Molecular mechanisms and drivers of pathogen emergence

Deepak Balasubramanian[1,2], Mario López-Pérez[1,2,3], Trudy-Ann Grant[1,2], C. Brandon Ogbunugafor[4], Salvador Almagro-Moreno[1,2*]

[1]Burnett School of Biomedical Sciences, University of Central Florida, Orlando FL 32816 USA

[2]National Center for Integrated Coastal Research, University of Central Florida, Orlando FL 32816 USA

[3]Evolutionary Genomics Group, División de Microbiología, Universidad Miguel Hernández, Apartado 18, San Juan 03550, Alicante, Spain

[4]Department of Ecology and Evolutionary Biology, Yale University, New Haven CT USA

[*]For correspondence: samoreno@ucf.edu (S. Almagro-Moreno)

Keywords: Pathogen emergence, horizontal gene transfer, pathoadaptive mutations, genomic background, virulence-adaptive polymorphisms.

Abstract

Pathogen emergence (PE) is a complex phenomenon with major public health implications. Over the past decades, numerous underlying mechanisms facilitating the emergence of pathogenic bacteria have been elucidated. In this review, we highlight the diverse molecular and environmental drivers associated with PE, with an emphasis on the interplay of canonical gene transfer mechanisms and the increasingly appreciated role of genetic variations, providing a more coherent picture of this process. Given the interactive and multifactorial nature of PE, we contend that the development of approaches that embrace the integration of these factors is indispensable in order to truly comprehend this complex phenomenon and develop strategies to mitigate this threat.

Pathogen emergence

Emerging and reemerging infectious diseases pose immense health and economic risks worldwide [1]. **Pathogen emergence** (PE) (see Glossary) is a multifactorial, complex phenomenon that is influenced by various genetic adaptations, ecological drivers and selective pressures imposed by the human host [2]. Understanding the specific elements associated with the process of PE has major public health implications and is critical for successful disease management and control [3]. Numerous extensively reviewed mechanisms associated with the emergence of **pathogens** have been elucidated over the past decades [4-8]. This knowledge has provided an initial picture of some of the factors influencing PE and the roles of these individual players in facilitating this non-linear phenomenon. Nonetheless, recent findings highlighting the interplay of these elements depict a more complex scenario in which these multifactorial interactions are an essential component towards the manifestation of PE. For decades, our knowledge on the molecular drivers of PE was mostly associated to the role of individual **mobile genetic elements** (MGEs) acquired by **horizontal gene transfer (HGT)** and the specific pathogenic traits they encoded. Recent findings contextualize horizontally acquired elements within the overall architecture of the genome and cell physiology, providing a more nuanced and interactive view of these drivers in the process of PE. For instance, plasmids in *Acinetobacter baumannii* encode transcriptional regulators that modulate the expression of virulence factors encoded on the ancestral chromosome, mediating host tissue tropism [9]. Furthermore, the development of novel bioinformatic frameworks and expanded computational resources are permitting advanced genomic analyses of bacterial populations and how they provide clues about the PE process.

These have led to discoveries that underscore the critical relevance of the genomic background of the bacteria acquiring the MGEs and the role of diverse adaptive mutations associated with PE, what we term **genetic variations associated with pathogen emergence (GVPEs)**. For example, **pathoadaptive mutations** occurring within the host play a major role in the acquisition of **virulence** potential in commensal bacteria such as *Escherichia* coli, leading to resistance to host antimicrobial peptides and complement among other virulence traits [10]. Besides, clinical strains of pathogens such as *Vibrio cholerae* possess a unique genomic background containing **virulence adaptive polymorphisms** (VAPs) that favor the emergence of pathogenic traits [11]. Furthermore, the preeminent role of environmental drivers in the emergence of bacterial pathogens is beginning to be elucidated and the relative influence of its individual components unraveled [12, 13]. For instance, there appears to be specific selective pressures that ecosystems generate to enable the emergence of pathogenic traits within environmental populations of facultative pathogens like *Vibrio vulnificus* [12].

In this review, we summarize the diverse molecular mechanisms and drivers that play a key role in PE placing an emphasis on the recent findings that reveal their interplay. Specifically, we describe the modes of **horizontal gene transfer** (HGT) that mediate the acquisition and expression of genes associated with virulence traits in pathogenic bacteria and how novel discoveries suggest a complex network of interactions between them. We highlight the recently appreciated role of GVPEs such as genomic **preadaptations** stemming from the environment or pathoadaptive mutations selected within the host in the fitness and survival of bacterial pathogens, and how abiotic and biotic drivers affect the selection of virulence traits. Finally, given the clearly multifactorial nature of PE, we contend that approaches that integrate the synergies and interactions between these mechanisms are critical to unravel this complex phenomenon, will help identify potential sources of emergent pathogens, and aid in the development of reliable therapeutics against these threats.

Horizontal gene transfer and pathogen emergence

The acquisition of virulence traits and colonization factors via HGT is a quintessential and well-established component of the process of PE **(Fig. 1a)**. In this section, we will discuss the association of these major molecular drivers of HGT with PE highlighting newly discovered roles of these mobile elements. Importantly, we will emphasize new developments that uncover the interplay between these widely studied MGEs and the recipient cell, and are helping define a more complex and accurate picture of PE.

Plasmids have long been associated with the acquisition of various critical **virulence determinants** and pathogenic traits such as a) toxins and secretion systems, b) colonization factors, c) immune evasion, and d) antibiotic resistance ([4], **Fig. 1a**). Plasmid acquisition plays a primary role in PE and has led to speciation events in several pathogens. For instance, the emergence and speciation of the three pathogenic *Yersinia* species from environmental isolates stems from plasmid acquisition events [14, 15]. Acquisition of virulence plasmid pYV by environmental *Yersinia* results in a virulent ancestral strain that can overcome the host defense mechanisms (type III secretion system, T3SS) and colonize host lymph tissues (YadA adhesin) [14]. Furthermore, the speciation of the systemic pathogen *Yersinia pestis* from the enteric *Yersinia* *pseudotuberculosis* was due to the acquisition of two key plasmids, pPla and pMT1, that play a critical role in host adaptation and virulence [16, 17]. Central to their role as major drivers of PE, plasmids encode the main toxins of several human pathogens such as the heat-stable and -labile enterotoxins of *Escherichia coli* ETEC, or the beta and epsilon toxins of *Clostridium perfringens* [18, 19].

Besides antibiotic resistance genes (**Box 1**), plasmids confer key niche adaptations to pathogens such as increased tissue tropism or expansion of their metabolic profile (e.g., usage of wider range of carbon sources). Pan-plasmidome studies on the nosocomial pathogen *Enterococcus faecium* reveal that clinical isolates have a much larger number of plasmids and of greater size than isolates from animal or environmental sources [20]. Interestingly, the core plasmidome of *E. faecium* clinical isolates, besides containing genes associated with resistance to antibiotics, encodes a varied array of genes involved in carbohydrate utilization (e.g., mannose, fructose or sorbose) that widens the niche specificity of the bacterium [20]. Recent studies underline the wider interactions of plasmids with the bacterial genome. Plasmid pAB5 of *A. baumannii* broadens tissue tropism by favoring colonization of the urinary tract over the lung, however, it does so by an unusual reverse regulation mechanism [9] (**Fig 1B**). Typically, transcriptional regulators on acquired plasmids do not control expression of chromosomal genes. Strikingly, pAB5 modulates bacterial chromosomal genes via transcriptional regulators encoded within the plasmid, including genes associated with adherence to urothelial cells [9] (**Fig 1B**). The modulation of chromosomal genes by regulators on plasmids is also observed in other bacteria such as *Chlamydia trachomatis*, suggesting that the phenomenon of cross-regulation is more widely prevalent in pathogens [21, 22]. Thus, despite being some of the best studied agents of HGT, recent discoveries show unexpected roles of plasmids in virulence regulation and host colonization.

