Mashhad University of Medical Sciences (MUMS)

Reviews in Clinical Medicine

Clinical Research Development Center Ghaem Hospital

Black holes", "Genome fluidity", and Evolution of Bacterial Species

Kobra Salimiyan Rizi (Ph.D)[1,2*], Hadi Farsiani (Ph.D)[2]

[1]*School of Medicine, Isfahan University of Medical Sciences, Isfahan 81744176, Iran.*
[2]*Antimicrobial Resistance Research Center, Department of Medical Bacteriology and Virology, Qaem University Hospital, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran.*

ARTICLE INFO

Article type
Review article

Article history
Received: 10 Jul 2022
Revised: 14 Aug 2022
Accepted: 21 Sep 2022

Keywords
Bacterial genome evolutionary
Black holes
Shigella, Rickettsiae
Mycobacterium leprae

ABSTRACT

The formation of the massive deletion of genes called "black holes," which are detrimental to a pathogenic lifestyle, provides a bacterial evolutionary route that permits a pathogen to augment virulence and host-adaptability. For example, cadaverine substrate has inhibitory effects on the enterotoxin activity of the Shigella species. Therefore, the encoding gene of the lysine decarboxylase enzyme (CadA gene) as an anti-virulence gene is deleted from the genome of the Shigella. This and other similar cases in bacteria can be used for antitoxin therapy. Hence, identification of the role of black holes in the pathogenic evolution of bacteria output could possibly lead to novel treatments of infectious diseases in human beings. Here, we reviewed different types of pathoadaptation mutations among pathogens. Cases of black holes among the important human bacterial pathogens included Shigella, Rickettsiae, Mycobacterium leprae, Burkholderia, Bordetella, and Chlamydia. We found that the most prevalent pathoadaptive pathway among bacteria was the gene inactivation or deletion route.

Please cite this paper as:
Salimiyan Rizi K, Farsiani F. Black holes", "Genome fluidity", and Evolution of Bacterial Species. Rev Clin Med. 2022;9(3): 146-155.

Introduction

In the microbial world, we are witnessing the emergence and the exploitation of new bacterial strains with different characteristics of similar species. The development of pathogenic bacterial strains from non-pathogenic ancestors (1) and resistant bacteria from sensitive strains (2) is the outcome of the genetic evolution in bacteria, resulting in the subsequent adaptation or fitness of bacteria with the novel niche.

The evolution of bacterial genomes occurred via a diversity of processes at single cell or population levels, such as horizontal gene transfer (HGT) (intragenic recombination) (3), homologous recombination (HR; by the three types of gene transfers: conjugation, transduction, and transformation) (4), vertical genetic transmission (parent-offspring) (5), and genetic mutations due to the acquisition and loss of functional accessory genes in bacterial genomes. The analysis of pan-genome (the whole collection of genes within a species; supragenome) pathogenic bacteria indicates dynamics changes and adaptability in the bacterial genome proportionate to changes in the host body and their immune system (6).

Nowadays, this whole-genome sequence-based bioinformatics approach (in 2005) is applied for genetic diversity and comparative genomic analyses of pathogenic bacteria. The bioinformatics tools or pipelines, such as Gene-Mark, RAST, BPGA, PGAweb, De-Bruijn graph, Acyclic vari-

***Corresponding author:** Kobra Salimiyan Rizi,*
School of Medicine, Isfahan University of Medical Sciences, Isfahan 81744176, Iran.
***E-mail:** salimian.k@gmail.com*
***Tel:** 9382657837*

ation graph, Variation graph (VG), BLAST, MLST, GET_HOMOLOGUES, PGAP, MLSA, MCH (a clustering algorithm), PanOCT, Pangloss, PanX, Panseq, Pan4Drafts, Piggy, Hidden Markov Models (HMM), and Roary, are examples of tools applied in this field regarding bacteria (7).

For example, Streptococcus pneumoniae and Streptococcus pseudopneumoniae have high genetic similarity (due to HGT and homologous recombination); therefore, the correct identification of the Mitis-group streptococci from Viridans-group streptococci is particularly difficult for the unreliable phenotypic or a genotypic biomarker. Gonzales et al. (2020) conducted a pan-genomic analysis to identify species-unique gene markers among these strains. Their study results exhibited 10 genes unique for S.

pneumoniae (in 100% strains) and nine specific genetic markers for S. pseudopneumoniae (in >95% strains) (8). However, "genome plasticity" is another basis for determination of genetic variability among bacterial population without pangenomics.

This phrase is utilized to characterize the mobile genetic elements (MGEs) or "mobilome" and hyper-variable genetic regions converting the bacterial whole genome into a dynamic issue (9). Bacterial microevolution as an evolutionary process includes the genetic changes inside each of the genomes of bacterial population. Over time, these modifications can be augmented or weakened due to the effects of both genetic drift and natural selective pressure. Genetic drift demonstrates the evolution that occurred via the demise and birth of cells in the bacterial population (10), and it works randomly on all genetic variants. In small populations, these drifts have strong effects (11), and in large populations, their effects become weaker.Occasionally, to increase the virulence of a pathogen, one or more certain genes within the bacterial genome must be deleted and inactivated to enhance its survival by increasing the compatibility of the pathogen with its susceptible host and causing complete expression of the pathogen phenotype. Pathoadaptive mutations are involved in generating black holes (Table 1).

Table 1. Examples of the pathoadaptive mutations in bacteria

Type of pathoadaptive mutation mechanism	Examples in bacteria	Ref
Amplification of gene copies	• Acriflavin resistance protein D (AcrD) in EPEC	(13)
	• Iha adhesin in EHEC	(14)
	• P-fimbrial genes in some UPEC	(15)
Gene duplication	• Shiga toxin 2 (stx2) coding gene in some EPEC O157:H7 strains	(16)
	• Cold-shock protein A in E. coli	(17)
Gene inactivation or deletion	• Lysine-decarboxylase coding gene (cadA) in Shigella	(18)
	• ompT gene, encoding a surface protease in Shigella	(19)
	• Genes involved in the sugar metabolism, lipid–, aa–, & nucleotide biosynthesis in Rickettsiae	(20)
	• The mce4 operon in the cholesterol catabolism of Mycobacterium leprae	(21)
	• The L-arabinose assimilation operon (araBAD) & the flagellar motor encoding genes in Burkholderia mallei	(22, 23)
	• Down-regulation of acrABC locus expression in Bordetella pertussis	(24)
	• The LPS of B. pertussis never contains an O-antigen domain due to the ------ wbm locus inactivation	(25)
	• Missing of a conventional fructose-1,6-bisphosphatase in Chlamydia --------- trachomatis	(26)
Gene variation by point mutations	• The mutated FimH adhesin in UPEC strains	(27)
	• The mutated Afa/Dr-family & class-5 fimbrial adhesins in DAEC and -- ETEC isolates, respectively.	(28, 29)

