, 309). ORF are depicted as rectangles: dotted grey, RNA *selC*; white, phage-like integrase gene; dark grey, mobility genes; light grey, all other PAI genes. See the text for details.

Hou (156) proposed a fourth hypothesis, in which the 3_ end of tRNA plays a major role. In his hybrid theory, the conserved CCA ends provide the initial site for integration by an integrase. The 3_ end of a tRNA

hybridizes to one strand of a duplex DNA during recombination. This stabilizes the separation of the DNA duplex for recombination . [20]

Whether this theory or one of the others is correct has yet to be elucidated. Nevertheless, it is apparent that phages and PAI use conserved genes as integration sites. These conserved genes might confer safety to the mobile genetic element that they can integrate in any genome of members of a given population. This need could be, in an evolutionary biology point of view, important to maintain pathogenicity factors in a bacterial population [20]

PAI OF GRAM-NEGATIVE PATHOGENS

Helicobacter pylori

H. pylori infects the mucosa of the stomach, an organ that has long been considered an environment too hostile for bacterial colonization. Infections with *H. pylori* are common and are often acquired in childhood, and acute infection can lead to chronic colonization of the gastric mucosa (for a recent review, see reference 342). This colonization usually leads to chronic gastritis, and subsequent forms of disease are dependent on host as well as on bacterial factors. In the majority of individuals with gastritis, the infection remains asymptomatic.[21]

However, patients with low or high production of gastric acid can develop gastric ulcer or duodenal ulcer, respectively. There is also a strong correlation between infection with *H. pylori* and development of mucosa-associated lymphoid tissue lymphoma and gastric cancer, resulting in the classification of *H. pylori* as a carcinogen. *H. pylori* organisms are curved, rod-shaped bacteria with a group of polar flagella and are covered by a membrane sheath.

Motility is an important virulence factor and enables the bacteria to penetrate the mucin layer of the gastric epithelium . The bacteria also produce urease. This enzyme catalyzes the formation of CO2 and ammonia that can neutralize the acidic pH in the vicinity of the bacteria. Cultivation of *H. pylori* requires a microaerophilic atmosphere and complex media. Clinical isolates of *H. pylori* have been classified into type I and type II strains, which are associated with different clinical outcomes ranging from gastric ulcer to asymptomatic colonization.

There are also various forms of intermediate virulence. Type I strains carry genes encoding both, the cytotoxins CagA and VacA, while type II strains contain *vacA* genes only (376). VacA is a secreted toxin that induces extensive vacuolation in epithelial cells, cell death, and destruction of epithelial integrity. The attachment of type I strains to gastric epithelial cells induces the synthesis and secretion of several chemokines, and the secretion of interleukin-8 (IL-8) is frequently assayed in model systems. It has also been observed that the infection of epithelial cells by *H. pylori* leads to dramatic rearrangements of the host cell actin cytoskeleton and the formation of pedestals that are reminiscent of EPEC-induced pedestals, as well as to changes in the gross morphology of host cells (hummingbird phenotype). These phenotypes are associated with alterations in the signal transduction pathways of the host cell and the presence of a tyrosine phosphorylated protein .

