Evolutionary Dynamics of Pathoadaptation Revealed by Three Independent Acquisitions of the VirB/D4 Type IV Secretion System in *Bartonella*

Alexander Harms[1,6], Francisca H.I.D. Segers[2,7], Maxime Quebatte[1], Claudia Mistl[1], Pablo Manfredi[1], Jonas Körner[1], Bruno B. Chomel[3], Michael Kosoy[4], Soichi Maruyama[5], Philipp Engel[2], and Christoph Dehio[1,*]

[1]Focal Area Infection Biology, Biozentrum, University of Basel, Switzerland

[2]Department of Fundamental Microbiology, University of Lausanne, Switzerland

[3]Department of Population Health and Reproduction, School of Veterinary Medicine, University of California Davis

[4]Bacterial Diseases Branch, Division of Vector-Borne Disease, Centers for Disease Control and Prevention, Fort Collins, Colorado

[5]Laboratory of Veterinary Public Health, Department of Veterinary Medicine, College of Bioresource Sciences, Nihon University, Tokyo, Japan

[6]Present address: Department of Biology, Center of Excellence for Bacterial Stress Response and Persistence, University of Copenhagen, Denmark

[7]Present address: Institute of Zoology, Johannes Gutenberg University Mainz, Germany

*Corresponding author: E-mail: christoph.dehio@unibas.ch.

Accepted: March 3, 2017

Data deposition: The genome sequences of all (re)sequenced or reannotated bartonellae have been deposited in the NCBI GenBank database under BioProject accession PRJNA360985 with accession numbers KY583505 (*B. ancashensis* 20.00), CP019789 (genome of *B. schoenbuchensis* R1), CP019790 (pVbh of *B. schoenbuchensis* R1), CP019489 (*B.* sp. 1-1c), CP019780 (*B.* sp. A1379B), MUYE00000000 (*B.* sp. AR15-3), CP019782 (*B.* sp. CDCskunk), CP019785 (*B.* sp. Coyote22sub2), CP019784 (*B.* sp. HoopaDog 114), CP019783 (*B.* sp. HoopaFox 11B), CP019787 (*B.* sp. JB15), CP019788 (*B.* sp. JB63), CP019786 (*B.* sp. Raccoon60), MUBG00000000 (*B.* sp. SikaDeer WD12.1), CP019781 (*B.* sp. SikaDeer WD16.2), and MUYW00000000 (*B. taylorii* IBS296). All sequence alignments are available from the corresponding author upon request.

Abstract

The α-proteobacterial genus *Bartonella* comprises a group of ubiquitous mammalian pathogens that are studied as a model for the evolution of bacterial pathogenesis. Vast abundance of two particular phylogenetic lineages of *Bartonella* had been linked to enhanced host adaptability enabled by lineage-specific acquisition of a VirB/D4 type IV secretion system (T4SS) and parallel evolution of complex effector repertoires. However, the limited availability of genome sequences from one of those lineages as well as other, remote branches of *Bartonella* has so far hampered comprehensive understanding of how the VirB/D4 T4SS and its effectors called Beps have shaped *Bartonella* evolution. Here, we report the discovery of a third repertoire of Beps associated with the VirB/D4 T4SS of *B. ancashensis*, a novel human pathogen that lacks any signs of host adaptability and is only distantly related to the two species-rich lineages encoding a VirB/D4 T4SS. Furthermore, sequencing of ten new *Bartonella* isolates from under-sampled lineages enabled combined *in silico* analyses and wet lab experiments that suggest several parallel layers of functional diversification during evolution of the three Bep repertoires from a single ancestral effector. Our analyses show that the Beps of *B. ancashensis* share many features with the two other repertoires, but may represent a more ancestral state that has not yet unleashed the adaptive potential of such an effector set. We anticipate that the effectors of *B. ancashensis* will enable future studies to dissect the evolutionary history of *Bartonella* effectors and help unraveling the evolutionary forces underlying bacterial host adaptation.

Key words: filamentation induced by cAMP, AMPylation, parallel evolution, bacterial effector.

© The Author(s) 2017. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Introduction

The successful infection of eukaryotic hosts by many bacteria depends on the subversion of host cell functioning by effector proteins that are translocated into target cells via macromolecular machineries like the type IV secretion system (T4SS) (Costa, et al. 2015; Gonzalez-Rivera et al. 2016). The evolution of these machineries from ancestors with functions in genuine bacterial processes is well established (Abby and Rocha 2012; Frank et al. 2005; Guglielmini et al. 2013). However, the evolutionary trajectories of secreted effectors have proven difficult to study, because they are frequently blurred by horizontal gene transfer (Brown and Finlay 2011; Burstein et al. 2016). One major exception is the α-proteobacterial genus *Bartonella* where two distinct effector sets secreted by the VirB/D4 T4SS have evolved from a single common ancestor and are thought to promote host adaptability (Engel et al. 2011; Saenz et al. 2007). These bacteria therefore provide an ideal setting to study the molecular basis of host adaptation and the evolution of bacterial effector proteins.

The characteristic stealth infection strategy of *Bartonella* is hallmarked by arthropod transmission and persistent, hemotropic infections in their respective mammalian reservoir hosts (Harms and Dehio 2012). Host-specific adaptation is critical for successful implementation of the *Bartonella* stealth infection strategy, because reservoir host infections are typically long-lasting but do not cause obvious disease symptoms. In contrast, infections of incidental hosts are often self-limiting but can be accompanied by considerable morbidity (Angelakis and Raoult 2014; Harms and Dehio 2012). The phylogeny of *Bartonella* is divided into four major lineages (L1–L4) of which L3 and L4 together comprise the overwhelming majority of different reservoir hosts (Engel et al. 2011). Conversely, *B. bacilliformis* (L1) is a strict human pathogen, whereas the species of L2 seem to be limited to ruminant hosts (Harms and Dehio 2012; Minnick et al. 2014). The diversity of species in L3 and L4 is the result of two parallel adaptive radiations that have been linked to the acquisition of the VirB/D4 T4SS and the vast potential of its secreted effectors to manipulate host cell functions (Engel et al. 2011; Saenz et al. 2007). Previous work showed that the VirB/D4 T4SS had been independently acquired in L3 and L4 together with a single primordial effector that, via parallel series of effector gene duplication and functional diversification, evolved into complex effector repertoires (Engel et al. 2011). The modular evolvability and functional diversity of these effector repertoires is thought to have enabled the surge of host adaptability that sparked the adaptive radiations of L3 and L4 (Engel et al. 2011; Guy et al. 2013). This view is supported by the observation that neither *B. bacilliformis* (L1) nor any species of L2 encode host-interacting T4SS or show any evidence of recent host shifts (Engel et al. 2011; Saenz et al. 2007). Surprisingly, a recent study on *B. ancashensis*, a new species of L1 that shares its human reservoir host with *B. bacilliformis*, reported the discovery of a VirB/D4 T4SS in the genome of this organism, but no further functional or evolutionary analyses were carried out (Hang et al. 2015).

All *Bartonella* effector proteins (Beps) share a bipartite secretion signal composed of a C-terminal Bep intracellular delivery (BID) domain and a positively charged tail (Schulein et al. 2005; Siamer and Dehio 2015). In addition, the far majority of Beps contain an N-terminal filamentation induced by cAMP (FIC) domain. The FIC—BID domain structure is thought to reflect the ancestral state of these effector proteins from which other effector architectures have derived throughout the course of evolution (Engel et al. 2011). FIC domains are enzymatic domains that mediate the posttranslational modification of target proteins via AMPylation, which typically results in target protein inactivation (Harms et al. 2016). The AMPylation activity of FIC domains is dependent on the integrity of an HPFX[D/E]GNGRXXR signature motif that forms part of the active site. Alternative sequences at the signature motif are indicative of divergent molecular activities, e.g., in case of the *Legionella pneumophila* effector AnkX that mediates target protein phosphocholination or the Doc toxin of bacteriophage P1 that is a kinase (Castro-Roa et al. 2013; Mukherjee et al. 2011). With regard to the Beps, it has been shown that Bep1 of *B. rochalimae* (L3) AMPylates a specific subset of small host GTPases (Dietz et al., in revision), that Bep2 of the same organism AMPylates the host intermediate filament protein vimentin (Pieles et al. 2014), and that BepA of *B. henselae* (L4) causes the AMPylation of two unknown host proteins of ca. 40–50 kDa (Palanivelu et al. 2010). Apart from effectors harboring a FIC domain, novel effector types have evolved independently in both L3 and L4. These derived effector architectures consist of one or several BID domains and/or tyrosine-containing motives phosphorylated by host kinases to subvert host cell functioning (Engel et al. 2011; Siamer and Dehio 2015).

In this work, we report the discovery of a third set of Beps linked to a VirB/D4 T4SS in *B. ancashensis*, a human pathogen of L1 that is clinically indistinguishable from its sister species *B. bacilliformis* (Blazes et al. 2013; Minnick et al. 2014; Mullins et al. 2013). We performed an extensive analysis of the Beps of *B. ancashensis*, L3, and L4, that revealed signs of functional diversification in all three effector repertoires. However, our results suggest that the Beps of *B. ancashensis* may not yet have explored the full adaptive potential as present in the other lineages. Furthermore, we present the genome sequences of eight additional isolates of L3 that enabled a comprehensive view on the Beps of this lineage for which only a few genomes have previously been available (Engel et al. 2011). Taken together, our analyses shed new light on the

Table 1

Genomic Features of *Bartonella* Genomes That Were (Re)Sequenced or Reannotated as Part of This Study