These genetic mutations show a genetic route for increasing the bacterial virulence without the occurrence of the HGT process of specific virulence factors. The genes, whose restoring interferes with the pathogenic lifestyle of pathogenic bacteria, are called "anti-virulence genes" (AVGs). Natural selective pressure can lead to deletion of large regions of the bacterial genome (aiming at amplifying

suppressing of variations), which contain anti-virulence genes producing "black holes" (genome deletions) within the pathogen genome (12).

1.1.Shigella species

One of the pathoadaptive changes at the level of virulence gene expression is involved with the genes encoding the type III secretion system (T3SS) and invasion effectors in Shigella species, which are under regulation of growth medium temperature (30). The significant trigger, concluding the expression of the virulence plasmid entry region genes [encoding the type III secretion system (T3SS); mxi-spa core, a pathogenicity island (PAI)] is a temperature shift to 37°C after take-up by the host. Non-permissive environmental conditions are any temperature below 37°C and pH 6.0 (31).This genus has four species: S. dysenteriae, S. flexneri, S. boydii, and S. sonnei.

Shigella is considered a paraphyletic member of the species E. coli via the parallel evolutionary process. Obtaining novel properties via acquisition of the virulence plasmid or the invasion plasmid (pINV) develops the ability to produce several toxins, siderophore aerobactin, and immune evasion in Shigella. In this convergent evolution, Shigella species are non-motile (dissimilar to most E. coli) due to inactivation of flagella coding genes and lack of fimbrial adhesins. Four metabolic routes observed in E. coli are known to be deleted or inactivated in all Shigella isolates, such as nadA/nadB, cadA, speG, and ompT pathways (32). Yang et al. indicated that each Shigella species included a single circular chromosome and a large virulence plasmid (33). This virulence plasmid with a 230 kilobase (kb) size, includes a 31 kb region of invasion genes that needs the invasion and intercellular spreading of these bacteria (Fig. 1) (34, 35).

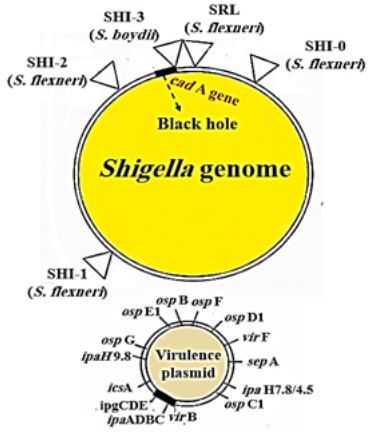

Figure1. The selected pathogenicity islands (SHI-0, SHI-1, SHI-2, SHI-3, SRL) location in the chromosome of *Shigella* & selected plasmid-encoded genes of *Shigella*. A closed dark square indicates the nearly location of the black hole in pINV(Torres AG, 2004).

The secreted proteins required for host cell entry are encoded by invasion plasmid antigen (ipa) genes, and the products of Mxi–spa genes (~20 numbers) are needed for ipa protein secretion (36). Specifically, the CadA metabolic pathway with the expression of lysine decarboxylase enzyme (LDC) in E. coli (>90 %) and some Enterobacterales are activated in conditions such as low external pH and the presence of their respective amino acid substrate (lysine). However, no isolates of enteroinvasive E. coli (EIEC) and Shigella, which are the bacterial etiology agents of bacillary dysentery, express LDC activity.

Early genetic and MLEE studies on Shigella and EIEC bacteria have shown their close genetic association with their commensal E. coli ancestor, as well as the terms of clinical features and pathogenesis. Furthermore, these observations indisputably demonstrate that Shigella spp. is the clone of E. coli (33, 37). The scientists speculated that probably in these strains the lack of activity and expression of the CadA gene could be important for their virulence. Cadaverine (1, 5-pentanediamine) is a polyamine produced by the LDC enzyme from the L-lysine substrate (Fig 2).

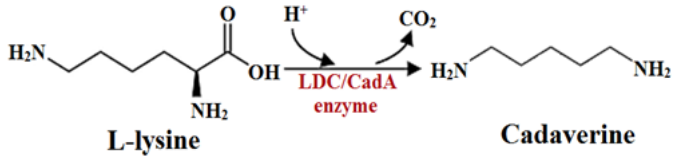

Figure 2. Converting of L-Lysine substrate into cadaverine by a CadA-mediated decarboxylation reaction (Park et al., 2017) (38)

Polyamines (cadaverine, spermidine, and putrescine) as small poly-cationic molecules in most cells are related to each other with various physiological reactions, such as translation process, gene regulation, resistance to stress, cellular proliferation, and differentiation. However, in Shigella, cadaverine blocks the activity of enterotoxins, and as a result, leads to a less severe disease.

It influences the invasive flow, inhibits the release of Shigella into the cytoplasm of infected cells, and blocks the migration of polymorphonuclear (PMNs)-type WBCs through intestinal epithelium cells (IECs). Although, higher spermidine levels associate via increased survival of Shigella within the infection of macrophages, the lack of cadaverine production enhances the pathogenic ability of this bacterium in human tissues. The high levels of spermidine are presented with the absence of the spermidine acetyltransferase (SAT) enzyme; the enzyme (coded via the speG gene) converting spermidine to its inert form, and N-acetylspermidine. The absence of expression in CadA and speG genes is the result of a parallel convergent evolutionary phenomenon that focuses on inactivating events ranging from point mutations to large deletions (black holes) in the bacterial whole genome (Fig 3).

