

Genomic features of the *Helicobacter pylori* strain PMSS1 and its virulence attributes as deduced from its *in vivo* colonisation patterns

Victoria Dyer,¹ Holger Brüggemann,²
Meike Sörensen,¹ Anja A. Kühn,³ Kirstin Hoffman,¹
Volker Brinkmann,¹ Maria del Mar Reines,¹
Stephanie Zimmerman,¹ Thomas F. Meyer^{1*} and
Manuel Koch¹

¹Department of Molecular Biology, Max Planck Institute for Infection Biology, Berlin, 10117, Germany.

²Department of Biomedicine, Aarhus University, Aarhus C, 8000, Denmark.

³Division of Gastroenterology, Infectiology and Rheumatology, Medical Department, Campus Benjamin Franklin, Charité, Berlin, 12200, Germany.

Summary

The human gastric pathogen *Helicobacter pylori* occurs in two basic variants, either exhibiting a functional *cagPAI*-encoded type-4-secretion-system (T4SS) or not. Only a few *cagPAI*-positive strains have been successfully adapted for long-term infection of mice, including the pre-mouse Sydney strain 1 (PMSS1). Here we confirm that PMSS1 induces gastric inflammation and neutrophil infiltration in mice, progressing to intestinal metaplasia. Complete genome analysis of PMSS1 revealed 1,423 coding sequences, encompassing the *cagPAI* gene cluster and, unusually, the location of the cytotoxin-associated gene A (*cagA*) approximately 15 kb downstream of the island. PMSS1 harbours three genetically exchangeable loci that are occupied by the *hopQ* coding sequences. *HopQ* represents a critical co-factor required for the translocation of CagA into the host cell and activation of NF- κ B via the T4SS. Long-term colonisation of mice led to an impairment of *cagPAI* functionality. One of the bacterial clones re-isolated at four months post-infection revealed a mutation in the *cagPAI* gene *cagW*, resulting in a

frame shift mutation, which prevented CagA translocation, possibly due to an impairment of T4SS function. Rescue of the mutant *cagW* re-established CagA translocation. Our data reveal intriguing insights into the adaptive abilities of PMSS1, suggesting functional modulation of the *H. pylori cagPAI* virulence attribute.

Introduction

The bacterium *Helicobacter pylori* colonises the gastric mucosa of ~50% of the human population. The majority of infected individuals remain asymptomatic but a significant proportion develop severe diseases, including peptic ulcers (~15%), gastric adenocarcinomas (~1%) or gastric MALT lymphomas (~0.1%). *H. pylori* can be subdivided depending on the presence or absence of the *cag* pathogenicity island (*cagPAI*) and the severity of *H. pylori*-induced disease is strongly increased in the presence of *cagPAI* (Kusters *et al.*, 2006). The *cagPAI* encodes a type IV secretion system (T4SS) and the effector protein cytotoxin-associated gene A (CagA). Upon binding to the host cell, *H. pylori* translocates CagA into the cytoplasm, where it is partially tyrosine-phosphorylated by members of the Abl and Src family kinases and can bind to several host factors, including SHP-2 phosphatase, altering host cell function (Higashi *et al.*, 2002). Moreover, the *H. pylori* T4SS has recently been shown to translocate ubiquitous Gram-negative bacterial metabolites of LPS synthesis into the host cells (Gall *et al.*, 2017; Stein *et al.*, 2017; Zimmermann *et al.*, 2017), where it is most likely ADP heptose that induces a vivid inflammatory response (Pfannkuch *et al.*, 2018). The *H. pylori* T4SS has also been reported to translocate peptidoglycan (Viala *et al.*, 2004).

The relative frequency of *cagPAI*-positive strains shows marked geographical variation, ranging from 0% for the African strain (hpAfrica2) to ~60% for the European strain (hpEurope) and ~95% for the East Asian strain (hspEastasia) (Olbermann *et al.*, 2010). Putative secondary loss of the *cagPAI* in European strains strongly suggests that

Accepted 5 September, 2018. *For correspondence. E-mail meyer@mpiib-berlin.mpg.de; Tel. +49 30 28 460 400; Fax +49 30 28 460 401.

positive selection pressure for *cagPAI* is outweighed by a negative selection pressure in regions with low benefit of the *cagPAI* for the pathogen. One potential reason for this phenomenon is the strong inflammatory response associated with *cagPAI*-positive strains that counter-selects for the T4SS (Philpott *et al.*, 2002). Similarly, this adaptive process may happen *in vivo*, as shown in both mice and rhesus macaques (Barrozo *et al.*, 2013).

Mice can easily be colonised with non-functional *cagPAI* strains such as Sydney Strain 1 (SS1) and Hp76 (Lee *et al.*, 1997; Gomez-Duarte *et al.*, 1998). However, very few published functional *cagPAI*-positive *H. pylori* strains are capable of colonising the rodent stomach – PMSS1 and LSH100 (Arnold *et al.*, 2011; Lowenthal *et al.*, 2009) for mice and B128 for gerbils, presenting a model of accelerated pathology (Israel *et al.*, 2001). While the PMSS1 seems to have the highest level of pathogenicity in mice, over time (3 months) re-isolated strains have been reported to lose the ability to translocate CagA (Arnold *et al.*, 2011). A putative structural protein of the T4SS machinery, CagY, was suggested to be responsible for this process (Barrozo *et al.*, 2013). Apart from *cagPAI*-encoded functions, two outer membrane proteins of *H. pylori* have been implicated in T4SS function: BabA and the virulence factor HopQ (Belogolova *et al.*, 2013; Ishijima *et al.*, 2011). This suggests that an adhesive component is important for CagA translocation and this notion has recently been corroborated by the identification of CEACAMs as receptors for HopQ on the epithelial host cell surface (Javaheri *et al.*, 2016; Königer *et al.*, 2016). There are two *hopQ* alleles, type I and type II, and each strain tends to encode only one. The *hopQ* type I is strongly correlated with the presence of the *cagPAI* as seen in the commonly used human *cagPAI*-positive strains P12 and G27 (Belogolova *et al.*, 2013).

To shed more light on the selective pressure acting on the functionality of the *cagT4SS in vivo* and to identify the essential factors of the T4SS machinery, we infected mice with PMSS1 and tested for *cagT4SS* functionality in re-isolates *ex vivo*. By sequencing the complete genomes of the mouse colonizing strain PMSS1 and two re-isolates, we identified three sites in the PMSS1 genome occupied by *hopQ* genes. In addition, we also revealed negative selection against *cagPAI*, resulting in a loss of ability to translocate CagA. This defect appears to be due to a mutation in the *cagW* gene, resulting in T4SS impairment.

Results

PMSS1 induces corporal intestinal metaplasia in C57BL/6 mice

It has previously been shown that mice infected with PMSS1 exhibit T4SS-dependent gastric cancer precursor

lesions three months post-infection (m.p.i.) (Arnold *et al.*, 2011). To corroborate this observation, we established a long-term *in vivo* infection study to compare the pathogenicity of the PMSS1 (Arnold *et al.*, 2011) and P12 (Schmitt and Haas, 1994) strains, both of which are proficient at translocating CagA *in vitro* (Arnold *et al.*, 2011; Odenbreit *et al.*, 2000). Histopathology scores of corporal tissue at two and four m.p.i. confirmed that PMSS1 but not P12 was able to induce and sustain a state of chronic inflammation in the majority of the mice (92% vs 46%, respectively) (Fig. 1A). H&E staining of gastric tissue sections showed a substantial increase in neutrophils (arrowheads) in the mucosal and submucosal regions of the corpus of PMSS1-infected mice four m.p.i (Fig. 1B), while infection with P12 caused only minor infiltration (Fig. 1B). Similarly, only PMSS1-infected mice showed evidence of Alcian blue-stained acidic mucins produced by goblet cells (arrows), indicative of glands that have undergone intestinal metaplasia, as well as glands with elongated neck regions indicative of hyperplasia (Fig. 1B). A slight to moderate increase in the overall mucosal thickness was observed in many of the infected mice; however, as a slightly thicker mucosa was occasionally also observed in the non-infected group, this in itself was not scored as hyperplasia.

PMSS1 undergoes negative selection pressure

Although colonisation of the gastric mucosa by PMSS1 is efficient (Supplementary Fig. 1), bacterial isolates cultured from infected mice have shown a reduced capacity for infection due to loss of functional pathogenicity, i.e. the ability to translocate CagA, at three m.p.i (Arnold *et al.*, 2011). To test if this phenotype also occurred in our mice and to further investigate the mechanism that leads to this functional impairment, we cultured bacterial clonal isolates from infected mice at two and four months. Clonal isolates were successfully cultured from 8 of the 14 PMSS1-infected mice, but none were obtained from any of the mice infected with P12. Subsequent infection of AGS cells for 3 h at MOI 100 showed that all eight clonal isolates were capable of expressing CagA; however, clonal isolates Iso1 and Iso6 were incapable of CagA translocation to the host cells (Fig. 2A). Corroborating this observation is the absence of the distinct cellular morphology known as the 'hummingbird' phenotype (arrowheads) after infection with the non-competent Iso6, which would indicate the presence of the translocated CagA (Fig. 2B), raising the question of how these isolates were altered *in vivo*. Scanning electron microscopy (SEM) images of the T4SS dysfunctional and functional isolates Iso6 and Iso8, as well as the wild type PMSS1 and P12 strains, all show the distinct