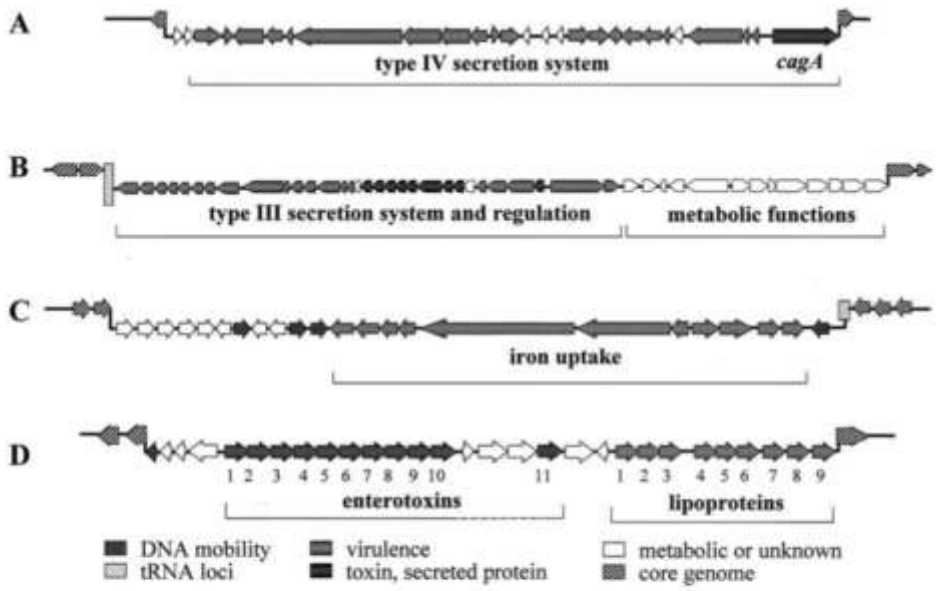

FIG. 5. Examples of PAI of various pathogens. The topology of PAI of various pathogens is depicted to demonstrate different features of PAI. The functional classes of the genes are as indicated in the figure. (A) The *cag* island of *H. pylori* harbors genes for a type IV secretion system (T4SS) (grey symbols) that can mediate the translocation of the effector protein CagA (dark grey) into eukaryotic cells modified from reference 92. (B) *Salmonella* SPI-2 has a mosaic structure. It has been defined as a genetic element of about 40 kb that is absent from the related species *E. coli*. Only a 25-kb portion is required for systemic infection and encodes a T3SS system (grey), secreted proteins (dark grey), and regulatory proteins (white). Another portion (15 kb) is not required for virulence and harbors genes for metabolic or unknown functions (light grey symbols), such as an enzyme system for alternative electron acceptors during anaerobic growth. Genes associated with mobility are indicated by dark dotted symbols. Modified from reference 134. (C) The HPI of *Y. enterocolitica* is an example of an unstable PAI. Several is elements are present within this PAI (dotted arrows). Genes in HPI encode an high-affinity iron uptake system (dark grey) that is important for the extracellular proliferation of the pathogen during colonization of the host. Modified from reference 45. (D) The *v*Sal PAI of MRSA is shown. A remarkable feature of PAI in *S. aureus* is the presence of a large number of genes with related functions, such as genes for enterotoxin (dark grey) or lipoproteins (grey). Modified from reference 9.

***cag* PAI.** Detailed analysis of the *cagA* loci in type I and type II strains indicated that the latter group showed deletions of a large chromosomal region. This locus had the typical characteristics of a PAI and was termed the *cag* PAI . Censini et al. characterized this locus and showed that the *cag* PAI had a size of 37 to 40 kb, flanked by direct repeats of 31 bp (Fig. 5A). The locus has a G_C content of 35%, in contrast to the 39% observed for the core genome . A gene for a tRNA has not been identified at the point of integration, but the *glr* gene (glutamate racemase) was disrupted by insertion of the PAI. There are no genes associated with DNA mobility within the *cag* PAI of type I strains. However, the presence of an IS *605* element within the *cag* PAI of strains with an intermediate virulence phenotype was observed. In strains of intermediate virulence, various forms of deletions with the *cag* PAI were detected, and in certain strains the locus was separated into two portions, referred to as *cag*I and *cag*II . These observations support a correlation between the

presence and integrity of the *cag* PAI and the severity of disease. Studies with a mouse model have shown that an association between *cag* PAI-negative *H. pylori* strains and *cag* PAI-positive strains that are mouse adapted and have modulated their ability to activate a proinflammatory response can better colonize mice than the parental strains do, indicating that the *cag* PAI of type I strains may become lost during colonization of infected animals . In addition to large deletions and chromosomal rearrangements of the *cag* PAI, there are indications that point mutations in the PAI genes result in ablation of CagA translocation and IL-8 induction . This effect can be explained by loss of function of the T4SS.[22]

Work by several groups has demonstrated that the translocation of CagA into target cells is required for these phenotypes . After translocation, CagA is typrosine phosphorylated and induces growth factor-like phenotypes in the host cell. SHP-2 (SRC homology 2 domain [SH2]-containing tyrosine phosphatase) was identified as a cellular target of CagA . It was observed that SHP-2 and CagA form a complex that could activate particular pathways and lead to actin polymerization and pedestal formation . Activation of SHP-2 by CagA might contribute to the abnormal proliferation and movement of gastric epithelial cells, thus contributing to the pathogenesis of *H. pylori* gastric infections. It has also been proposed that phosphorylated CagA may trigger the transcription of nuclear genes, which may explain the increased frequency of gastric cancer in patients infected with *cagA*-positive *Helicobacter* strains . Sequence analysis of the *cag* PAI indicated that a T4SS is encoded by this locus and that CagA is a translocated substrate of the secretion apparatus. Only 6 of the 27 to 29 predicted open reading frames (ORF) in the *cag* PAI show significant sequence similarity to components of the