Organism	Reservoir Host	Lineage	Size (bp)[a]	Contigs[a]	GC %[a]	CDS	Analysis	Source of Strain/Sequence
B. ancashensis 20.00	Human (*Homo sapiens*)	1	1,467,789	1	37.1	1,251	reannotated	ATCC BAA-2694/ (Hang et al. 2015)
B. schoenbuchensis R1	Roe deer (*Capreolus capreolus*)	2	1,672,626/55,761	1/1	37.9/35.2	1,555/84	resequenced	Dehio collection/this study
B. sp. Sika Deer WD12.1	Sika deer (*Cervus nippon*)	2	1,739,423/49,553	2/1	37.9/34.9	1,565/71	sequenced	(Sato et al. 2012b)/this study
B. sp. Sika Deer WD16.2	Sika deer (*Cervus nippon*)	2	1,762,996	1	37.6	1,526	sequenced	(Sato et al. 2012b)/this study
B. sp. CDC_skunk	Stripped skunk (*Mephitis mephitis*)	3	1,615,223	1	36.1	1,374	sequenced	this study/this study
B. sp. JB15	Japanese badger (*Meles anakuma*)	3	1,494,018	1	35.3	1,283	sequenced	(Sato et al. 2012a)/this study
B. sp. JB63	Japanese badger (*Meles anakuma*)	3	1,493,693	1	35.3	1,255	sequenced	(Sato et al. 2012a)/this study
B. sp. Hoopa Fox 11B	Gray fox (*Urocyon cinereoargenteus*)	3	1,579,438	1	36	1,316	sequenced	(Henn et al. 2007)/this study
B. sp. Hoopa Dog 114	Dog (*Canis canis*)	3	1,571,582	1	35.9	1,348	sequenced	(Henn et al. 2007)/this study
B. sp. A1379B	Red fox (*Vulpes vulpes*)	3	1,541,976	1	35.8	1,289	sequenced	(Henn et al. 2009)/this study
B. sp. Coyote22sub2	Coyote (*Canis latrans*)	3	1,561,331	1	35.9	1,325	sequenced	(Henn et al. 2009)/this study
B. sp. Racoon60	Raccoon (*Procyon lotor*)	3	1,615,700	1	36.1	1,400	sequenced	(Henn et al. 2009)/this study
B. sp. AR15-3	American Red Scquirrel (*Tamiasciurus hudosonicus*)	3	1,630,082	2	35.9	1,447	resequenced	Dehio collection/ (Engel et al. 2011)
B. sp. 1-1c	Rat (*Rattus norvegicus*)	3	1,600,621	1	35.9	1,358	resequenced	Dehio collection/ (Engel et al. 2011)
B. taylorii IBS325	Vole (*Microtus* sp.)	4	1,677,768	194	38.8	1,336	sequenced	Dr. Y. Piemont, Dr. R. Heller/this study

NOTE.—Relevant features of the *Bartonella* genomes sequenced, resequenced, or reannotated as part of this study. A comprehensive list of all *Bartonella* genomes used throughout this work is presented in supplementary table S1, Supplementary Material online.

[a]The first number refers to the main chromosome, the second number refers to plasmids.

evolutionary trajectories of *Bartonella* effector proteins and highlight their value as a model for the evolution of bacterial pathogenesis.

Materials and Methods

Bacterial Strains and Growth Conditions

Escherichia coli strains were routinely grown in LB medium. Plasmids were moved into *Bartonella* by conjugation using an *E. coli* donor strain with chromosomal RP4 machinery (Jonas Körner, unpublished). All bartonellae were routinely grown for 3–7 days on nutrient agar containing 5% defibrinated sheep blood in a water-saturated atmosphere with 5% CO_2 at 35 °C. Most species and isolates were grown on heart infusion agar (Oxoid), but *B. birtlesii* and *B. taylorii* were cultured on Columbia agar (Oxoid) and the two isolates from Sika deer were cultured on tryptic soy agar (Oxoid). Derivatives of pCD366 were selected with kanamycin ad 50 μg ml^{-1} (*E. coli*) or ad 30 μg ml^{-1} (*Bartonella*), and the *rpsL* genotype of the *B. henselae* Houston-1 lab strain RSE247 was selected with streptomycin ad 100 μg ml^{-1}.

Bartonella Isolates and Genome Sequences

Bartonella organisms were from our laboratory collection or were obtained from different sources as listed in table 1. *Bartonella taylorii* IBS296 isolated from a vole (*Microtus* sp.) was obtained from Dr. Yves Piemont and Dr. Remy Heller. *Bartonella* sp. CDCskunk was isolated from a stripped skunk (*Mephitis mephitis*) at the Janos Biosphere Reserve, Mexico. Genome sequences of relevant *Bartonella* species and other organisms were downloaded from the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov; last accessed March 9, 2017). The genome sequences of additional *Bartonella* isolates were determined as described below. Comprehensive information regarding all genomes

used in this study is assembled in supplementary table S1, Supplementary Material online.

Plasmid Construction

Plasmids were constructed using standard restriction-based procedures. Derivatives of pCD366 (an RSF1010 derivative encoding promoterless *gfpmut2* downstream of a multiple cloning site (Seubert et al. 2003) were used for P*virB2::gfpmut2* promoter probes as described previously and designated the pAH196 series (Quebatte et al. 2010). All oligonucleotide primers are listed in supplementary table S2, Supplementary Material online. All vectors and details of their construction are listed in supplementary table S3, Supplementary Material online.

Genome Sequencing, Assembly, and Annotation

Genomic DNA was isolated from about 0.25 g of bacterial material using the Power Soil DNA isolation kit (Mobio). SMRT sequencing (Pacific Biosciences) and assembly with HGAP (Chin et al. 2013) were carried out at the Yale Center for Genome Analysis. For all newly sequenced genomes, we obtained 1–3 contigs representing the entire chromosome of a given species. In case of complete genomes (i.e., one contig with overlapping ends), the origin of replication was set at the change in sign of the GC skew and annotations were subsequently carried out using the Integrated Microbial Genomes (IMG) platform (Markowitz et al. 2014). The genome of *B. taylorii* IBS296 was sequenced separately (see below). Beps were identified and manually annotated in all genomes based on the presence of a BID domain secretion signal as described below. Frameshifts were identified in coding regions using the Microbial Genome Submission Check tool on NBCI (https://www.ncbi.nlm.nih.gov/genomes/frameshifts/frameshifts.cgi; last accessed March 9, 2017). Frameshifts close to homopolymeric sequence stretches were corrected manually by inserting or deleting a single base in the homopolymeric stretch. These stretches are prone to SMRT sequencing errors and their corrections resulted in all cases in a single continuous ORF. Inconsistencies identified in the previously published genome of *B. ancashensis* (Hang et al. 2015) were confirmed by resequencing the corresponding locus using the Microsynth Barcode Economy Run service.

Sequencing of *Bartonella Taylorii* IBS296

Bartonella taylorii IBS296 was sequenced at the Quantitative Genomics Facility of the Department of Biosystems Science and Engineering (D-BSSE) of the ETH Zürich in Basel, Switzerland. Two read sets were generated from a 50 bp single-read run on a HiSeq2000 and a 150 bp paired-end run on a MiSeq sequencer (Illumina), achieving an average coverage above 400×. Mapping against the genome sequence of the *B. taylorii* 8TBB reference strain using Bowtie2 (Langmead and Salzberg 2012) was complemented by a *de novo* assembly of mis and unmapped reads using SOAPdenovo2 (Luo et al. 2012).

Identification and Protein Domain Annotation of *Bartonella* Effector Proteins

Bep and *biaA* genes in the genomes of all organisms shown in figure 1 were systematically identified *in silico* using BLAST implemented in Geneious v9.1.5 with amino acid query sequences against the translation of a local genome sequence database in all six reading frames (tBLASTn). Beps were identified with the bipartite secretion signal (terminal BID domain and C-terminal tail) of all effectors of *B. henselae* (L4) and *B.* sp. 1-1c (L3) where the effectors had been annotated previously (Engel et al. 2011; Schulein et al. 2005). BiaA orthologs were identified analogously with BLAST searches using the homologous FicA antitoxins VbhA of *B. schoenbuchensis* and YeFicA of *Y. enterocolitica* as query sequences (Harms et al. 2015).

FIC domains were manually annotated in all effectors using the protein analysis and classification tool of InterPro (http://www.ebi.ac.uk/interpro/; last accessed March 9, 2017) implemented in Geneious v9.1.5. BID domains of all effectors were manually annotated based on BLAST searches with all BID domains (including internal ones) of the effectors of *B. henselae* (L4) and *B.* sp. 1-1c (L3). Eukaryotic-like tyrosine phosphorylation motifs were identified and annotated in all effectors using ScanSite3 (http://scansite3.mit.edu/; last accessed March 9, 2017) with medium stringency and 0.25% fidelity cutoff. Sequence logos were determined using WebLogo (http://weblogo.berkeley.edu/; last accessed March 9, 2017).

Phylogenetic Analyses of *Bartonella* Species

The maximum likelihood phylogeny of figure 1 was created from 29 bartonellae and 5 outgroup species based on a concatenate alignment of the nucleotide sequence (fig. 1) or amino acid sequence (supplementary fig. S1*B*, Supplementary Material online) of 509 core gene orthologs. Orthologous gene families were identified using the OrthoFinder software (Emms and Kelly 2015). From the resulting 4981 protein families (excluding singletons), we selected families which contained exactly one gene from each genome in our analysis. This resulted in 494 core genes for tree construction, to which we added 14 ribosomal protein genes that are duplicated in *B. bacilliformis*, and the housekeeping gene *groEL* which was clustered by OrthoFinder with multiple orthologs for some outgroup species. For each of these additional genes, we used the homolog that was most similar to its counterpart in the other strains. We aligned the protein sequences with ClustalW2 (Larkin et al. 2007). The nucleotide sequences were aligned in codons according to the protein alignments with an in-house Python script. From all alignments, we removed the columns with gaps in over 50% of the sequences with trimAl v1.2 (Capella-Gutierrez et al. 2009).