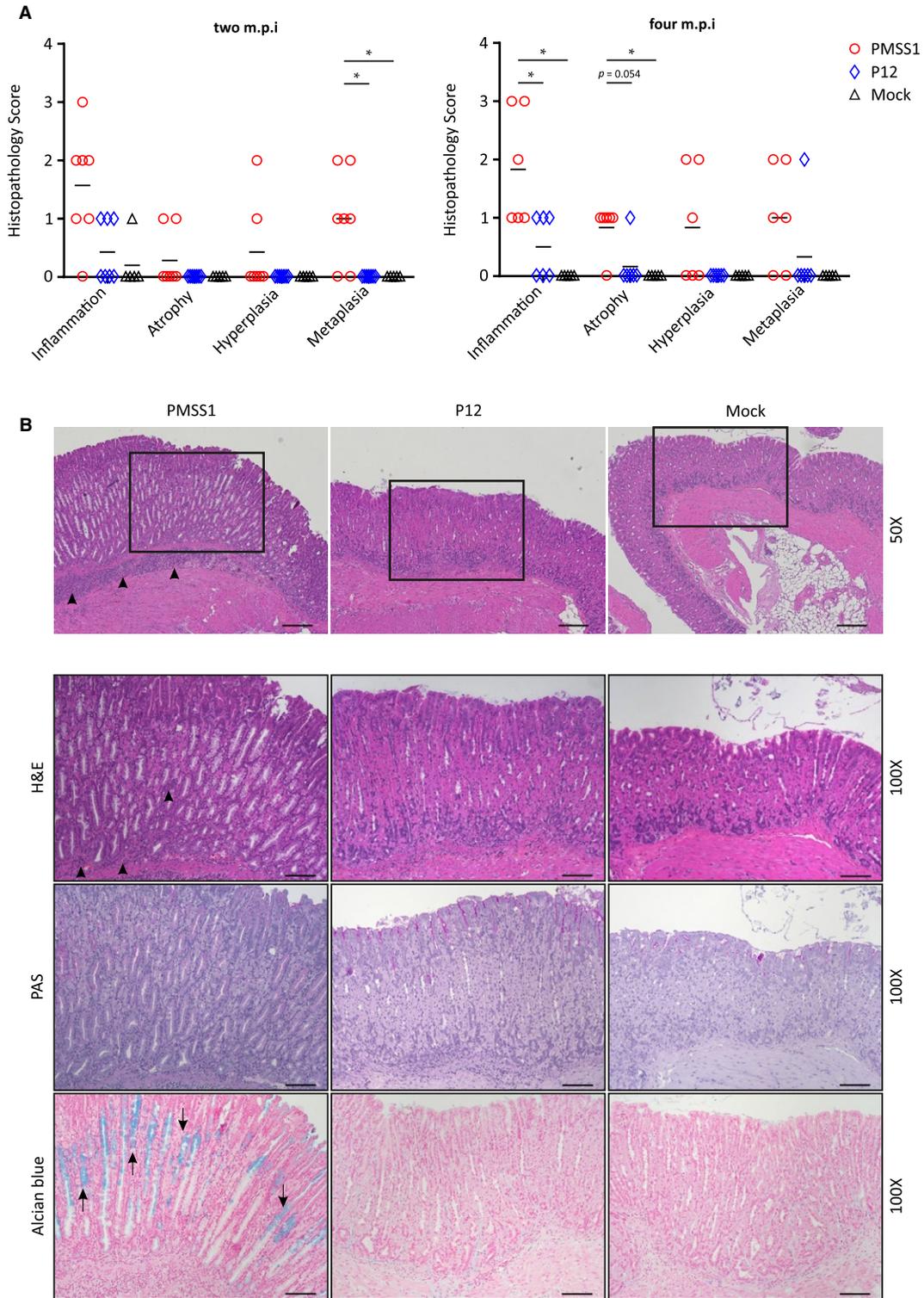


Fig. 1. PMSS1 induces corporal intestinal metaplasia *in vivo*.

A. The corporal tissue from C57BL/6 mice mock infected or infected with either PMSS1 or P12 at 2 and 4 m.p.i. was scored in a blinded manner for inflammation, atrophy, hyperplasia and metaplasia on a scale of none (0) to marked (4). Each point represents one mouse; bars depict the mean; * $p < 0.01$.

B. Corporal tissue sections at 4 m.p.i. showing hyperplastic glands and intestinal metaplasia (black arrows) only in PMSS1-infected mice. Micrographs show one representative mouse from each group. Scale bars at 50X and 100X represent 200 μm and 100 μm respectively.

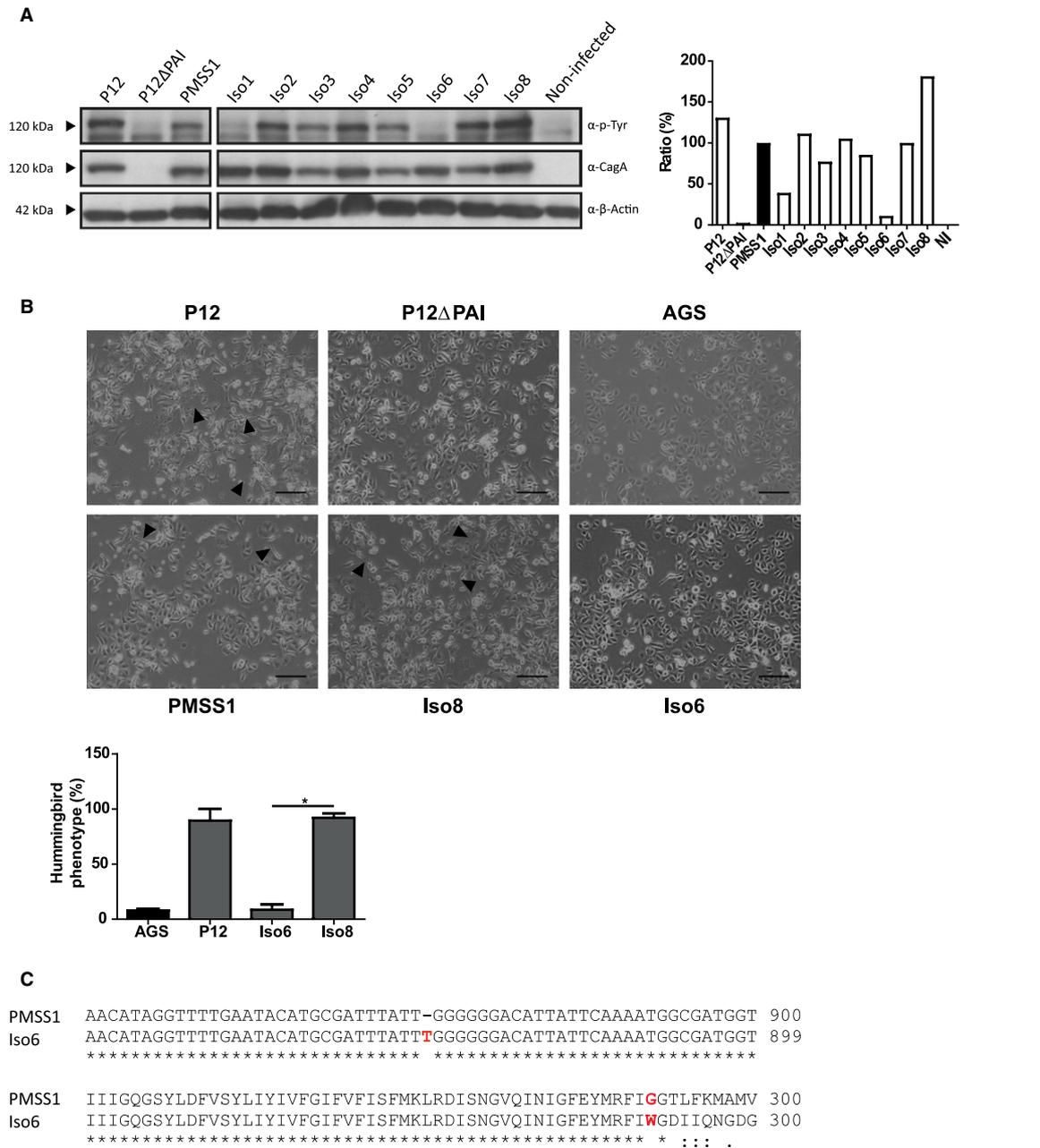


Fig. 2. PMSS1 retains colonisation efficacy by undergoing negative selection pressure.

A. AGS cells were infected for 3 h (MOI 100) with either the wild-type *H. pylori* strains, PMSS1 or P12; the *in vivo* PMSS1 bacterial isolates or the *cagPAI*-deficient strain P12ΔPAI. Immunoblotting detected the expression of CagA at approximately 118 kDa for the PMSS1 strains and at approximately 120 kDa for P12. Interestingly, translocated CagA (p-Tyr) was found to occur in all bacterial strains except in isolates Iso1 and Iso6, and the *cagPAI* mutant strain P12ΔPAI. Anti-β-actin was incorporated as a loading control. Right panel: quantification of translocated CagA.

B. Corroboration of translocated CagA was observed in AGS cells exhibiting the hummingbird phenotype (black arrowheads). After 24-h infection (MOI 10), the translocation competent bacterial isolate (Iso8) demonstrated the characteristic phenotype, with the filopodia exhibiting the same protrusion lengths as the original PMSS1 inoculating strain, while the non-competent isolate (Iso6) failed to do so, as did the non-competent P12 PAI deletion mutant; scale bar 100 μm. Right panel: quantification of the two isolates compared to non-infected (AGS) and the P12 control.