T4SS of other bacteria, and the contribution of other genes to formation of a translocation apparatus is not clear (Fig. 5A).[23]

Systematic mutagenesis approaches to analysis of the functions of the genes in the *cag* PAI were performed by Fischer et al. and Selbach et al. . The individual inactivation of 27 putative genes and phenotypic analysis identified a subset of 17 genes that are required for the translocation of CagA into host cells and a subset of 14 genes that are required for the stimulation of IL-8 synthesis in host cells. Although the assembly of T4SS is not understood in full detail, these observations indicate that the majority of genes within the *cag* PAI are required for the formation of a functional T4SS, by encoding either structural components or protein important for the assembly and regulation of the system. Neither approach resulted in the identification of mutant strains that were deficient in CagA translocation but capable of inducing IL-8 secretion. These observations indicate the absence of a further translocated protein responsible for IL-8 induction within the *cag* PAI or a direct effect of the T4SS in IL-8 induction. The secretion of VacA is not dependent on the *cag* PAI, and so far no further proteins translocated by *cag* PAI-encoded T4SS have been identified. The observation that the *cag* PAI is absent or partially deleted in *H. pylori* strains with low virulence might suggest that the function of the *cag* PAI-encoded T4SS is not compatible with a long-lasting colonization of the gastric epithelium. The inflammatory response elicited by *H. pylori* after contact-dependent translocation could lead either to a clearance of the infection or to a severe immunopathology. However, epidemiological data indicate that the frequency of *cag*_ isolates of *H. pylori* is much higher in the Asian population than in the Western population, indicating that further host and pathogen factors are involved in colonization . The genome of *H. pylori* is characterized by a

high flexibility, and an extremely high frequency of recombination was observed . The DR flanking the *cag* PAI probably function as sites for recombination and deletion of the locus.[24].

CHAPTER 3

Discussion

Although we live in a sea of seemingly harmless microbes, a small number of bacteria possess the ability to cause disease. A fundamental question in the study of microbial pathogenesis is to determine what genetic components enable a pathogen to cause disease, while closely related organisms remain relatively harmless. Much progress has been made in the last few years in our understanding of microbial diseases, especially characterizing the genetic elements that encode virulence factors. However, only recently has it become apparent that diverse virulence factors are genetically clustered together selectively in pathogen genomes.

In 1994, Jörg Hacker and colleagues coined the phrase "pathogenicity island" (PAI) (Blum et al., Infect Immun. *62*, 606–614, 1994) to describe a region of 70 kb of DNA in uropathogenic *E. coli*'s chromosome that encoded virulence factors (including α-hemolysin), yet was not present in nonpathogenic *E. coli*. They showed that this region of DNA was inserted in the *selC* tRNA region. It is now known that uropathogenic *E. coli* contains at least four PAIs inserted into various tRNA sites, ranging in size from 40 kb to 190 kb, encoding diverse virulence factors such as fimbriae, iron utilization, and α-hemolysin. Two of the PAIs are flanked by direct repeats, indicative of genetic mobility.

The PAIs of uropathogenic *E. coli* exemplify the composition and definition of pathogenicity islands. PAIs, by definition, encode one or

more virulence factors such as adhesins, invasions, or secretion systems needed to export virulence factors. PAIs are present in the genome of pathogenic bacteria but absent from nonpathogenic members of the same or closely related species. Although originally defined as chromosomal elements, it is now apparent that regions of virulence plasmids also constitute PAIs. The size of a PAI is usually relatively large, ranging from 10 to 200 kb. Small insertion sequences encoding single virulence factors are considered islets, although they obviously share some properties with PAIs.

A hallmark of PAIs is that their G+C DNA content is usually quite different from the remainder of the genome, indicating horizontal gene transfer as a mechanism of acquisition. Because of their potential genetic mobility, they are often, but not always, flanked by direct repeats. Their site of insertion is often associated with tRNA genes, which are sites of integration of foreign DNA and attachment sites for bacteriophages. Finally, PAIs are often unstable DNA regions. Like most definitions in biology, there are exceptions to each criteria, and despite their similar functional characteristics, PAIs are actually quite heterogeneous.