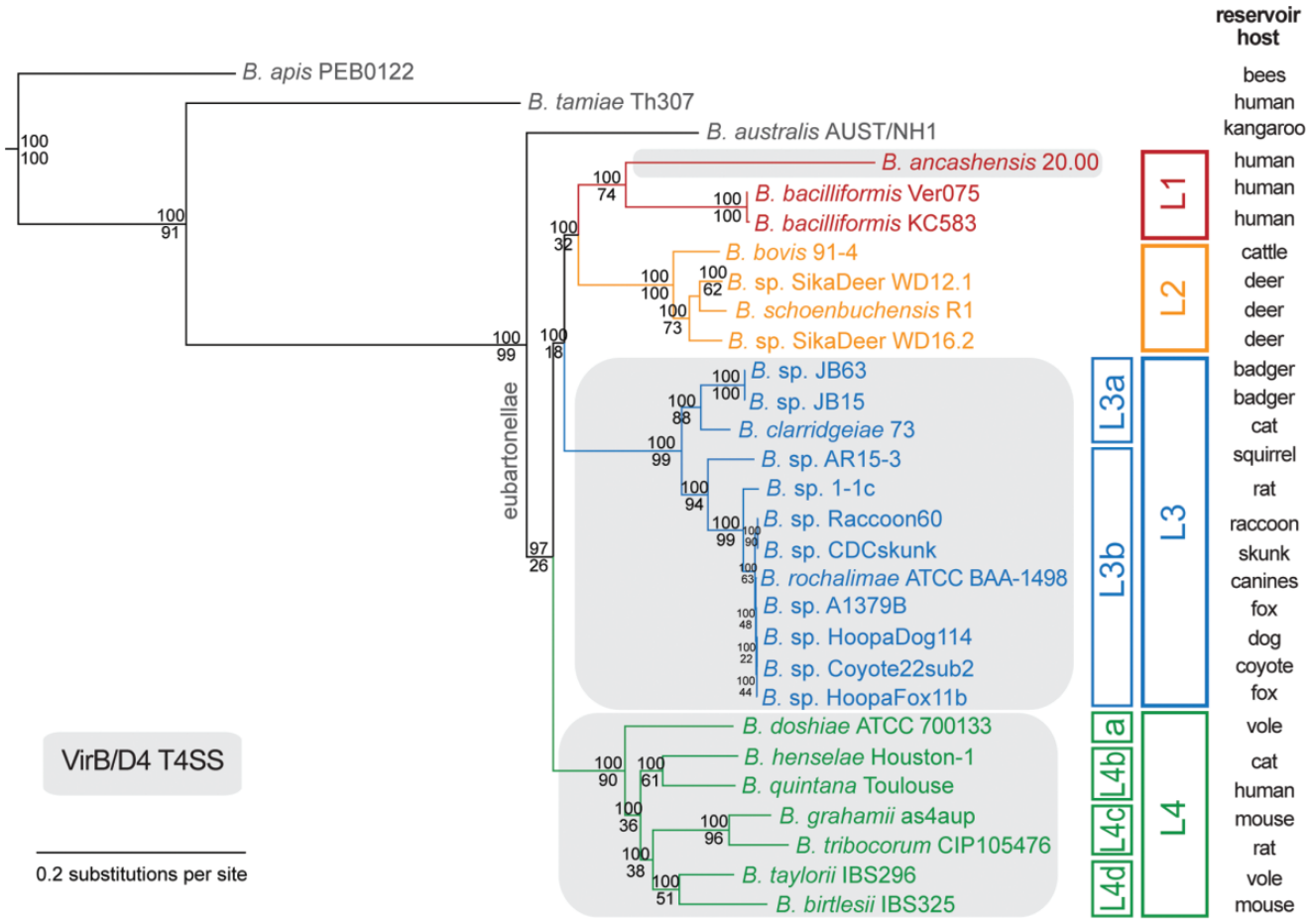

Fig. 1.—Comprehensive phylogeny of the genus *Bartonella*. Maximum likelihood phylogeny of *Bartonella* based on the nucleotide sequence alignment of 509 concatenated core genes. Branch labels show bootstrap support (top; 100 replicates) and % support by single gene phylogenies (bottom). Lineages encoding a VirB/D4 T4SS are highlighted by grey shading. Reservoir hosts are shown on the right-hand side of the illustration (simplified; for details, see table 1 and supplementary table S1, Supplementary Material online). A representation of the phylogeny with full outgroup as well as an analogous phylogeny based on concatenated protein sequence are shown in supplementary figures S1*A* and S1*B*, Supplementary Material online.

The best model of evolution for each gene was predicted with Prottest-3.4 (Darriba et al. 2011) and MEGA v6.0 (Tamura et al. 2013) for protein and nucleotide sequences, respectively. Invariable site models were not taken into consideration as recommended by the author of RAxML (Stamatakis 2014). We inferred single gene trees for all protein and DNA sequence alignments with RAxML v8.0.0. The trimmed protein and DNA alignments were concatenated with an in-house Python script. Next, genome-wide trees were inferred with RAxML v8.0.0 using the model of evolution that overall was the best model of evolution for most single gene alignments (GTR + G for DNA; JTT + G for protein). To assess the robustness of the phylogenies, we did 100 bootstrap replicates and counted the number of single gene trees that were congruent with the genome-wide phylogenies at each branch using an in-house Python script and the package APE v3.4 (Paradis et al. 2004) in R v3.2.0.

Whole-Genome Alignments

Whole genome alignments were generated as pairwise tBLASTx files using the DoubeACT tool (http://www.hpa-bioinfotools.org.uk; last accessed March 9, 2017) and visualized using the genoPlotR package in R v3.2.0 (Guy et al. 2010).

Phylogenetic Analyses of VirB/D4 T4SS and Beps

For phylogenies of *Bartonella* effector proteins, Bep sequences of all species shown in figure 1 were aligned using MAFFT implemented in Geneious v9.1.5. This alignment was manually curated to remove regions of largely gapped or obviously nonhomologous sequence. The optimal model for maximum likelihood analysis was determined using ProtTest3 and found to be JTT in all cases (Darriba et al. 2011). Maximum likelihood phylogenies were constructed from the alignments using

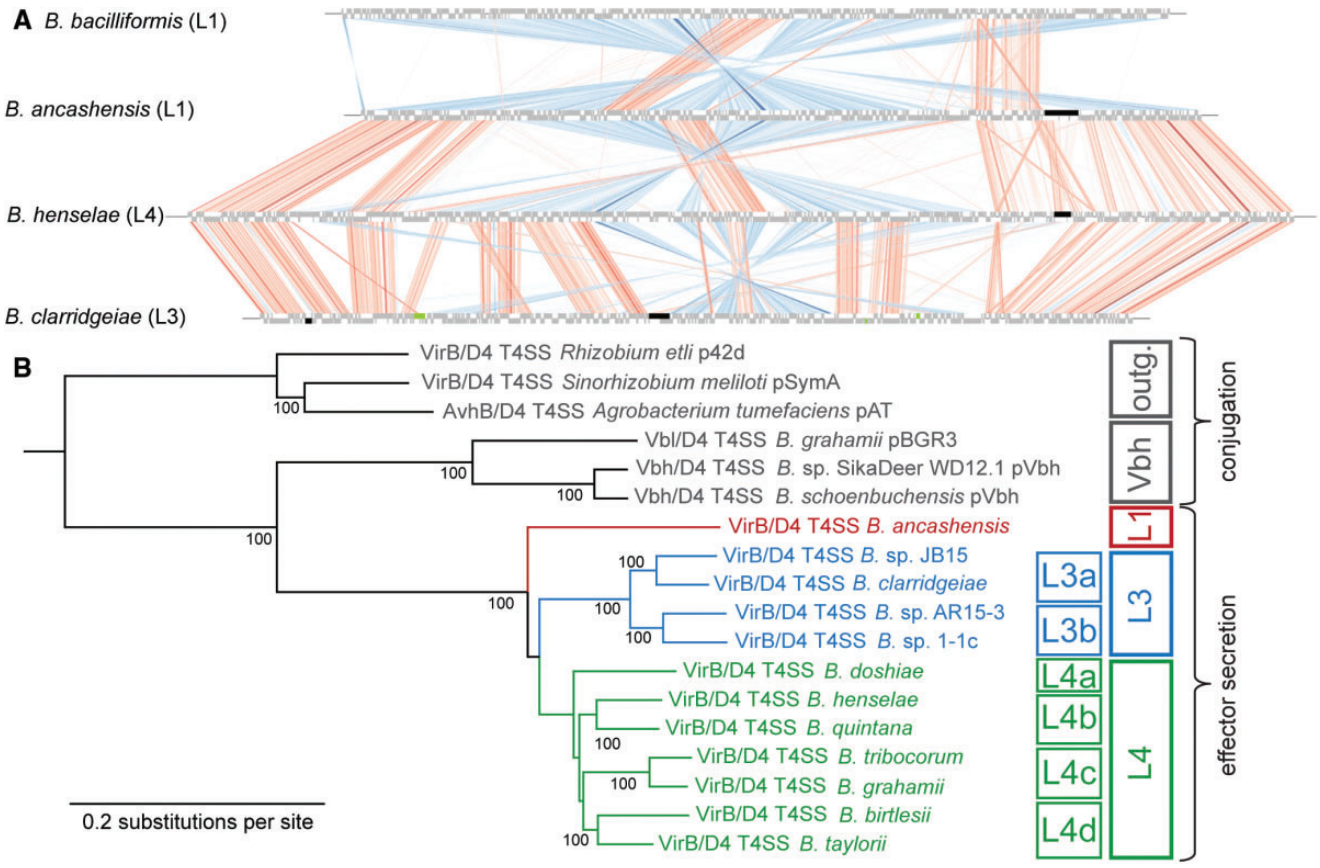

Fig. 2.—The VirB/D4 T4SS of *B. ancashensis* was acquired independently from its homologs in L3 and L4. (*A*) Whole-genome alignment of *B. bacilliformis* KC583 (lacking any VirB/D4 T4SS), *B. ancashensis*, *B. clarridgeiae* (representing L3), and *B. henselae* (representing L4) with grey bands indicating tblastx hits. *virB/bep* and *bep* loci are colored in black and green, respectively. Although the *virB/D4/bep* locus of *B. ancashensis* is clearly at a distinct position from any related locus in L3, it is found close to the *virB/D4/bep* locus of L4 at a different position of the secretion system cassette (SSC) region that encodes various horizontally acquired virulence factors (Guy et al. 2013) (more details shown in supplementary fig.S2, Supplementary Information online). (*B*) Maximum likelihood phylogeny of the *Bartonella* VirB/D4 T4SS and related machineries based on concatenated protein alignments of all components (with rhizobial conjugation systems as outgroup). Branch numbers represent bootstrap support when >80 (of 100 replicates).

PhyML implemented in Geneious v9.1.5 with standard settings and 100 bootstraps. Outgroups were chosen as indicated either from effectors of a different phylogenetic lineages or from homologous regions of conjugative relaxases. The VirB/D4 T4SS phylogeny of figure 2*B* was generated similarly from a concatenate of alignments of all single components of the machinery.

Flow Cytometric Analysis of Gene Expression

The expression of GFP promoter fusions in *B. henselae* carrying pCD366-derived reporter plasmids was probed using flow cytometry as described previously (Quebatte et al. 2010; Quebatte et al. 2013). In brief, bacteria were grown on Columbia blood agar plates at 35 °C and 5% CO_2 for 2 days. The bacteria were resuspended in M199 (Gibco) supplemented with 10% fetal calf serum (Amimed) to a final OD_{600} of 0.008 and incubated in 48-well plates in a humidified atmosphere at 35 °C and 5% CO_2. Expression of the *gfpmut2* reporter was measured as GFP fluorescence of ca. 25,000 cells using a FACSCalibur flow cytometer (BD Biosciences) with an excitation at 488 nm. Data analysis was performed using the FlowJo software (Flowjo LLC).