C. Clustal Omega comparison of the single base insertion occurring in Iso6, which results in the frame shift mutation that subsequently hinders CagA translocation by rendering the T4SS dysfunctional. [Colour figure can be viewed at wileyonlinelibrary.com]

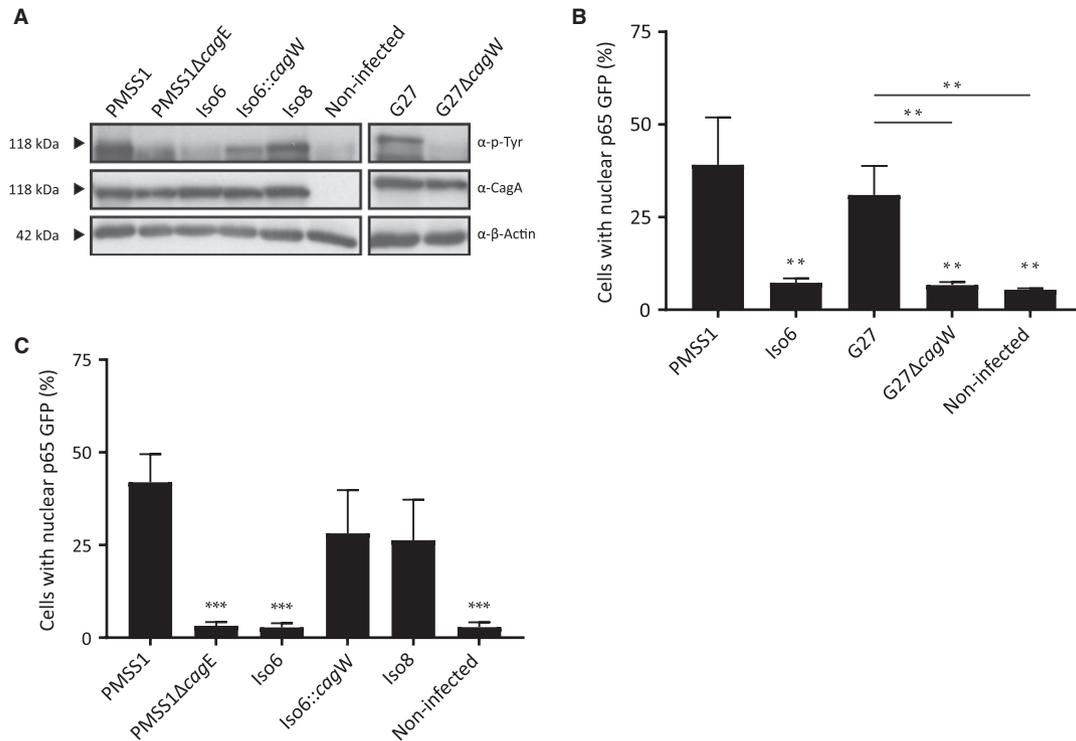


Fig. 3. Rescuing *cagW* restores the functionality of the T4SS.

A. Expression and translocation of CagA in AGS cells infected for 3 h (MOI 100) was assessed using immunoblotting. All strains were capable of expressing CagA; however, the loss of ability to translocate CagA observed for Iso6 was reversed by rescuing the *cagW* gene in Iso6::*cagW*. The *H. pylori* strain G27 and the corresponding mutant G27Δ*cagW* served as additional controls.

B and C. To confirm that CagA translocation is restored, NF-κB activation was assessed by infecting SiB02 cells with the same strains as in A. for 45 min (MOI 100). The cells were fixed and p65 expression was acquired by automated microscopy and analysed using ScanR software. Data represent mean ± SD of three biological replicates, ** $p < 0.01$, *** $p < 0.001$ compared to PMSS1, one-way ANOVA.

'screw like' morphology and the presence of the T4SS pili (arrowheads) (Supplementary Fig. 2A).

CagW is essential for CagA translocation

Since Iso1 and Iso6 were unable to translocate CagA, we hypothesised that the strains had acquired a mutation that impaired the functionality of the T4SS. We sequenced the complete genomes of Iso6 and Iso8 (which contains a functional T4SS) in addition to PMSS1, and found a single thymine insertion within the *cagW* gene in Iso6 but not Iso1 (Supplementary Fig. 3A). This single nucleotide insertion caused a frame shift mutation (Fig. 2C), resulting in a truncated protein that impairs the assembly of the T4SS (Fig. 3A). We sequenced two further clones isolated from the same mouse as Iso6 and found that both of them harboured the same mutation in the *cagW* gene at base position 871 (Supplementary Fig. 3B), indicating that a single T4SS mutant had out-competed the entire population. Interestingly, despite this, there was a pronounced infiltration of neutrophils and intestinal metaplasia within the corporal region of this particular mouse (Supplementary Fig. 3C). To confirm that this single mutation was responsible for

preventing CagA translocation, we rescued the *cagW* gene in Iso6 by replacing it with the wild-type *cagW* gene of strain PMSS1. We then assessed the ability of the two strains to express and translocate CagA (Fig. 3A) and activate NF-κB (Fig. 3B and 3C) in infected AGS cells. As positive and negative controls we used the CagA translocation competent isolate 8 and PMSS1 parent strains, as well as the non-translocation competent PMSS1Δ*cagE* and G27Δ*cagW* mutant (Belogolova *et al.*, 2013) respectively. As observed for Iso6, G27Δ*cagW* was also unable to translocate CagA or activate NF-κB. In contrast, these functions were restored in the rescued Iso6::*cagW* mutant. In summary, our data show that CagW is necessary for T4SS functionality and further substantiates that NF-κB activation depends on a functional T4SS.

General features of the PMSS1 genome

To investigate the genetic changes underlying the loss of pathogenicity observed in some re-isolates, we assembled the draft genome of the PMSS1 parent strain and compared it to that recently published by Draper *et al.* (Draper *et al.*, 2017). Comprising 1,590,017 base pairs (bp), the circular chromosome of the PMSS1 draft

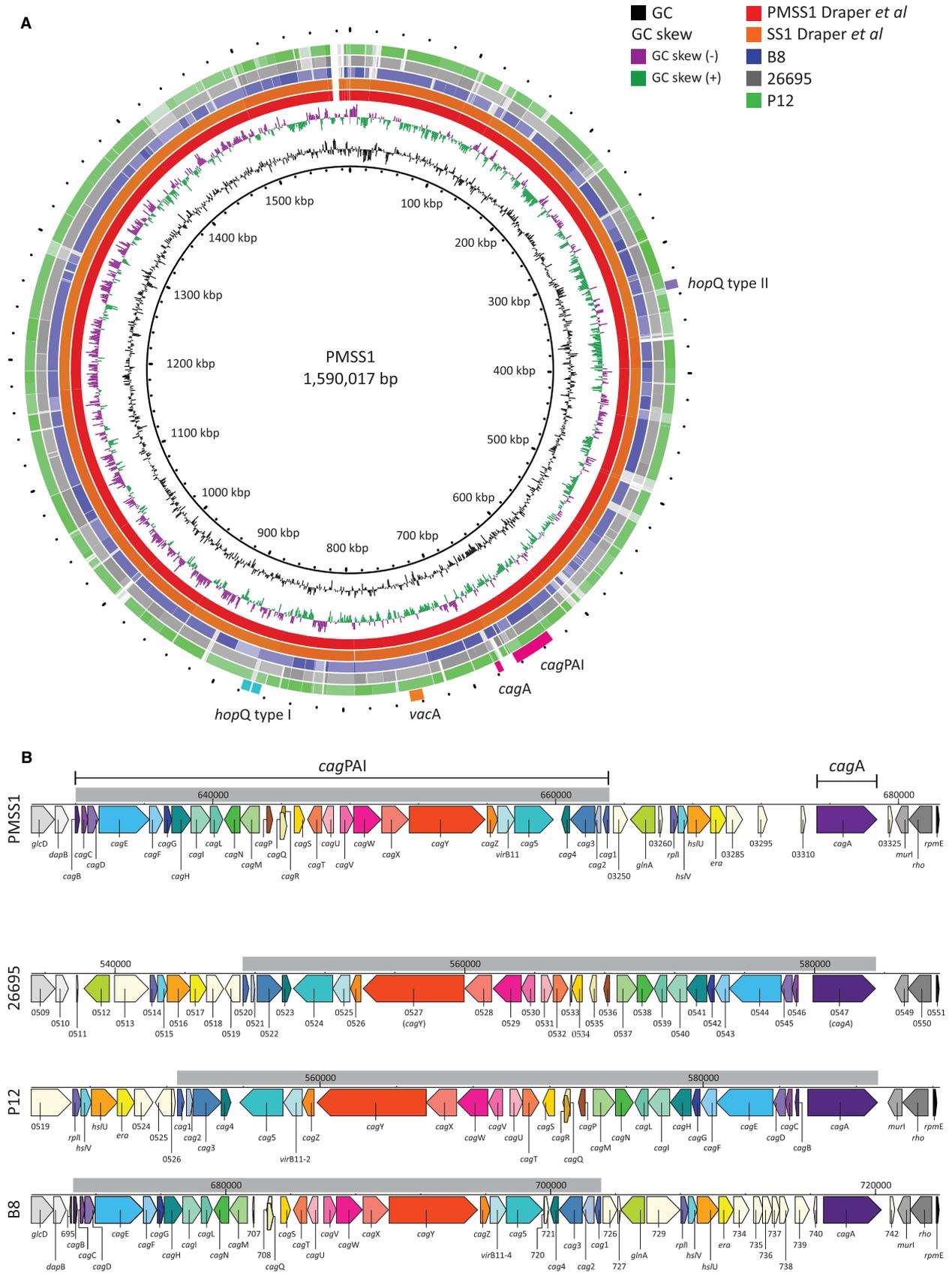


Fig. 4. *H. pylori* strain PMSS1 encodes a functional *cagPAI*.

A. Comparative analysis of the PMSS1 genome (inner concentric ring) to the following *H. pylori* genomes (outer concentric rings): PMSS1 (red) and SS1 (orange) (Draper *et al.*, 2017), the Mongolian gerbil-adapted strain B8 (blue) (Farnbacher *et al.*, 2010), the reference genome 26695 (grey) and the *H. pylori* strain P12 (green). The third and fourth concentric rings depict the GC content (black) and the GC skew, *i.e.* genes coded on either the forward (green) or reverse (purple) strand. Several regions of genome variability (white gaps) were identified between PMSS1 and the genomes of B8, 26695 and P12. The gap located at 12 o'clock corresponds to the encoded plasmid. The following gene features are encoded by PMSS1: *vacA* and *cagPAI*; however, the oncogene *cagA* is located ~40 kb downstream of the *cagPAI* gene cluster.

B. Comparative analysis of the *cagPAI* region (shaded grey) encoded by PMSS1 (Track 1) to the reference genome 26695 (Track 2) and the *H. pylori* strains P12 (Track 3) and B8 (Track 4). Genes are annotated in accordance with their GenBank gene feature names and orthologues are depicted in identical colours across the strains. Hypothetical genes are shown in pale yellow.