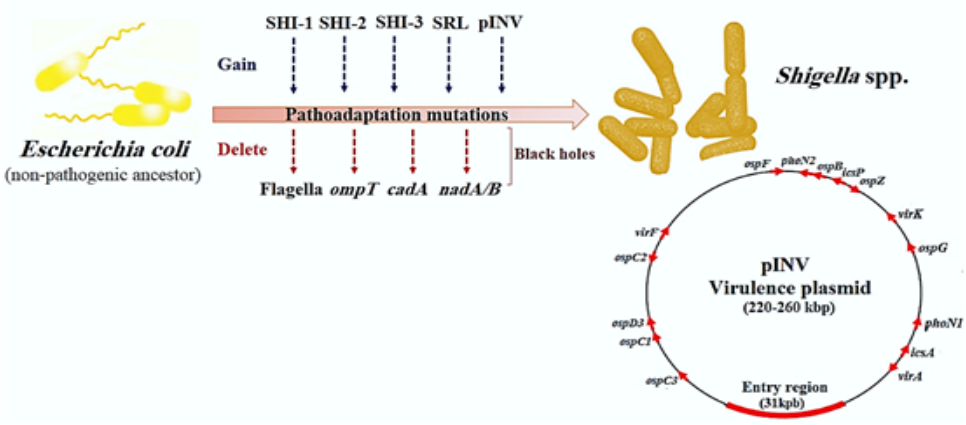

Figure 3. Major genetic events (genome fluidity) resulting in the convergent evolution of Shigella spp. (or even EIEC pathotypes) from the commensal harmless E.coli ancestor (39, 40)

The evidence suggests that the genetic plasticity of the Shigella virulence plasmid is interceded by intra- and inter-molecular events among insertion sequences. In addition, the proof that cadaverine can block enterotoxin activity may lead to rather general models for toxin activity or entry to host cells and offers a road for antitoxin therapy in the treatment of infectious diseases.

1.2.Rickettsiae species

Numerous obligate, intracellular pathogenic microorganisms have small genomes of <1.5 Mb (reductive genetic evolution). Rickettsiae (order: Rickettsiales, family: Rickettsiaceae) species are arthropod-borne obligate intracellular gram-negative bacteria that are pathogenic for several vertebrate hosts like human beings. They have diverged into various phylogenetic groups, including the typhus group (TG), spotted fever group (SFG), transitional group (TRG), and ancestral non-virulent group (AG) (Table 2).

Table 2. A brief data of Rickettsia spp. and O. tsutsugamushi

Group	Species (some examples)	Genome size (bp)	Diseases	Ref
TRG	R. felis Δ	1487240	SF rickettsioses	(47)
	R. akari	1231204	Rickettsial pox	(45)
SFG	R. conorii	1268755	Mediterranean SF	(48)
	R. rickettsii	1257710	Rocky mountain SF	(49)
AG	R. canadensis	1159772	----------	(50)
	R. bellii €	1522076		(51)
TG	R. prowazekii	1111523	Epidemic typhus	(52)
	R. typhi	1111496	Murin (Endemic) typhus	(53)
O. tsutsugamushi	----------	2127052*	Scrub typhus	(45)

Abbreviations: O. tsutsugamushi; Orientia tsutsugamushi. * Taxon-ID: 640427127. Δ R. felis has two plasmids. € The earliest diverging species among known Rickettsiae is R. bellii and can remain alive inside phagocytic amoebae niche.

The species causing severe rickettsioses show a relatively small chromosome size (0.8–2.3 Mbp) and without plasmids (with the exception of R. felis) (41, 42). The comparative genetic analysis results by Fournier et al. [2009] have demonstrated that virulence in Rickettsia was not related to obtaining foreign DNA, but it could be due to the genome decay process. Ogata et al. in their study [2001], by comparing R. prowazekii and R. conorii genomes to each other, found that the genomes of the most powerful pathogens did not become larger with the acquisition of genes related to the virulence, rather they became more decayed (43). This small size genome of Rickettsiae is due to the reductive evolution during endosymbiosis. In fact, compatibility with the intracellular niche or endo-symbiont viability style of the family Rickettsiae is based on the genome destruction ongoing process of reducing genes (Fig. 4).

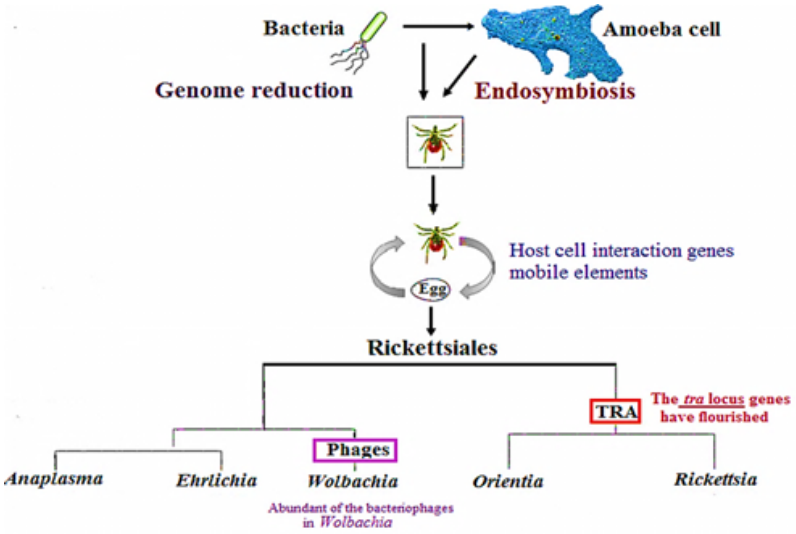

Figure 4. The genome evolution of Rickettsiales (Darby A et al, 2007)(54). Adaptation to the eukaryotic cells caused superfluous bacterial genes, generating an overall reduction in genome size (pseudogene-riddled genome)

They discard genes affected in metabolisms supplied by their eukaryotic host cells (pseudogene-riddled genome) (44). Rickettsia plasmids are considered a mirror of the evolutionary data of these bacteria (e.g., horizontal accession of genes, genome reductive mechanisms, and duplication processes) (http://img.jgi.doe.gov/cgi-bin/m/main.cgi online data base) (45).

The completely intact type IV secretion system (T4SS) structures that was demonstrated in all Rickettsia members confirms the evolutionary role of plasmids (46).

As numerous amoeba-associated bacteria or their relatives are real pathogens for human beings, it has been suggested that amoebae act as evolutionary "training grounds" and give bacteria the power to infect higher eukaryotes cells (51).