Results

Comprehensive Phylogeny of *Bartonella* Suggests Three Separate Instances of VirB/D4 T4SS Acquisition

We determined the genome sequences of eight new isolates of L3 and resequenced a number of additional *Bartonella* species (table 1; see also supplementary table S1, Supplementary Material online, for a comprehensive list of bartonellae used in this study). Our assemblies yielded for many genomes a single, circular chromosome with a GC skew typical for bacteria. Genome sizes and GC content were in the range of previously

sequenced *Bartonella* species (Engel et al. 2011; Guy et al. 2013; Saenz et al. 2007). Based on 509 core genes of these genomes and the genomes of additional bartonellae as well as five rhizobial outgroup species, we constructed a robust maximum-likelihood phylogeny (fig. 1 and supplementary fig. S1, Supplementary Material online). The topology of this phylogeny is largely similar to previous phylogenies based on smaller sets of taxa (Guy et al. 2013; Segers et al. 2017; Zhu et al. 2014). Generally, the phylogeny shows that the genus *Bartonella* comprises two deeply rooting lineages (*B. apis* and *B. tamiae*) and the radiating lineage of the eubartonellae (Segers et al. 2017; Zhu et al. 2014). The eubartonellae are split into the previously described four phylogenetic lineages L1, L2, L3, and L4, and *Bartonella australis*. Monophyly of these major phylogenetic groups is well supported by the single gene phylogenies, suggesting that horizontal gene transfer and recombination did not greatly affect the topology of the phylogeny. *B. australis* was placed as the earliest diverging lineage of the eubartonellae in accordance with a previous analysis (Guy et al. 2012) and thus cannot be assigned to any of the lineages L1-4. Within the eubartonellae, L3 shows a bifurcated organization with one sublineage being formed by *B. clarridgeiae* and two newly sequenced isolates from Japanese badgers (*Meles anakuma*) and a second sublineage consisting of *B.* sp. AR15-3, *B.* sp 1.-1-c, *B. rochalimae*, and diverse newly sequenced isolates that are closely related to *B. rochalimae*. For simplicity, we named these sublineages L3a and L3b, respectively (fig. 1).

The phylogeny confirms that *Bartonella* L4 is composed of three discernible sublineages plus *B. doshiae* that occupies a more basal position (fig. 1) (Guy et al. 2012; Zhu et al. 2014). Similar to the sublineages of L3, we classified the four subgroups of L4 as L4a-d. As expected from previous phylogenetic analyses (Mullins et al. 2013), *B. ancashensis* is robustly placed within L1, but sits on a remarkably long branch that indicates considerable evolutionary distance to its sister species *B. bacilliformis*. Furthermore, the phylogeny confirms that L1 and L2 form a monophyletic clade as proposed by Guy et al. (2013). None of the sequenced strains of these two lineages encode a VirB/D4 T4SS except for *B. ancashensis*. When this machinery was acquired in L1 and which sets of effector proteins are encoded in *B. ancashensis* has so far remained elusive.

Three Separate VirB/D4 T4SS Machineries in *Bartonella* Form a Monophyletic Group Closely Related to the Vbh/D4 T4SS

A comparison of the chromosomal loci encoding the VirB/D4 T4SS of L3, L4, and *B. ancashensis* shows that these machineries are encoded at distinct positions in the genome (fig. 2*A* and supplementary fig. S2, Supplementary Material online). We therefore hypothesized that the VirB/D4 T4SS of *B. ancashensis* may have a distinct evolutionary history than the homologous machineries present in L3 and L4. In order to test this hypothesis, we generated a phylogeny based on all proteins that form the VirB/D4 T4SS and also included closely related machineries mediating bacterial conjugation (Saenz et al. 2007; Schulein et al. 2005) (fig. 2B). This tree shows that the VirB/D4 T4SS machineries of *Bartonella* form a monophyletic clade that is the sister group of the conjugative Vbh/D4 T4SS of *Bartonella* (Saenz et al. 2007; Schulein et al. 2005). The tree also shows that the VirB/D4 T4SS of *B. ancashensis* forms a separate deep-branching lineage, which supports the hypothesis that this system is evolutionary distant from the VirB T4SS systems in L3 and L4. Interestingly, the topology of the VirB/D4 T4SS tree for L3 and L4 is congruent with the species phylogeny (fig. 1), suggesting that these machineries coevolved with the host genome after their original acquisition.

Three Bep Repertoires Evolved by Parallel Series of Gene Duplication and Diversification from a Single Common Ancestor

We next studied the secreted effectors of the three *Bartonella* VirB/D4 T4SS with a particular focus on the question if the effector genes of *B. ancashensis* had evolved independently, as in the case of L3 and L4, or if they are direct orthologs of one of the other two lineages. For this purpose, we identified all Beps of *B. ancashensis*, the newly sequenced genomes of L3, and the other bartonellae shown in figure 1 (see details in Materials and Methods). *B. ancashensis* encodes at least eighteen *bona fide* effectors that were numbered according to the index of the *locus_tag* of the effector genes in the previously published genome sequence of this organism (Hang et al. 2015). Direct comparison of the genetic organization of *virB/D4/bep* genes in the three lineages (including all newly sequenced organisms) revealed that the *virB/D4* and *bep* genes are encoded in a single genomic island in *B. ancashensis* similar to the loci of L4, but different from the more dispersed organization that is found in L3 genomes (fig. 3*A*) (Engel et al. 2011).

A maximum likelihood phylogeny based on the conserved BID domain secretion signal of the Beps was generated to study the evolutionary relationship of the three effector sets with related sequences found in conjugative relaxases as outgroup (fig. 3*B* and supplementary fig. S3*A*, Supplementary Material online). The tree topology shows that the Beps of *B. ancashensis* form a monophyletic sister clade of the effectors of L3 and L4. Although this topology mirrors the phylogeny of the VirB/D4 T4SS machinery, it is not supported by the bootstrap analysis (i.e., values < 80), likely due to the high degree of sequence divergence within *Bartonella* effectors and compared with conjugative relaxases. Therefore, we additionally generated an unrooted phylogeny of all Beps excluding the distantly related relaxase sequences. This tree supports the monophyly of *B. ancashensis* effectors (fig. 3C and

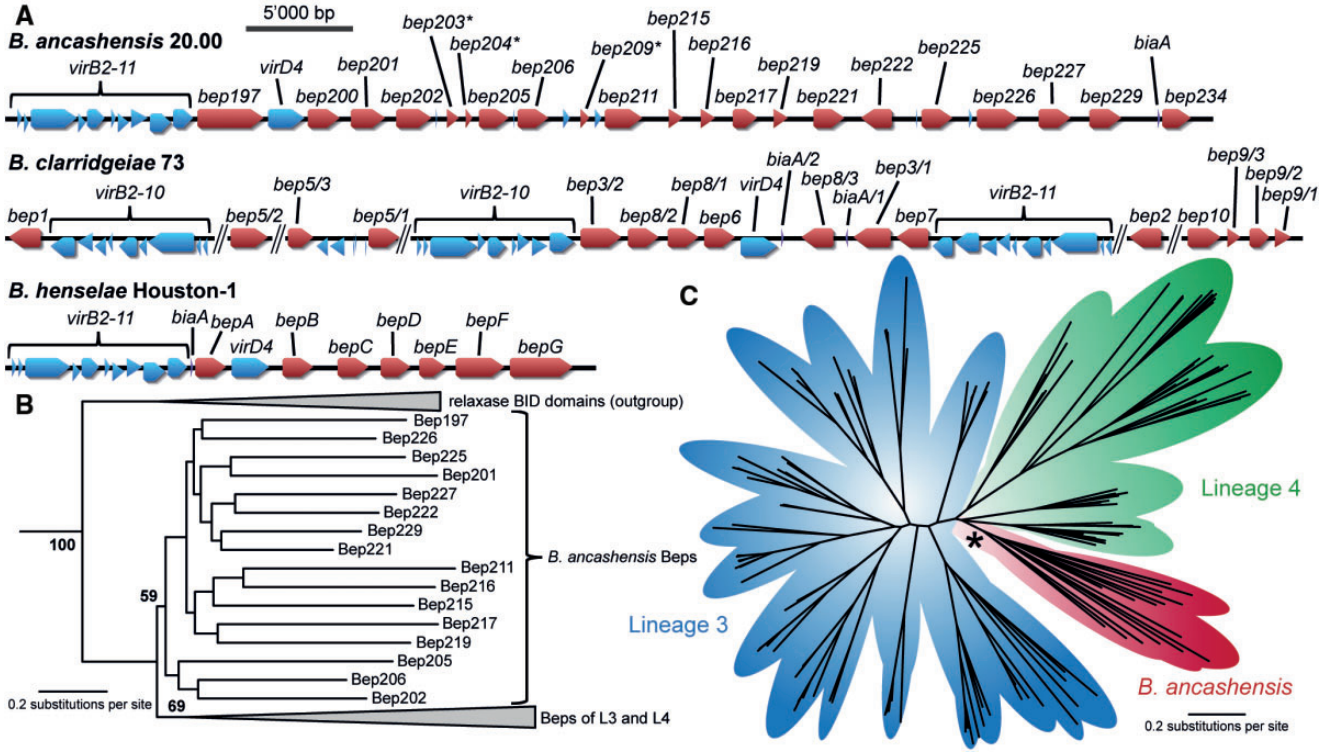

Fig. 3.—The phylogeny of *Bartonella* effectors reveals that the repertoire of *B. ancashensis* evolved independently. (*A*) Genetic architecture of *virB/D4/bep* loci of *B. ancashensis*, *B. clarridgeiae* (L3), and *B. henselae* (L4). Very short effector-like genes of *B.ancashensis* that may be gene fragments are marked with an asterisk. (*B*) Maximum likelihood phylogeny of Beps encoded by the organisms shown in figure 1 based on their terminal BID domain and C-terminal region and including the homologous part of rhizobial relaxases as outgroup to root the tree. Bootstrap values (of 100 replicates) are shown for the clade formed by all Beps as well as the monophyletic groups formed by effectors of *B. ancashensis* or L3/L4 species. (*C*) Maximum likelihood phylogeny created similar to (*B*), but without outgroup and thus based on a longer sequence alignment. The phylogeny strongly supports the monophyly of *B. ancashensis* Beps (89/100 bootstraps; highlighted with an asterisk). The full phylogenies of (*B*) and (*C*) with all labels are shown as supplementary figures S3*A* and S3*B*, Supplementary Information online.

supplementary fig. S3*B*, Supplementary Material online) with bootstrap values >80. These results imply that – beyond the two well-described lines of effector evolution in L3 and L4—a third Bep repertoire evolved independently from a single common ancestor linked to the VirB/D4 T4SS of *B. ancashensis*.