Table 1. Comparative analysis of the PMSS1 general genome features.

<i>Helicobacter pylori</i> strain	PMSS1	26695	P12
Accession number	AZBR00000000	NC_000915	NC_011498
Genome size (bp)	1,590,017	1,667,867	1,673,813
GC content (%)	39.0	38.9	38.8
Coding sequences	1423	1576	1486
Average length of CDS (bp)	968	954	958
Coding regions (%)	86.6	90.2	89.7
Annotated CDS	1122	1091	1111
Hypothetical CDS	313	499	
Pseudogenes	12	0	93
Genomic islands	<i>cagPAI</i> <i>comB</i>	<i>cagPAI</i> <i>comB</i>	<i>cagPAI</i> <i>comB</i>
Plasmid	Yes	No	Yes
rRNA	3		6
tRNA	36	36	36

genome is slightly smaller than the *H. pylori* reference genome 26695 (1,667,867 bp) and the *H. pylori* strain P12 (1,673,813 bp), with an average GC content of 39.0% and 1,423 predicted coding sequences (CDS), of which 1,122 are annotated. Encoded within the PMSS1 draft genome are the genes comprising the *cag* pathogenicity island, the cytotoxin-associated gene *cagA* and vacuolating cytotoxin gene *vacA* (Fig. 4A). A precise overview of the general genome sequence features is listed in Table 1. The draft genome sequence of the PMSS1 strain has been deposited into the National Centre for Biotechnology Information (NCBI) database and accessible under GenBank® accession number AZBR00000000. The *cagPAI* gene cluster encoded by PMSS1 was found to comprise 28 genes, spanning 31,096 bp. Seventeen of the genes encoded in this cluster are required for the assembly and function of the T4SS apparatus, which is responsible for the translocation of CagA (Fischer *et al.*, 2001) and HBP (Zimmermann *et al.*, 2017) and also peptidoglycan (Viala *et al.*, 2004) into the host cell cytoplasm. The *cagPAI* region of PMSS1 is coded on the leading strand, comparable to most *H. pylori* strains. However, the gene cluster is notably inverted and the *cagA* gene lacking. Instead, *cagA* is located approximately 15 kb downstream and bordered by the 3' end of the glutamate racemase gene with the same orientation as observed in the *H. pylori* genomes 26695 and P12 (Fig. 4B). Notably,

our results shown in Fig. 2 indicate that the location of *cagA* outside of the *cagPAI* region affects neither gene expression nor subsequent translocation (Fig. 2B), although the CagA protein translocated by PMSS1 is smaller than that of the P12 strain (~118 vs. 121 kDa). Although the orientation of the *cagPAI* region and location of *cagA* in PMSS1 is rather unusual, it has been previously observed in the Mongolian gerbil-adapted *H. pylori* strain B8 (Fig. 4B), another persistent coloniser of rodents (Farnbacher *et al.*, 2010), which is also able to actively express and translocate CagA (Lind *et al.*, 2014).

We further compared our PMSS1 sequence (GenBank: AZBR00000000) with the genome of the PMSS1 published by Draper *et al.* (2017) (GenBank: CP018823), which was sequenced using the PacBio technology. A single nucleotide polymorphism (SNP) analysis was carried out using the program Parsnp (Treangen *et al.*, 2014). The core chromosome of the two PMSS1 sequences contained five SNPs in total (Supplementary Fig. 4A). Two SNPs were found in A- or T-rich repetitive intergenic regions and three SNPs were non-synonymous substitutions found within open reading frames of the genes for citrate synthase (*gltA*; HPYLPMS1_00027), ferric uptake regulation protein (*fur*; HPYLPMS1_00405) and iron-regulated outer membrane protein (HPYLPMS1_01439); each SNP resulted in a single amino acid change.

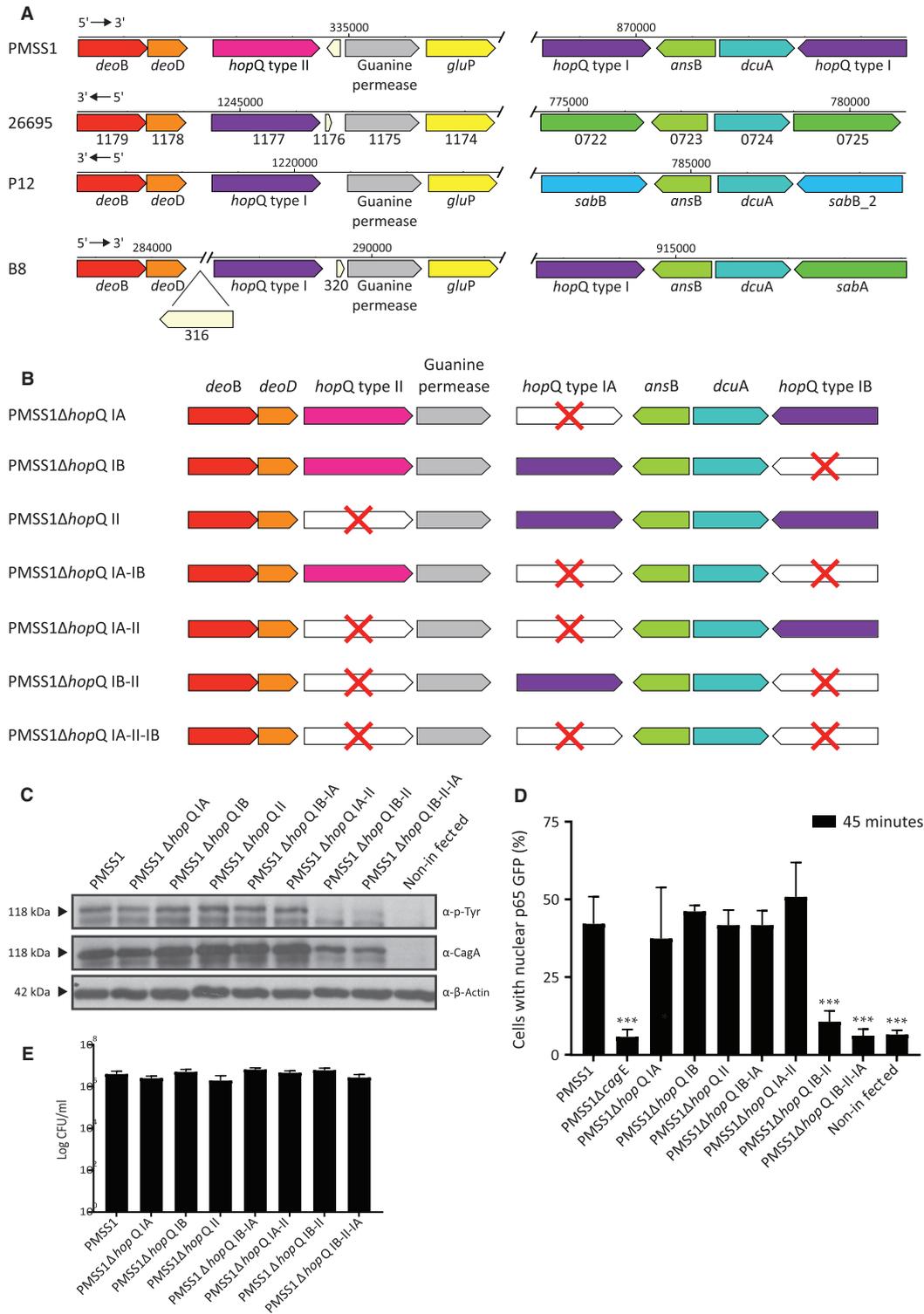


Fig. 5. HopQ influences T4SS-mediated NF- κ B activation.

A. Comparative analysis of the triple *hopQ* locus encoded by PMSS1 (Track 1) to the *H. pylori* reference genome 26695 (Track 2) and the *H. pylori* strains P12 (Track 3) and B8 (Track 4). The 26695 and P12 gene regions have been inverted to better illustrate the genetic recombination between the *sabA*, *sabB* and *hopQ* genes. Genes are annotated in accordance with their GenBank® gene feature names and orthologues are depicted in identical colours across the strains. Hypothetical genes are coloured in pale yellow.

B. Seven PMSS1 Δ *hopQ* deletion mutants were constructed by successively replacing each *hopQ* allele with either a *Kan*-, *Cam*- or *ErmE*-resistant cassette.

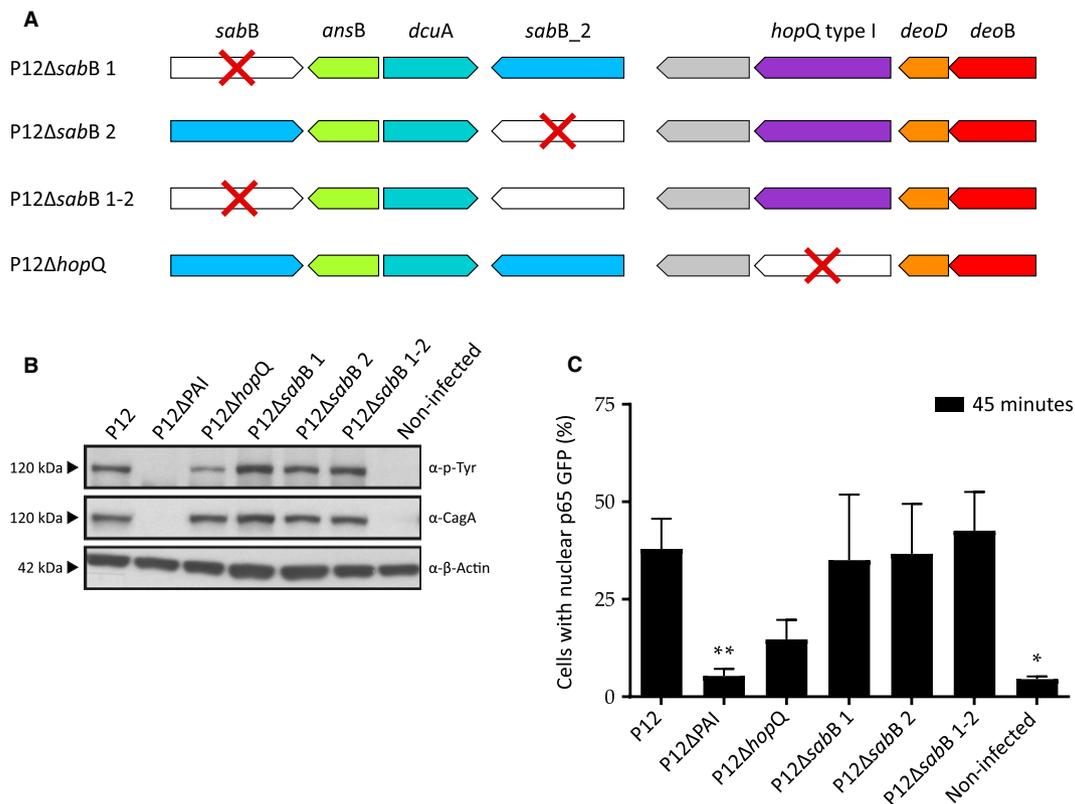
C. Western blotting analysis of AGS cells infected for 3 h (MOI 100) detected protein at ~120 kDa with anti-phospho-tyrosine (PY-99), which detects phosphorylated CagA within the cells, and anti-CagA, which detects CagA expression.