However, the genome size of R. prowazekii is highly decreased, but the genes encoding recombination and repair proteins are preserved within it, which are necessary for protection against the host immune response (55).

Its genome structure exhibits many genes with high similarity to orthologs in intracellular bacteria from ameobae. The process of gene loss has been a relatively prevalent strategy in the evolution of Rickettsiales genomes, which has been found to occur in Rickettsia endosymbiont Ixodes scapularis (REIS) (56). Moreover, an integrative conjugative mobile element called RAGE (Rickettsiales amplified genetic element) has been observed to proliferate in the chromosome, and even in plasmids in Rickettsiae. REIS encodes nine conserved RAGEs. These consist of %35 of the total genome, producing REIS as one of the major plastic and repetitive bacterial mobile elements. The process of gene loss in metabolic pathways in Rickettsia usually occurs in genes coding enzymes involved in sugar metabolism, lipid biosynthesis, nucleotide synthesis, and amino acid synthesis. Therefore, they depend on a permissive host cell for nutrition and building bio-blocks.

1.3.Mycobacterium leprae

Mycobacterium leprae is the etiology agent of leprosy (or Hansen's disease) in the majority of human cases. Another species, which is called "M. lepromatosis" is responsible for the minor of human leprosy cases (57). Chronic infection leprosy is basically a disease of peripheral nerve system (PNS) and skin; however, it affects bones, eyes, and the mucosa of the upper respiratory tract. These bacteria are obligate pathogens that are not cultivable in the outer of the host cell medium (cell-free media). Schwann cells in nerves and the macrophage cells in the skin (Langerhans

cells; LCs) are two types of cells that M. leprae has shown tropism for them Mycobacterium leprae is a slow-growing intracellular Mycobacterium, and the average incubation period of the disease is approximately 5-4 years. Nevertheless, M. lepromatosis has an affinity with endothelial cells and can cause vasculitis or necrotic erythema. It usually causes diffuse lepromatous leprosy (DLL).

According to the single nucleotide polymorphism (SNP) typing method, M. leprae is divided into four distinctive genotypes, which are associated with a diverse human population (58).

diverse human population (58). An extremely reduction in the genome of M. leprae has been seen (1537 of 2977 ancestral genes were lost) that almost %40 of the genes were inactivated to convert pseudogenes (59). Pseudogenes are molecular fossils as well as inactive reading frames via functional counterparts in the tubercle bacillus (60). The genome size of M. lepromatosis is approximately 3.22 Mb, which was %1.6 smaller than that of M. leprae genome (~3.27 Mb). This reductive evolution is distinctive among pathogenic bacteria (Table3).

Table 3. Comparison of genomic properties of obligate pathogens

Species	Mycobacterium leprae	Rickettsia prowazekii	Chlamydia trachomatis	Burkholderia mallei	Bordetella pertussis
Genome size (bp)	3268203	1111523	1044282	5780166**	4086189
GC content	57.79¬%	29.1%	41.3%	~ 68.6%	67.3%
ORFs	1614	834	~ 874–890	5535	3816
Encoding genome	49.5 %	75.4%	88.5%	68%	82%
Plasmid(s)	NR	NR	4793bp	NR	41268 bp€
Pseudogene(s)	27% (n=1133)	24%	n=14–18	n=5–305''	9.4% (n=359)

beta plasmid 1-Abbreviations: ORFs: Open reading frames, NR: not reported. ** It has two chromosomes € The IncP (71) (2012) .Hayden H & et al .pBP1

The evidence confirms the strict intracellular nature of this pathogen. Mycobacterium leprae has missed many genes for cholesterol catabolism,

siderophore production, a part of the oxidative pathway, majority of microaerophilic and anaerobic respiratory chains, and several catabolic routes (Fig. 5).

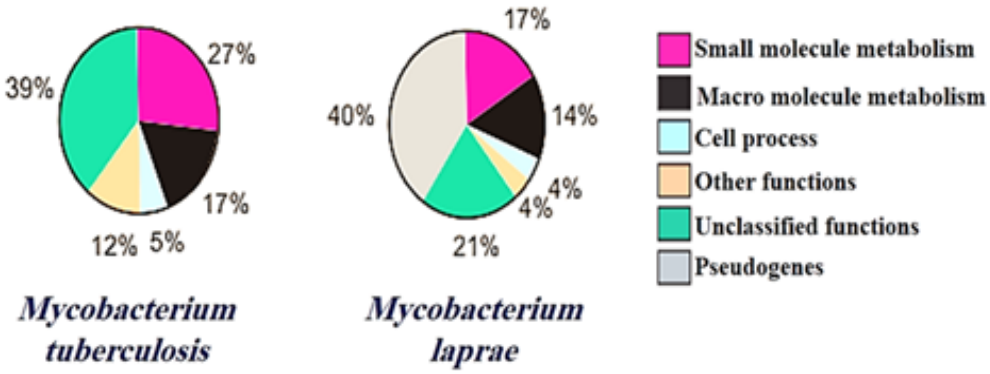

Figure 5. The percentage of the entire potential ORFs dedicated to main cellular functions in the leprosy bacillus and M.tb (data annotated via Cole et al., 1998 & 2001 from the databases of the M. leprae and M. tb genome projects) (62-66). The gene reduction and genome degradation can be seen in the M. leprae genome

For instance, the mce4 operon encoding for a sterol lipid transport net is established in other mycobacteria like M. tuberculosis. This bacterium uses host glucose supplies as the carbon store to biosynthesize most of its amino acids (61). Extensive recombination processes among dispersed repetitive sequences lead to genome downsizing and the current mosaic arrangement in M. leprae.

1.4.Burkholderia species

Burkholderia mallei is a facultative aerobic, bipolar, gram-negative rod-shaped bacterium belonging to the family of Burkholderiaceae. This species causes "glanders" disease in Equidae. In addition, B. mallei generally leads to chronic infections (sometimes fatal), and acute infection in humans and horses (23).