Integrative and conjugative elements (ICEs). Also known as conjugative transposons, ICEs typically encode genes that provide resistance to antibiotics, help outcompete other bacteria during colonization and enhance survival within the host [23-25]. For example, analysis of over 180,000 bacterial genomes identified several ICEs in *E. coli* and *Pseudomonas aeruginosa* that encode carbapenemases (antibiotic resistance), siderophores (iron acquisition) and bacteriocins (kill competing bacteria) [23]. In other cases, ICEs directly enhance survival within the host environment. ß-lactam antibiotics induce oxidative stress in the bacterial cells, in addition to their effect on cell wall biosynthesis [26]. Resistance to oxidative stress provided by the ICE-ßox of *Legionella pneumophilia* leads to increased survival in the presence of ß-lactam antibiotics and enhances replication within host macrophages [24].

Recent research shows that ICEs can also cooperate for the dissemination of plasmids within bacterial populations, suggesting a wider impact of these MGEs in PE [25] **(Fig 1C)**. Strains of *Klebsiella pneumoniae* that cause invasive disease harbor an ICE element, ICEKp1, encoding a complete type IV secretion system (T4SS) that can mobilize plasmids. A second ICE, ICEKp2, co-occurs with ICEKp1 in some clinical isolates and encodes a Mob2 ATPase [25]. Surprisingly, ICEKp2 enhances the efficiency of plasmid mobilization driven by ICEKp1 as Mob2 ATPase complements the activity of T4SS from ICEKp1, thereby enabling efficient plasmid transfer to neighboring cells [25] **(Fig 1C)**. Novel discoveries regarding ICE elements, such as aiding the dissemination of other MGEs, highlight how the interplay between molecular players foster the emergence of pathogenic traits.

Insertion sequences (IS) and transposons. Unlike ICEs and plasmids that physically transfer virulence-associated genes, the impact of **IS elements** on the emergence of pathogenic traits is primarily via gene inactivation, translocation and inversion that reshape the genome leading to modified gene expression [27]. This inherently limits the identification of the specific mechanisms that these elements have on PE [27]. Nonetheless, even though their exact function in numerous pathogens is not clear, the accumulation of these elements in the bacteria is often associated with PE. Accumulation of IS elements impacts numerous key traits such as virulence regulation, antibiotic resistance, or metabolic adaptations that increase bacterial fitness in the host [27]. The transition of *E. faecium* from a human commensal to a nosocomial pathogen is driven by accumulation of IS elements on its chromosome [28]. Genomic comparisons of pathogenic clades of *E. faecium* against non-pathogenic ones show that over 55% of the acquired elements by the nosocomial clades are IS elements [28, 29]. This widespread distribution leads to genetic rearrangements and facilitate genome plasticity, which ultimately contributes to host adaptation [29]. Inpatient evolution of clinical isolates of *P. aeruginosa* reveals that successful colonization and dissemination within the human host is largely shaped by IS elements such as the 2.9 kb CLJ-ISL3 [30]. This element modulates several key pathogenic traits including host cell adhesion, production of hydrogen cyanide and antimicrobial resistance [30]. Interestingly, host-adapted isolates of *P. aeruginosa* that display enhanced cytotoxicity and antibiotic resistance profiles harbor about 7 times more copies of CLJ-ISL3, in addition to 36 other IS elements [30, 31]. To date, it is not entirely clear how the accumulation of IS elements favors host adaptation and other virulence traits, or whether there are secondary mechanisms involved in addition to gene inversion and translocation events. Whereas IS elements impact virulence adaptive traits, **transposons** appear to primarily transfer genes encoding antibiotic resistance [32]. The accessory roles of antibiotic resistance in PE are explored in **Box 1**. The extent of the interactions between IS elements, transposons and other MGEs in the emergence of pathogenic bacteria remains poorly understood. Overall, by modulating key traits associated with virulence and antimicrobial resistance due to their accumulation in the genome, IS elements and transposons represent a prevailing and highly flexible means that is strongly associated with the emergence of human pathogens.

Pathogenicity islands (PAIs) are large mobile genomic islands confined to pathogenic isolates that typically encode genes associated with host colonization **(Fig. 1)** [33-35]. The pandemic El Tor strains of *V. cholerae* harbor four major PAIs [36]. One of them, the *Vibrio* pathogenicity island 1 (VPI-1), encodes the toxin-coregulated pilus, an essential intestinal colonization factor that mediates microcolony formation in the intestine [37]. A second PAI, VPI-2, encodes two major virulence traits: a neuraminidase, which unmasks the receptors of the cholera toxin, and sialic acid utilization genes, which confer a competitive advantage in the intestine [38, 39]. Interestingly, there appears to be cross-talk between the two PAIs, as the excision of both VPIs for transfer to other cells, is concerted by one recombination directionality factor encoded within VPI-2 [40]. Some PAIs also contribute to the survival and dissemination of pathogens within the human host such as the *Salmonella* pathogenicity islands (SPIs) [41, 42]. Serovars of *Salmonella* *enterica* encode two T3SSs within SPI-1 and SPI-2. The SPI-1-encoded T3SS, mediates host cell adhesion of *Salmonella* and is essential for GI tract infections, whereas the SPI-2-encoded T3SS plays a role both in intestinal and systemic infections and is essential for intracellular growth and maintenance of small colony variants [41-43]. Initially, the T3SS of SPI-2 only encoded two effector proteins (SseF and SseG) but subsequent independent HGT events mediated by prophages and plasmids at other genomic loci, have widened its repertoire of effectors to as many as 28 proteins in some serovars [42]. Therefore, PAIs appear to function as modular toolboxes playing a role in PE by encoding critical traits associated with host colonization.

Bacteriophages are a major vehicle for the acquisition of bacterial toxins in numerous pathogens. Bacteriophages Sp5 and Sp15, which encode the Shiga toxin, mediated the emergence of new virulent and epidemic clones of *E. coli* 0157:H7 [44]. Cholera toxin, the source of the diarrhea associated with cholera, was acquired by *V. cholerae* via lysogenic conversion of the filamentous CTXϕ phage [45]. Several *Staphylococcus aureus* toxins (e.g., exfoliative toxin A, enterotoxin A) and immune modulators were also acquired by phage lysogenization [46, 47]. Prophages can also enhance host colonization and fitness of numerous pathogens. In chronic wound infections caused by *P. aeruginosa*, acquisition of prophages from closely-related strains results in increased adaptation to the host [48]. The disruption of genes as a result of prophage integration ultimately leads to higher biofilm formation by *P. aeruginosa* and enhanced colonization [48]. Similarly, the *E. faecalis* strain V583 encodes a composite prophage, øV1/7, that infects and lyses related enterococcal strains and provides a competitive advantage during colonization of the mammalian intestine [49]. Interestingly, phages can also mediate transfer of PAIs within and between species and genera. Phages ø12 and øSLT can transfer the *S. aureus* PAI SaPObov5 from *S. aureus* to *Listeria monocytogenes* and non-*aureus* strains, conferring the ability to coagulate human plasma and aiding the infection process [50]. Therefore, besides influencing various aspects of virulence and host colonization, phages also play an important role in the transfer of other MGEs, further increasing their relevance as major agents shaping the emergence and evolution of bacterial pathogens.

Integrons are genetic elements encoded within the bacterial chromosome that mediate the efficient capture and expression of exogenously acquired genes [7]. Although integrons primarily capture genes associated with antibiotic resistance and toxin-antitoxin genes, some of their cargo can also drive host colonization and virulence [51-53]. The production of capsular polysaccharide (CPS) is critical for *V. vulnificus* pathogenicity. One of the four CPS loci encoded in the *V. vulnificus* genome are located within its integron. Inactivation of this locus render clinical strains uncapsulated and avirulent [51]. Furthermore, integrons also form an important component in shaping *V. cholerae* pathogenicity as all clinical strains harbor a superintegron (a large integron island) that encodes genes involved in secondary metabolism and cell surface modification [52, 54]. Specifically, they encode a) CPS biosynthetic proteins that have been directly implicated in virulence, b) plasmid *Achromobacter* secretion factors that facilitates toxin secretion, and c) lipocalins that enable host colonization [52]. Similarly, integrons in several ESKAPE pathogens encode antibiotic resistance genes and contribute to their pathogenic potential [53]. Overall, integrons represent a unique group of elements that link cell-mediated acquisition and integration of foreign DNA, and expression of the acquired trait.