More targeted phylogenies of L3 and L4 Beps enabled us to study the phylogenetic positions of the effectors encoded in the newly sequenced genomes (supplementary figs. S4 and S5, Supplementary Material online). Most Beps of the new L3 isolates fell within one of the ortholog groups that had been previously identified within this lineage (Engel et al. 2011). However, our enhanced resolution of the internal diversity of L3 also revealed new aspects. Previously, *B. clarridgeiae* had been the only representative of a deep-branching sublineage of L3, and some differences between its effector repertoire and the Beps of other L3 species were apparent but could not be interpreted with certainty (Engel et al. 2011). Our new genome sequences both of L3a and L3b show that there are systematic differences between the effector repertoires of the two sublineages of L3: whereas L3b bartonellae encode each one ortholog of Bep4 and typically two closely related effectors classified as Bep7, the species of L3a encode no Bep4 and only one Bep7 (fig. 4*A* and supplementary fig. S4, Supplementary Material online). Instead, they feature two additional paralogs of Bep8 (classified as Bep8/2 and Bep8/3) and one additional paralog of Bep3 (classified as Bep3/2). Furthermore, there are differences regarding the genetic organization of Bep9 (see below). Conceptually similar variations in the effector repertoire have long been known for L4 where, e.g., well-studied *B. henselae* encodes BepG, a unique effector composed of four BID domains (see below), while this organism and closely related species in L4b lack orthologs of the otherwise ubiquitous FIC–BID effector BepI (supplementary fig. S5, Supplementary Material online) (Engel et al. 2011; Saenz et al. 2007). Our observations strongly suggest that continuous pathoadaptation is driving the diversification of effector repertoires of L3 as well as L4.

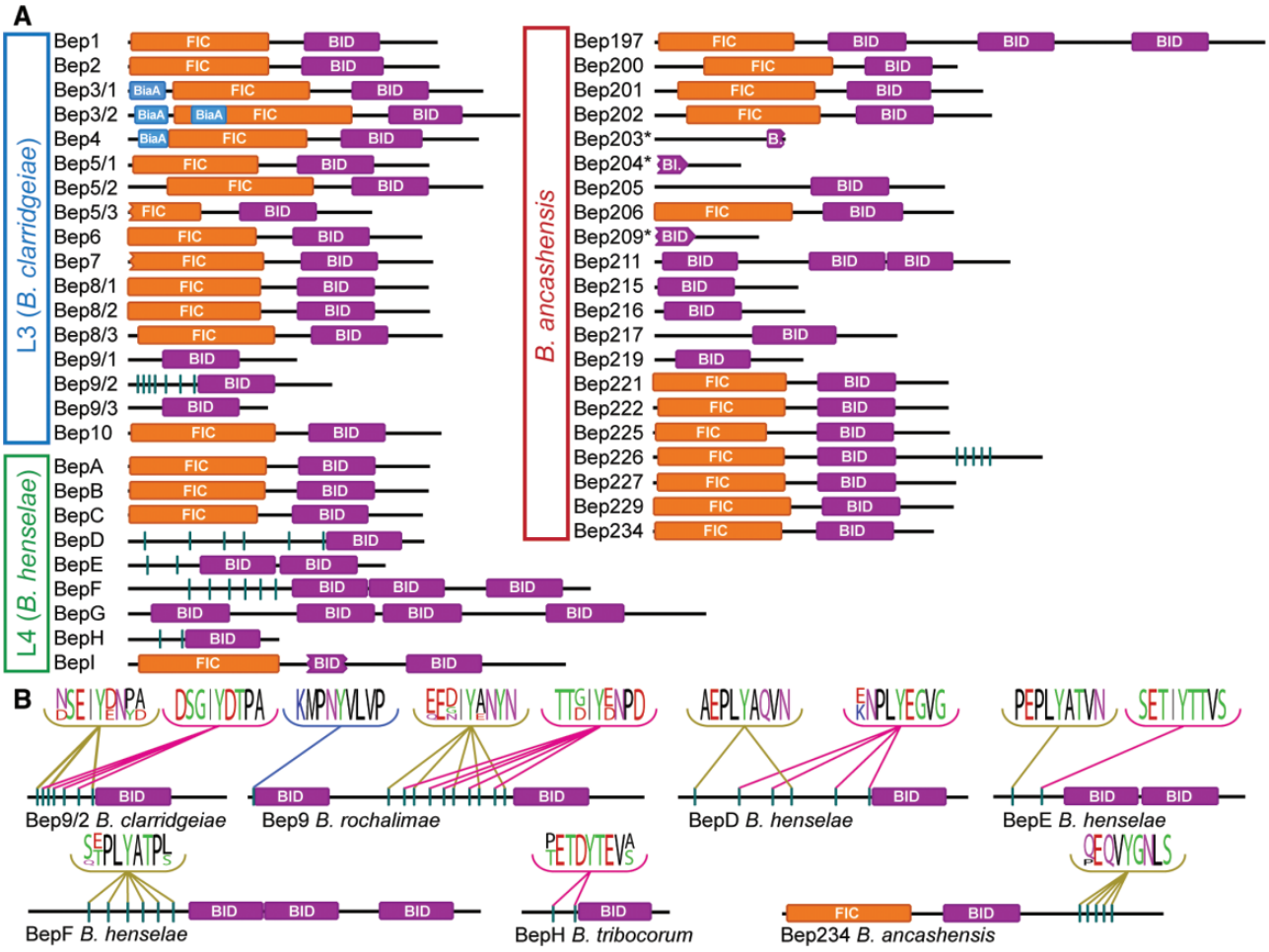

Fig. 4.—Domain architectures and tyrosine phosphorylation motifs of the three Bep repertoires. (*A*) The domain architecture of all orthologous groups of Beps found in *B. ancashensis*, L3, and L4 is shown in comparison. FIC domains are displayed in orange, BID domains are displayed in purple, BiaA-like modules are shown in blue. Predicted tyrosine phosphorylation motifs are shown as vertical lines in cyan. The illustrations are based on representative organisms (*B. clarridgeiae* for L3 and *B. henselae* for L4), and effectors absent in these species are shown from *B. rochalimae* (Bep4; L3) of *B. tribocorum* (BepH and BepI; L4). (*B*) The consensus sequences of tyrosine phosphorylation motifs are shown for Bep226 and representative effectors of L3 (*B. clarridgeiae*/*B. rochalimae*) and L4 (*B. henselae*/*B. tribocorum*).

Protein Domain Architecture of *B. Ancashensis* Beps

A deeper analysis of the protein domain architecture of the Beps that we identified in the newly sequenced organisms of L3 and L4 confirmed that all isolates encode effector repertoires fully orthologous to Beps of bartonellae in the same phylogenetic sublineage but different from those of other sublineages or lineages (supplementary figs. S4 and S5, Supplementary Material online). Apart from the minor variations described above, all species of L3 encode a set of around ten orthologous Beps with FIC-BID domain architecture as it had been reported for the original four genomes of this lineage (Engel et al. 2011; Bep1-8 and Bep10; fig. 4*A* and supplementary fig. S4, Supplementary Material online). Additionally, the species of L3b encode a single, rather large Bep with eukaryotic-like tyrosine phosphorylation motifs (Bep9), whereas three smaller effectors (one with tyrosine phosphorylation motifs) harboring closely homologous BID domains are found at the same locus in L3a (classified as Bep9/1-3; fig. 4*A* and supplementary fig. S4, Supplementary Material online).

Like most Beps encoded by species of L3 and L4, the far majority of Beps of *B. ancashensis* display the characteristic and ancestral FIC-BID domain architecture of the Beps (fig. 4*A*) (Engel et al. 2011). Furthermore, the effector repertoire contains several effectors consisting only of one or more BID domains in a way similar to, e.g., BepG of L4. Additionally, one effector (Bep197) contains three BID domains following an N-terminal FIC domain, a domain organization not found in any other effector of *Bartonella*. The presence of these effectors

with several BID domains suggest that in *B. ancashensis*, just like it had been shown for BepA or BepE of *B. henselae* (Okujava et al. 2014; Pulliainen et al. 2012), BID domains have secondarily evolved functions in host interaction on top of their role as part of the Beps' secretion signal.

Tyrosine Phosphorylation Motifs in Beps of *B. Ancashensis*

One major discovery in the effector repertoires of L3 and L4 was the independent evolution of effectors containing tandem arrays of eukaryotic-like tyrosine phosphorylation motifs that are modified by host kinases after translocation (Engel et al. 2011; Selbach et al. 2009). Upon phosphorylation, these Beps recruit diverse host proteins to form signaling platforms that reorganize host cell signaling in favor of the pathogen, similar to what has been shown for ancestrally unrelated effectors with tyrosine phosphorylation motifs of other pathogens (Hayashi et al. 2013; Selbach et al. 2009). Because these motifs have not been studied by direct experimentation in most effectors of *Bartonella*, we used ScanSite to predict them in all Beps (fig. 4; see details in Materials and Methods) (Obenauer et al. 2003). Apart from occasional single and weak hits, tyrosine phosphorylation motifs of Beps of L3 and L4 were exclusively found in orthologous groups of effectors that had already been known to harbor such motifs and that contain only BID domains as additional features (Beps D/E/F/H in L4 and Bep9 in L3; fig. 4) (Engel et al. 2011; Saenz et al. 2007). Interestingly, one single effector of *B. ancashensis*, Bep226, also displays an array of predicted tyrosine phosphorylation motifs, but these are found as an extension at the far C-terminus of an effector with otherwise regular FIC-BID domain architecture (fig. 4). The sequence logos of the predicted tyrosine phosphorylation motifs of Bep226 and the Beps of L3 or L4 are clearly distinct from each other (fig, 4*B*) (Engel et al. 2011). We therefore conclude that, similar to the Bep repertoires as a whole, the tyrosine phosphorylation motifs of *Bartonella* effectors evolved *de novo* at least three times, i.e., at least once in each of the three lineages L1, L3, and L4.

Analysis of Bep FIC Domains in *B. Ancashensis* Suggests Functional Diversification during Effector Evolution

Among the FIC domains of L3 and L4 Beps, only the orthologs of Bep1, Bep2, and Bep3 consistently display a canonical HPFX[D/E]GNGRXXR active site motif that confers AMPylation activity (figs. 5 and 6) of specific host target proteins. Such enzymatic activity has indeed been shown for Bep1 (Dietz *et al.*, in revision) and Bep2 (Pieles et al. 2014). Other orthologous groups of effectors share conserved, yet noncanonical Fic signature motives like, e.g., the large majority of Bep4 orthologs that have a cysteine at the position of the catalytic histidine of FIC domains (fig. 5*A*). We suggest that these noncanonical (but conserved) active site motifs may enable catalytic activities other than AMPylation as the conserved molecular function of these effectors. A third group of Bep FIC domains like those of BepA/B or Bep5 do not show consistent sequence motifs at the FIC domain active site, suggesting that these orthologs either have different molecular functions or may have no enzymatic activity at all and instead act, e.g., as target binders (fig. 5*A* and supplementary fig. S6, Supplementary Material online). Interestingly, the FIC domains of effectors of *B. ancashensis* mostly display a canonical active site motif, suggesting that they primarily contribute to pathogenesis by host protein AMPylation (fig. 5*B*).

