Phasevarion Mediated Epigenetic Gene Regulation in Helicobacter pylori

Yogitha N. Srikhanta1,2,*, Rebecca J. Gorrell2,3,4,*, Jason A. Steen5, Jayde A. Gawthorne1, Terry Kwok4, Sean M. Grimmond5, Roy M. Robins-Browne2,3, Michael P. Jennings1,6

1 Institute for Glycomics, Griffith University, Gold Coast, Queensland, Australia, 2 Department of Microbiology and Immunology, University of Melbourne, Melbourne, Victoria, Australia, 3 Murdoch Childrens Research Institute, Royal Children’s Hospital, Melbourne, Victoria, Australia, 4 Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Victoria, Australia, 5 Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia

Abstract

Many host-adapted bacterial pathogens contain DNA methyltransferases (mod genes) that are subject to phase-variable expression (high-frequency reversible ON/OFF switching of gene expression). In Haemophilus influenzae and pathogenic Neisseria, the random switching of the modA gene, associated with a phase-variable type III restriction modification (R-M) system, controls expression of a phase-variable regulon of genes (a “phasevarion”), via differential methylation of the genome in the modA ON and OFF states. Phase-variable type III R-M systems are also found in Helicobacter pylori, suggesting that phasevarions may also exist in this key human pathogen. Phylogenetic studies on the phase-variable type III modH gene revealed that there are 17 distinct alleles in H. pylori, which differ only in their DNA recognition domain. One of the most commonly found alleles was modH5 (16% of isolates). Microarray analysis comparing the wild-type P12modH5 ON strain to a P12ΔmodH5 mutant revealed that six genes were either up- or down-regulated, and some were virulence-associated. These included flaA, which encodes a flagella protein important in motility and hopG, an outer membrane protein essential for colonization and associated with gastric cancer. This study provides the first evidence of this epigenetic mechanism of gene expression in H. pylori. Characterisation of H. pylori modH phasevarions to define stable immunological targets will be essential for vaccine development and may also contribute to understanding H. pylori pathogenesis.

Introduction

The host-adapted pathogen Helicobacter pylori is the most common cause of bacterial infection worldwide [1,2] and is an important etiologic agent of gastritis, peptic ulcers, and gastric cancer [3,4]. Unless treated, colonization usually persists for life, indicating that H. pylori is well adapted to the gastric environment.

In order to adapt its physiology to its environment and ensure survival, H. pylori has evolved molecular mechanisms for generating genetic variation [5]. One mechanism is phase-variation, which is the high frequency reversible on/off switching of gene expression. Phase-variation is commonly mediated by mutations in simple tandem DNA repeats in the open reading frame or promoter region of genes encoding surface expressed virulence determinants [6]. The independent, random switching of these genes results in phenotypically diverse populations that can rapidly adapt to host environments and evasion of immune responses [7]. While phase-variation is typically associated with genes encoding surface structures, several host-adapted bacterial pathogens, including H. pylori, have DNA methyltransferases (mod genes) associated with type III restriction modification (R-M) systems that contain simple tandem DNA repeats which have been proven to phase vary [5,8,9].

R-M systems are ubiquitous in bacteria and confer protection to the bacterial host against invasion by foreign DNA [10]. R-M systems are classified into three groups: Types I, II or III on the basis of their subunit composition, DNA cleavage position, sequence-specificity and co-factor requirements [11]. Type III systems are composed of a methyltransferase (modification, mod) gene and an endonuclease (restriction, res) gene, whose products form a two-subunit enzyme – Mod and Res [12]. In Type III systems Res must form a complex with Mod to be functional [13], although, Mod can function independently of Res [14]. The Mod subunit contains several conserved motifs in the N- and C-terminal regions and the central region contains the DNA-recognition domain that dictates sequence specificity [15].

We have recently shown that in three human pathogens (Haemophilus influenzae [16], Neisseria gonorrhoeae and Neisseria meningitidis [17]) the random switching of the modA gene controls expression of a phase-variable regulon of genes (a “phasevarion”), via differential methylation of the genome in the modA ON and OFF states. In this study we investigated the modH gene, a phase-
variable DNA methytransferase of \textit{H. pylori}, to determine if it plays a role in gene regulation.

\textbf{Results}

\textbf{Seventeen modH alleles present in \textit{H. pylori}}

To investigate if the \textit{modH} gene associated with a type III R-M system of \textit{H. pylori} behaves as a phasevarion [16,17], we first carried out phylogenetic analysis of \textit{modH}. We have previously reported that \textit{Helicobacter} strains, like the pathogenic \textit{Neisseria}, contain multiple phase-variable type III R-M systems [9]. We defined one of these phase-variable type III R-M systems as \textit{modH}. As with \textit{H. influenzae} and the pathogenic \textit{Neisseria} [16,17], for each \textit{mod} gene of \textit{H. pylori} there are distinct alleles that differ only in their DNA recognition domain ([15]; see Figure 1A). Comparison of the fully sequenced and annotated \textit{H. pylori} genomes available at the time revealed that there were four distinct alleles of \textit{modH} based on differences in their DNA recognition domain [9]. The \textit{modH} gene contains tracts of simple tandem guanosine repeats that mediate phase-variation of \textit{mod} gene expression.