Genetic variations associated with pathogen emergence

In addition to virulence traits encoded within horizontally acquired DNA, other driving forces that are essential for host colonization are what we term genetic variations associated with pathogen emergence (GVPEs). These encompass a) preadaptations associated with the genetic makeup of the bacterial cells, and b) mutations that occur within the host. Specifically, in this section, we will explore the roles of two GVPEs: **virulence-adaptive polymorphisms** in the genomic background and **pathoadaptive mutations** in the process of pathogen emergence.

Virulence-adaptive polymorphisms and genomic background. Pathogenic bacteria encode preadaptations as part of their genomic background that are essential to successfully colonize the human host and play a major role in PE. These preadaptations have been historically overlooked as they are independent from the horizontal acquisition of genetic material, enhance environmental survival, and, unlike PMs, occur in the environment prior to host selective pressures. For instance, using comparative analyses of environmental and clinical strains of *V. cholerae*, we identified the presence of genes that circulate in environmental populations and encode **virulence adaptive polymorphisms (VAPs)** [11]. VAPs are allelic variations in core genes that confer preadaptations to virulence, typified by *ompU*, which encodes an outer membrane porin that plays a critical role in bacterial pathogenesis [11]. Unlike strains harboring non-VAP alleles of *ompU*, those with VAP-encoding alleles of this gene exhibit phenotypes associated with clinical outcomes such as antimicrobial resistance, indicating that these adaptations to virulence are present in their genomic background prior to host colonization [11]. Expression of *ompU* is regulated by the master virulence regulator

ToxR, which is also encoded within the core genome [55]. Interestingly, ToxR is required for the expression of the virulence regulator ToxT, encoded within VPI-1, and CT, encoded within the CTXø phage, demonstrating common regulatory pathways between VAPs and MGEs encoding virulence genes [56, 57]. Other studies have explored the relationship between the genomic background of *E. coli* strains and the retention and expression of acquired virulence factors [58]. These analyses have identified two distinct genomic profiles in pathogenic strains: a) an 'ancestral background' that allows expression of factors associated with mild/chronic diarrhea and that is found in most *E. coli*, and b) a 'derived background' that favors expression of factors with more severe pathologies, and that is found in EHEC, ETEC, and *Shigella*/EIEC [58]. Furthermore, newly emergent hypervirulent strains of *Klebsiella* that cause pyogenic liver abscesses, possess a unique genomic background that is distinct from the non-hypervirulent strains [59]. Interestingly, an interplay between the genomic background and adaptive mutations of drug-resistant *Mycobacterium tuberculosis* enhances the ability of this pathogen to infect and proliferate within susceptible hosts as well as increase its transmissibility [60, 61]. Finally, genomic analyses of environmental strains of *V. vulnificus* revealed that some isolates encode genes associated with the emergence of pathogenic traits such as serum resistance, suggesting that preadaptations to the host emerge from their unique genomic makeup [12, 13]. Overall, it is becoming increasingly clear that the specific genomic background of a given strain is essential for the emergence of pathogenic traits and its ability to colonize the host. This might explain the uneven distribution of virulent clones in various pathogenic species as certain strains within them might not encode the necessary preadaptations to emerge as human pathogens.

Pathoadaptive mutations and intra-host evolution. Besides a particular set of preadaptations in the genomic background of emergent pathogens, pathoadaptive mutations that arise within the host also confer virulence-adaptive traits and are a critical component of PE. Typically, pathoadaptive mutations are associated with tissue tropism, immune avoidance or metabolic adaptations [62-65]. The central role of pathoadaptive mutations in PE is exemplified by the speciation of pathogenic *Shigella* species from their closely related non-pathogenic commensal *E. coli* **(Fig. 2A)** [62]. Commensal *E. coli* require CadA-mediated lysine decarboxylase activity for growth in the intestine, however, CadA blocks the activity of the *Shigella* toxins. Convergent pathoadaptive mutations in *cadA* occurred independently in various *Shigella* species allowing for toxin activity and contributing to the emergence of pathogenic *Shigella* from commensal *E. coli* [62] **(Fig. 2A)**. Interestingly, the genes encoding the shiga toxins were likely acquired by phage-mediated transduction, suggesting interplay between pathoadaptive mutations and regulation of MGE-acquired virulence genes [66]. Pathoadaptive mutations also mediate tissue specificity and niche adaptation in pathogenic *E. coli* **(Fig. 2B)**. Over 90% of commensal *E. coli* express the mannose-sensitive type 1 fimbriae, a central mediator of the intestinal colonization process that binds the tri-mannose residues [67]. Variations in *fimH* (A27V) reduce the affinity to tri-mannose but increases it to mono-mannose residues, which are prevalent in the urothelial glycoproteins [63]. These pathoadaptive mutations in the *fimH* gene result in enhanced adhesion to the urinary tract epithelia and change tissue tropism in *E. coli*, ultimately allowing the transition from an intestinal commensal to a UTI pathogen (**Fig. 2B**).

Pathoadaptive mutations are also associated with the emergence of other major virulence phenotypes such as immune avoidance or resistance to host defense mechanisms. For instance, commensal *Mycobacterium abscessus* isolates adapted to the cystic fibrosis (CF) lung exhibit an excessive accumulation of nonsynonymous mutations in 30 genes [64]. In addition to other pathoadaptive mutations, isolates harboring mutations in the cell wall biosynthetic gene *ubiA* and global regulators such as *phoR* and *crp*/*fur* demonstrate enhanced survival in macrophages **(Fig. 2C)** [64]. However, pathoadaptive mutations in *phoR* and genes of the GBL locus reduce bacterial survival on fomites and consequently their transmission fitness, highlighting the fitness costs associated with pathoadaptive mutations in non-host environments [64] (also, **Box 2**). Interestingly, eight other loci of *M. abscessus* show conserved mutations across multiple patients suggesting convergent evolutionary pathways towards an analogous phenotype [64]. In non-pathogenic *E. coli*, pathoadaptive mutations in the LPS transporter components LptD (G580S) and LptE (T95I) result in increased production of outer membrane vesicles, and resistance to host complement and various antimicrobials [10]. Niche adaptation mediated by pathoadaptive mutations can result in hypervirulence in some pathogens such as *P. aeruginosa* (**Fig. 2D)** [65]. In the lungs of CF patients, the bacterium acquires mutations in the regulators of mucoidy (*mucA* and *algU*) that result in the characteristic hypermucoid cells seen in the CF lung. This enables *P. aeruginosa* to resist oxidative stress and evade phagocytosis (**Fig. 2D)** [65]. PMs facilitating niche adaptation and other virulence traits have also been identified in pathogens such as *Francisella tularensis* and group A streptococci [68, 69], stressing the relevance of these host adaptations in the process of PE. Interestingly, there is an interplay between pathoadaptive mutations in the core genome and transcriptional regulators encoded within MGEs that aid bacterial adaptation in the host [70]. SrfN is a *Salmonella* inner membrane protein encoded on the core genome and contributes to bacterial proliferation during systemic infection [70]. Pathoadaptive mutations in the promoter region of *srfN* enables SsrB, which is encoded within a PAI, to regulate *srfN* expression [70]. This event likely shaped host-pathogen interactions by promoting chronic *Salmonella* infections [70, 71].