Burkholderiapseudomalleicauses“melioidosis” in humans and animals. Burkholderia mallei has two chromosomes. They include chromosome 3.5)1Mb) and chromosome 2.3) 2 Mb). In fact, B. mallei with a facultative intracellular lifestyle has a downsized genome for descending from a soil-dwelling pathogen (B. pseudomallei). This species evolved divergently from B. pseudomallei due to selective deletions and a reduction in the genome of B. pseudomallei. This extensive genomic reduction probably resulted in its inability to live outside the host and converted B. mallei to an obligate mammalian pathogen (intracellular pathogen). In most cases, B. mallei is considered an equine-adapted B. pseudomallei clone. Although, the core-genome of B. mallei species is smaller than that of B. pseudomallei, the variable gene sets are larger (67).

The genetic arrangement in multi-chromosome bacteria is not random. The gene distribution in these bacteria indicated that most of the genes essential for basic life activities, such as central metabolism and bacterial cell growth, were generally located in the primary (one) chromosome. The secondary (or other) chromosome consists of few essential genes and contains niche-specific genes and adaptation. B. mallei chromosome 1 contains genes coding the exopolysaccharide capsule, type IV pili, and type III & IV secretion systems. Chromosome 2 carries genes coding for LPS biosynthesis, capsule production and bacterial metabolism.

B. mallei is non-motile (non-flagellated) unlike B. pseudomallei, owing to an IS element (named IS407A) disrupting the fliP gene, the flagellar biosynthetic protein gene, and a frameshift mutation in one of the genes encoding bacterial flagellum motor (motB) (68 ,23).

In B. mallei, the L-arabinose assimilation operon (araBAD operon) is lost (22). Burkholderia mallei as a ‘young pathogen’ has a specific high rate and diverse mobile genetic elements (69). Perhaps, this issue can be described by fast evolution events that could enable modifications, such as selection pressures in novel habitats, limitations in the population history, and weakly selection against repetitive elements owing to the reduced effective population size. It is notable that the mutations in few pseudogenes presumably cause B. mallei to be non-motile and non-flagellated, while the other chemotaxis and motility genes are not yet decayed (70).

1.5.Bordetella species

Another example of bacterial evolution via black hole occurrence is Bordetella pertussis (a human-only pathogenic). This small, aerobic, fastidious, gram-negative encapsulated and coccobacillus shape is the causative agent of the highly contagious ‘pertussis’ or whooping cough disease (71). From a genetic taxonomic viewpoint, the “classical” bordetellae, which includes B. pertussis, B. parapertussis, and B. bronchiseptica, create an evolutionary monophyletic group with a limited genetic diversity (Fig. 6).

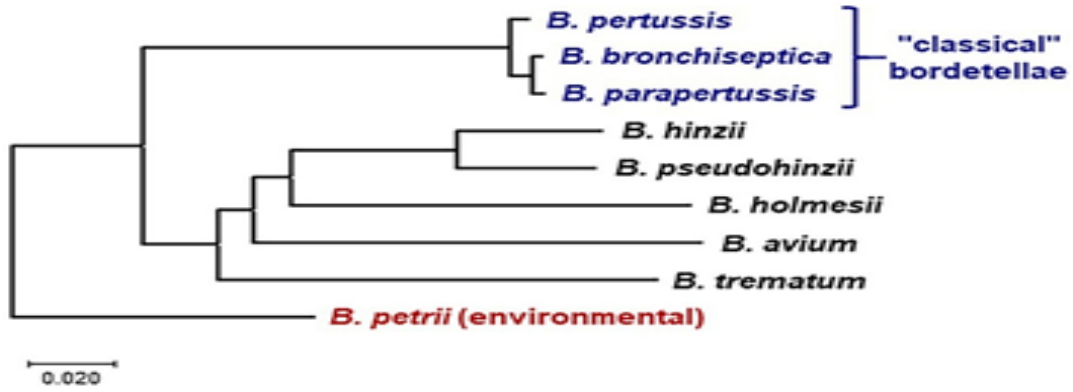

Figure 6. Whole genome phylogeny of Bordetella by the neighbor-joining tree (Kumar et al, 72) (2018) based on pairwise average nucleotide identities (ANIs) (Yoon et al., 73) (2017).

However, they have shown different host-adaptation strategies. Examples of the non-classical bordetellae are B. holmesii, B. avium, and B. trematum. Closely genetic relatedness between B. pertussis via eight other species of the Bordetella genus based on the results of DNA-DNA hybridization, MLEE, comparative 16S rDNA sequence analysis, microarrays, and comparative genome hybridization, methods was obtained. However, B. bronchiseptica has recently been considered the ancestor of B. parapertussis and B. pertussis. Loss of DNA segments (ISE-mediated deletion events) and intra-genomic recombination increase with recombination among IS element repeats, particularly the IS481 element with ≥240 copies per genome, play a crucial role

in B. pertussis evolution, and this species loses over 20% of its genome when diverging from B. bronchiseptica (74). Most of the lost or inactivated genes are involved in the membrane transport, small-molecule metabolism, and regulation of gene expression, and in synthesis of surface structures. The classical bordetellae share numerous significant virulence factors, such as pertussis toxin (PT), tracheal cytotoxin (TCT), pertactin (PRN), filamentous hemagglutinin (FHA), and adenylate cyclase-haemolysin (AC-Hly). The type-VI secretion system (T6SS) in B. bronchiseptica is a species-specific factor that is lost in B. pertussis and may not be functional in B. parapertussis due to absent subsets of genes or pseudogenes in this locus. In Bordetella, loss of genes was more notable than gene acquisition during evolution (75).

The significant sensitivity to hydrophobic molecules and fatty acids of B. pertussis is due to mutations in the acr genetic locus, which is the species-specific pathoadaptation. The AcrABC efflux system is responsible for palmitate secretion in Bordetella species. Unlike B. bronchiseptica that presents a high-activity efflux system, B. pertussis has a poorly functional AcrABC efflux system, and perhaps that is why this species is highly susceptible to the palmitate. The low activity AcrABC may give adequate protection from hydrophobic molecules inside the B. pertussis host niche, but inhibits excretion of important metabolites (24).

Furthermore, the LPS of B. pertussis never contain an O-antigen domain (a homopolymer of 2, 3-dideoxy-2, 3-di N-acetylgalactos aminuronic acid) unlike the LPS of B. bronchiseptica and B. parapertussis. According to genetic researches, the DNA fragment direct downstream of the wlb locus (named wbm locus) is needed for the O-antigen biosynthesis in both B. bronchiseptica and B. parapertussis. However, this genetic region is substituted via an IS in B. pertussis and results in the deletion of wbm(Δwbm) (25).