To investigate whether additional alleles of \textit{modH} are present in \textit{H. pylori}, and to examine the distribution of \textit{modH} alleles and their repeat sequence type and number, sequence analysis of a genetically diverse set of \textit{H. pylori} clinical isolates was performed. This analysis revealed that all strains examined contained the \textit{modH} gene, with thirteen additional \textit{modH} alleles observed, bringing to seventeen the total number of \textit{modH} alleles observed (Table 1, Figure 1B). Here we define \textit{modH} alleles of the same group as sharing more than 90% amino acid identity with other group members in a global pairwise alignment. In previous studies of \textit{mod} genes inter-allelic diversity in the DNA recognition domain is very high with less than 30% amino acid identity shared by DNA recognition domain regions of different groups [16,17]. In contrast in comparison of the \textit{modH} alleles, there are two groups of alleles (\textit{modH}5/\textit{modH}11/\textit{modH}12 and \textit{modH}3/\textit{modH}10) that share large regions of high identity, separated by regions with little or no identity, a pattern suggestive of recombination. The most frequently occurring \textit{modH} alleles observed in the strain survey are \textit{modH}3 (16% of isolates tested) and \textit{modH}5 (16% of isolates tested). Sequencing of the repeat region of the \textit{modH} alleles revealed that the repeat numbers vary in length between 9 and 15 bp in different strains, resulting in the \textit{modH} gene being in-frame (ON) or out-of-frame (OFF) for expression; consistent with phase-variation of the \textit{modH} gene in this collection of strains (Table 1). Sequence analysis of the corresponding \textit{res} genes in the strain collection revealed that strains F32 (\textit{modH}2), and 52 (\textit{modH}11) have nonsense mutations resulting in truncation of the \textit{res} gene, while strain 51 (\textit{modH}5) has a base pair missing resulting in a frame-shift mutation (Figure 1A, Table 1). Furthermore, for strains 2047 (\textit{modH}2), L101 (\textit{modH}3) and GN760 (\textit{modH}11) a full-length \textit{res} gene could not be amplified (Table 1).

Representatives of all 17 \textit{modH} alleles were compared by multiple sequence alignment. Figure 1B illustrates the diversity seen throughout the DNA recognition domain of the \textit{modH} alleles at the nucleotide level. There are several short regions of high similarity within the DNA recognition domain. The mosaic pattern observed in the alignment schematic (Figure 1B) suggests that large segments have been deleted or replaced via genetic recombination. To examine recombination within the DNA recognition domains in more detail and to determine its contribution to \textit{modH} allele diversity, we undertook all versus all \textit{BLASTn} searches using the 13 representative \textit{modH} sequences (Figure 1C). The number of reciprocal exchanges identified gave a clear indication that the \textit{modH} alleles have recombined in the past. By using this approach we identified new relationships between the \textit{modH} alleles. Some alleles were found to have undergone recombination more readily than others to generate new alleles. For example, \textit{modH}5, \textit{modH}11 and \textit{modH}12 share near identical 5’ and 3’ regions, but each has a different central fragment suggesting recent origin from an ancestral allele (Figure 1C).

Analysis of differentially expressed genes in the \textit{H. pylori} \textit{modH5} phasevarion

To determine whether phase-variation of the \textit{modH} allele in \textit{H. pylori} resulted in changes in global gene expression, we conducted a study using \textit{H. pylori} strain P12, which expresses the most common \textit{modH}5 allele. The \textit{modH}5 allele was inactivated by insertion of a cat cassette to make the mutant strain P12\textit{modH}5. Wild-type P12\textit{modH}5 ON and P12\textit{ΔmodH}5 were compared by microarray analysis using \textit{H. pylori} P12 genome arrays (Methods). Six genes were found to be differentially expressed by a ratio of 1.6-fold or more, with 2 genes up-regulated in P12\textit{ΔmodH}5 relative to wild-type and 4 genes down-regulated. These data confirmed that \textit{modH}5 phase-variation has an influence on gene expression (Table 2). One gene with an increased expression of 2.4 fold in the \textit{modH} mutant encodes the surface exposed protein, HopG (also known as HopY), a potential vaccine candidate [18]. HopG is required for colonization [19] and may be associated with gastric cancer [20]. Two genes associated with flagella showed increased expression in the \textit{modH} ON strain. FlaA is the major component of the flagellar filament [21,22] and is required for normal motility, which is essential for colonization [23] and the establishment of persistent infection [24]. HPP12_904 is homologous to the flagellar hook-length control protein FlIK that is also essential for motility [25].

Quantitative real time PCR (QRT-PCR) confirmed that \textit{flaA} and HPP12_904 were expressed at a higher level in the wild-type P12\textit{modH}5 ON parent strain compared to P12\textit{ΔmodH}5, while hopG and HPP12_0255 were expressed at a higher level in the P12\textit{ΔmodH}5 mutant compared to P12\textit{modH}5 ON (Table 2). QRT-PCR that compared the wild-type P12\textit{modH}5 ON (G\textsubscript{10}) strain to a P12\textit{modH}5 OFF strain (G\textsubscript{0}/G\textsubscript{4}), also confirmed the microarray results (Table 2).