Future directions

***Other mechanisms influencing PE*.** In addition to the canonical modes of HGT described above, other gene transfer mechanisms might play significant roles in the emergence of bacterial pathogens [72]. Potential candidates include membrane vesicles and gene transfer agents [73, 74]. Membrane vesicles can carry DNA in addition to other cytoplasmic content [75, 76]. Interestingly, membrane vesicles from pathogenic *E. coli* O157:H7 strains harbor genes encoding virulence determinants that result in enhanced cytotoxicity of non-pathogenic *E. coli* JM109 cells [77]. Nonetheless, there are only limited studies investigating membrane vesicles as mechanisms of virulence gene transfer [75, 77-79]. Furthermore, the extent of their involvement in the transfer and acquisition of virulence traits in bacterial communities such as in biofilms remains largely unexplored [80]. Bacteriophage-derived gene transfer agents encompass another potential mode of HGT [74, 81]. The role of these elements, which exclusively package random segments of bacterial DNA and inject them into recipient cells, in PE remains unknown [74, 81]. Future research will uncover the impact of these non-canonical HGT mechanisms in shaping bacterial PE and potentially reveal novel ones that could further explain the large extent of DNA mobility in bacterial populations.

Fitness costs of virulence and the boundaries of PE. There is increasing evidence that the acquisition and expression of virulence traits is associated with substantial fitness costs (**Box 2**). The relevance of these fitness trade-offs can be exemplified by the number of mechanisms associated with preventing the indiscriminate uptake and expression of exogenous genetic material. For example, **xenogeneic silencers (XS)** such as HN-S in *E. coli*, Lsr2 in *Mycobacterium tuberculosis*, or Rok in *Bacillus* spp. negatively regulate the expression of acquired virulence genes [82-85]. Recent studies reveal an interplay between XS proteins, the core genome and horizontally acquired genes. H-NS in *V. cholerae* binds to the promoters of the genes encoding two virulence regulators: ToxR, encoded in the ancestral chromosome, and ToxT, encoded within the PAI VPI-1 [86, 87]. Overall, it appears that bacterial cells have devised sophisticated "check and balance" mechanisms that prevent indiscriminate uptake and expression of foreign DNA. However, the molecular cues that dictate whether to silence, degrade, or express specific incoming genes and how this affects PE remains to be further elucidated.

Interplay of molecular mechanisms and environmental drivers. As exemplified above, PE is a multifactorial phenomenon that is not dictated by individual HGT processes or GVPEs, but by the interaction between various molecular and ecological drivers acting on the bacterial cell in both the environment and the human host (**Box 3**). To date, even though we have amassed a large body of knowledge regarding the mechanistic details of individual drivers of PE, the *interplay* between the different molecular and ecological elements influencing this phenomenon are still beginning to be elucidated (**Fig. 3**). The examples above highlight some fascinating insights that have been recently gained highlighting the combined effect of several molecular processes in shaping PE. Furthermore, examples of the interplay between multiple MGEs have been identified and seem to challenge our conventional understanding of HGT mechanisms. For instance, PAIs and plasmids that do not encode their own mobilization determinants can be transferred between strains by generalized transduction mediated by bacteriophages, suggesting a sophisticated orchestration between different MGEs [88, 89]. Nonetheless, the dynamics and potential synergy between the horizontally acquired elements, GVPEs and environmental drivers, both in their ecosystem and the human host, have barely been addressed. Our recent findings on *V. vulnificus* demonstrate that specific biotic (e.g., bacterial community structure) and abiotic factors (e.g., physicochemical variables) in their natural environment appear to generate selective pressures facilitating the emergence of strains with pathogenic potential in the population [12]. How these environmental pressures affect the acquisition and distribution of MGEs or GVPEs remains to be addressed. Furthermore, the extent of the role of the human host in subsequently creating new selection bottlenecks that synergistically lead to the emergence of pathogenic traits remains a critical area of study towards a comprehensive understanding of PE.

Concluding remarks

Pathogen emergence is a complex and multifactorial phenomenon. In this review, we describe novel developments in this field including recent findings that demonstrate the complex interactions that MGEs establish with the bacterial chromosome and other horizontally acquired elements. Furthermore, we expose the critical role of GVPEs, such as pathoadaptive mutations or VAPs, in the development of pathogenic traits in emerging pathogens. Given the critical public health relevance of PE, integrated studies of detailing the interactions of these elements require the development of holistic and multidisciplinary approaches that align with the “One Health” initiative and examine the potentially synergistic effect of the various mediators driving this complex phenomenon. Elucidating how the factors associated with the emergence and transmission of human pathogens interact and lead to non-linear outcomes is critical in order to establish effective surveillance strategies and reliable therapeutic treatments.

Acknowledgements

The authors would like to thank the anonymous reviewers for their thoughtful comments and suggestions. This work was supported by an NSF CAREER award (#2045671) to SAM.

References

1. Bloom, D.E. *et al.* (2017) Emerging infectious diseases: A proactive approach. Proc Natl Acad Sci U S A 114 (16), 4055-4059.
2. Martinez, J.L. (2013) Bacterial pathogens: from natural ecosystems to human hosts. Environ Microbiol 15 (2), 325-33.
3. Bonneaud, C. *et al.* (2019) Understanding the emergence of bacterial pathogens in novel hosts. Philos Trans R Soc Lond B Biol Sci 374 (1782), 20180328.
4. Rodriguez-Beltran, J. *et al.* (2021) Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. Nat Rev Microbiol 19 (6), 347-359.
5. Penades, J.R. *et al.* (2015) Bacteriophage-mediated spread of bacterial virulence genes. Curr Opin Microbiol 23, 171-8.
6. Botelho, J. and Schulenburg, H. (2021) The role of integrative and conjugative elements in antibiotic resistance evolution. Trends Microbiol 29 (1), 8-18.
7. Ghaly, T.M. *et al.* (2020) The peril and promise of integrons: Beyond antibiotic resistance. Trends Microbiol 28 (6), 455-464.
8. Arnold, B.J. *et al.* (2021) Horizontal gene transfer and adaptive evolution in bacteria. Nat Rev Microbiol.
9. Di Venanzio, G. *et al.* (2019) Urinary tract colonization is enhanced by a plasmid that regulates uropathogenic *Acinetobacter baumannii* chromosomal genes. Nat Commun 10 (1), 2763.
10. Kaito, C. *et al.* (2020) Non-pathogenic *Escherichia coli* acquires virulence by mutating a growth-essential LPS transporter. PLoS Pathog 16 (4), e1008469.
11. Shapiro, B.J. *et al.* (2016) Origins of pandemic *Vibrio cholerae* from environmental gene pools. Nat Microbiol 2, 16240.
12. Lopez-Perez, M. *et al.* (2021) Ecological diversification reveals routes of pathogen emergence in endemic *Vibrio vulnificus* populations. Proc Natl Acad Sci U S A 118 (40).
13. Lopez-Perez, M. *et al.* (2019) Evolutionary model of cluster divergence of the emergent marine pathogen *Vibrio vulnificus*: From genotype to ecotype. mBio 10 (1).
14. Iriarte, M. and Cornelis, G.R. (1996) Molecular determinants of *Yersinia* pathogenesis. Microbiologia 12 (2), 267-80.
15. Demeure, C.E. *et al.* (2019) *Yersinia pestis* and plague: an updated view on evolution, virulence determinants, immune subversion, vaccination, and diagnostics. Genes Immun 20 (5), 357-370.
16. Smiley, S.T. *et al.* (2019) *Yersinia pestis* Pla protein thwarts T cell defense against plague. Infect Immun 87 (5).
17. Prentice, M.B. *et al.* (2001) *Yersinia pestis* pFra shows biovar-specific differences and recent common ancestry with a *Salmonella enterica* serovar Typhi plasmid. J Bacteriol 183 (8), 2586-94.
18. Croxen, M.A. and Finlay, B.B. (2010) Molecular mechanisms of *Escherichia coli* pathogenicity. Nat Rev Microbiol 8 (1), 26-38.
19. Freedman, J.C. *et al.* (2015) *Clostridium perfringens* type A-E toxin plasmids. Res Microbiol 166 (4), 264-79.
20. Arredondo-Alonso, S. *et al.* (2020) Plasmids shaped the recent emergence of the major nosocomial pathogen *Enterococcus faecium*. mBio 11 (1).