Chlamydia trachomatis

Chlamydia trachomatis is an obligate intracellular gram-negative bacterium belonging to the Chlamydiaceae family. It causes trachoma, an infectious blinding eye disease, and sexually transmitted infections, including Lymphogranuloma venereum (LGV) (77). The name Chlamydozoa, meaning "cloak and animal," was basically suggested to display how the infectious form of the microorganism, now known as an elementary body (EB), appeared cloaked inside an intracellular structure known as an inclusion (78).

This bacterium has a biphasic intracellular lifecycle including an infectious extracellular elementary body (EB) and an intracellular replicative reticulate body (RB). Specifically, C. trachomatis can be divided into two biovars: (i) trachoma biovars that are limited to the epithelial cells from the eye and or genital tract, and (ii) LGV biovars that cause invasive infections like STIs. Since C.

trachomatis has an obligate intracellular lifestyle, it needs to uptake nutrients or anabolic precursors from the host cell. This is usually shown in a reduction in the metabolic capability encoded by the obligate parasite, and a concurrent decrease in the genome size. Chlamydia trachomatis has a highly conserved small genome of nearly 1.04 Mb and harbors a plasmid of approximately 7 kb (79).

The presence of small-sized plasmids in Chlamydia is necessary for the shuffle genetic capability between strains and species. A low number of pseudogenes exist in the genomes of all chlamydial species. The earliest stages of Chlamydia evolution include an immense genome reduction upon converting into an obligate intracellular parasite. The reductive genome adaption in Chlamydia includes gene sets associated with amino acid synthesis, nucleotide assembly, and other activities involved in free-living growth. For example, according to the several genetic bioinformatics analysis, a conventional fructose-1,6-bisphosphatase (FBPase/FDPase) appears to be missing in C. trachomatis (because a gluconeogenic route is not functional in intracellular C. trachomatis) (26).

Moreover, the further remarkability of chlamydiae is considered the energy parasitism for host cells due to the exchange of ADP for the host-derived ATP. Recently, a novel vision for the genomic diversity of deeply branching chlamydial lineages has been presented based on the data of metagenome-assembled genomes (MAGs). As the chlamydial core genome reduces, the accessory genome expands strongly by adding more genomes from deep-branching lineages (80).

In addition, the genome analysis studies into species of the Chlamydia genus offers genetic exchange as the origin of the add-guaAB gene cluster and copN that are settled in the plasticity zone (PZ) as an area of heterogeneity among species (81).

2.Conclusion

Multiple pathogenic bacteria with a small genome size have evolved from bacteria via larger genomes through genetic reductive processes. Bacteria, such as M. leprae, C. trachomatis, R. prowazekii, Burkholderia, Bordetella, and Shigella, under the selection force and host-adaptation, led to gene loss and reductive genome evolution. A factor that can influence and constrain this type of evolution is the controversy of the function of the target gene with bacterial virulence as well as the degree of bacterial

compatibility within the host (anti-virulence and anti-adaptability genes). The massive genetic deletions occurring in these cases are called "black holes". This evolutionary route will enable the pathogen to intensify the virulence and its pathoadaptibility.

Conflict of interest

The authors declare no conflicts of interest.