\textbf{Discussion}

We recently confirmed gene regulation as a function of phase-variable type III R-M systems in the human pathogens \textit{H. influenzae} [16] and pathogenic \textit{Neisseria} [17], thus defining a new paradigm in bacterial gene regulation “the phase-variable regulon; “phasevarioni” [9]. In this study we investigated whether the
Figure 1. Sequence analysis of the 17 H. pylori modH alleles. (A) Diagrammatic representation of the mod and res genes of H. pylori. The methyltransferase gene (mod) and restriction endonuclease (res) genes and the repeat region that mediates phase-variation are indicated. Type III R-M system conserved motifs are also shown: in mod, the catalytic region (DPPY), the AdoMet binding pocket (FxGxG), and the DNA recognition domain (DRD); in res, the ATP binding motif (TGxGKT), and ATP hydrolysis motif (DEPH) and the endonuclease domain. A black circle indicates the position of a nonsense mutation or frame-shift mutation in res. (B) The variable regions for each of the 17 modH alleles in the multiple sequence
alignment were aligned in ClustalW and visualised with JalView using the overlay feature. The nucleotides are represented as vertical bars colored according to consensus identity (dark blue >90% identity; light blue >50% identity; white <50% identity or gap). The modH alleles were from the following H. pylori strains (listed in Table 1); modH1 BH13, modH2 1061, modH3 11637, modH4 GN760, modH5 2A, modH6 3A, modH7 CHP7, modH8 CHP2, modH9 CHP4, modH10 219, modH11 GN760, modH12 L252, modH13 L264, modH14 SouthAfrica7, modH15 Gambia 94/2A, modH16 Cuz20 and modH17 908. To generate the amino acid sequences of the DRD region for this comparison the modH genes were translated, starting and ending with the residues corresponding to amino acid residues 214 and 449, respectively, with reference to the sequence of the H. pylori P12 type III methyltransferase (gb ACJ08645.1). (C) The 17 modH alleles are shown as coloured lines. BLASTn matches longer than 20 nucleotides and >90% identity between the 17 modH alleles were mapped as a box onto the corresponding allele in the appropriate colour. Detailed information on each individual coloured box is provided in Figure S1 and Table S1. (D) Diagrammatic representation of the tBLASTn match between modH3 and modH10. The nucleotides are represented as vertical bars (dark blue >90% identity; light blue >50% identity; white <50% identity or gap). The numbers below the figure indicate the percent identity as defined by BLASTn for the area between the double-headed arrows.

doi:10.1371/journal.pone.0027569.g001

Phase-variable type III R-M systems described in H. pylori serve a similar regulatory function.

H. pylori contain a number of phase-variable type II and type III R-M systems [9,26,27], as well as several active orphan type methyltransferases [28,29,30], suggesting that these methyltransferases may have functions other than restriction, such as gene regulation [31]. Additionally, H. pylori strains have been found to have inactive type III and II res genes, indicating that DNA restriction may not be the function of phase-variable R-M systems [8,31].

Furthermore, by putative phase-variable mod gene associated with a type II R-M system in H. pylori (M.HpyAIV) was shown to influence gene expression of katA [31]. The M.HpyAIV gene has also been associated with the induction of a more robust host response in mice, suggesting an involvement in gene regulation [32].

In type III R-M systems DNA sequence specificity is conferred by the Mod subunit [15]. Differences in the modA DNA recognition domain have previously been observed in H. influenzae with 17 distinct modA alleles defined in this organism [33,34]. In pathogenic Neisseria, we identified three distinct modA alleles and two distinct modB alleles [17]. Our phylogenetic studies on the modH gene of a collection of H. pylori clinical isolates revealed that there are 17 distinct modH alleles based on differences in their DNA recognition domain. In pathogenic Neisseria, strains with the same DNA recognition domain regulated the same set of genes, while those with different mod alleles regulated the expression of different sets of genes [17]. Seventeen distinct modH alleles in H. pylori suggest that 17 distinct phasevarions exist. ModH3 and modH5 were the most frequent alleles observed in clinical isolates.

Further studies on the modH5 phasevarion were conducted using microarray to compare H. pylori strain PI2, which expresses the modH5 allele, to a mutant strain. Of the genes regulated by the modH5 phasevarion, two encode proteins that have important roles in motility, FlaA and FlhK. Motility is an essential factor for the colonization and persistence of H. pylori in the human stomach [24] and therefore flagella have an important role in virulence. In addition, H. pylori FlaA has low intrinsic capacity to activate innate immunity via the Toll-like receptor 5 [35,36,37]. Therefore, altered expression of flagella may be advantageous for the adaptation of H. pylori to alternate host environments and in evading the host immune response. The gene encoding the essential outer membrane protein for colonization, HopG [19] was also found to be regulated by the modH5 phasevarion. Bacterial adherence mediated by HopG and outer membrane proteins is thought to play an important role in the colonization of the gastric epithelium by H. pylori [38], making HopG an attractive vaccine target [18]. Hence phasevarion mediated phase-variation of hopG has the potential to mediate escape from the host immune response.

Only a relatively small number of genes were found to be under the control of the modH5 phasevarion. This may be the full extent of the regulon, or only be a sub-set of the regulon due to the analysis being done under standard in vitro culture conditions. Differences in gene expression can only be detected if the genes in question are being expressed. Using different physiologically relevant conditions, such as specific pH conditions that reflect the gastric environment, may result in more genes being found to be under the influence of the modH5 phasevarion.

Here we provide evidence for a role for phase-variable mod genes associated with type III R-M systems in gene regulation in H. pylori. Although we cannot exclude another as yet undescribed role for these modH phasevarion in H. pylori biology, we have confirmed phasevarion mediated epigenetic mechanism of gene expression does operate in H. pylori. Further characterisation of this phasevarion will contribute to an improved understanding of H. pylori pathogenesis and may guide vaccine development for this important human pathogen by defining stably expressed immunological targets in modH5 strains. Based on our previous studies [9] it is likely that other H. pylori modH alleles also function as phasevarions and regulate gene expression. Future studies in H. pylori involving gene regulation, host/pathogen interactions or vaccine development need to control for the potential for modH phase variation to alter global gene expression.