PAIs represent a major paradigm shift in evolutionary theory. Previously, pathogens were thought to have evolved in minor steps, slightly modifying existing factors until they gained a new function that was useful in virulence. By enabling the organism to exploit a new host niche, this selective advantage ensured that the virulence factor was maintained. However, PAIs have forced us to rethink pathogen evolution, realizing that changes occurred in quantum leaps, corresponding to acquisition of PAIs (E. Groisman and H. Ochman, Cell *87*, 791–794, 1996; C. Lee, Infect. Agents Dis. *5*, 1–7, 1996). We also now realize that evolution by leaps is not limited to virulence factors. Instead, acquisition of large

segments of DNA occurs throughout the microbial world, transferring genes encoding many functions often not involved in pathogenesis. For example, approximately 17% of the *E. coli* K12 genome appears to be acquired by recent horizontal transfer. Since these regions do not encode virulence factors, they are instead called "genomic islands." *E. coli* O157:H7, the causative agent of food poisoning from hamburger, has approximately 20% more DNA than its nonpathogenic K12 relative inserted in large segments throughout the chromosome, again indicating that horizontal transfer of large amounts of DNA is an active process in the microbial world. Indeed, it is thought that *E. coli* O157:H7 is a recently evolved pathogen which arose from acquiring a pathogenicity island which encodes a type III secretion system and effectors (the locus of enterocyte effacement), and a small prophage which encodes Shiga toxin. An attractive hypothesis is that microbes possess a "core genome," to which various genetic upgrades are added, providing the organism with extra factors that confer an additional selective advantage, such as the ability to cause disease. Thus PAIs provide us with tools to study bacterial evolution, which can then be extended to many mobile genetic events.

Once entering a new host, further PAI evolution occurs. This might include deletion or alteration of direct repeats that were needed for successful integration (which would "lock in" the PAI), or other modifications to improve gene expression (such as preferential codon alterations). Additionally, many PAIs are regulated by genetic elements distant to the PAI, or even on another genetic element such as a plasmid. Presumably, once acquired, pathogens fine tune the regulation of PAI expression to coincide with infection, often linking it to expression of

other preexisting virulence factors, thereby yielding a coordinated response.

PAIs have been described predominantly in Gram-negative pathogens, but this is probably because the pathogenicity mechanisms of these pathogens are better described. However, Gram-positive pathogens such as *Listeria* also contain PAIs. Although it is unknown whether PAIs occur in eukaryotic pathogens such as parasites or fungi, one would predict that they are probably present, due to the genetic flexibility of these organisms.

With the increasing number of pathogen genome sequences available, PAIs represent an obvious region within which to identify new virulence factors. However, PAIs are not always apparent in a genomic sequence. For example, *Chlamydia* species encode several genes of a type III secretion system (found in several PAIs), yet the G+C content of the *Chlamydia* type III genes is similar to the rest of its genome. One possibility is that this type III secretion is the ancestor of other type III systems, or, more likely, the G+C content of the type III system happens to be the same as *Chlamydia*.

In a recently released book entitled *Pathogenicity Islands and Other Mobile Virulence Elements*, a detailed summary of what is known about PAIs is presented by many of the best researchers studying these sequences. Each chapter describes the PAIs in a particular pathogen, written by definitive experts on that particular pathogen. In addition, there is an excellent introduction by the editors in the first chapter in which they put forward their definition of PAIs (as outlined above) and other useful thoughts about PAIs, including their composition, origin, and mechanisms of genetic mobility. They also include a list comparing various PAIs. Throughout the various chapters, many authors also offer their own

definitions of PAIs, which are generally similar to the editors' version, but still offer differences that challenge any complete definition, reflecting the heterogeneity of PAIs. Perhaps the best definition of a PAI is found at the end of the first chapter: “A pathogenicity island is what a good microbiologist terms a pathogenicity island” (pg. 10), although defining a “good microbiologist” is a separate issue…

The book begins with a thoughtful and historical perspective of PAIs and mobile genetic elements provided by Stanley Falkow in the foreword. Falkow discusses the way *Salmonella* has evolved by sequentially acquiring two type III secretion systems—the first to invade host cells (SPI-1) and the second to live within macrophages (SPI-2)—in addition to the acquisition of at least three other PAIs.