D. SiB02 cells were infected with the PMSS1 Δ *hopQ* mutants (MOI 100) for 45 min, then p65 was acquired by automated microscopy and analysed using ScanR software. Data represent mean \pm SD of three biological replicates, *** p < 0.001 compared to PMSS1, one-way ANOVA.

E. AGS cells were infected with PMSS1 Δ *hopQ* mutants (MOI 100) for 90 minutes and bacterial adherence determined by CFUs.

Apart from the 26 sequence gaps in our PMSS1 sequence, which are all closed in the PacBio-sequenced PMSS1 chromosome by Draper *et al.* (2017), one other difference was noted concerning the *cagA* gene. In both the PMSS1 chromosomal sequences, the *cagA* gene was located downstream of *cagPAI*. However, in the sequence published by Draper *et al.*, the *cagA* gene was copied four times (Draper *et al.*, 2017; Jang *et al.*, 2017); in contrast, we observed only one copy of *cagA*. (Supplementary

Fig. 4B). This could be due to a contraction in the clonal isolate we sequenced. It was reported that the PMSS1 strain represents a heterogeneous population in terms of the numbers of *cagA* copies (Jang *et al.*, 2017); isolates carried from zero to four copies of *cagA* were arranged as direct repeats. An alternative explanation could be the different sequencing approaches used in our study: it cannot be excluded that the Illumina-sequencing data generated in our study misassembled the *cagA*-containing

**Fig. 6.** CagA translocation is independent of *sabB*.

A. Representation of the three *sabB* deletion mutants.

B. Western blotting analysis of AGS cells infected for 3 h (MOI 100) detected protein at approximately 120 kDa with anti-phospho-tyrosine (PY-99), which detects phosphorylated CagA within the cells, and anti-CagA, which detects CagA expression.

C. SiB02 cells were infected with the P12 Δ *sabB* mutants (MOI 100) for 45 min, then p65 was acquired by automated microscopy and analysed using ScanR software. Data represent mean \pm SD of three biological replicates, ** p < 0.01, * p < 0.05 compared to PMSS1, one-way ANOVA.

large repetitive sequence. Even in the PacBio-sequenced PMSS1 chromosome, the number of *cagA* repeats could not be conclusively determined ('Coverage levels suggested either three or four copies') (Draper *et al.*, 2017).

HopQ type IB and *HopQ* type II are required for NF- κ B activation

Other characterised strains encode a *hopQ* locus at the *omp27* locus (corresponding to HP1177 in strain 26695), as well as *sabA* or *sabB* at positions HP0722 and HP0725. Occasionally one of these is replaced by *hopQ* (Cao and Cover, 2002; Talarico *et al.*, 2012). In contrast, we observed here that PMSS1 encodes a unique triple *hopQ* locus – two copies of *hopQ* type I and one copy of *hopQ* type II (Fig. 5A) – in addition to *hopZ* and *oipA*, which are required for adherence. To identify whether all three of the *hopQ* genes encoded by PMSS1 play a role in T4SS-dependent CagA translocation and NF- κ B activation, as previously shown in other *H. pylori* strains (Belogolova *et al.*, 2013), we constructed seven PMSS1 Δ *hopQ* mutants (three single knockouts, three double knockouts and one triple knockout) (Fig. 5B) and infected AGS cells. Neither the single knockouts nor the PMSS1 Δ *hopQ* IA-IB or PMSS1 Δ *hopQ* IA-II double knockouts showed altered CagA translocation, as determined by immunodetection of tyrosine-phosphorylated CagA at ~118 kDa (Fig. 5C). This observation was confirmed by infecting SiB02 cells (Bartfeld *et al.*, 2010), which stably express p65-GFP and thus enable the detection of p65 translocation to the nucleus as a readout for NF- κ B activation. After infection, the cells showed comparable levels of nuclear p65 GFP localisation to that observed after infection with the wild-type PMSS1 (Fig. 5D). In contrast, the double knockout mutant PMSS1 Δ *hopQ* IB-II and the triple knockout PMSS1 Δ *hopQ* IB-II-IIA were neither able to translocate CagA nor activate NF- κ B (Fig. 5B,C). Although it has been suggested that HopQ affects the adherence to human cells (Bonsor *et al.*, 2018), we observed no difference in the adhesion to AGS cells with any of the strains (Fig. 5E). This is in congruence with our own previous observations where we suggested that HopQ functions specifically in T4SS docking rather than in conferring attachment of the whole bacteria to the epithelium (Belogolova *et al.*, 2013). Taken together, we conclude that in PMSS1 one copy of either *hopQ* type IB or *hopQ* type II is required for full CagA translocation, while *hopQ* type IA alone is not sufficient.

SabB is not required for CagA translocation

As previously shown, the double locus of *hopQ* type I in PMSS1 corresponds to the double locus of *sabB* in the P12 strain (Fig. 5A) (Talarico *et al.*, 2012). Due to the

paralogous structure of these genes, we wanted to determine whether *sabB* plays a similar role in the functionality of the T4SS, and thus has a comparable function to HopQ. AGS and SiB02 cells were infected with three P12 Δ *sabB* mutants (two single mutants and one double mutant) (Fig. 6A). There was no effect of SabB knockout on T4SS functionality as measured by the western blot analysis of CagA phosphorylation (Fig. 6B), and no significant difference in the activation of NF- κ B compared to the wild-type P12 strain (Figure 6C). We thus conclude that SabB plays no role in *cagT4SS* functionality and is instead likely to play a *cagPAI*-independent role as an adhesin.

Discussion

Due to the extensive co-evolution and concomitant adaptation of *H. pylori* with its human host (Linz *et al.*, 2007), the use of animal models has been problematic when it comes to experimental reproduction of the natural disease progression of the infection. Functional deficiencies of commonly used *H. pylori* strains, such as colonisation efficacy (*e.g.* P12 versus PMSS1), and the presence or absence of the *cagPAI* have often impeded new insights into the authentic processes of human pathogenesis. This delineates the need for further improvement of the existing *in vivo* infection models. Here we performed a mouse *in vivo* study using the *cagPAI*-positive, T4SS functional PMSS1 strain to analyse how long-term infection affects both the pathogen and the host. We observed that the effective colonisation of the murine gastric mucosa resulted in the induction of chronic inflammation and subsequent corporal intestinal metaplasia as early as two months *p.i.*, which was maintained at four months *p.i.* Interestingly, the PMSS1 strain occasionally became modified during *in vivo* infection.

The *cagPAI* region appears to be genetically unstable, as evidenced by losses and acquisitions during human migration in an evolutionary context (Gressmann *et al.*, 2005). Mechanistically, this region is susceptible to genetic rearrangement and disruption, the most notable being the insertion of element IS605, which causes the gene cluster to break up into two regions, *cag* I and *cag* II (Censini *et al.*, 1996). This insertion has been linked to variability in *H. pylori* virulence (Deguchi *et al.*, 2004; Lai *et al.*, 2013). The *cagPAI* gene cluster encodes 28 and possibly as many as 32 genes (Censini *et al.*, 1996), 22 of which are required for the assembly of the T4SS (Backert *et al.*, 2000; Odenbreit *et al.*, 2000; Fischer *et al.*, 2001). Following the translocation of CagA, tyrosine phosphorylation by Abl and Src family kinases induces the activation of pro-inflammatory signalling pathways. Systematic mutation of the 27 genes comprising the *cagPAI* of strain

26695 identified 17 gene components required for CagA translocation, including *cagW*, which encodes an inner membrane protein (Fischer *et al.*, 2001). Salama *et al.* (2004) constructed a transposon mutant library for the *H. pylori* G27 genome, which we subsequently screened for mutants unable to induce NF- κ B. We identified 59 mutants that showed a decrease in the T4SS-mediated nuclear localisation of p65, 3 of which were located within *cagW* (Belogolova *et al.*, 2013). Here we demonstrate that individual PMSS1 clonal isolates taken at different *in vivo* infection time points had undergone host-induced negative selection to abrogate CagA translocation. For the bacterial isolate Iso6, this alteration was identified as a single nucleotide insertion in *cagW*, resulting in a frame shift mutation that prevented T4SS assembly and CagA translocation. It appears that the overall instability of *cagPAI* in *H. pylori* is indicative of an ambivalent advantage of the encoded T4SS, both in the context of human evolution and during chronic infection of a host organism, notably also mice.