Materials and Methods

Bacterial strains and growth conditions

H. pylori strains were routinely grown from glycerol stocks for 2 days on GC agar (Oxoid, Basingstoke, UK) plates supplemented with 10% (v/v) horse serum (Invitrogen Corp, Carlsbad, CA), vitamin mix and antibiotics (nystatin, 20 mg/ml; trimethoprim, 2.5 mg/ml; vancomycin, 10 mg/ml) in a microaerobic atmosphere as described previously [39]. Plates for cultivation of mutant strains were further supplemented with chloramphenicol (4 mg/ml for routine culture, 10 mg/ml for selection of transformants).

DNA manipulation and analysis

All enzymes were sourced from New England Biolabs. Sequencing was performed on PCR products using QiaQuick PCR purification kit (Qiagen) and Big-Dye (Perkin Elmer) sequencing kits. Data were analysed using MacVector v11.0 (Accelrys).

ModH alleles (formerly called ModC [9]) were classified as ModH1-13 according to all-versus-all global pairwise amino acid alignments of the ModH DRD region and a within-group minimum identity cut-off of 90%. Global pairwise alignments were calculated with a dynamic programming technique as implemented in Jalview (http://www.jalview.org/) (alignment parameters: BLOSUM62 substitution matrix, gap-open penalty of 12, gap-extend penalty of 2). Inter-allele comparisons were carried out using all versus all BLASTn and BLASTp comparisons of representative ModH DRD alleles using stand-alone NCBI BLAST without filters (version 2.2.18). Amino acid and nucleotide
| Strain   | modH allele | G tract repeat number | Origin               | Sequence source (Accession number) |
|----------|-------------|-----------------------|----------------------|-------------------------------------|
| L203     | modH1       | 13 (ON)               | The Netherlands      | This study (HQ734252)               |
| BH13     | modH1       | 12 (OFF)              | Brazil               | This study (HQ734242)               |
| 35A      | modH1       | 8 (OFF)               | Norway               | CP002096.1 (HMPREF4655_20115)       |
| PeCan4   | modH1       | 9 (OFF)               | Cancer Patient       | CP002074.1 (HPPC_07455)             |
| J99      | modH1       | 11 (OFF)              | USA                  | NC_000921.1 (hpj1411)              |
| 2047d    | modH2       | 12 (OFF)              | The Netherlands      | This study (HQ734240)               |
| 1061     | modH2       | 12 (OFF)              | Canada               | This study (HQ734238)               |
| 163(A)   | modH2       | 11 (OFF)              | Brazil               | This study (HQ734234)               |
| CHP1     | modH2       | 13 (ON)               | Australia            | This study (HQ734243)               |
| 26695    | modH2       | 12 (OFF)              | UK                   | NC_000915.1 (HP1522)               |
| 2022     | modH3       | 15 (OFF)              | The Netherlands      | This study (HQ734239)               |
| L101d    | modH3       | 9 (OFF)               | The Netherlands      | This study (HQ830157)               |
| 11637    | modH3       | 11 (OFF)              | Australia            | This study (HQ734241)               |
| L71      | modH3       | 10 (ON)               | The Netherlands      | This study (HQ734251)               |
| 551      | modH3       | 12 (OFF)              | Mouse-adapted strain| This study (HQ830158)               |
| F30      | modH3       | 14 (OFF)              | East Asia            | BAJ57491.1 (HPF30_1394)             |
| S1       | modH3       | 14 (OFF)              | Korea                | CP000012.1 (KHP_1374)              |
| 83       | modH3       | 12 (OFF)              | USA                  | CP002065.1 (HMPREF0462_1519)       |
| HPAG_1   | modH4       | 13 (ON)               | Sweden               | CP000241.1 (HPAG1_1393)            |
| 1134     | modH4       | 14 (OFF)              | Canada               | This study (HQ734257)               |
| Sat464   | modH4       | 10 (ON)               | Peru                 | CP00207.1 (HPSAT_07320)            |
| 98-10    | modH4       | 11 (OFF)              | Japan                | ABSX01000015.1 (HP9810_885g17)     |
| HPG27    | modH4       | 11 (OFF)              | Italy                | CP00173.1 (HPG27_1444)             |
| 2025     | modH5       | 15 (OFF)              | The Netherlands      | This study (JN974761)               |
| CHP5     | modH5       | 12 (OFF)              | Australia            | This study (HQ734246)               |
| KC7617   | modH5       | 11 (OFF)              | Canada               | This study (HQ734250)               |
| 5A       | modH5       | 10 (ON)               | The Netherlands      | This study (HQ734233)               |
| P12      | modH5       | 10 (ON)               | Germany              | CP001217.1 (HP12_1497)             |
| 2A       | modH5       | 9 (OFF)               | The Netherlands      | This study (HQ734231)               |
| HPB8     | modH5       | 13 (ON)               | Gerbil-adapted strain| NC_014257 (HPB8_7)                 |
| India7   | modH5       | 14 (OFF)              | India                | CP002331.1 (HPIN_07505)            |
| L251     | modH5       | 9 (OFF)               | The Netherlands      | This study (HQ734253)               |
| 3A       | modH6       | 12 (OFF)              | The Netherlands      | This study (HQ734232)               |
| L2624    | modH6       | 12 (OFF)              | The Netherlands      | This study (HQ734256)               |
| Shi470   | modH6       | 10 (ON)               | Asia/South America   | CP001072.2 (HPSh_07815)            |
| F32"     | modH6       | 12 (OFF)              | East Asia            | BAJS8990.1 (HPF32_1408)            |
| F57      | modH6       | 11 (OFF)              | East Asia            | BAJ60509.1 (HPF57_1435)            |
| Lithuania75 | modH6 | 12 (OFF) | Lithuania            | CP002334.1 (HPLT_07575)          |
| 758TM    | modH6       | 9 (OFF)               | Canada               | This study (HQ734236)               |
| CHP7     | modH7       | 15 (OFF)              | Australia            | This study (HQ734248)               |
| A5620    | modH7       | 10 (ON)               | Canada               | This study (JN974762)               |
| CHP2     | modH8       | 11 (OFF)              | Australia            | This study (HQ734244)               |
| CHP4     | modH9       | 13 (ON)               | Australia            | This study (HQ734245)               |
| L80      | modH10      | 11 (OFF)              | The Netherlands      | This study (JN974763)               |
| 219      | modH10      | 9 (OFF)               | Brazil               | This study (HQ734235)               |
| SIM180   | modH10      | 9 (OFF)               | Peru                 | NC_014560.1 (HPSIM_07770)          |
| CHP6     | modH10      | 10 (ON)               | Australia            | This study (HQ734247)               |
| GN780"   | modH11      | 11 (OFF)              | Canada               | This study (HQ734237)               |
| F16      | modH11      | 12 (OFF)              | East Asia            | AP011940.1 (HPF16_1417)            |
| 52"      | modH11      | 11 (OFF)              | Korea                | CP001680.1 (HPKB_1423)             |
sequences were aligned using ClustalX (version 2.0.11). Multiple alignments were viewed and edited in Jalview (42). The GenBank accession numbers are HQ734231–HQ734257, HQ830157–HQ830158 and JN974761–JN974763.