The second chapter discusses the various methods for detecting mobile genetic elements encoding virulence factors, and determining gene expression during infections. The remainder of the book moves through the better studied PAIs of human, animal, and plant pathogens, including a few chapters on other mobile genetic elements. Common themes are apparent throughout the book, including type III secretion systems in human pathogens such as *E. coli*, *Yersinia*, *Shigella*, *Salmonella*, and in the plant pathogen *Pseudomonas syringae*. Type IV secretion is discussed in the context of the *Helicobacter pylori cag* PAI and the Ti plasmid DNA transfer system of *Agrobacterium*. Virulence cassettes common to different pathogens are mentioned throughout the book, including islets, islands, and even archipelagos! The *prfA* virulence gene cluster of *Listeria monocytogenes* has been a rich source of *Listeria* virulence factors, and this region, as well as the internalin multigene family, is also discussed.

One of the more sophisticated mechanisms of virulence factor mobility is that found in *Vibrio cholerae*. As discussed in chapter 9, this pathogen

encodes cholera toxin, a critical toxin needed for cholera disease, as part of a filamentous phage. The receptor for the cholera toxin phage is the toxin-coregulated pilus (TCP). Ironically, TCP is actually encoded as part of another filamentous phage that constitutes the *V. cholerae* pathogenicity island (VPI). Discovery of these two phages and the virulence factors they encode has had a major impact on our understanding of *V. cholerae* pathogenesis and the evolution of cholera outbreaks.

The last chapter discusses conjugative transposons, which are similar to PAIs in that they integrate foreign DNA sequences that can excise from the chromosome, circularize, and then transfer to another bacterium.

Because this is the first comprehensive book on PAIs, it represents a refreshing view of bacterial pathogenesis. Most reviews focus on either the disease, the virulence factors that cause disease, or the secretion systems used to export the virulence factors. By thinking about these pathogens from a PAI perspective, new insights into pathogen origin and genetic organization become apparent. A potential criticism of the book is that there is little speculation about where the PAIs came from before landing in their current hosts. The origin of PAIs, and the factors encoded within them (such as type III secretion systems) remains unknown. As we move into the postgenomic era in microbiology, understanding PAIs becomes increasingly important, since many new PAIs will be identified. This book is written at an advanced level, yet presented in a logical order and easy to follow. In addition to being the authoritative reference on pathogenicity islands, it will serve as a valuable resource for students and scholars interested in bacterial pathogenesis, bacterial evolution, mobile genetic elements, and genomics.

Conclusions

The cdt positive isolates association with stx 1, and hly plasmid genes could lead us to the idea that these virulence genes are associated with mobile genetic elements, which help them to spread among a wide spectrum of bacteria. The existence of prophages in almost all types of CDT producer strains could be the evidence of recombination events between bacterial chromosome, plasmids, and phages that bring the existing diversity among the strains.

Recommendation

Pathogenicity islands (PAIs) are distinct genetic elements on the chromosomes of a large number of bacterial pathogens. PAIs encode various virulence factors and are normally absent from non-pathogenic strains of the same or closely related species.

(PAIs) are discrete DNA segments of ~10 to >100 kbp that encode virulence factors and other accessory proteins, but no essential proteins. They reside within larger genetic units, usually the chromosomes, occasionally within plasmids or bacteriophages.

The function of PAIs carry genes encoding one or more virulence factors, including, but not limited to, adhesins, toxins, or invasins. They may be located on a bacterial chromosome or may be transferred within a plasmid Pathogenicity Islands and Their Role in Bacterial Virulence and Survival

Q Infections caused by bacterial pathogens are still a significant problem in modern medicine . Therefore , the identification of the factors that are related to the infections and the understanding of the processes involved in the evolution of pathogenic bacteria from their nonpathogenic progenitors is an important subject of research . It has long been known that acquisition of virulence deter infants by horizontal gene transfer is