References

1. Cesbron S, Briand M, Essakhi S, et al. Comparative genomics of pathogenic and nonpathogenic strains of Xanthomonas arboricola unveil molecular and evolutionary events linked to pathoadaptation. Frontiers in plant science. 2015;6:1126.
2. Davies J, Davies D. Origins and evolution of antibiotic resistance. Microbiology and molecular biology reviews. 2010;74:417-433.
3. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. nature. 2000;405:299-304.
4. Srividhya K, Alaguraj V, Poornima G, et al. Identification of prophages in bacterial genomes by dinucleotide relative abundance difference. PLoS One. 2007;2:e1193.
5. Hall JP, Brockhurst MA, Harrison E. Sampling the mobile gene pool: innovation via horizontal gene transfer in bacteria. Philosophical Transactions of the Royal Society B: Biological Sciences. 2017;372:20160424.
6. Kim Y, Gu C, Kim HU, et al. Current status of pan-genome analysis for pathogenic bacteria. Current opinion in biotechnology. 2020;63:54-62.
7. Tettelin H, Medini D. The pangenome: Diversity, dynamics and evolution of genomes. 2020.
8. Gonzales-Siles L, Karlsson R, Schmidt P, et al. A pangenome approach for discerning species-unique gene markers for identifications of Streptococcus pneumoniae and Streptococcus pseudopneumoniae. Frontiers in cellular and infection microbiology. 2020;10:222.
9. Costa SS, Guimarães LC, Silva A, et al. First Steps in the Analysis of Prokaryotic Pan-Genomes. Bioinformatics and Biology Insights. 2020;14:1177932220938064.
10. Lassalle F, Didelot X. Bacterial Microevolution and the Pangenome. The Pangenome. 2020:129.
11. Didelot X, Walker AS, Peto TE, et al. Within-host evolution of bacterial pathogens. Nature Reviews Microbiology. 2016;14:150-162.
12. Maurelli AT. Black holes, antivirulence genes, and gene inactivation in the evolution of bacterial pathogens. FEMS microbiology letters. 2007;267:1-8.
13. Fukiya S, Mizoguchi H, Tobe T, et al. Extensive genomic diversity in pathogenic Escherichia coli and Shigella strains revealed by comparative genomic hybridization microarray. Journal of bacteriology. 2004;186:3911-3921.
14. Yin X, Wheatcroft R, Chambers JR, et al. Contributions of O island 48 to adherence of enterohemorrhagic Escherichia coli O157: H7 to epithelial cells in vitro and in ligated pig ileal loops. Applied and environmental microbiology. 2009;75:5779-5786.
15. Mobley HL, Donnenberg MS, Hagan EC. Uropathogenic Escherichia coli. EcoSal Plus. 2009;3.
16. Kusumoto M, Nishiya Y, Kawamura Y, et al. Identification of an insertion sequence, IS1203 variant, in a Shiga toxin 2 gene of Escherichia coli 0157: H7. Journal of bioscience and bioengineering. 1999;87:93-96.
17. Yamanaka K, Fang L, Inouye M. The CspA family in Escherichia coli: multiple gene duplication for stress adaptation. Mol Microbiol. 1998;27:247-255.
18. Day WA, Jr., Fernández RE, Maurelli AT. Pathoadaptive mutations that enhance virulence: genetic organization of the cadA regions of Shigella spp. Infection and immunity. 2001;69:7471-7480.
19. Nakata N, Tobe T, Fukuda I, et al. The absence of a surface protease, OmpT, determines the intercellular spreading ability of Shigella: the relationship between the ompT and kcpA loci. Mol Microbiol. 1993;9:459-468.
20. Diop A, Raoult D, Fournier P-E. Paradoxical evolution of rickettsial genomes. Ticks and tick-borne diseases. 2019;10:462-469.
21. Borah K, Girardi KdCdV, Mendum TA, et al. Intracellular Mycobacterium leprae utilizes host glucose as a carbon source in Schwann cells. Mbio. 2019;10:e02351-19.
22. Holden MT, Titball RW, Peacock SJ, et al. Genomic plasticity of the causative agent of melioidosis, Burkholderia pseudomallei. Proceedings of the National Academy of Sciences. 2004;101:14240-14245.
23. Lu R, Popov V, Patel J, et al. Burkholderia mallei and Burkholderia pseudomallei stimulate differential inflammatory responses from human alveolar type II cells (ATII) and macrophages. Frontiers in cellular and infection microbiology. 2012;2:165.
24. MacArthur I, Belcher T, King JD, et al. The evolution of Bordetella pertussis has selected for mutations of acr that lead to sensitivity to hydrophobic molecules and fatty acids. Emerging microbes & infections. 2019;8:603-612.
25. Preston A, Maskell D. Invited review: The molecular genetics and role in infection of lipopolysaccharide biosynthesis in the Bordetellae. Journal of endotoxin research. 2001;7:251-261.
26. Mehlitz A, Eylert E, Huber C, et al. Metabolic adaptation of Chlamydia trachomatis to mammalian host cells. Molecular microbiology. 2017;103:1004-1019.
27. Sokurenko EV, Chesnokova V, Dykhuizen DE, et al. Pathogenic adaptation of Escherichia coli by natural variation of the FimH adhesin. Proceedings of the National Academy of Sciences of the United States of America. 1998;95:8922-8926.
28. Servin AL. Pathogenesis of Afa/Dr diffusely adhering Escherichia coli. Clinical microbiology reviews. 2005;18:264-292.
29. 29. Chattopadhyay S, Tchesnokova V, McVeigh A, Kisiela DI, Dori K, Navarro A, et al. Adaptive evolution of class 5 fimbrial genes in enterotoxigenic Escherichia coli and its functional consequences. The Journal of biological chemistry. 2012;287:6150-6158.
30. Dorman CJ, Porter ME. The Shigella virulence gene regulatory cascade: a paradigm of bacterial gene control mechanisms. Molecular microbiology. 1998;29:677-684.
31. Le Gall T, Mavris M, Martino MC, et al. Analysis of virulence plasmid gene expression defines three classes of effectors in the type III secretion system of Shigella flexneri. Microbiology. 2005;151:951-962.
32. Prosseda G, Di Martino ML, Campilongo R, et al. Shedding of genes that interfere with the pathogenic lifestyle: the Shigella model. Research in microbiology. 2012;163:399-406.
33. Yang F, Yang J, Zhang X, et al. Genome dynamics and diversity of Shigella species, the etiologic agents of bacillary dysentery. Nucleic acids research. 2005;33:6445-6458.
34. Hale T, Oaks E, Formal S. Identification and antigenic characterization of virulence-associated, plasmid-coded proteins of Shigella spp. and enteroinvasive Escherichia coli. Infection and immunity. 1985;50:620-629.
35. Oaks EV, Hale T, Formal S. Serum immune response to Shigella protein antigens in rhesus monkeys and humans infected with Shigella spp. Infection and immunity. 1986;53:57-63.
36. Porter M, Dorman C. Differential regulation of the plasmid-encoded genes in the Shigella flexneri virulence regulon. Molecular and General Genetics MGG. 1997;256:93-103.
37. Pupo GM, Karaolis D, Lan R, et al. Evolutionary relationships among pathogenic and nonpathogenic Escherichia coli strains inferred from multilocus enzyme electrophoresis and mdh sequence studies. Infection and immunity. 1997;65:2685-2692.
38. Park SH, Soetyono F, Kim HK. Cadaverine production by using cross-linked enzyme aggregate of Escherichia coli lysine decarboxylase. Journal of microbiology and biotechnology. 2017;27:289-296.
39. Venkatesan MM, Goldberg MB, Rose DJ, et al. Complete DNA sequence and analysis of the large virulence plasmid of Shigella flexneri. Infection and immunity. 2001;69:3271-3285.
40. Pasqua M, Michelacci V, Di Martino ML, et al. The intriguing evolutionary journey of enteroinvasive E. coli (EIEC) toward