21. Patton, M.J. *et al.* (2018) Plasmid negative regulation of CPAF expression is Pgp4 independent and restricted to invasive *Chlamydia trachomatis* biovars. mBio 9 (1).e02164-17
22. Coulson, G.B. *et al.* (2015) Transcriptome reprogramming by plasmid-encoded transcriptional regulators is required for host niche adaption of a macrophage pathogen. Infect Immun 83 (8), 3137-45.
23. Botelho, J. *et al.* (2020) Comprehensive genome data analysis establishes a triple whammy of carbapenemases, ICEs and multiple clinically relevant bacteria. Microb Genom 6 (10).
24. Flynn, K.J. and Swanson, M.S. (2014) Integrative conjugative element ICE-ßox confers oxidative stress resistance to *Legionella pneumophila in vitro* and in macrophages. mBio 5 (3), e01091-14.
25. Farzand, R. *et al.* (2019) ICEKp2: description of an integrative and conjugative element in *Klebsiella pneumoniae*, co-occurring and interacting with ICEKp1. Sci Rep 9 (1), 13892.
26. Dwyer, D.J. *et al.* (2015) Unraveling the physiological complexities of antibiotic lethality. Annu Rev Pharmacol Toxicol 55, 313-32.
27. Vandecraen, J. *et al.* (2017) The impact of insertion sequences on bacterial genome plasticity and adaptability. Crit Rev Microbiol 43 (6), 709-730.
28. Leavis, H.L. *et al.* (2007) Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of *E. faecium*. PLoS Pathog 3 (1), e7.
29. Bayjanov, J.R. *et al.* (2019) *Enterococcus faecium* genome dynamics during long-term asymptomatic patient gut colonization. Microb Genom 5 (7), 1-11.
30. Sentausa, E. *et al.* (2020) Insertion sequences drive the emergence of a highly adapted human pathogen. Microb Genom 6 (9), 1-15.
31. Elsen, S. *et al.* (2014) A type III secretion negative clinical strain of *Pseudomonas aeruginosa* employs a two-partner secreted exolysin to induce hemorrhagic pneumonia. Cell Host Microbe 15 (2), 164-76.
32. Partridge, S.R. *et al.* (2018) Mobile genetic elements associated with antimicrobial resistance. Clin Microbiol Rev 31 (4), 1-60.
33. Hochhut, B. *et al.* (2005) Pathogenicity islands and their role in bacterial virulence and survival. Contrib Microbiol 12, 234-254.
34. Dobrindt, U. *et al.* (2004) Genomic islands in pathogenic and environmental microorganisms. Nat Rev Microbiol 2 (5), 414-24.
35. Hacker, J. and Carniel, E. (2001) Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. EMBO Rep 2 (5), 376-81.
36. Islam, M.T. *et al.* (2017) Emergence, ecology and dispersal of the pandemic generating *Vibrio cholerae* lineage. Int Microbiol 20 (3), 106-115.
37. Kumar, A. *et al.* (2020) *Vibrio* Pathogenicity Island-1: The master determinant of cholera pathogenesis. Front Cell Infect Microbiol 10, 561296.
38. Jermyn, W.S. and Boyd, E.F. (2002) Characterization of a novel *Vibrio* pathogenicity island (VPI-2) encoding neuraminidase (*nanH*) among toxigenic *Vibrio cholerae* isolates. Microbiology (Reading) 148 (Pt 11), 3681-3693.
39. Jermyn, W.S. and Boyd, E.F. (2005) Molecular evolution of *Vibrio* pathogenicity island-2 (VPI-2): mosaic structure among *Vibrio cholerae* and *Vibrio mimicus* natural isolates. Microbiology (Reading) 151 (Pt 1), 311-322.

40. Carpenter, M.R. *et al.* (2015) Pathogenicity island cross talk mediated by recombination directionality factors facilitates excision from the chromosome. J Bacteriol 198 (5), 766-76.
41. Lou, L. *et al.* (2019) *Salmonella* Pathogenicity Island 1 (SPI-1) and its complex regulatory network. Front Cell Infect Microbiol 9, 270.
42. Jennings, E. *et al.* (2017) *Salmonella* SPI-2 type III secretion system effectors: molecular mechanisms and physiological consequences. Cell Host Microbe 22 (2), 217-231.
43. Schultz, B.M. *et al.* (2021) New insights on the early interaction between typhoid and non-typhoid *Salmonella* serovars and the host cells. Front Microbiol 12, 647044.
44. Svab, D. *et al.* (2015) A novel transducible chimeric phage from *Escherichia coli* O157:H7 Sakai strain encoding Stx1 production. Infect Genet Evol 29, 42-7.
45. Waldor, M.K. and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. Science 272 (5270), 1910-4.
46. van Wamel, W.J. *et al.* (2006) The innate immune modulators staphylococcal complement inhibitor and chemotaxis inhibitory protein of *Staphylococcus aureus* are located on beta-hemolysin-converting bacteriophages. J Bacteriol 188 (4), 1310-5.
47. Wirtz, C. *et al.* (2009) Transcription of the phage-encoded Panton-Valentine leukocidin of *Staphylococcus aureus* is dependent on the phage life-cycle and on the host background. Microbiology (Reading) 155 (Pt 11), 3491-3499.
48. Marshall, C.W. *et al.* (2021) Rampant prophage movement among transient competitors drives rapid adaptation during infection. Sci Adv 7 (29).
49. Duerkop, B.A. *et al.* (2012) A composite bacteriophage alters colonization by an intestinal commensal bacterium. Proc Natl Acad Sci U S A 109 (43), 17621-6.
50. Chen, J. *et al.* (2015) Intra- and inter-generic transfer of pathogenicity island-encoded virulence genes by *cos* phages. ISME J 9 (5), 1260-3.
51. Smith, A.B. and Siebeling, R.J. (2003) Identification of genetic loci required for capsular expression in *Vibrio vulnificus*. Infect Immun 71 (3), 1091-7.
52. Boucher, Y. *et al.* (2011) Local mobile gene pools rapidly cross species boundaries to create endemicity within global *Vibrio cholerae* populations. mBio 2 (2).
53. Chen, D.Q. *et al.* (2018) Integron mediated bacterial resistance and virulence on clinical pathogens. Microb Pathog 114, 453-457.
54. Chun, J. *et al.* (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. Proc Natl Acad Sci U S A 106 (36), 15442-7.
55. Crawford, J.A. *et al.* (1998) Analysis of ToxR-dependent transcription activation of *ompU*, the gene encoding a major envelope protein in *Vibrio cholerae*. Mol Microbiol 29 (1), 235-46.
56. Champion, G.A. *et al.* (1997) A branch in the ToxR regulatory cascade of *Vibrio cholerae* revealed by characterization of *toxT* mutant strains. Mol Microbiol 23 (2), 323-31.
57. Krukonis, E.S. *et al.* (2000) The *Vibrio cholerae* ToxR/TcpP/ToxT virulence cascade: distinct roles for two membrane-localized transcriptional activators on a single promoter. Mol Microbiol 38 (1), 67-84.