one of the major driving forces in the emergence and evolution of new pathogens. Furthermore , our knowledge of the organization of the bacterial genome has greatly increased within the last few years due to the availability of more than 120 completely sequenced eubacterial genomes , including those of almost all pathogenic bacteria , which has introduced a new area of pathogen research . It has become evident that the typical bacterial genome consists of a conserved ' core gene pool ' comprising genes that encode essential structural features and fundamental metabolic pathways , and a flexible gene pool that is more variable and encodes functions only advantageous under specific growth conditions . Core genes are characterized by a relatively homogenous G + C content and they are normally encoded in stable regions of the chromosome that are conserved in their orga negation in closely related species . In contrast , the flexible gene pool comprises variable regions of the chromosome and various mobile genetic elements such as plasmids , bacteriophages , IS Clements and transposons , conjugative trans poisons , integrin’s and superintegrons that are transferred between different organisms by the means of natural transformation , transduction or conjugation . Many of the genes encoding toxins , adhesions , secretion systems , invasions or other virulence - associated factors have been found to be encoded by mobile genetic elements . Furthermore , the analysis of the genomes

PAI OF GRAM-NEGATIVE PATHOGENS

Helicobacter pylori H. pylori infects the mucosa of the stomach, an organ that has long been considered an environment too hostile for bacterial colonization. Infections with H. pylori are common and are often acquired in childhood, and acute infection can lead to chronic colonization of the gastric mucosa.

There are several other types of PAL of Gram_Negative Pathogens:

Pseudomonas aeruginosa, shigella spp., yersinia spp., Vibrio cholerae, salmonella, E_coli

PAI OF GRAM-POSITIVE PATHOGENS

Listeria monocytogenes

Only two of the six Listeria species currently recognized cause listeriotic, L. monocytogenes and L. Ivanovo. Whereas L. monocytogenes can infect humans and a wide range of animals including mammals and birds, L. ivanovii is mainly pathogenic for ruminants. Listeriosis is a food-borne infection, whose serious complications include sepsis and meningitis. It occurs mainly in immunocompromised patients, such as transplant recipients, infants and elderly persons receiving chemotherapy, individuals with diabetes or liver disease, and patients with human immunodeficiency virus infections. There are several other types of PAL of Gram_Positive Pathogens: Staphylococcus aureus, streptococcus spp., Enter cocas facials.

Conclusions

Pathognsity islands contribute to virulence and survival of pathogens in several ways.First, the acquisition of Pathognsity islands has been described as ‘evolution in quantum
leaps’, because they often carry more than one virulence or fitness determinant , These Pathognsity islands -encoded factors enable the bacterium to colonize novel niches
in the eukaryotic host and facilitate the adaptation to the respective environmental conditions. This increase of fitness gives an advantage over coresiding
bacteria. Furthermore, the genome of many bacterial pathogens contains more than one Pathognsity islands that encode important virulence factors, thereby determining the
capability to cause disease. In addition, to ensure coordinated expression of virulence or virulence-related genes that are located on GEIs, a tight connection
to regulatory networks of the bacterium has evolved, as well as a link of islandencoded regulators to genes encoded elsewhere in the genome. Finally, an
ongoing mobilization and transfer of Pathognsity islands as well as reorganization, partial or complete deletion of existing Pathognsity islands affect long-term (macro-) as well as short-term (micro-) evolution of pathogenic bacteria.

Recommendation

the development of computational approaches for finding the relationship between donor and recipient genomes through PAIs might also be very useful. The corresponding visualization tools for revealing their

relationships might be needed so that researchers can use them to study gene transfer mechanisms.