We recently discovered that the target specificity of one *Bartonella* effector protein with AMPylation activity is largely determined by the residues at the tip of an elongated β-sheet (the "flap") (Dietz et al., in revision). Interestingly, the orthologs of Bep1 and Bep2 (AMPylating small GTPases and vimentin, respectively) each consistently display very similar motifs at the tip of the flap, whereas paralogous Beps generally have different motifs (fig. 5*B*) (Dietz et al., in revision) (Pieles et al. 2014). The Bep FIC domains of *B. ancashensis* show a wide variety of sequences at this position, suggesting that they AMPylate different host proteins and indicating that functional diversification and not gene dosage effects drive maintenance of a large effector repertoire in this organism (fig. 5*B*).

FicA Antitoxin Homologs (BiaA) Found at *Bep* Loci and Fused to Bep FIC Domains

It was previously shown that the enzymatic activities of FIC domains are regulated by a small inhibitory module that obstructs binding of the ATP substrate and that is encoded as part of a separate small protein in a group of FIC domain proteins to which the Beps belong (Engel et al. 2012; Harms et al. 2016). These small proteins were originally described as FicA antitoxins that control the AMPylation activity of the FicT toxins in the frame of a classical bacterial toxin-antitoxin module (Harms et al. 2015). Moreover, previous work had introduced a FicA homolog encoded at a *bep* locus as BiaA (Pieles et al. 2014). Surprisingly, our analyses showed that all bartonellae with VirB/D4 T4SS invariably encode one (or rarely two) copies of BiaA at a *virB/D4/bep* locus (fig. 6; see also fig. 3*A*). All these BiaA orthologs have conserved the glutamate that is characteristic of FicA orthologs and is essential for the inhibition of FIC domain activities, suggesting that they are functional and may regulate the activity of Bep FIC domains (fig. 6) (Engel et al. 2012). Furthermore, we discovered that all orthologs of Bep3 and Bep4 carry an N-terminal extension that is homologous to BiaA (figs. 4*A* and 7; see also supplementary fig. S7, Supplementary Material online). It therefore appears that the FIC domain activity of Beps is tightly controlled by several layers of regulation.

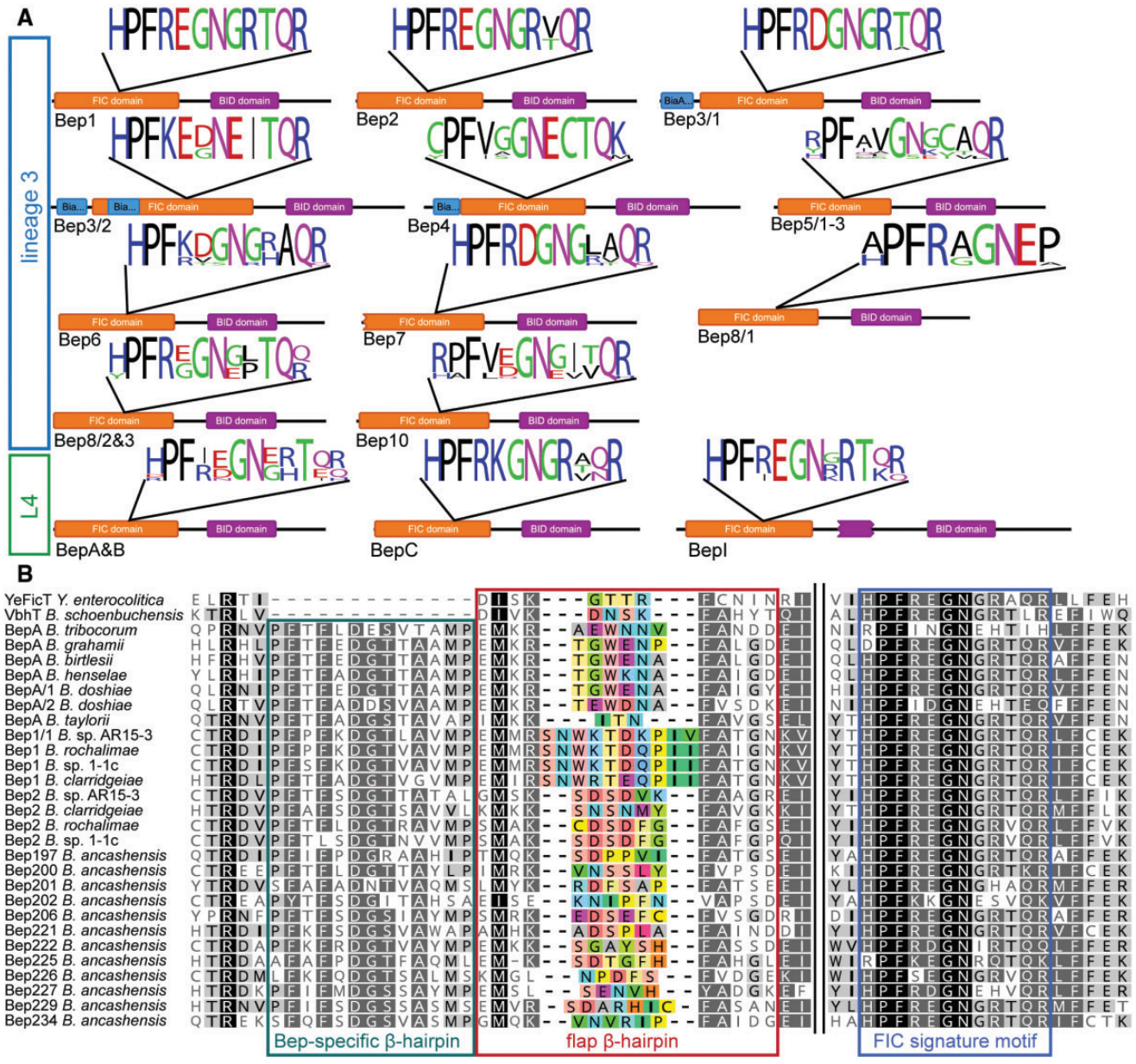

Fig. 5.—FIC domains of *B. ancashensis* Beps show strong signs of functional diversification. (*A*) The sequence logos of FIC signature motifs are shown for the different orthologous groups of Beps of L3 and L4 (see all signature motifs in supplementary fig. S6, Supplementary Information online). (*B*) Alignment of *B. ancashensis* Bep FIC domains with representatives of L3 and L4 as well as FicT toxins. The variable region around the tip of the flap is highlighted with a red frame and amino acid coloring. The position of the Fic signature motif at the active site is marked with a blue frame. Notably, all Beps share a short β-hairpin right ahead of the β-sheet that is not part of the FIC domain core and absent from FicT toxins (cyan) (Palanivelu et al. 2010).

Regulation of All Three Instances of the VirB/D4 T4SS by the BatRS Two-Component System

Previous work in the *B. henselae* model has shown that expression of the *virB* operon is driven from P*virB2*. This promoter is controlled by the BatR/BatS two-component system, the *Bartonella* ortholog of a conserved rhizobial regulator of host interactions (Quebatte et al. 2010). We therefore suspected that all instances of the VirB/D4 T4SS in *Bartonella* may be controlled by the BatR/BatS two-component system, despite the fact that the chromosomal integration of the secretion machinery and the evolution of the effector genes seem to have occurred independently. Strikingly, sequences similar to the known binding site of BatR in P*virB2* of *B. henselae* were found in the P*virB2* promoters of all strains (fig. 8*A*). We therefore tested the induction of P*virB2::gpfmut2* promoter fusions in a host-like

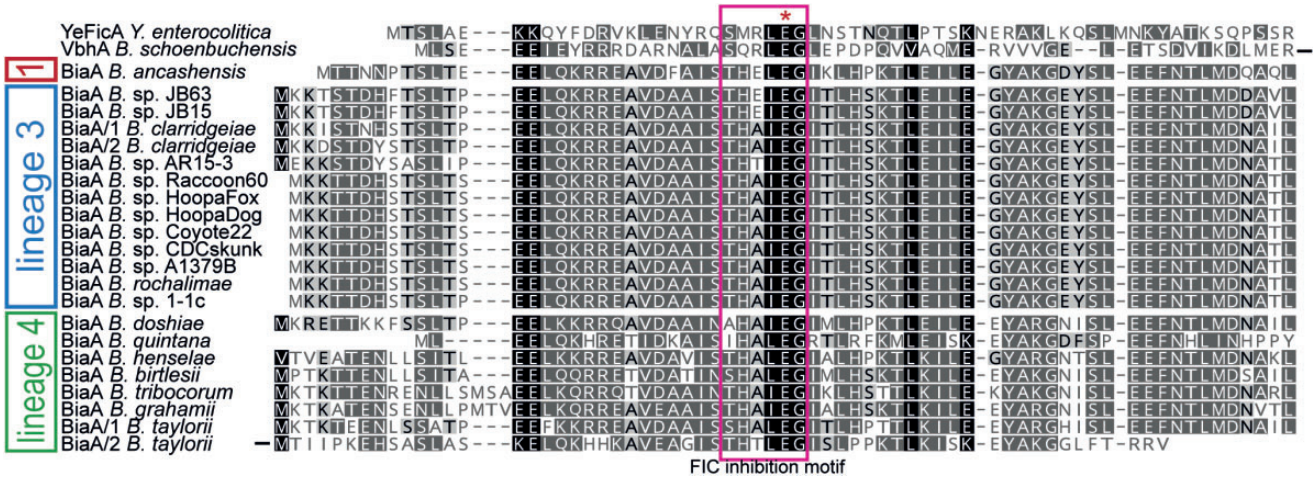

Fig. 6.—FicA antitoxin homologs (BiaA) found at *bep* loci in all bartonellae. Alignment of BiaA orthologs with the antitoxins VbhA of *B. schoenbuchensis* and YeFicA of *Yersinia enterocolitica*, The FIC inhibition motif [S/T]xxxE[G/N] is highlighted by a purple box. Notably, the glutamate essential for the regulation of FIC domain functioning is conserved in all orthologs (marked by an asterisk).