A major driving force that might act against the maintenance of a functional *cagPAI* is the huge inflammatory potential associated with its functional presence (Backert and Naumann, 2010). Numerous processes have been put forward to explain its role in NF- κ B activation and recent observation provided a clear insight into the central role of a heptose biphosphate sugar molecule, HBP (Gall *et al.*, 2017; Stein *et al.*, 2017), in the activation of a novel host pattern recognition pathway (Zimmermann *et al.*, 2017). This pathway is not only fully dependent on a functional *cagPAI* secretion, but also on certain determinants of bacterial LPS synthesis (Stein *et al.*, 2017; Zimmermann *et al.*, 2017), and leads to the formation of TIFAsomes (Gaudet *et al.*, 2015) even in the absence of CagA translocation into the host cells (Zimmermann *et al.*, 2017). These observations suggest that the inflammatory molecule HBP acts as a pathogen-associated molecular pattern (PAMP) that enables the host cells to sense a microbial attack (Koch *et al.*, 2012; Zimmermann *et al.*, 2017). In this light, inflammation caused by the co-secreted molecule HBP may be an unwanted complication of CagA translocation, which is believed to exert multiple functions to the pathogen's advantage (Bauer *et al.*, 2013; Tammer *et al.*, 2007; Tan *et al.*, 2011), explaining the ambivalent benefit and evolutionary instability of the *cagPAI*. As only two mice harboured isolates that had lost the ability to translocate CagA, it is not possible to draw any significant conclusions about whether this adaptation is driven by increased inflammation in the gastric mucosa. The mouse harbouring isolate 6 did indeed show pronounced immune infiltration and metaplasia at four months, while the mouse harbouring isolate 1 did not show any immune infiltration or other pathology at two months. However, it remains uncertain at what point the

mutation arose and how long it would take for the pathology to subside again.

Recently, the PMSS1 strain was shown to encode multiple copies of *cagA* (Jang *et al.*, 2017), which can be expanded and contracted – although the underlying mechanism remains unknown. Moreover, it was shown that an increase in copy number directly correlates with an increase in CagA expression and, concomitantly, in pathogenesis (Jang *et al.*, 2017). This novel feature was also observed by Draper *et al.* (2017) who recently published the complete genome sequences of the PMSS1 and SS1 strains. In the present study, we too sequenced the complete genome of PMSS1 and identified two other notable rearrangements affecting the *cagPAI* gene cluster: (1) an inversion of the *cagPAI* region and (2) a remote location of the *cagA* gene approximately 40 kb downstream of the *cagPAI*, the latter of which was also observed by Draper *et al.* In contrast, we observed only one copy of *cagA*. This could be due to a contraction in the clonal isolate we sequenced or one copy being missed due to the standard sequencing approach utilised (Jang *et al.*, 2017). The general arrangement of the *cagPAI* region in PMSS1 we observed is very similar to that of the gerbil colonising strain B8 (Farnbacher *et al.*, 2010); however, in neither strain does this unique arrangement appear to alter or influence the expression and subsequent translocation of CagA.

Effective colonisation of the gastric mucosa by *H. pylori* involves a number of different genes encoded by the outer membrane Hop family (*babA*, *sabA*, *alpA/B*, *hopZ*, *hopQ* and *oipA*), which bind to receptors expressed on the surface of the gastric epithelium. It appears that of those adhesins, only HopQ has a particular function in facilitating the activation of the *cagPAI* secretion system, possibly by supporting a kind of localised adherence (Belogolova *et al.*, 2013). Interestingly, HopQ was subsequently identified to bind to the CEACAM receptors (Javaheri *et al.*, 2016; Königer *et al.*, 2016), whereas all other Hop family adhesins are thought to preferentially interact with carbohydrates (Aspholm *et al.*, 2006). Interestingly, the murine orthologue of the CEACAM1 receptor does not bind HopQ, due to differences in the amino acid distribution. In the present study, we observed an efficient colonisation of the mouse gastric mucosa with PMSS1, which lacks the adhesins BabA and SabA/B but instead encodes a triple *hopQ* locus at the known SabA/SabB loci. Deletion of both the *hopQ* type II and the second copy of the *hopQ* type I gene prevented the translocation of CagA, indicating a direct interaction with the T4SS. As *hopQ* and *sabB* are paralogues, the finding that in PMSS1 the double *sabB* locus is directly replaced with the double *hopQ* type I locus suggests that the *sabB* also interacts with the T4SS. However, CagA translocation remained unaffected when we deleted *sabB*, demonstrating that there is

a connection between HopQ, the T4SS and CagA translocation that is independent of *sabA/B* and *babA*.

Our observations shed light on numerous issues concerning the nature of the interaction between HopQ and the T4SS and the host-induced negative selection pressure. They also suggest that there are alternative factors driving inflammation and cancerogenesis after CagA translocation is disabled.

Experimental procedures

Bacterial strains and cultivation

The following *H. pylori* strains were used in this study: the *cagPAI*-positive strain P12 was obtained from our in-house collection (Belogolova *et al.*, 2013), first mentioned by Schmitt *et al.* (1994), and G27 (Salama *et al.*, 2004); the mouse colonising strain pre-mouse Sydney Strain 1 (PMSS1) (Arnold *et al.*, 2011) first described by Lee *et al.* (1997) (strain collection numbers P511, P409 and P504) and the mutants P12 Δ PAI, P12 Δ hopQ and G27 Δ cagW (strain collection numbers P387, P508 and P461) (Belogolova *et al.*, 2013). Bacteria were cultured from frozen glycerol stocks on GC agar plates (Oxoid, Wesel, Germany), supplemented with 5% horse serum (Oxoid, Wesel, Germany); vancomycin (10 μ g ml⁻¹); trimethoprim (5 μ g ml⁻¹); and nystatin (1 μ g ml⁻¹) under microaerobic conditions (85% N₂, 10% CO₂ and 5% O₂) at 37 °C for 48 h. *H. pylori* mutant strain P12 Δ PAI was cultivated on agar plates supplemented with kanamycin (8 μ g ml⁻¹), while P12 Δ hopQ and G27 Δ cagW were grown on agar plates supplemented with chloramphenicol (4 μ g ml⁻¹). For complete genome sequencing and *in vivo* infection studies, bacteria were cultured overnight in brain-heart infusion (BHI) medium supplemented with 10% foetal calf serum, vancomycin (10 μ g ml⁻¹), trimethoprim (5 μ g ml⁻¹) and nystatin (1 μ g ml⁻¹) under microaerophilic conditions (CampyPak®) at 37 °C and 110 revolutions per minute (rpm). The PMSS1 Δ hopQ and P12 Δ sabB mutants were constructed by replacing the *hopQ* or *sabB* genes with kanamycin-, chloramphenicol- or erythromycin-resistant gene cassettes.

Gene deletion in *H. pylori* P12 and replacement of the frame shift *cagW* gene with the WT gene of strain PMSS1

Deletion of genes (*hopQ IA, IB, II* and *sabB*) and gene replacement (*cagW*^{mut} to *cagW*^{WT}) was performed as described by Belogolova *et al.* (2013). Briefly, for deletion of genes, the upstream region (0.5 kb) and the downstream region (0.5 kb) of the target gene were amplified using the primer combinations A/B and C/D respectively. Up and downstream fragments were gel-purified, digested with *Bam*HI and ligated. After column purification of the ligation mixture, a PCR was carried out using the ligation product as a template with the primer combinations A/D. The 1 kb products of up and downstream fragments were cloned into pGEM-T-easy vector and separated by the *KanR*-, *ErmE*- or *Cam*-resistant cassettes at the *Bam*HI restriction site. For *cagW* gene replacement,

only the downstream region (0.5 kb) and the complete wild-type (WT) *cagW* gene (1.6 kb) were amplified. The amplified gene and the downstream fragment were digested with *Bam*HI, ligated and amplified as described before. The 2.1 kb product of *cagW*^{WT} ligated to the downstream fragment was cloned into pGEM-T-easy vector and separated by the *KanR* gene at the *Bam*HI restriction site. All plasmids were subsequently transformed into the dam-negative *E. coli* strain GM2199. For gene knockout, isolated plasmid DNA from positive *E. coli* clones was used to transform *H. pylori* PMSS1 WT (strain collection no P504). For gene replacement, isolated plasmid DNA from positive *E. coli* clones was used to transform *H. pylori* Iso6 (strain collection P570). Transformed *H. pylori* were selected on antibiotic-containing agar plates. The gene knockout or gene replacement was confirmed by PCR amplification using primers spanning DNA up- and downstream of sites targeted by primers A and D and the insert confirmed by Sanger sequencing.

Bacterial isolation and colonisation quantification

Serial dilutions of homogenised gastric tissue were plated out in triplicate on GC agar plates supplemented with 5% defibrinated horse blood (Oxoid, Wesel, Germany) containing 0.2% cyclodextrin, amphotericin B (8 μ g ml⁻¹), bacitracin (200 mg/ml), cefsulodin (5 μ g ml⁻¹), nalidixic acid (10 μ g ml⁻¹), 2.5 U ml⁻¹ polymyxin B, trimethoprim (5 μ g ml⁻¹) and vancomycin (100 μ g ml⁻¹) under microaerobic conditions at 37 °C for five days. Colonies exhibiting morphology representative of PMSS1 (pale with an entire margin, ~2 mm ϕ) were quantified. Individual clones successfully cultured from PMSS1-infected mice were selected and preserved as frozen glycerol stocks.

Cell culture and *H. pylori* infection

AGS (ATC® CRL-1739™ human gastric adenocarcinoma cell line) cells were cultured in Gibco® RPMI 1640 medium (Life Technologies, Camarillo, CA, USA) supplemented with 10% FCS (Biochrom, Berlin, Germany) and maintained in a humidified environment at 37 °C and 5% CO₂. 48 h prior to infection, cells were seeded in 12-well plates (TPP, Trasadingen, Switzerland) and serum starved 24 h later. AGS infections were carried out for three hours at MOI 100 unless stated otherwise.

Adherence CFU

AGS cells were infected with *H. pylori* for 90 min at MOI 100. After removing the non-adherent bacteria by washing with RPMI, the adherent bacteria were recovered by detaching cells in RPMI supplemented with 0.5% saponin (Sigma-Aldrich Cat No. S4521). Serial dilutions were plated out in triplicate onto GC agar plates supplemented with the corresponding antibiotics. Bacterial colonies were enumerated and the CFU/ml determined.