mod and res specific PCR

The modH gene and resH gene were amplified and sequenced using the primers listed in Table 3. H. pylori clinical isolates were used as templates (Table 1). The reaction was performed in 50 μl using KOD (Novagen) reagents, and 1 unit of KOD DNA polymerase with the following cycling conditions for the modH gene: 30 cycles of 94°C for 30 sec, 50°C for 30 sec, 70°C for 1 min and 1 cycle of 72°C for 5 min with 5 μM of the primer pair HP_MODHF1 and HP_MODHR5. A 693 bp region containing the DNA recognition domain (603 bp downstream of the DNA recognition region was 90% identical at the amino acid and nucleotide level to the modH allele. A strain was defined as having a particular modH allele if the DNA recognition region was 90% identical at the amino acid and nucleotide level to the modH allele. Refer to Figure 1. The resH gene contains a nonsense mutation (strain F32, nucleotide 1587 and strain 52, nucleotide 537 change from G to T) or missing base pair resulting in a frame-shift mutation (strain 51, nucleotide 2054).

Table 1. Differentially expressed genes in H. pylori wild-type P12 modH5 ON versus the mutant strain P12ΔmodH5.

| Strain | modH allele | G tract repeat number | Origin | Sequence source |
|--------|-------------|-----------------------|--------|----------------|
| CHP8   | modH11      | 10 (ON)               | Australia | This study (HQ734249) |
| L252   | modH12      | 10 (ON)               | The Netherlands | This study (HQ734254) |
| L264   | modH13      | 13 (OFF)              | The Netherlands | This study (HQ734255) |
| SouthAfrica7 | modH14    | 9 (OFF)               | South Africa | CP002336.1 (HPSA_07265) |
| Gambiat94/24 | modH15    | 9 (OFF)               | Gambia | CP002332.1 (HPGAM_08025) |
| Cuz20  | modH16      | 11 (OFF)              | Peru    | CP002076.1 (HPCU_07650) |
| 908    | modH17      | 12 (OFF)              | Africa  | CP002184.1 (hp908_1508) |

*Number and expression state of poly-guanosine repeats within the modH allele; in-frame (ON) or out-of-frame (OFF).

Table 2. Differentially expressed genes in H. pylori wild-type P12 modH5 ON versus the mutant strain P12ΔmodH5.

| Gene ID | Gene Name            | Ratio | B-Stat | QRT-PCR |
|---------|----------------------|-------|--------|---------|
| Reduced expression in H. pylori strain P12 modH5 mutant |
| HPP12_1497 | type III R-M system methyltransferase | -5.94 | 6.350 |
| HPP12_0609 | flagellin A          | -1.97 | 3.583 | 2.00 ± 0.609 |
| HPP12_0904 | hypothetical protein | -1.79 | 0.831 | -1.92 ± 0.615 |
| HPP12_0870 | flagellar hook protein Flg | -1.57 | 0.118 |
| Increased expression in H. pylori strain P12 modH5 mutant |
| HPP12_0255 | hypothetical protein | 1.80 | 0.667 | 4.72 ± 0.686 |
| HPP12_0253 | outer membrane protein HopG | 2.38 | 1.631 | 3.12 ± 0.427 |

*The genes listed are either down- or up-regulated in the H. pylori P12ΔmodH5 mutant strain. The identity of the gene is indicated with the gene ID in the annotation of the H. pylori P12 genome [42].