REFERENCE

1. Pallen, M.J.; Wren, B.W. Bacterial pathogenomics. *Nature* **2007**, *449*, 835–842.
2. Koskiniemi, S.; Sun, S.; Berg, O.G.; Andersson, D.I. Selection-driven gene loss in bacteria. *PLoS Genet.* **2012**, *8*, e1002787.
3. Maurelli, A.T. Black holes, antivirulence genes, and gene inactivation in the evolution of bacterial pathogens. *FEMS Microbiol. Lett.* **2007**, *267*, 1–8.
4. Penn, K.; Jenkins, C.; Nett, M.; Udwary, D.W.; Gontang, E.A.; McGlinchey, R.P.; Foster, B.; Lapidus, A.; Podell, S.; Allen, E.E.; *et al*. Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. *ISME J.* **2009**, *3*, 1193–1203
5. Hacker, J.; Bender, L.; Ott, M.; Wingender, J.; Lund, B.; Marre, R.; Goebel, W. Deletions of chromosomal regions coding for fimbriae and hemolysins occur *in vitro* and *in vivo* in various extraintestinal Escherichia coli isolates. *Microb. Pathog.* **1990**, *8*, 213–225.
6. Blum, G.; Falbo, V.; Caprioli, A.; Hacker, J. Gene clusters encoding the cytotoxic necrotizing factor type 1, Prs-fimbriae and alpha-hemolysin form the pathogenicity island II of the uropathogenic Escherichia coli strain J96. *FEMS Microbiol. Lett.* **1995**, *126*, 189–195.
7. Blum, G.; Ott, M.; Lischewski, A.; Ritter, A.; Imrich, H.; Tschape, H.; Hacker, J. Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an Escherichia coli wild-type pathogen. *Infect. Immun.* **1994**, *62*, 606–614.
8. Swenson, D.L.; Bukanov, N.O.; Berg, D.E.; Welch, R.A. Two pathogenicity islands in uropathogenic Escherichia coli J96: Cosmid cloning and sample sequencing. *Infect. Immun.* **1996**, *64*, 3736–3743.
9. McDaniel, T.K.; Jarvis, K.G.; Donnenberg, M.S.; Kaper, J.B. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc. Natl. Acad. Sci. USA* **1995**,
10. *92*, 1664–1668.
11. Billington, S.J.; Sinistaj, M.; Cheetham, B.F.; Ayres, A.; Moses, E.K.; Katz, M.E.; Rood, J.I.

12. Identification of a native Dichelobacter nodosus plasmid and implications for the evolution of the vap regions. *Gene* **1996**, *172*, 111–116.
13. Censini, S.; Lange, C.; Xiang, Z.; Crabtree, J.E.; Ghiara, P.; Borodovsky, M.; Rappuoli, R.; Covacci, A. Cag, a pathogenicity island of Helicobacter pylori, encodes type I-specific and disease-associated virulence factors. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 14648–14653.
14. Fetherston, J.D.; Perry, R.D. The pigmentation locus of Yersinia pestis KIM6+ is flanked by an insertion sequence and includes the structural genes for pesticin sensitivity and HMWP2. *Mol. Microbiol.* **1994**, *13*, 697–708.
15. Du, P.; Yang, Y.; Wang, H.; Liu, D.; Gao, G.F.; Chen, C. A large scale comparative genomic analysis reveals insertion sites for newly acquired genomic islands in bacterial genomes. *BMC Microbiol.* **2011**, *11*, 135.
16. Karlin, S. Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes. *Trends Microbiol.* **2001**, *9*, 335–343.
17. Vernikos, G.S.; Parkhill, J. Interpolated variable order motifs for identification of horizontally acquired DNA: Revisiting the Salmonella pathogenicity islands. *Bioinformatics* **2006**, *22*, 2196–2203.
18. Rajan, I.; Aravamuthan, S.; Mande, S.S. Identification of compositionally distinct regions in genomes using the centroid method. *Bioinformatics* **2007**, *23*, 2672–2677.
19. Waack, S.; Keller, O.; Asper, R.; Brodag, T.; Damm, C.; Fricke, W.F.; Surovcik, K.; Meinicke, P.; Merkl, R. Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. *BMC Bioinforma.* **2006**, *7*, 142.
20. **Akopyants, N. S., S. W. Clifton, D. Kersulyte, J. E. Crabtree, B. E. Youree, C. A. Reece, N. O. Bukanov, E. S. Drazek, B. A. Roe, and D. E. Berg.** 1998. Analyses of the *cag* pathogenicity island of *Helicobacter pylori.* Mol. Microbiol. **28:**37–53.
21. **Alfano, J. R., A. O. Charkowski, W. L. Deng, J. L. Badel, T. Petnicki-Ocwieja, K. van Dijk, and A. Collmer.** 2000. The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants. Proc. Natl. Acad. Sci. USA **97:**4856–4861.
22. **Alfano, J. R., and A. Collmer.** 1997. The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death. J. Bacteriol. **179:**5655–5662.
23. **Al-Hasani, K., I. R. Henderson, H. Sakellaris, K. Rajakumar, T. Grant, J. P. Nataro, R. Robins-Browne, and B. Adler.** 2000. The *sigA* gene which is borne on the she pathogenicity island of *Shigella flexneri* 2a encodes an exported cytopathic protease involved in intestinal fluid accumulation. Infect. Immun.
24. virulence community-acquired MRSA. Lancet **359:**1819–1827.