P65 translocation assay

The monoclonal p65-GFP expressing AGS cell line SiB02 (Bartfeld *et al.*, 2010) was cultured for *H. pylori* infection as described above. Infections were carried out for 45 minutes at MOI 100 after which the cells were fixed with ice-cold methanol containing Hoechst 33342 at a final dilution of 1:3000 for 20 minutes at 4 °C. Quantification of p65 was performed by automated microscopy as previously described (Bartfeld *et al.*, 2010). Briefly, nine images were taken for each well and the activated nuclei quantified using the Scan^{AR} image analysis software (Olympus).

Immunoblotting

Protein samples were lysed directly in 2× Laemmli buffer (4% (w/v) SDS; 20% (w/v) glycerol, 10% 2-mercaptoethanol, 120 mM Tris-HCl, pH 6.8 and 0.02% bromophenol blue) and denatured by boiling at 95 °C for 5 min. Fifteen microlitres of sample per well were loaded on a 7.5% SDS-PAGE gel. Proteins were transferred to a polyvinylidene difluoride (PVDF) membrane via semi-dry blotting techniques at 60 mA and then blocked for 1 hour in 3% bovine serum albumin (BSA) at room temperature. Membranes were probed overnight with either mouse anti-phospho-Tyr primary antibody (SC-7020, Santa Cruz®, TX, USA) or rabbit anti-CagA primary antibody (SC-25766, Santa Cruz®, TX, USA) at a final dilution of 1:1000. Secondary HRP-labelled anti-mouse or anti-rabbit antibodies (NA931 and NA934, GE Healthcare, Germany) were used at a final dilution of 1:3000. The proteins were detected via chemiluminescence. Anti-β-actin antibody (A5411, Sigma-Aldrich®, Munich, Germany) was used as a loading control at a final dilution of 1:5000. Quantification was performed using ImageJ.

Scanning electron microscopy

For the imaging of *H. pylori*, AGS cells were seeded onto 12-mm glass coverslips and infected with *H. pylori* for 90 minutes at MOI 25. Briefly, cells were fixed with 2.5% glutaraldehyde (Science Services, Munich, Germany, Cat. No. 15949), contrasted with 0.5% osmium tetroxide and 1% tannic acid. The samples were dehydrated using a graded ethanol series and critical-point dried. The sections were evaporation-coated with 3 nm of platinum/carbon and analysed with a Leo 1550 field emission scanning electron microscope (Carl Zeiss, Oberkochen, Germany).

Animals

Six-week-old female C57BL/6 mice (n=38) were purchased from Charles River Laboratories (Sulzfeld, Germany) and housed in individually ventilated cages, with access to food and water *ad libitum* on a 12-hour light/12-hour dark cycle. Prior to infection, all mice received 100 µl NaHCO₃ (100 mM) via oral gavage. Bacterial strains were cultured as described and resuspended in PBS at ~10⁹/100 µl. On two alternate days, two groups of mice were infected via oral gavage with 100 µl of either PMSS1 (n=14) or P12 (n=14), while the remaining group (n=10) was mock infected with PBS.

Mice were sacrificed by cervical dislocation and the stomach resected. Following the removal of the forestomach, an incision was performed along the greater curvature and the stomach was dissected into four longitudinal sections for downstream applications. All *in vivo* experiments were approved (G 0205/12) by the Landesamt für Gesundheit und Soziales (Berlin, Germany) and conducted in accordance with the stipulated animal ethics legal regulations.

Histopathology and scoring

Gastric tissue was fixed in 4% paraformaldehyde for 24 h, embedded in paraffin, cut into 1–2 µm sections, dewaxed and stained for haematoxylin and eosin (H&E). Additional staining of neutral mucins by periodic acid-Schiff (PAS) and acid mucins by Alcian Blue was performed. The gastric tissue was scored in a blinded manner for the following features: presence of *Helicobacter*, level of leukocyte infiltration, atrophy, hyperplasia and metaplasia according to a modified version of the Sydney Scoring System (Dixon *et al.*, 1996). The Sydney system grades human gastritis and was modified for experimental gastritis.

Genome sequencing

The *H. pylori* strain PMSS1 and the *in vivo* clones Iso6 and Iso8 were cultured overnight as described above. Total genomic DNA (gDNA) was isolated using the Qiagen Genomic DNA isolation kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The purity and quality of the gDNA were assessed on a 1% agarose gel and with a nanodrop apparatus (Thermo Fisher Scientific, Wilmington, DE, USA). Extracted DNA of the re-isolates Iso 6 and Iso 8 was used to prepare Nextera XT shotgun libraries for the Genome Analyzer II (Illumina, San Diego, CA, USA) with a 112-bp paired-end sequencing run. For strain PMSS1, genome sequencing was carried out with a hybrid approach using the 454 GS-FLX system with titanium chemistry (Roche Life Science, Mannheim, Germany) and the Genome Analyzer IIx, resulting in a lower contig number and higher sequence accuracy. Libraries were prepared according to the manufacturer's protocol at the Göttingen Genomics Laboratory, Germany. Raw reads were quality controlled with FastQC v0.11.2 (<https://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>) and subsequently trimmed using Trimmomatic 0.32 (<https://www.usadellab.org/cms/?page=trimmomatic>) to remove sequences with quality scores lower than 20 (Illumina 1.9 encoding) and remaining adaptor sequences. De novo assembly was carried out using the SPAdes v3.5 software (Bankevich *et al.*, 2012), resulting in: PMSS1, 26 contigs, comprising 1,590,017 bp; Iso 6, 53 contigs (1,580,715 bp) and Iso 8, 55 contigs (1,578,391 bp). The GenBank accession numbers of the draft genome sequences are AZBR00000000 (PMSS1), AZBQ00000000 (Iso 6) and AZBS00000000 (Iso 8).

Gene prediction, annotation and comparative genome analysis

Open reading frames (ORFs) and tRNAs were identified and annotated using the RAST server (Aziz *et al.*, 2008). For comparative genome analysis, the programs ACT (<https://www.sanger.ac.uk/science/tools/artemis-comparison-tool-act>) and BRIG (Alikhan *et al.*, 2011) were used, resulting in the identification of genome differences respectively. The following *H. pylori* reference genomes were taken from GenBank to build a BRIG genome comparison: strain 26695 [BioProject: PRJNA233] and strain P12 [BioProject: PRJNA32291].

Statistical analysis

All graphs were prepared using GraphPad Prism 7 software (GraphPad Software Inc., La Jolla, CA, USA). *In vivo* data were analysed using the Kruskal-Wallis and Wilcoxon Rank Sum tests using STATA14 software package (STATA Corp., College Station, TX, USA). *In vitro* data are presented as mean \pm SD. Data were considered significant if $p < 0.05$.

Acknowledgements

The authors wish to thank Dr January Weiner for creating the *cagPAI* and outer membrane protein comparison graphics and Dr Rike Zietlow for editing the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

VD, MS, KH and SZ carried out the experiments; VB carried out SEM imaging; AK carried out histopathological staining and scoring; HB assembled the PMSS1 draft genome sequence; HB, MK and TFM designed and supervised the experiment; TFM and VD wrote the manuscript.

References

- Alikhan, N.F., Petty, N.K., Ben Zakour, N.L. and Beatson, S.A. (2011) BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*, **12**, 402.
- Arnold, I.C., Lee, J.Y., Amieva, M.R., Roers, A., Flavell, R.A., Sparwasser, T., *et al.* (2011) Tolerance rather than immunity protects from *Helicobacter pylori*-induced gastric preneoplasia. *Gastroenterology*, **140**, 199–209.
- Aspholm, M., Kalia, A., Ruhl, S., Schedin, S., Arnqvist, A., Linden, S., *et al.* (2006) *Helicobacter pylori* adhesion to carbohydrates. *Methods in Enzymology*, **417**, 293–339.
- Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., *et al.* (2008) The RAST server: rapid annotations using subsystems technology. *BMC Genomics*, **9**, 75.
- Backert, S. and Naumann, M. (2010) What a disorder: proinflammatory signaling pathways induced by *Helicobacter pylori*. *Trends in Microbiology*, **18**, 479–486.
- Backert, S., Ziska, E., Brinkmann, V., Zimny-Arndt, U., Fauconnier, A., Jungblut, P.R., *et al.* (2000) Translocation of the *Helicobacter pylori* CagA protein in gastric epithelial cells by a type IV secretion apparatus. *Cellular Microbiology*, **2**, 155–164.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., *et al.* (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, **19**, 455–477.
- Barrozo, R.M., Cooke, C.L., Hansen, L.M., Lam, A.M., Gaddy, J.A., Johnson, E.M., *et al.* (2013) Functional plasticity in the type IV secretion system of *Helicobacter pylori*. *PLoS Pathogens*, **9**, e1003189.
- Bartfeld, S., Hess, S., Bauer, B., Machuy, N., Ogilvie, L.A., Schuchhardt, J., *et al.* (2010) High-throughput and single-cell imaging of NF-kappaB oscillations using monoclonal cell lines. *BMC Cell Biology*, **11**, 21.
- Bauer, B., Wex, T., Kuester, D., Meyer, T. and Malfertheiner, P. (2013) Differential expression of human beta defensin 2 and 3 in gastric mucosa of *Helicobacter pylori*-infected individuals. *Helicobacter*, **18**, 6–12.
- Belogolova, E., Bauer, B., Pompaiah, M., Asakura, H., Brinkman, V., Ertl, C., *et al.* (2013) *Helicobacter pylori* outer membrane protein HopQ identified as a novel T4SS-associated virulence factor. *Cellular Microbiology*, **15**, 1896–1912.
- Bonsor, D.A., Zhao, Q., Schmidinger, B., Weiss, E., Wang, J., Deredge, D., *et al.* (2018) The *Helicobacter pylori* adhesin protein HopQ exploits the dimer interface of human CEACAMs to facilitate translocation of the oncoprotein CagA. *The EMBO Journal*, **37**, pii: e98664.
- Cao, P. and Cover, T.L. (2002) Two different families of hopQ alleles in *Helicobacter pylori*. *Journal of Clinical Microbiology*, **40**, 4504–4511.
- Censini, S., Lange, C., Xiang, Z., Crabtree, J.E., Ghiara, P., Borodovsky, M., *et al.* (1996) *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. *Proceedings of the National Academy of Sciences*, **93**, 14648–14653.
- Deguchi, R., Igarashi, M., Watanabe, K. and Takagi, A. (2004) Analysis of the *cag* pathogenicity island and IS605 of *Helicobacter pylori* strains isolated from patients with gastric cancer in Japan. *Alimentary Pharmacology and Therapeutics*, **20**(Suppl 1), 13–16.
- Dixon, M.F., Genta, R.M., Yardley, J.H. and Correa, P. (1996) Classification and grading of gastritis. *The American Journal of Surgical Pathology*, **20**, 1161–1181.
- Draper, J.L., Hansen, L.M., Bernick, D.L., Abedrabbo, S., Underwood, J.G., Kong, N., *et al.* (2017) Fallacy of the unique genome: sequence diversity within single *Helicobacter pylori* strains. *MBio*, **8**, pii: e02321-16.
- Farnbacher, M., Jahns, T., Willrodt, D., Daniel, R., Haas, R., Goesmann, A., *et al.* (2010) Sequencing, annotation, and comparative genome analysis of the gerbil-adapted *Helicobacter pylori* strain B8. *BMC Genomics*, **11**, 335.