pathogenicity. Frontiers in microbiology. 2017;8:2390.
41. Fournier P-E, El Karkouri K, Leroy Q, et al. Analysis of the Rickettsia africae genome reveals that virulence acquisition in Rickettsia species may be explained by genome reduction. BMC genomics. 2009;10:1-15.
42. Gillespie JJ, Beier MS, Rahman MS, et al. Plasmids and rickettsial evolution: insight from Rickettsia felis. PLoS One. 2007;2:e266.
43. Ogata H, Audic S, Renesto-Audiffren P, et al. Mechanisms of evolution in Rickettsia conorii and R. prowazekii. Science. 2001;293:2093-2098.
44. El Karkouri K, Kowalczewska M, Armstrong N, et al. Multi-omics analysis sheds light on the evolution and the intracellular lifestyle strategies of spotted fever group Rickettsia spp. Frontiers in microbiology. 2017;8:1363.
45. Sachman-Ruiz B, Quiroz-Castañeda RE. Genomics of Rickettsiaceae: an update. Farm Anim Dis Recent Omic Trends New Strategies Treat. 2018;83.
46. Cascales E, Christie PJ. The versatile bacterial type IV secretion systems. Nature Reviews Microbiology. 2003;1:137-149.
47. Ogata H, Renesto P, Audic S, et al. The genome sequence of Rickettsia felis identifies the first putative conjugative plasmid in an obligate intracellular parasite. PLoS biology. 2005;3:e248.
48. Eremeeva ME, Madan A, Shaw C, et al. New Perspectives on Rickettsial Evolution from New Genome Sequences of Rickettsia, particularly R. canadensis, and Orientia tsutsugamishi. ANNALS-NEW YORK ACADEMY OF SCIENCES. 2005;1063:47.
49. Clark TR, Noriea NF, Bublitz DC, et al. Comparative genome sequencing of Rickettsia rickettsii strains that differ in virulence. Infection and immunity. 2015;83:1568-1576.
50. Myers WF, Wisseman Jr CL. The taxonomic relationship of Rickettsia canada to the typhus and spotted fever groups of the genus Rickettsia. Rickettsiae and rickettsial diseases Academic Press, Inc, New York, NY. 1981:313-325.
51. Ogata H, La Scola B, Audic S, et al. Genome sequence of Rickettsia bellii illuminates the role of amoebae in gene exchanges between intracellular pathogens. PLoS genetics. 2006;2:e76.
52. Andersson JO, Andersson S. Genome degradation is an ongoing process in Rickettsia. Molecular biology and evolution. 1999;16:1178-1191.
53. McLeod MP, Qin X, Karpathy SE, et al. Complete genome sequence of Rickettsia typhi and comparison with sequences of other rickettsiae. Journal of bacteriology. 2004;186:5842-5855.
54. Darby AC, Cho N-H, Fuxelius H-H, et al. Intracellular pathogens go extreme: genome evolution in the Rickettsiales. TRENDS in Genetics. 2007;23:511-520.
55. 55. Moran NA. Microbial minimalism: genome reduction in bacterial pathogens. Cell. 2002;108(5):583-6.
56. Georgiades K, Merhej V, El Karkouri K, et al. Gene gain and loss events in Rickettsia and Orientia species. Biology direct. 2011;6:1-14.
57. Donoghue HD, Taylor GM, Mendum TA, et al. The distribution and origins of ancient leprosy. Hansen's Disease The Forgotten and Neglected Disease. 2018.
58. Chavarro-Portillo B, Soto CY, Guerrero MI. Mycobacterium leprae's evolution and environmental adaptation. Acta tropica. 2019;197:105041.
59. ORÉ NAEH, TEKAIA FR, TI AT, WHEELER PR, CHURCHER C, DA V ID HS, et al. The decaying genome of Mycobacterium leprae. Lepr Rev. 2001;72:387-398.
60. Han XY, Silva FJ. On the age of leprosy. PLoS neglected tropical diseases. 2014;8:e2544.
61. Marques MAM, Berrêdo-Pinho M, Rosa TL, et al. The essential role of cholesterol metabolism in the intracellular survival of Mycobacterium leprae is not coupled to central carbon metabolism and energy production. Journal of bacteriology. 2015;197:3698-3707.
62. Cole S, Eiglmeier K, Parkhill J, et al. Massive gene decay in the leprosy bacillus. Nature. 2001;409:1007-1011.
63. Wixon J. Featured organism: reductive evolution in bacteria: Buchnera sp., Rickettsia prowazekii and Mycobacterium leprae. Comparative and functional genomics. 2001;2:44-48.
64. Honore N, Bergh S, Chanteau S, et al. Nucleotide sequence of the first cosmid from the Mycobacterium leprae genome project: structure and function of the Rif-Str regions. Molecular microbiology. 1993;7:207-214.
65. Cole S. Why sequence the genome of Mycobacterium tuberculosis? Tubercle and Lung Disease. 1996;77:486-490.
66. Vissa VD, Brennan PJ. The genome of Mycobacterium leprae: a minimal mycobacterial gene set. Genome biology. 2001;2:1-8.
67. Losada L, Ronning CM, DeShazer D, et al. Continuing evolution of Burkholderia mallei through genome reduction and large-scale rearrangements. Genome biology and evolution. 2010;2:102-116.
68. Song H, Hwang J, Yi H, et al. The early stage of bacterial genome-reductive evolution in the host. PLoS pathogens. 2010;6:e1000922.
69. Mira A, Pushker R, Rodriguez-Valera F. The Neolithic revolution of bacterial genomes. Trends in microbiology. 2006;14:200-206.
70. Nierman WC, DeShazer D, Kim HS, et al. Structural flexibility in the Burkholderia mallei genome. Proceedings of the National Academy of Sciences. 2004;101:14246-14251.
71. Hayden HS, Lim R, Brittnacher MJ, et al. Evolution of Burkholderia pseudomallei in recurrent melioidosis. PLoS One. 2012;7:e36507.
72. Kumar S, Stecher G, Li M, et al. MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology and evolution. 2018;35:1547.
73. Yoon S-H, Ha S-M, Kwon S, et al. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. International journal of systematic and evolutionary microbiology. 2017;67:1613.
74. Parkhill J, Sebaihia M, Preston A, et al. Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella parapertussis and Bordetella bronchiseptica. Nature genetics. 2003;35:32-40.
75. Park J, Zhang Y, Buboltz AM, et al. Comparative genomics of the classical Bordetella subspecies: the evolution and exchange of virulence-associated diversity amongst closely related pathogens. BMC genomics. 2012;13:1-17.
76. Linz B, Ivanov YV, Preston A, et al. Acquisition and loss of virulence-associated factors during genome evolution and speciation in three clades of Bordetella species. BMC genomics. 2016;17:1-17.
77. Collingro A, Tischler P, Weinmaier T, Penz T, Heinz E, Brunham RC, et al. Unity in variety—the pan-genome of the Chlamydiae. Molecular biology and evolution. 2011;28(12):3253-70.
78. 78. Clarke IN. Evolution of Chlamydia trachomatis. Annals of the New York Academy of Sciences. 2011;1230:E11-E18.
79. Seth-Smith HM, Harris SR, Persson K, et al. Co-evolution of genomes and plasmids within Chlamydia trachomatis and the emergence in Sweden of a new variant strain. Bmc Genomics. 2009;10:1-10.
80. Collingro A, Köstlbacher S, Horn M. Chlamydiae in the Environment. Trends in Microbiology. 2020;28:877-888.
81. Gomes JP, Bruno WJ, Nunes A, et al. Evolution of Chlamydia trachomatis diversity occurs by widespread interstrain recombination involving hotspots. Genome research. 2007;17:50-60.