58. Escobar-Paramo, P. *et al.* (2004) A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. Mol Biol Evol 21 (6), 1085-94.
59. Struve, C. *et al.* (2015) Mapping the evolution of hypervirulent *Klebsiella pneumoniae*. mBio 6 (4), e00630.
60. Borrell, S. and Gagneux, S. (2011) Strain diversity, epistasis and the evolution of drug resistance in *Mycobacterium tuberculosis*. Clin Microbiol Infect 17 (6), 815-20.
61. Cohen, T. and Murray, M. (2004) Modeling epidemics of multidrug-resistant *M. tuberculosis* of heterogeneous fitness. Nat Med 10 (10), 1117-21.
62. Day, W.A., Jr. *et al.* (2001) Pathoadaptive mutations that enhance virulence: genetic organization of the *cadA* regions of *Shigella* spp. Infect Immun 69 (12), 7471-80.
63. Hommais, F. *et al.* (2003) The FimH A27V mutation is pathoadaptive for urovirulence in *Escherichia coli* B2 phylogenetic group isolates. Infect Immun 71 (6), 3619-22.
64. Bryant, J.M. *et al.* (2021) Stepwise pathogenic evolution of *Mycobacterium abscessus*. Science 372 (6541).
65. Marvig, R.L. *et al.* (2013) Genome analysis of a transmissible lineage of *Pseudomonas aeruginosa* reveals pathoadaptive mutations and distinct evolutionary paths of hypermutators. PLoS Genet 9 (9), e1003741.
66. Unkmeir, A. and Schmidt, H. (2000) Structural analysis of phage-borne *stx* genes and their flanking sequences in shiga toxin-producing *Escherichia coli* and *Shigella dysenteriae* type 1 strains. Infect Immun 68 (9), 4856-64.
67. Schembri, M.A. *et al.* (2000) Functional flexibility of the FimH adhesin: insights from a random mutant library. Infect Immun 68 (5), 2638-46.
68. Aziz, R.K. *et al.* (2010) Microevolution of group A streptococci *in vivo*: capturing regulatory networks engaged in sociomicrobiology, niche adaptation, and hypervirulence. PLoS One 5 (4), e9798.
69. Larsson, P. *et al.* (2009) Molecular evolutionary consequences of niche restriction in *Francisella tularensis*, a facultative intracellular pathogen. PLoS Pathog 5 (6), e1000472.
70. Osborne, S.E. *et al.* (2009) Pathogenic adaptation of intracellular bacteria by rewiring a *cis*-regulatory input function. Proc Natl Acad Sci U S A 106 (10), 3982-7.
71. Lawley, T.D. *et al.* (2006) Genome-wide screen for *Salmonella* genes required for long-term systemic infection of the mouse. PLoS Pathog 2 (2), e11.
72. Garcia-Aljaro, C. *et al.* (2017) Beyond the canonical strategies of horizontal gene transfer in prokaryotes. Curr Opin Microbiol 38, 95-105.
73. Toyofuku, M. *et al.* (2019) Types and origins of bacterial membrane vesicles. Nat Rev Microbiol 17 (1), 13-24.
74. Quebatte, M. and Dehio, C. (2019) *Bartonella* gene transfer agent: Evolution, function, and proposed role in host adaptation. Cell Microbiol 21 (11), e13068.
75. Fulsundar, S. *et al.* (2014) Gene transfer potential of outer membrane vesicles of *Acinetobacter baylyi* and effects of stress on vesiculation. Appl Environ Microbiol 80 (11), 3469-83.
76. Lee, E.Y. *et al.* (2009) Gram-positive bacteria produce membrane vesicles: proteomics-based characterization of *Staphylococcus aureus*-derived membrane vesicles. Proteomics 9 (24), 5425-36.

77. Yaron, S. *et al.* (2000) Vesicle-mediated transfer of virulence genes from *Escherichia coli* O157:H7 to other enteric bacteria. Appl Environ Microbiol 66 (10), 4414-20.
78. Bitto, N.J. *et al.* (2021) *Staphylococcus aureus* membrane vesicles contain immunostimulatory DNA, RNA and peptidoglycan that activate innate immune receptors and induce autophagy. J Extracell Vesicles 10 (6), e12080.
79. Jiang, Y. *et al.* (2014) Membrane vesicles of *Clostridium perfringens* type A strains induce innate and adaptive immunity. Int J Med Microbiol 304 (3-4), 431-43.
80. Dell'Annunziata, F. *et al.* (2021) Gene transfer potential of outer membrane vesicles of Gram-negative bacteria. Int J Mol Sci 22 (11).
81. Segers, F.H. *et al.* (2017) Genomic changes associated with the evolutionary transition of an insect gut symbiont into a blood-borne pathogen. ISME J 11 (5), 1232-1244.
82. Pina-Iturbe, A. *et al.* (2020) Horizontally acquired homologs of xenogeneic silencers: modulators of gene expression encoded by plasmids, phages and genomic islands. Genes (Basel) 11 (2).
83. Bartek, I.L. *et al.* (2014) *Mycobacterium tuberculosis* Lsr2 is a global transcriptional regulator required for adaptation to changing oxygen levels and virulence. mBio 5 (3), e01106-14.
84. Lippa, A.M. *et al.* (2021) H-NS-like proteins in *Pseudomonas aeruginosa* coordinately silence intragenic transcription. Mol Microbiol 115 (6), 1138-1151.
85. Williams McMackin, E.A. *et al.* (2019) H-NS family members MvaT and MvaU regulate the *Pseudomonas aeruginosa* type III secretion system. J Bacteriol 201 (14).
86. Kazi, M.I. *et al.* (2016) ToxR antagonizes H-NS regulation of horizontally acquired genes to drive host colonization. PLoS Pathog 12 (4), e1005570.
87. Stone, J.B. and Withey, J.H. (2021) H-NS and ToxT inversely control cholera toxin production by binding to overlapping DNA sequences. J Bacteriol 203 (18), e0018721.
88. Humphrey, S. *et al.* (2021) Staphylococcal phages and pathogenicity islands drive plasmid evolution. Nat Commun 12 (1), 5845.
89. O'Brien, F.G. *et al.* (2015) Origin-of-transfer sequences facilitate mobilisation of non-conjugative antimicrobial-resistance plasmids in *Staphylococcus aureus*. Nucleic Acids Res 43 (16), 7971-83.
90. Rossolini, G.M. *et al.* (2014) Update on the antibiotic resistance crisis. Curr Opin Pharmacol 18, 56-60.
91. Quintiliani, R., Jr. and Courvalin, P. (1996) Characterization of Tn1547, a composite transposon flanked by the IS16 and IS256-like elements, that confers vancomycin resistance in *Enterococcus faecalis* BM4281. Gene 172 (1), 1-8.
92. Botelho, J. *et al.* (2017) Characterization of the pJB12 plasmid from *Pseudomonas aeruginosa* reveals Tn6352, a novel putative transposon associated with mobilization of the *blaVIM*-2-harboring In58 integron. Antimicrob Agents Chemother 61 (5).
93. Weigel, L.M. *et al.* (2007) High-level vancomycin-resistant *Staphylococcus aureus* isolates associated with a polymicrobial biofilm. Antimicrob Agents Chemother 51 (1), 231-8.
94. Farrugia, D.N. *et al.* (2013) The complete genome and phenome of a community-acquired *Acinetobacter baumannii*. PLoS One 8 (3), e58628.