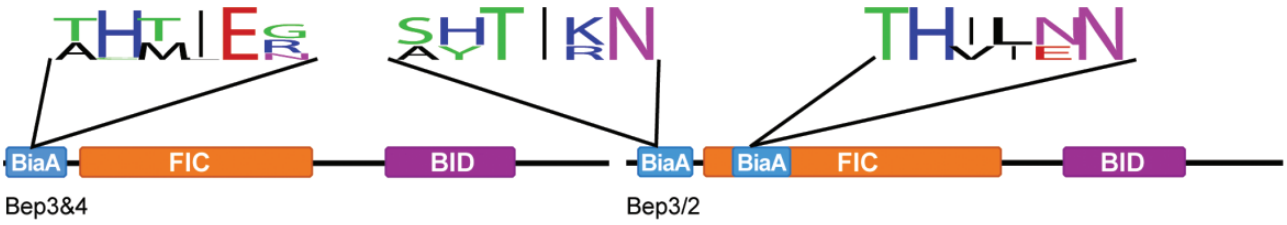

Fig. 7.—BiaA-like modules of Bep3 and Bep4 orthologs. BiaA modules were found in the N-terminal regions of all orthologs of Bep3 and Bep4 (see also fig. 4*A*). Note that the Bep3/2 effector, a basal member of the Bep3 clade (see fig. 4 and supplementary fig. S4, Supplementary Information online) has even two such modules, one at the very N-terminus and one inserted into the FIC domain. The consensus sequence of FIC domain inhibition motifs is shown for relevant orthologous groups. The inhibitory glutamate is generally conserved, but replaced by a positively charged residue in Bep3/2.

environment and found that (i) all instances of the *Bartonella* VirB/D4 T4SS are induced in this experimental setup and that (ii) this induction was invariably dependent on BatR (fig. 8*B*) (Quebatte et al. 2010). Conversely, no change in expression or any BatR-dependency was found for the P*vbhB2* promoter of the conjugative Vbh/D4 T4SS of *B. schoenbuchensis* (fig. 8*B*). Interestingly, the whole population of bacteria induced P*virB2* of L3 or *B. ancashensis* in this host-like environment, but—as shown previously for *B. henselae* (Quebatte et al. 2010)—only part of the population responded for P*virB2* of all sublineages of L4. It therefore seems that this lineage has acquired an additional level of regulation that controls expression of the VirB/D4 T4SS, possibly involving signaling through the second messenger (p)ppGpp as shown previously for P*virB2* expression of *B. henselae* (Quebatte et al. 2013).

Discussion

In this work, we presented several lines of evidence that the VirB/D4 T4SS of *B. ancashensis* and its Beps are the product of a third, independent series of chromosomal integration and effector gene amplification/diversification events on top of the two that had been described previously in L3 and L4 (figs. 1–5, supplementary figs. S1–S5, Supplementary Material online; Engel et al. 2011). It is striking that the three instances of VirB/D4 T4SS and Beps exhibit many common features including details of effector evolution (like the emergence of tyrosine phosphorylation motifs) or regulation by the BatR/BatS two-component system (figs. 4 and 8). Although the emergence of similar effector sets is likely the result of parallel evolution (Engel et al. 2011), it is possible that regulation by the BatR/BatS two-component system may have already been a feature of the ancestral VirB/D4 T4SS that harbored a single primordial effector gene.

With the sequencing of eight additional strains from L3, our study sheds light on the evolution and diversity of Beps in this group of *Bartonella* spp. Despite the fact that the sequenced strains were isolated from different host species in different countries, we only found little sequence divergence compared with the previously sequenced genomes. Although we cannot

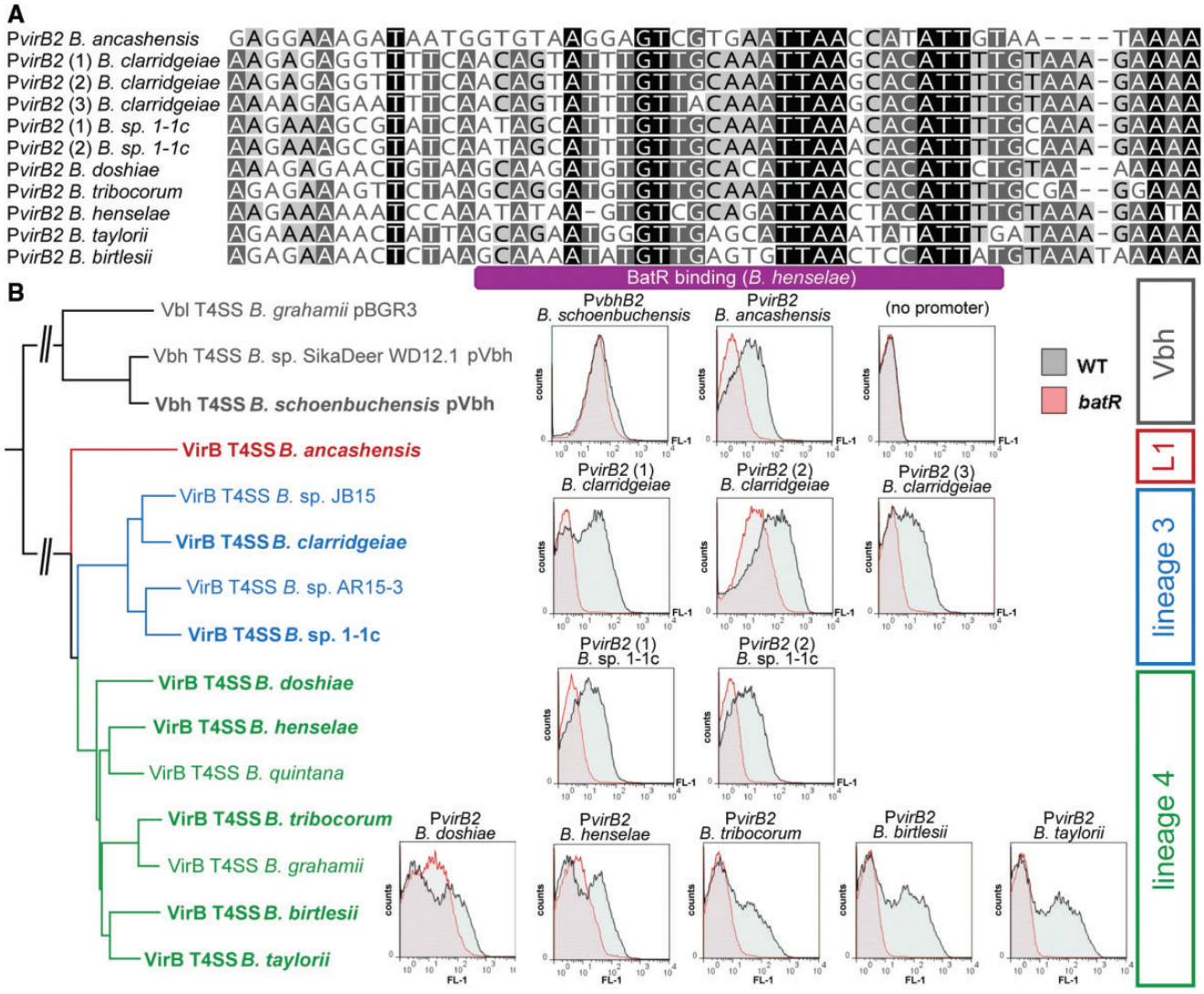

Fig. 8.—The BatR/BatS two-component system controls all three instances of the VirB/D4 T4SS. (*A*) The alignment of *virB2* upstream regions (around −300 bp) of *B. ancashensis*, L3, and L4 reveals similarities to the known BatR binding site in P*virB2* of *B. henselae* (−306 bp to − 277 bp [Quebatte et al. 2010]). (*B*) The BatR-dependent induction of P*virB2* of all *Bartonella* lineages and sublineages (organisms highlighted in bold in the tree adapted from fig. 2*B*) was probed using flow cytometry after incubation in a setup mimicking the host environment (see Materials and Methods).

fully rule out sampling biases, these results suggest that the genetic diversity within this lineage is limited compared with L4. Regarding the Bep repertoires, we found that the different orthologous groups of FIC domain containing effectors display extensive diversification of the active site motifs and conspicuous sequence variability at a site involved in target recognition (fig. 5 and supplementary fig. S6, Supplementary Material online). It seems therefore likely that different orthologous groups of Beps generally display different biological functions and sometimes distinct molecular activities, just as it has been worked out for several Beps of *B. henselae* (Siamer and Dehio 2015).

Interestingly, most effectors of *B. ancashensis* have conserved a canonical Fic signature motif at the active site, indicating that they may still largely act as AMP-transferases but modify diverse host targets (fig. 6 and supplementary fig. S6, Supplementary Information online). We therefore hypothesize that, in three lineages of *Bartonella*, Beps have independently explored the characteristic molecular plasticity of FIC domains and evolved a wide range of biochemical activities and target binding features to enable elaborate modulation of host cell functioning (Harms et al. 2016). Future studies should experimentally address the molecular functions of Bep FIC domains in order to elucidate the basis of their diversification.

Our analyses showed that *B. ancashensis* encodes one effector (Bep226) with an array of predicted eukaryotic-like tyrosine phosphorylation motifs. However, their positioning at

the C-terminus of a FIC-BID effector is not like anything found in the other lineages (fig. 4). In the absence of evidence for horizontal gene transfer and given that the consensus sequence of these motifs is largely different from those of other Beps (fig. 4*B*), we conclude that they have arisen three times independently as a true innovation in all three Bep repertoires. This finding highlights the critical role that Beps with tyrosine phosphorylation motifs may play as signaling platforms in host cells (Selbach et al. 2009), though no biological role for these motifs in any Bep has yet been uncovered. Previous work on Beps of L3 or L4 stayed descriptive regarding the phosphorylation itself and only studied phenotypes that appeared to be unrelated to these motifs (Engel et al. 2011; Okujava et al. 2014). Future studies should therefore aim to comparatively assign molecular and biological functions to the tyrosine phosphorylation motifs of Beps of all three lineages in order to understand the evolutionary driving forces behind the repeated emergence of this effector type.

Every instance of VirB/D4 T4SS and Beps in *Bartonella* is invariably associated with BiaA proteins, i.e., homologs of FicA antitoxins (figs. 3*A* and 6). Bep3 and Bep4 orthologs additionally harbor BiaA-like modules at their N-terminus (figs. 4*A* and 7, supplementary fig. S7, Supplementary Information online). Previous work reported that the coexpression of BiaA can improve the yield of ectopic Bep expression and inhibit the functioning of Bep FIC domains (Pieles et al. 2014), similar to what had been shown for FicA antitoxins of FicTA modules (Harms et al. 2015). It is therefore tempting to speculate that BiaA proteins may serve as a chaperone for Beps with FIC domains prior to their translocation into host cells, because type IV secretion is generally thought to involve substrate unfolding (Cabezon et al. 2015). Analogous proteins are well-known for effector-secreting type III secretion systems and also not uncommon for other host-targeting T4SS (Gonzalez-Rivera et al. 2016). Similarly, such a chaperoning function could explain the presence of BiaA-like modules encoded at the N-terminus of some Beps. However, in contrast to solitary BiaA proteins, these would be translocated into host cells as part of the Bep polypeptide. Given that the molecular function of FicA antitoxins is to manipulate binding of the ATP substrate (Engel et al. 2012), we speculate that these BiaA-like modules may guide or modulate the enzymatic activities of FIC domains.