Table 2. Differentially expressed genes in H. pylori wild-type P12 modH5 ON versus the mutant strain P12ΔmodH5.

| Gene ID | Gene Name       | Ratio | B-Stat | QRT-PCR |
|---------|-----------------|-------|--------|---------|
| Reduced expression in H. pylori strain P12 modH5 mutant |
| HPP12_1497 | type III R-M system methyltransferase | -5.94 | 6.350 |
| HPP12_0609 | flagellin A | -1.97 | 3.583 | 2.00 ± 0.609 |
| HPP12_0904 | hypothetical protein | -1.79 | 0.831 | -1.92 ± 0.615 |
| HPP12_0870 | flagellar hook protein Flg | -1.57 | 0.118 |
| Increased expression in H. pylori strain P12 modH5 mutant |
| HPP12_0255 | hypothetical protein | 1.80 | 0.667 | 4.72 ± 0.686 |
| HPP12_0253 | outer membrane protein HopG | 2.38 | 1.631 | 3.12 ± 0.427 |

*The genes listed are either down- or up-regulated in the H. pylori P12ΔmodH5 mutant strain. The identity of the gene is indicated with the gene ID in the annotation of the H. pylori P12 genome [42].

Table 3. Summary of downstream modH ORF variation.

| Strain                | modH allele | G tract repeat number | Origin          | Sequence source |
|-----------------------|-------------|-----------------------|-----------------|----------------|
| CHP8                  | modH11      | 10 (ON)               | Australia       | This study (HQ734249) |
| L252                  | modH12      | 10 (ON)               | The Netherlands | This study (HQ734254) |
| L264                  | modH13      | 13 (OFF)              | The Netherlands | This study (HQ734255) |
| SouthAfrica7          | modH14      | 9 (OFF)               | South Africa    | CP002336.1 (HPSA_07265) |
| Gambiat94/24          | modH15      | 9 (OFF)               | Gambia          | CP002332.1 (HPGAM_08025) |
| Cuz20                 | modH16      | 11 (OFF)              | Peru            | CP002076.1 (HPCU_07650) |
| 908                   | modH17      | 12 (OFF)              | Africa          | CP002184.1 (hp908_1508) |

*The strain was defined as having a particular modH allele if the DNA recognition region was 90% identical at the amino acid and nucleotide level to the modH allele. Refer to Figure 1. The resH gene contains a nonsense mutation (strain F32, nucleotide 1587 and strain 52, nucleotide 537 change from G to T) or missing base pair resulting in a frame-shift mutation (strain 51, nucleotide 2054).
Table 3. Primers used for modH allele study.

| Primer   | Nucleotide Sequence 5’-3’ |
|----------|---------------------------|
| HP_MODHF1| GGATAAGAGTACAAATAAGAATTACGTG |
| HP_MODHF2| CTCTACGAGGCCGTAATTTAGGCCTT |
| HP_MODHF3| CCAATGAAAGGTATTAAATAA |
| HP_MODHF4| ACTCAATACTTTTATGGATTTAGG |
| HP_MODHF5| GAGGAATTAAGAAGGCTTAGTAC |
| HP_MODHF6| GGGCCCTTATCTTGCTCCAG |
| HP_MODHREPEATF | GGGGGGACCGAACCCGCGAT |
| HP_MODHR1 | GTTTTAAACCCATCTTATTGG |
| HP_MODHR2 | CCGTCTGTTTAGCAAATCTTATAG |
| HP_MODHR3 | GATAATGCGTCTATTATCTAC |
| HP_MODHR4 | CTGGACGAGAATTAGGACGCCC |
| HP_MODHR5 | CTACCCCCATATCTTAAATGCCC |
| HP_RESHF1 | GGGGATTAAAGATAGGATTTAG |
| HP_RESHR1 | GTCCTGATGAAACATTAGAG |
| HP_RESHR7 | CTTTTTATGCTGTAATACCCGAAAC |
| HP_0253F | CTGGCAGCGACATTITATG |
| HP_0253R | CCCAAGTGGTACCCGCTAT |
| HP12_0255F | TGGTGATAACGCTCAGATAGC |
| HP12_0255R | GCTAGAATAATATGCTATTAC |
| HP12_0609F | TAGTTCAAGCGCAGCCAGGTT |
| HP12_0609R | TGGTGAAACGGCTGATAAGC |
| HP12_0904F | AACGCTAAAGGCGAAAACCC |
| HP12_0904R | GAGGTGTTGTTGCAGTTATTG |
| 16S | AGCGGATGCTGCGGATATTAC |
| 16SR | TCCTCTGTCGAATGTACG |
| RGRM4F | ATGCAAATAAAGAAGATTGGT |
| RGRM4R | CTACCCCCCAACTTTTAAACCGC |
| RGRM4Fmut | CGGATCGTGGGGAATTAGAATAAGTAG |
| RGRM4Rmut1 | CGGATCGTGGGGAATTAGAATAAGTAG |
| RGRM4mut2 | CGGATCAGGGTACGAGGACCCATACAAAGATTG |
| RGRM4mut3 | CGGATCAGGGTACGAGGACCCATACAAAGATTG |

Underlined sequences represent introduced BamHI restriction sites. doi:10.1371/journal.pone.0027569.t003

region of modH containing the DNA binding region and poly-G tract, with a choromaphenicol resistance cassette (cat) lacking a transcriptional terminator [39]. The cloned ORF was inversely amplified using primer pair RGRM4Fmut/RGRM4Rmut1 to excise the 1480 bp region and introduce BamHI sites, which were used to ligate BamHI-cut cat. The second mutant, P12modH5 OF, carried an insertion in the polyG tract to alter it to N11, resulting in a frame-shift mutation into an “OFF” phase. The cloned ORF was inversely amplified using primer pair RGRM4Fmut2/RGRM4Rmut3 to insert a silent thymidine residue into the poly-G tract, changing G10 to 5’-GGGGGTGTTGGG-3’, excising a 270 bp region, and introducing BamHI sites for ligation with BamHI-cut cat. Both mutagenesis constructs carried the cat cassette in the same orientation as modH. For transformation of H. pylori P12, the linear mutagenesis cassettes were amplified using primer pair RGRM4F/RGRM4R. Naturally competent P12 was transformed using the purified PCR product as described previously [39] and transformants were screened by sequence analysis.