95. Bartell, J.A. *et al.* (2019) Evolutionary highways to persistent bacterial infection. Nat Commun 10 (1), 629.
96. Tarafder, A.K. *et al.* (2020) Phage liquid crystalline droplets form occlusive sheaths that encapsulate and protect infectious rod-shaped bacteria. Proc Natl Acad Sci U S A 117 (9), 4724-4731.
97. Bottery, M.J. *et al.* (2019) Temporal dynamics of bacteria-plasmid coevolution under antibiotic selection. ISME J 13 (2), 559-562.
98. Choi, Y. *et al.* (2020) Comparison of fitness cost and virulence in chromosome- and plasmid-mediated colistin-resistant *Escherichia coli*. Front Microbiol 11, 798.
99. Sturm, A. *et al.* (2011) The cost of virulence: retarded growth of *Salmonella* Typhimurium cells expressing type III secretion system 1. PLoS Pathog 7 (7), e1002143.
100. Diard, M. *et al.* (2013) Stabilization of cooperative virulence by the expression of an avirulent phenotype. Nature 494 (7437), 353-6.
101. Rietsch, A. and Mekalanos, J.J. (2006) Metabolic regulation of type III secretion gene expression in *Pseudomonas aeruginosa*. Mol Microbiol 59 (3), 807-20.
102. Vasanthakrishnan, R.B. *et al.* (2015) PrfA regulation offsets the cost of *Listeria* virulence outside the host. Environ Microbiol 17 (11), 4566-79.
103. Delavat, F. *et al.* (2017) The hidden life of integrative and conjugative elements. FEMS Microbiol Rev 41 (4), 512-537.
104. Gaillard, M. *et al.* (2008) Host and invader impact of transfer of the *clc* genomic island into *Pseudomonas aeruginosa* PAO1. Proc Natl Acad Sci U S A 105 (19), 7058-63.
105. Lee, S.M. *et al.* (2007) Fitness cost of staphylococcal cassette chromosome *mec* in methicillin-resistant *Staphylococcus aureus* by way of continuous culture. Antimicrob Agents Chemother 51 (4), 1497-9.
106. Combe, M. *et al.* (2017) Global and local environmental changes as drivers of Buruli ulcer emergence. Emerg Microbes Infect 6 (4), e21.
107. Raeside, C. *et al.* (2014) Large chromosomal rearrangements during a long-term evolution experiment with *Escherichia coli*. mBio 5 (5), e01377-14.
108. Mathee, K. *et al.* (1999) Mucoid conversion of *Pseudomonas aeruginosa* by hydrogen peroxide: a mechanism for virulence activation in the cystic fibrosis lung. Microbiology (Reading) 145 (Pt 6), 1349-1357.
109. Sitkiewicz, I. (2018) How to become a killer, or is it all accidental? Virulence strategies in oral streptococci. Mol Oral Microbiol 33 (1), 1-12.
110. Matz, C. and Kjelleberg, S. (2005) Off the hook--how bacteria survive protozoan grazing. Trends Microbiol 13 (7), 302-7.
111. Diaz-Munoz, S.L. and Koskella, B. (2014) Bacteria-phage interactions in natural environments. Adv Appl Microbiol 89, 135-83.
112. Jurkevitch, E. and Jacquet, S. (2017) *Bdellovibrio* and like organisms: outstanding predators! Med Sci (Paris) 33 (5), 519-527.
113. Erken, M. *et al.* (2013) The rise of pathogens: predation as a factor driving the evolution of human pathogens in the environment. Microb Ecol 65 (4), 860-8.
114. Hosseinidoust, Z. *et al.* (2013) Predation in homogeneous and heterogeneous phage environments affects virulence determinants of *Pseudomonas aeruginosa*. Appl Environ Microbiol 79 (9), 2862-71.

115. Bruggemann, H. *et al.* (2006) Adaptation of *Legionella pneumophila* to the host environment: role of protein secretion, effectors and eukaryotic-like proteins. Curr Opin Microbiol 9 (1), 86-94.
116. Park, J.M. *et al.* (2020) Combinatorial selection in amoebal hosts drives the evolution of the human pathogen *Legionella pneumophila*. Nat Microbiol 5 (4), 599-609.
117. Steinberg, K.M. and Levin, B.R. (2007) Grazing protozoa and the evolution of the *Escherichia coli* O157:H7 Shiga toxin-encoding prophage. Proc Biol Sci 274 (1621), 1921-9.
118. Matz, C. *et al.* (2011) Acquired type III secretion system determines environmental fitness of epidemic *Vibrio parahaemolyticus* in the interaction with bacterivorous protists. PLoS One 6 (5), e20275.
119. Lysenko, E.S. *et al.* (2010) Within-host competition drives selection for the capsule virulence determinant of *Streptococcus pneumoniae*. Curr Biol 20 (13), 1222-6.

Glossary

Horizontal gene transfer (HGT): Movement of genetic information between bacteria that can confer in some cases virulence traits to the recipient cell.

Genetic variations associated with pathogen emergence (GVPEs): Allelic variations and mutations leading to preadaptations stemming from the environment or adaptations occurring within the host that favor the emergence of pathogenic traits.

Mobile genetic elements (MGEs): Segments of DNA that mediate the movement of genes within and between bacteria. Includes plasmids, pathogenicity islands (PAIs), transposons, integrative conjugative elements (ICEs), and prophages. Many encode functions associated with pathogen emergence such as colonization factors, metabolic enzymes, or toxins.

Pathoadaptive mutations: Intra-host mutational changes that enhance virulence or adaptation to the host without acquisition of additional virulence factors.

Pathogen: An organism that can colonize and damage a host.

Pathogen emergence (PE): A multifactorial non-linear phenomenon that culminates with the ability of a bacterium to effectively colonize and harm the human host.

Preadaptation: Genetic variation stemming from the environment that confers virulence-associated traits that enhance host colonization such as resistance to host antimicrobials.

Virulence: The ability of a pathogen to cause damage to a host.

Virulence adaptive polymorphisms: Allelic variations in core genes that confer preadaptations to virulence and circulate in environmental populations.

Virulence determinants: Traits in pathogens that directly influence pathogenicity and are typically associated with processes aiding virulence, host colonization, cell-associated factors, or secreted ones.

Xenogeneic silencers: Proteins that bind to and repress the expression of AT-rich sequences typically found on acquired DNA. Xenogeneic silencers play a major role in PE by preventing indiscriminate expression of horizontally acquired virulence genes.

Text Boxes

Box 1. Role of antibiotic resistance in pathogen emergence

Antibiotic resistance (AR), although not considered a virulence factor *per se*, can complement pathogenicity in major ways. Exposure to antibiotics and acquisition of resistance typically occurs in nosocomial settings or the environment [90]. Both settings can be a significant driving force in the emergence of resistance. Even though the different molecular drivers of PE discussed here can mediate their acquisition, some like transposons or ICE elements primarily transfer genes associated with AR. For instance, most of the known cargo genes of transposons are related to AR, including resistance to vancomycin in *E. faecalis* [91], and ß-lactams such as carbapenems in *P. aeruginosa* [92]. Plasmids mediate in-patient emergence of vancomycin-resistant *S. aureus* (VRSA) in polymicrobial biofilms comprising *E. faecalis*, *E. faecium*, *P. aeruginosa*, *Micrococcus* spp., and *Morganella morganii* [93]. Similarly, several ICEs can spread AR among bacteria, such as the multidrug RND efflux pump-encoding *ebyCAB* genes on the 63.3 kb ICE element Tn*4371* in MDR *Stenotrophomonas maltophilia* [23, 32]. Nosocomial selection bottlenecks are believed to have shaped PAI-mediated resistance in *A. baumannii* to over six different classes of antibiotics, heavy metals and quaternary ammonium compounds [94]. Additionally, pathoadaptive mutations can shape intra-host appearance of AR. For instance, in *P. aeruginosa*, mutations in genes encoding antibiotic efflux pump repressors and other genes involved in resistance to multiple classes of antibiotics drive the *in vivo* emergence of multi-drug resistant strains [95]. A fascinating mode of phage-mediated AR are the phage liquid crystal ‘armors’ formed by the capsids of filamentous phages that encapsulate cells and promote antibiotic tolerance [96].

Overall, given the ubiquitousness of HGT in bacteria and wide availability of AR genes in bacterial populations the problem of antibiotic resistance will likely remain associated with that of PE for the coming decades.

Box 2. Fitness costs associated with virulence

The expression of virulence factors in a non-host environment is associated with fitness costs. The coordinated expression of virulence genes adds to the energy expenditure in terms of resources, additional regulators and chaperones [97]. This imposes substantial fitness costs to the bacterial cell such as lower growth rates [97]. The fitness costs associated with the acquisition of antibiotic resistance has been widely examined [6, 98], however, a limited set of studies have focused on the trade-offs inherent to the emergence of traits associated with pathogenesis. For instance, *in vitro* expression of the type 3 secretion system in *S*. Typhimurium is correlated with a 2-fold reduction in its growth rate [99]. In some cases, virulence genes are expressed only by some cells within a population (bistable expression), as seen in *S.* Typhimurium and *P. aeruginosa*, resulting in a slow-growing subpopulation that expresses virulence genes, and a phenotypically avirulent fast-growing subpopulation [100, 101]. The associated costs likely explain why the expression of virulence factors is tightly regulated and occurs only in specific stages of the infection process. For instance, the constitutive expression of virulence genes in *Listeria monocytogenes* outside the host not only leads to severely impaired growth in rich media, but also a reduced ability of the bacteria to survive in soil environments [102].