- Fischer, W., Puls, J., Buhrdorf, R., Gebert, B., Odenbreit, S. and Haas, R. (2001) Systematic mutagenesis of the *Helicobacter pylori* cag pathogenicity island: essential genes for CagA translocation in host cells and induction of interleukin-8. *Molecular Microbiology*, **42**, 1337–1348.
- Gall, A., Gaudet, R.G., Gray-Owen, S.D. and Salama, N.R. (2017) TIFA signaling in gastric epithelial cells initiates the cag type 4 secretion system-dependent innate immune response to *Helicobacter pylori* infection. *MBio*, **8**, pii: e01168-17.
- Gaudet, R.G., Sintsova, A., Buckwalter, C.M., Leung, N., Cochrane, A., Li, J., *et al.* (2015) Cytosolic detection of the bacterial metabolite HBP activates TIFA-dependent innate immunity. *Science*, **348**, 1251–1255.
- Gomez-Duarte, O.G., Lucas, B., Yan, Z.X., Panthel, K., Haas, R. and Meyer, T.F. (1998) Protection of mice against gastric colonization by *Helicobacter pylori* by single oral dose immunization with attenuated *Salmonella typhimurium* producing urease subunits A and B. *Vaccine*, **16**, 460–471.
- Gressmann, H., Linz, B., Ghai, R., Pleissner, K.P., Schlapbach, R., Yamaoka, Y., *et al.* (2005) Gain and loss of multiple genes during the evolution of *Helicobacter pylori*. *PLoS Genetics*, **1**, e43.
- Higashi, H., Tsutsumi, R., Muto, S., Sugiyama, T., Azuma, T., Asaka, M., *et al.* (2002) SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein. *Science*, **295**, 683–686.
- Ishijima, N., Suzuki, M., Ashida, H., Ichikawa, Y., Kanegae, Y., Saito, I., *et al.* (2011) BabA-mediated adherence is a potentiator of the *Helicobacter pylori* type IV secretion system activity. *Journal of Biological Chemistry*, **286**, 25256–25264.
- Israel, D.A., Salama, N., Arnold, C.N., Moss, S.F., Ando, T., Wirth H.P. *et al.* (2001) *Helicobacter pylori* strain-specific differences in genetic content, identified by microarray, influence host inflammatory responses. *Journal of Clinical Investigation*, **107**, 611–620.
- Jang, S., Su, H., Blum, F.C., Bae, S., Choi, Y.H., Kim, A., *et al.* (2017) Dynamic expansion and contraction of cagA copy number in *Helicobacter pylori* Impact Development of Gastric Disease. *MBio*, **8**, pii: e01779-16.
- Javaheri, A., Kruse, T., Moonens, K., Mejias-Luque, R., Debraekeleer, A., Asche, C.I., *et al.* (2016) *Helicobacter pylori* adhesin HopQ engages in a virulence-enhancing interaction with human CEACAMs. *Nature Microbiology*, **2**, 16189.
- Koch, M., Mollenkopf, H.J., Klemm, U. and Meyer, T.F. (2012) Induction of microRNA-155 is TLR- and type IV secretion system-dependent in macrophages and inhibits DNA-damage induced apoptosis. *Proceedings of the National Academy of Sciences*, **109**, E1153–E1162.
- Königer, V., Holsten, L., Harrison, U., Busch, B., Loell, E., Zhao, Q., *et al.* (2016) *Helicobacter pylori* exploits human CEACAMs via HopQ for adherence and translocation of CagA. *Nature Microbiology*, **2**, 16188.
- Kusters, J.G., van Vliet, A.H. and Kuipers, E.J. (2006) Pathogenesis of *Helicobacter pylori* infection. *Clinical Microbiology Reviews*, **19**, 449–490.
- Lai, C.H., Perng, C.L., Lan, K.H. and Lin, H.J. (2013) Association of IS605 and cag-PAI of *Helicobacter pylori* isolated from patients with gastrointestinal diseases in Taiwan. *Gastroenterology Research and Practice*, **2013**, 356217.
- Lee, A., O'Rourke, J., De Ungria, M.C., Robertson, B., Daskalopoulos, G. and Dixon, M.F. (1997) A standardized mouse model of *Helicobacter pylori* infection: introducing the Sydney strain. *Gastroenterology*, **112**, 1386–1397.
- Lind, J., Backert, S., Pfliegerer, K., Berg, D.E., Yamaoka, Y., Sticht, H., *et al.* (2014) Systematic analysis of phosphotyrosine antibodies recognizing single phosphorylated EPIYA-motifs in CagA of Western-type *Helicobacter pylori* strains. *PLoS One*, **9**, e96488.
- Linz, B., Balloux, F., Moodley, Y., Manica, A., Liu, H., Roumagnac, P., *et al.* (2007) An African origin for the intimate association between humans and *Helicobacter pylori*. *Nature*, **445**, 915–918.
- Lowenthal, A.C., Hill, M., Sycuro, L.K., Mehmood, K., Salama, N.R. and Ottemann, K.M. (2009) Functional analysis of the *Helicobacter pylori* flagellar switch proteins. *Journal of Bacteriology*, **191**, 7147–7156.
- Odenbreit, S., Puls, J., Sedlmaier, B., Gerland, E., Fischer, W. and Haas, R. (2000) Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. *Science*, **287**, 1497–1500.
- Olbermann, P., Josenhans, C., Moodley, Y., Uhr, M., Stamer, C., Vauterin, M., *et al.* (2010) A global overview of the genetic and functional diversity in the *Helicobacter pylori* cag pathogenicity island. *PLoS Genetics*, **6**, e1001069.
- Pfannkuch, L., Hurwitz, R., Traulsen, J., Kosma, P., Schmid, M. and Meyer, T. (2018). ADP heptose, a novel pathogen associated molecular pattern associated with *Helicobacter pylori* type 4 secretion. bioRxiv doi: <https://doi.org/10.1101/405951>.
- Philpott, D.J., Belaid, D., Troubadour, P., Thiberge, J.M., Tankovic, J., Labigne, A., *et al.* (2002) Reduced activation of inflammatory responses in host cells by mouse-adapted *Helicobacter pylori* isolates. *Cellular Microbiology*, **4**, 285–296.
- Salama, N.R., Shepherd, B. and Falkow, S. (2004) Global transposon mutagenesis and essential gene analysis of *Helicobacter pylori*. *Journal of Bacteriology*, **186**, 7926–7935.
- Schmitt, W. and Haas, R. (1994) Genetic analysis of the *Helicobacter pylori* vacuolating cytotoxin: structural similarities with the IgA protease type of exported protein. *Molecular Microbiology*, **12**, 307–319.
- Stein, S.C., Faber, E., Bats, S.H., Murillo, T., Speidel, Y., Coombs, N., *et al.* (2017) *Helicobacter pylori* modulates host cell responses by CagT4SS-dependent translocation of an intermediate metabolite of LPS inner core heptose biosynthesis. *PLoS Pathogens*, **13**, e1006514.
- Talarico, S., Whitefield, S.E., Fero, J., Haas, R. and Salama, N.R. (2012) Regulation of *Helicobacter pylori* adherence by gene conversion. *Molecular Microbiology*, **84**, 1050–1061.
- Tammer, I., Brandt, S., Hartig, R., König, W. and Backert, S. (2007) Activation of Abl by *Helicobacter pylori*: a novel kinase for CagA and crucial mediator of host cell scattering. *Gastroenterology*, **132**, 1309–1319.
- Tan, S., Noto, J.M., Romero-Gallo, J., Peek, R.M. Jr and Amieva, M.R. (2011) *Helicobacter pylori* perturbs iron

- trafficking in the epithelium to grow on the cell surface. *PLoS Pathogens*, **7**, e1002050.
- Treangen, T.J., Ondov, B.D., Koren, S. and Phillippy, A.M. (2014) The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biology*, **15**, 524.
- Viala, J., Chaput, C., Boneca, I.G., Cardona, A., Girardin, S.E., Moran, A.P., *et al.* (2004) Nod1 responds to peptidoglycan delivered by the *Helicobacter pylori* cag pathogenicity island. *Nature Immunology*, **5**, 1166–1174.
- Zimmermann, S., Pfannkuch, L., Al-Zeer, M.A., Bartfeld, S., Koch, M., Liu, J., *et al.* (2017) ALPK1- and TIFA-dependent innate immune response triggered by the *Helicobacter pylori* type IV secretion system. *Cell Reports*, **20**, 2384–2395.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.