RNA extraction
Triplicate cultures of H. pylori strain wild-type P12modH5 ON, P12modH5 OFF and the P12ΔmodH5 mutant, were grown to exponential phase (optical density at 550 nm = 0.9) in BHI broth (Oxoid) supplemented with 10% (v/v) fetal bovine serum (Invitrogen), vitamin mix and vancomycin. Bacterial cells were stabilized using RNAprotect Bacteria Reagent (Qiagen) prior to RNA extraction and approximately 50 µg of total RNA was prepared from each sample using the RNeasy Maxi Kit according to the manufacturer’s instructions (Qiagen). The integrity and concentration of RNA were determined via micro-fluidic analysis on a bio-analysyer (Agilent Technologies).

Microarray analysis
Custom Agilent 8×15 k oligonucleotide microarrays (Agilent, CA, USA) were designed based on the publicly available sequence of H. pylori P12 (NC_011498) using E-array (Agilent, CA, USA). Reverse transcription reactions were performed in 40 µl volumes, containing 10 µg total RNA, 300 ng random hexamers, 0.5 mM dNTPs and 300 U SuperScript III Reverse Transcriptase (Invitrogen) at 42°C for 2.5 h. RNA contamination was removed from the cDNA by the addition of NaOH followed by column purification (Qiagen minElute, Qiagen). A total of 1 µg of purified cDNA was labeled using KREATech Cy3-ULS (KREATech, The Netherlands), and 625 ng was used to hybridize Agilent 8×15 k microarrays as per the manufacturer’s instructions.

Hybridized arrays were scanned on an Agilent Genepix G2563BA scanner, and features were extracted using Feature Extraction V9.5 (Agilent, CA, USA). Analysis was performed using LIMMA [40] as follows. Background correction was applied, spots from duplicate probes were averaged and log transformed. Between-array quantile normalization was then applied to the log transformed spot intensities. A moderated t-test on the normalized log intensities was performed to identify differentially expressed genes and the False Discovery Rate (FDR) used to control for multiple testing. Genes were ranked using the B-statistic (B-stat) method where both fold change and variance of signals in replicates are used to determine the likelihood that genes are truly differentially expressed. A threshold in the B-stat of 0.0 was adopted as genes with a B score >0 have a >50% probability of being truly differentially expressed [41]. All experimental data are available online at the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) submission number GSE26759. All data is MIAME compliant.

Quantitative Real-Time PCR
Oligonucleotides (Table 3) were designed using Primer Express 1.0 software (ABI Prism; PE Biosystems) and are named according to the open reading frame (ORF) being amplified. All real-time PCR reactions were performed in a 25 µl mixture containing 1 in 5 dilution of cDNA preparation (5 µl), 10× SYBR Green buffer (PE Applied Biosystems) and 3 µM of each primer. We used 16S RNA as the control in each quantitative PCR. Amplification and detection of specific products were performed with the ABI Prism 7700 sequence-detection system (PE Applied Biosystems) with the following cycle profile: 95°C for 10 min, followed by 45 cycles of 95°C for 15 sec and 60°C for 1 min. Data were analyzed with ABI prism 7700 (version 1.7) analysis software. Relative gene expression between the P12ΔmodH5 mutant and wild-type P12modH5 ON and P12modH5 OFF and wild-type P12modH5 ON was determined using the 2ΔΔCT relative quantification method.
Supporting Information

Figure S1 Diagrammatical representation of the 17 modH alleles of *H. pylori*. BLASTn was used to identify reciprocal exchanges between the modH DNA recognition domains of the following *H. pylori* strains (listed in Table 1): modH1 BH13, modH2 1061, modH7 11637, modH4 1134, modH5 2A, modH6 3A, modH7 CHP7, modH3 CHP2, modH9 CHP4, modH10 219, modH11 GN760, modH12 L252, modH13 L264, modH14 SouthAfrica7, modH15 Gambria 94/24, modH16 Cuz20 and modH17 908. Each unique modH DNA recognition domain is represented as a coloured box. BLASTn matches longer than 20 nt and >90% identity were mapped on to the corresponding allele in the appropriate colour. The number above the coloured boxes corresponds to Table S1 that contains details of the start and stop positions of each exchange. The nucleotide positions correspond to the DNA recognition domain only.

Table S1 Details of matches shown diagrammatically in Figure 1 (coordinates shown in Figure S1).

Acknowledgments

We are indebted to Drs J. Kusters and A van Vliet (Erasmus Medical Center, Rotterdam, The Netherlands) for providing us with clinical isolates of *H. pylori* clinical isolates from the Netherlands, Canada and Brazil.

Author Contributions

Conceived and designed the experiments: MPJ. Performed the experiments: YNS RJG JAS JAG. Analyzed the data: YNS RJG JAS JAG. Contributed reagents/materials/analysis tools: MPJ RRB SMG TK. Wrote the paper: YNS MPJ. Obtained funding: MPJ RRB SMG. Revised the paper: YNS MPJ RRB RJG.