In addition to bistable expression, bacteria have devised other ways to offset the inherent fitness costs of untimely virulence gene expression while still allowing for the emergence of pathogenic traits. Besides virulence genes, some ICEs express other genes that are beneficial to the bacterium such as metabolic enzymes, potentially offsetting the fitness costs associated with ICE maintenance [103]. For instance, the 103 kb ICE*clc*, which is widely distributed in Proteobacteria, does not confer a fitness cost even in the absence of selective pressure [104]. Several reasons could explain this reduced fitness costs, including expression from a constitutively active chromosomal promoter or the presence of its own regulators instead of dependence on the host [104]. Similarly, in contrast to the hospital-acquired MRSA, the more transmissible community-acquired MRSA strains encode several adaptations that result in a reduced fitness burden during host colonization and infection [105]. Thus, the intricate act of balancing this cost-benefit equation is an integral part in the emergence of pathogenic traits in bacteria.

Box 3. Biotic and abiotic drivers of pathogen emergence

The underlying ecological drivers that lead to the development of virulence traits in bacteria represent a critical component of the process of PE. Numerous biotic and abiotic factors found in the natural environment of the bacterium and/or the human host have been directly associated with the emergence of bacterial pathogens.

Certain ***abiotic factors*** drive adaptations in the environment that ultimately affect fitness in the host. In emergent pathogens such as *Mycobacterium ulcerans*, changes to the pH and oxygen levels in aquatic habitats, as well as increased temperature, favors biofilm formation and host transmission [106]. Environmental stressors also trigger IS element-associated genomic inversions, deletions, and duplications [107], enhancing fitness and the emergence of pathogenic traits. The oxidative stress induced by reactive oxygen species can influence the expression of virulence-associated phenotypes both in the environment and the host. Outside of the human host, *P. aeruginosa* can change to a hypermucoid phenotype upon exposure to reactive oxygen species due to modulation of the mismatch repair mechanism [108]. A similar response is also seen in the lungs of CF patients (**Fig. 2D**), affecting fitness in the host [65]. The accessibility to nutrients within the host is a major determinant of infection by some bacteria. For instance, the infection of the oral tract by pathogenic streptococci requires access to carbohydrates in the oral cavity and the concerted action of multiple biotic and abiotic factors [109].

Several ***biotic factors*** in the environment have also been identified that play a role in facilitating the PE process including environmental predation by heterotrophic protists, bacteriophages and predatory bacteria [110-112] (**Fig 3)**. The evolution of a variety of defense mechanisms against these threats play the dual role as virulence factors in humans [113]. For example, lytic phage-resistant variants of *P. aeruginosa* demonstrate enhanced production of key virulence factors, increased resistance to phagocytosis and cytotoxicity to keratinocytes [114]. In some cases, bacteria evolve specific features that increase survival in non-human hosts and ultimately contribute to pathogenicity in humans. This concept of convergent evolution of virulence is exemplified in *L. pneumophila* and the Dot/Icm type IV secretion system [115]. The Dot/Icm system evolved as a result of *L. pneumophila* growth in multiple amoebal hosts but plays a key role in bacterial survival within alveolar macrophages [116]. Similar examples include the shiga-like toxins produced by many pathogenic *E. coli* (STEC), which help bacterial cells evade the ciliate *Tetrahymena thermophila* [117], and both type III secretion systems of *Vibrio parahaemolyticus*, which are required for survival against protists in aquatic environments [118]. Competition against microorganisms from the host microbiota also contributes to the evolution of pathogenic strains, affecting both invasiveness and immune avoidance. *Strep. pneumoniae*, *H. influenzae*, and *Neisseria meningitidis* are all inhabitants of the nasal tract where competition between two or more species selects for virulence [119]. *H. influenzae* triggers host opsonization of the other bacteria in the milieu. In response to this, pneumococci produce a capsule that not only protects against phagocytosis but also enhances its invasiveness and pathogenicity [119].

Figure legends

Figure 1. Horizontal gene transfer mechanisms and interplay between MGEs associated with pathogen emergence. (A) Acquisition of genes via HGT represents a fundamental driver of the process of PE. Genes encoding pathogenic traits (red arrows) can be acquired by bacteria on various elements such as plasmids, insertion sequences (IS), transposons (Tn), pathogenicity islands (PAIs), integrative and conjugative elements (ICEs), or bacteriophages (phages). These genes can encode a variety of virulence factors such as toxins, adhesins or secretion systems, or favor adaptation by modulating the expression of virulence traits. Integrons present on the chromosome enable successful integration and expression of exogenous virulence genes. (B) Plasmid pAB5 modulates the expression of chromosomal genes in *A. baumannii*. Strains that do not harbor pAB5 exhibit enhanced colonization of lung epithelia, likely mediated by *cup1* and *cup2* gene clusters, which encode the CUP pili, a T6SS system, and genes involved in biofilm formation. Strains that acquire pAB5 exhibit enhanced adherence to urinary epithelia. This phenotype is due to modulation of chromosomal gene expression by regulators encoded on pAB5. The regulators encoded in pAB5 repress the expression of *cup1* and *cup2*, while activating the expression of *ompA*, *ompW*, and others. (C) Cooperation of two ICEs enhances plasmid mobilization in *K. pneumoniae*. Clinical strains of *K. pneumoniae* harbor an ICE, ICEKp1, which encodes a fully functional T4SS involved in low efficiency plasmid mobilization. A subset of clinical strains harbors a second ICE, ICEKp2 encoding a Mob2 ATPase. The Mob2 ATPase encoded on ICEKp2 complements the function of the T4SS encoded on ICEKp1, enhancing the efficiency of plasmid transfer to recipient cells.

Figure 2. **Pathoadaptive mutations play a major role in pathogen emergence**. Pathoadaptive mutations occurring within the host play a major role in the emergence of virulence-associated traits, such as toxin production, tissue tropism or immune avoidance. (A) After acquisition of the STX (Shiga) toxin by commensal *E. coli* (*stx*$^{-}$ *cadA*$^{+}$), the accumulation of PMs is associated with the emergence of pathogenic *Shigella* (*stx*$^{+}$ *cadA*$^{-}$) in the intestinal epithelium. Convergent pathoadaptive mutations in *cadA* allows activity of STX facilitating pathogenesis of *Shigella* spp. (B) Pathoadaptive mutations in the fimbrial *fimH* gene (A27V) of commensal intestinal *E. coli* determine tissue tropism in this species, as they aid in the emergence of urovirulent strains that preferentially bind to the urinary epithelium. (C) In the cystic fibrosis lung, pathoadaptive mutations are associated with the emergence of pathogenic strains of commensal *M. abscessus*. Pathoadaptive mutations in the cell wall biosynthetic *ubiA* gene and several transcriptional regulators facilitate enhanced survival within macrophages and allow bacterial dissemination. (D) Pathoadaptive mutations in the alginate regulatory genes *mucA* and *algU* of *P. aeruginosa* result in a hypermucoid phenotype that facilitates resistance to reactive oxygen species and aid phagocyte evasion. These mutations are required in the switch from acute to chronic infection lifestyles by this pathogen.

Figure 3. Molecular mechanisms and drivers of pathogen emergence. This schematic highlights factors that influence bacterial pathogen emergence both in the environment and within the human host: (A) molecular mechanisms such as horizontal gene transfer (HGT), and genetic variations associated with pathogen emergence (GVPEs) and (B) biotic and abiotic drivers. To date, the interplay between the different

molecular and ecological elements influencing this phenomenon remain poorly understood.