An adaptive radiation, i.e., the rapid diversification of an ancestral species into an array of divergently adapted derivatives, can occur upon encountering unexplored ecological niches and/or upon acquisition of an evolutionary key innovation that enables the organism to explore niches that were previously not accessible (Berner and Salzburger 2015). In case of L3 and L4 of *Bartonella*, it has previously been argued that the hemotropic lifestyle and vector transmission released these bacteria from major resource competition, so that the availability of niches would not be limiting as long as vector transmission is possible (Engel et al. 2011; Saenz et al. 2007). Conversely, an evolutionary key innovation such as enhanced host adaptability following the acquisition of a VirB/D4 T4SS and the evolution of complex effector sets could well explain the adaptive radiations of L3 and L4. The discovery of a VirB/D4 T4SS and Beps in *B. ancashensis* seemed to challenge this model, because this organism is only distantly related to *Bartonella* lineages exhibiting enhanced host adaptability (fig. 1). However, our *in silico* analyses suggest that the functional diversification of *B. ancashensis* Beps is less advanced those of L3 and L4 and largely relies on the ancestral type of *bona fide* AMP-tranferases with FIC-BID domain architecture (figs. 4 and 5, supplementary fig. S6, Supplementary Material online). This effector set may therefore lack certain key features that are essential to exploit the full adaptive potential of type IV secretion in the context of *Bartonella* infections and that had been acquired by the Beps of L3 and L4 during their parallel functional diversification. Possibly, the Beps of *B. ancashensis* may therefore represent an evolutionary state ancestral to the effector sets of L3 and L4. However, the paucity of information regarding the prevalence and pathogenesis of *B. ancashensis* make it difficult to speculate in what way the evolution of its Beps is linked to the evolution and the infection strategy of this organism. A role of these effectors during infections of human patients with *B. ancashensis* seems likely given the overall similarities to Beps of L3 and L4, though the few available reports of infections with *B. ancashensis* do not mention differences in disease manifestation to *B. bacilliformis* (Blazes et al. 2013; Mullins et al. 2013). Future work should therefore comparatively dissect the molecular functions and biological roles of the three effector sets in order to understand the actual nature of the evolutionary key innovation that sparked the adaptive radiations of L3 and L4. We anticipate that these studies will also give new insight into the evolutionary dynamics of bacterial effector proteins and the molecular basis of bacterial host adaptation in general.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

The authors are grateful to Dr Remy Heller and Dr Yves Piemont (†) for *B. taylorii* IBS296. Lynn Osikowicz and Andres Lopez-Perez are acknowledged for laboratory work associated with the isolation of *B.* sp. CDCskunk. This work was supported by the Swiss National Science Foundation (grant 310030B-149886) and by a European Research Council Advanced Investigator Grant (grant FICModFun 340330) to C.D. P.E was funded by the Swiss National Science Foundation (grant 31003A_160345). B.B.C. was funded in part by the Center for Companion Animal Health, School of Veterinary Medicine, University of California, Davis and the American Kennel Club Canine Health Foundation.

Literature Cited

Abby SS, Rocha EP. 2012. The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. PLoS Genet. 8:e1002983.

Angelakis E, Raoult D. 2014. Pathogenicity and treatment of *Bartonella* infections. Int J Antimicrob Agents 44:16–25.

Berner D, Salzburger W. 2015. The genomics of organismal diversification illuminated by adaptive radiations. Trends Genet. 31:491–499.

Blazes DL, et al. 2013. Novel *Bartonella* agent as cause of verruga peruana. Emerg Infect Dis. 19:1111–1114.

Brown NF, Finlay BB. 2011. Potential origins and horizontal transfer of type III secretion systems and effectors. Mob Genet Elements 1:118–121.

Burstein D, et al. 2016. Genomic analysis of 38 *Legionella* species identifies large and diverse effector repertoires. Nat Genet. 48:167–175.

Cabezon E, Ripoll-Rozada J, Pena A, de la Cruz F, Arechaga I. 2015. Towards an integrated model of bacterial conjugation. FEMS Microbiol Rev. 39:81–95.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973.

Castro-Roa D, et al. 2013. The Fic protein Doc uses an inverted substrate to phosphorylate and inactivate EF-Tu. Nat Chem Biol. 9:811–817.

Chin CS, et al. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569.

Costa TR, et al. 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nat Rev Microbiol. 13:343–359.

Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27:1164–1165.

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 16:157.

Engel P, et al. 2012. Adenylylation control by intra- or intermolecular active-site obstruction in Fic proteins. Nature 482:107–110.

Engel P, et al. 2011. Parallel evolution of a type IV secretion system in radiating lineages of the host-restricted bacterial pathogen *Bartonella*. PLoS Genet. 7:e1001296.

Frank AC, Alsmark CM, Thollesson M, Andersson SG. 2005. Functional divergence and horizontal transfer of type IV secretion systems. Mol Biol Evol. 22:1325–1336.

Gonzalez-Rivera C, Bhatty M, Christie PJ. 2016. Mechanism and function of type IV secretion during infection of the human host. Microbiol Spectr. 4(3). doi: 10.1128/microbiolspec.VMBF-0024-2015.

Guglielmini J, de la Cruz F, Rocha EP. 2013. Evolution of conjugation and type IV secretion systems. Mol Biol Evol. 30:315–331.

Guy L, Kultima JR, Andersson SG. 2010. genoPlotR: comparative gene and genome visualization in R. Bioinformatics 26:2334–2335.

Guy L, et al. 2012. A genome-wide study of recombination rate variation in *Bartonella henselae*. BMC Evol Biol. 12:65.

Guy L, et al. 2013. A gene transfer agent and a dynamic repertoire of secretion systems hold the keys to the explosive radiation of the emerging pathogen *Bartonella*. PLoS Genet. 9:e1003393.

Hang J, et al. 2015. Complete genome sequence of *Bartonella ancashensis* strain 20.00, isolated from the blood of a patient with verruga peruana. Genome Announc. 3(6):pii: e01217-15.

Harms A, Dehio C. 2012. Intruders below the radar: molecular pathogenesis of *Bartonella* spp. Clin Microbiol Rev. 25:42–78.

Harms A, Stanger FV, Dehio C. 2016. Biological diversity and molecular plasticity of FIC domain proteins. Annu Rev Microbiol. 70:341–360.

Harms A, et al. 2015. Adenylylation of gyrase and topo IV by FicT toxins disrupts bacterial DNA topology. Cell Rep. 12:1497–1507.

Hayashi T, Morohashi H, Hatakeyama M. 2013. Bacterial EPIYA effectors: where do they come from? What are they? Where are they going?. Cell Microbiol. 15:377–385.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.

Henn JB, et al. 2009. *Bartonella rochalimae* in raccoons, coyotes, and red foxes. Emerg Infect Dis. 15:1984–1987.

Henn JB, et al. 2007. Gray foxes (*Urocyon cinereoargenteus*) as a potential reservoir of a *Bartonella clarridgeiae*-like bacterium and domestic dogs as part of a sentinel system for surveillance of zoonotic arthropod-borne pathogens in northern California. J Clin Microbiol. 45:2411–2418.

Larkin MA, et al. 2007. Clustal W and clustal X version 2.0. Bioinformatics 23:2947–2948.

Luo R, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. Gigascience 1:18.

Markowitz VM, et al. 2014. IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res. 42:D560–D567.

Minnick MF, et al. 2014. Oroya fever and verruga peruana: bartonelloses unique to South America. PLoS Negl Trop Dis. 8:e2919.

Mukherjee S, et al. 2011. Modulation of Rab GTPase function by a protein phosphocholine transferase. Nature 477:103–106.

Mullins KE, et al. 2013. Molecular typing of "*Candidatus* Bartonella ancashi," a new human pathogen causing verruga peruana. J Clin Microbiol. 51:3865–3868.

Obenauer JC, Cantley LC, Yaffe MB. 2003. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic Acids Res. 31:3635–3641.

Okujava R, et al. 2014. A translocated effector required for *Bartonella* dissemination from derma to blood safeguards migratory host cells from damage by co-translocated effectors. PLoS Pathog. 10:e1004187.

Palanivelu DV, et al. 2010. Fic domain catalyzed adenylylation: insight provided by the structural analysis of the type IV secretion system effector BepA. Protein Sci. 20(3):492–499.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290.

Pieles K, Glatter T, Harms A, Schmidt A, Dehio C. 2014. An experimental strategy for the identification of AMPylation targets from complex protein samples. Proteomics 14:1048–1052.

Pulliainen AT, et al. 2012. Bacterial effector binds host cell adenylyl cyclase to potentiate Galphas-dependent cAMP production. Proc Natl Acad Sci U S A. 109:9581–9586.

Quebatte M, et al. 2010. The BatR/BatS two-component regulatory system controls the adaptive response of *Bartonella henselae* during human endothelial cell infection. J Bacteriol. 192:3352–3367.

Quebatte M, Dick MS, Kaever V, Schmidt A, Dehio C. 2013. Dual input control: activation of the *Bartonella henselae* VirB/D4 type IV secretion system by the stringent sigma factor RpoH1 and the BatR/BatS two-component system. Mol Microbiol. 90:756–775.

Saenz HL, et al. 2007. Genomic analysis of *Bartonella* identifies type IV secretion systems as host adaptability factors. Nat Genet. 39:1469–1476.

Sato S, et al. 2012a. Isolation and phylogenetic analysis of *Bartonella* species from wild carnivores of the suborder Caniformia in Japan. Vet Microbiol. 161:130–136.

Sato S, et al. 2012b. Prevalence and genetic diversity of *Bartonella* species in sika deer (*Cervus nippon*) in Japan. Comp Immunol Microbiol Infect Dis. 35:575–581.

Segers FHID, Kešnerová L, Kosoy M, Engel P. 2017. Genomic changes associated with the evolutionary transition of an insect gut commensal into a blood-borne pathogen. ISME J. (in press).

Schulein R, et al. 2005. A bipartite signal mediates the transfer of type IV secretion substrates of *Bartonella henselae* into human cells. Proc Natl Acad Sci U S A. 102:856–861.

Selbach M, et al. 2009. Host cell interactome of tyrosine-phosphorylated bacterial proteins. Cell Host Microbe. 5:397–403.

Seubert A, Hiestand R, de la Cruz F, Dehio C. 2003. A bacterial conjugation machinery recruited for pathogenesis. Mol Microbiol. 49:1253–1266.

Siamer S, Dehio C. 2015. New insights into the role of *Bartonella* effector proteins in pathogenesis. Curr Opin Microbiol. 23:80–85.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 30:2725–2729.
Zhu Q, Kosoy M, Olival KJ, Dittmar K. 2014. Horizontal transfers and gene losses in the phospholipid pathway of *Bartonella* reveal clues about early ecological niches. Genome Biol Evol. 6:2156–2169.

Associate editor: Rotem Sorek