References

1. Moss SF, Seed S (2003) *Helicobacter pylori*.Curr Opin Infect Dis 16: 445–451.
2. Plikat RK Jr., Crabtree JE (2006) *Helicobacter infection and gastric neoplasia*. J Pathol 208: 233–248.
3. Blaser MJ (1997) Ecology of *Helicobacter pylori* in the human stomach. J Clin Invest 100: 759–762.
4. Ernst PB, Gold BD (2000) The disease spectrum of *Helicobacter pylori*: the immunopathogenesis of gastrointestinal ulcer and gastric cancer. Annu Rev Microbiol 54: 615–640.
5. de Vries N, Duin D., Kuijper EJ, Pot RG, Wiesenekker P, et al. (2002) A site-specific ATP hydrolysis. Embo J 14: 2958–2966.
6. van Belkum A, Scherer S, van Alphen L, Verbrugh H (1998) Short-sequence repeats in prokaryotic genomes. Microb Med Mol Biol Rev 62: 273–293.
7. Moxon ER, Thaler DS (1997) Microbial genetics. The tinkerer’s evolving toolbox. Nature 387: 659, 661–652.
8. Fox KL, Srikhanta YN, Jennings MP (2007) Phase variable type III restriction-modification systems of host-adapted bacterial pathogens. Mol Microbiol 65: 1375–1379.
9. Srikhanta YN, Fox KL, Jennings MP (2010) The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. Nat Rev Microbiol 8: 196–206.
10. Birck S. (1995) Biochemistry of DNA restriction. Microbiol Rev 57: 434–450.
11. Boyer HW (1971) DNA restrictions and modification mechanisms in bacteria. Annu Rev Microbiol 25: 153–176.
12. Rousmael AA, Birck S. (2002) Complex restriction enzymes: NTP-driven molecular motors. Biochimie 84: 1047–1059.
13. Meisel A, Mackeldanz P, Birck S., Kruger DH, Schroeder C (1995) Type III restriction endonucleases translocate DNA in a reaction driven by recognition site-specific ATPase. Embo J 14: 2953–2960.
14. Bach B, Reiser J, Perrott V (1979) Methylation and cleavage sequences of the EcoP1 restriction-modification enzyme. J Mol Biol 120: 143–163.
15. Hambel M, Suri B, Rao DN, Hornby DP, Eberle H, et al. (1980) Type III DNA restriction and modification systems EcoP1 and EcoP5. Nucleotide sequence of the EcoP1 opeon, the EcoP1 mod gene and some EcoP1 mod mutants. J Mol Biol 200: 23–29.
16. Srikhanta YN, Maguire TL, Stacey KJ, Grimmond SM, Jennings MP (2005) The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. Proc Natl Acad Sci U S A 102: 5547–5551.
17. Srikhanta YN, Dowidowicz SJ, Edwards JL, Falsetta ML, Wu HJ, et al. (2009) Phasevarions mediate random switching of gene expression in pathogenic bacteria. Proc Natl Acad Sci U S A 106: 4314–4319.
18. Peck B, Otkamp M, Neu A, Niederweis M, Hendt E, et al. (2001) Characterization of four members of a multigene family encoding outer membrane proteins of *Helicobacter pylori*. J Bacteriol 183: 443–450.
19. Skoglund A, Bjorkholm B, Nilsson C, Andersson A, Jernberg C, et al. (2007) Transcriptional phase variation of a type III restriction-modification system in *Helicobacter pylori* colonized germ-free transgenic mice by *Helicobacter pylori* strains of two pathological origins: *Helicobacter pylori* Colonization of gastroduodenal ulcer and gastric cancer. Annu Rev Immunol 25: 3278–3288.
20. Giannakis M, Chen SL, Karam SM, Engstrand L, Gordon JI (2008) *Helicobacter pylori* evolution during progression from chronic atrophic gastritis to gastric cancer and its impact on gastric stem cells. Proc Natl Acad Sci U S A 105: 4358–4363.
21. Suerbaum S, Josenhans C, Labigne A (1993) Cloning and genetic characterization of the *Helicobacter pylori* flaA gene by polymerase chain reaction. J Bacteriol 175: 3278–3286.
39. Gorrell RJ, Yang J, Kusters JG, van Vliet AH, Robins-Browne RM (2005) Restriction of DNA encoding selectable markers decreases the transformation efficiency of Helicobacter pylori. FEMS Immunol Med Microbiol 44: 213–219.

40. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3: Article3.

41. Smyth GK, Yang YH, Speed T (2003) Statistical issues in cDNA microarray data analysis. Methods Mol Biol 224: 111–136.

42. Fischer W, Windhager L, Rohrer S, Zeiller M, Karrholz A, et al. (2010) Strain-specific genes of Helicobacter pylori: genome evolution driven by a novel type IV secretion system and genomic island transfer. Nucleic Acids Res 38: 6089–6101.
Author/s:
Srikhanta, YN; Gorrell, RJ; Steen, JA; Gawthorne, JA; Kwok, T; Grimmond, SM; Robins-Browne, RM; Jennings, MP

Title:
Phasevarion Mediated Epigenetic Gene Regulation in Helicobacter pylori

Date:
2011-12-05

Citation:
Srikhanta, Y. N., Gorrell, R. J., Steen, J. A., Gawthorne, J. A., Kwok, T., Grimmond, S. M., Robins-Browne, R. M. & Jennings, M. P. (2011). Phasevarion Mediated Epigenetic Gene Regulation in Helicobacter pylori. PLOS ONE, 6 (12), https://doi.org/10.1371/journal.pone.0027569.

Persistent Link:
http://hdl.handle.net/11343/264154

File Description:
Published version

License:
CC BY