RESEARCH ARTICLE

Comparative Genomics Revealed Multiple Helicobacter pylori Genes Associated with Biofilm Formation In Vitro

Eric Hong Jian Wong1*, Chow Goon Ng2*, Eng Guan Chua3, Alfred Chin Yen Tay3, Fanny Peters3, Barry J. Marshall3,4, Bow Ho2*, Khean Lee Goh5, Jamuna Vadivelu1*, Mun Fai Loke1

1 Department of Medical Microbiology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia, 2 Department of Microbiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore, 3 The Marshall Centre for Infectious Diseases Research and Training, School of Pathology and Laboratory Medicine (M502), University of Western Australia, Perth, Australia, 4 UM Marshall Centre, High Impact Research Building, University of Malaya, Kuala Lumpur, Malaysia, 5 Department of Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

* These authors contributed equally to this work.
¤ Current address: Singapore Precision Medicine Centre Pte Ltd, Singapore, Singapore, Singapore

*jamuna@um.edu.my

Abstract

Background
Biofilm formation by Helicobacter pylori may be one of the factors influencing eradication outcome. However, genetic differences between good and poor biofilm forming strains have not been studied.

Materials and Methods
Biofilm yield of 32 Helicobacter pylori strains (standard strain and 31 clinical strains) were determined by crystal-violet assay and grouped into poor, moderate and good biofilm forming groups. Whole genome sequencing of these 32 clinical strains was performed on the Illumina MiSeq platform. Annotation and comparison of the differences between the genomic sequences were carried out using RAST (Rapid Annotation using Subsystem Technology) and SEED viewer. Genes identified were confirmed using PCR.

Results
Genes identified to be associated with biofilm formation in H. pylori includes alpha (1,3)-fucosyltransferase, flagellar protein, 3 hypothetical proteins, outer membrane protein and a cag pathogenicity island protein. These genes play a role in bacterial motility, lipopolysaccharide (LPS) synthesis, Lewis antigen synthesis, adhesion and/or the type-IV secretion system (T4SS). Deletion of cagA and cagPAI confirmed that CagA and T4SS were involved in H. pylori biofilm formation.
Conclusions

Results from this study suggest that biofilm formation in *H. pylori* might be genetically determined and might be influenced by multiple genes. Good, moderate and poor biofilm forming strain might differ during the initiation of biofilm formation.

Introduction

*Helicobacter pylori* is strongly associated with gastroduodenal diseases, such as chronic gastritis, peptic ulcer, duodenal ulcer and gastric cancer [1]. In fact, this bacteria has been classified as a group I carcinogen by the International Agency for Research on Cancer [2]. Although *H. pylori* is susceptible to most antibiotics *in vitro*, only few antibiotics can be used for eradicating *H. pylori in vivo*, e.g. amoxicillin, clarithromycin, metronidazole and tetracycline [3]. Combination therapy is needed for successful eradication of *H. pylori*, Increasing prevalence of antibiotic resistance in *H. pylori* is problematic since it is one of the important causes of therapy failure [4]. The ability of *H. pylori* to form biofilm *in vitro* and also *in vivo* has been demonstrated [5-7]. *H. pylori* formed biofilm within *in vivo* environment especially gastric mucosa had been well demonstrated by previous study [7]. The formation of biofilm within *in vivo* environment was believed to be one of the factors which cause *H. pylori* eradication failure [7].

According to Donlan *et al.* (2002) [8], biofilms are defined as complex microbial ecosystems adherent to each other and/or to surface or interface. It is a microbially derived sessile community characterized by cells that are embedded in a matrix of extracellular polymeric substances that they have produced, and exhibit an altered phenotype with respect to growth rate and gene transcription [8]. Biofilms are important in bacterial pathogenesis as they plays a central role in helping microbes survive or spread within the host. This is because the biofilm matrix acts as shield, protecting bacteria from host defenses and antibiotics [9,10]. Bacterial biofilm can be 10 to 1000 times less susceptible to antimicrobial substances than the same organism in suspension [11]. Biofilm formation makes it difficult for antibiotics to reach the bacteria, underneath aiding the development of antibiotic resistance and increase the risk of treatment failure [8]. Therefore, biofilms are often associated with chronic infections [12]. Cammarota *et al.* demonstrated that *H. pylori* biofilm on the gastric mucosa epithelium was associated with eradication failure and N-acetylcysteine pretreatment before antibiotic therapy was effective in overcoming antibiotic resistance [13,14].

The ability of biofilm formation in bacteria is genetically controlled [15-19]. Based on previous studies, one of the most common genes associated with biofilm formation in bacteria (including *H. pylori, S. mutans* and *L. monocytogenes*) is the *luxS* gene [15-17]. Apart from *luxS* gene, few other genes have also been reported to be associated with biofilm formation in different bacteria [18,19]. According to Yoshida and Kuramitsu (2002) [20], *dkg, sgp* and *com* also play a role in *S. mutans* biofilm formation. Wu and colleagues [21] revealed that *treC* and *sugE* affect biofilm formation in *K. pneumoniae* by modulating capsular polysaccharide synthesis production. Taken together, biofilm formation in bacteria may be regulated by multiple genes working together.

Despite extensive studies conducted on *H. pylori* biofilm formation, precisely how genetic differences of different *H. pylori* strains influence biofilm formation remains not well-established. The objective of this study is to identify genes associated with *H. pylori* biofilm formation.
Materials and Methods

H. pylori strains

H. pylori standard strain J99 (ATCC 700824) was obtained from the American Type Culture Collection (USA). The H. pylori clinical strains used in this study were from the H. pylori collection of the Helicobacter Research Laboratory at the University of Malaya (Malaysia).

H. pylori strain 26695 (ATCC 70392) cagA and cagPAI deletion mutants were constructed using the PCR-based method as described by Chalker et al. [22] and Tan and Berg (2004) [23] in A/Prof. Ho Bow’s laboratory at the Department of Microbiology (National University of Singapore, Singapore).

Culturing of the H. pylori and biofilm

H. pylori were cultured on non-selective chocolate blood agar plate and incubated in a humidified incubator with 10% CO2 at 37°C for three days. The 3-days old H. pylori plate cultures were harvested with a sterile cotton swab and emulsified in brain heart infusion (BHI) broth (Oxoid, Hampshire, UK) supplemented with 1% β-cyclodextrin (Sigma-Aldrich, St. Louis, MO, USA) and 0.4% yeast extract (Oxoid, Hampshire, UK). After 3 days incubation, the total bacterial population of each of the cultures was standardized with the same bacterial turbidity value (bacterial turbidity = 0.042). Following which, 500 μl of bacterial suspension was then inoculated into each well of a 24-well tissue culture plate (BD Falcon, Franklin Lakes, NJ, USA). BHI broth without H. pylori served as a negative control. The growth of the biofilm was observed at the day 3, day 5, day 7 and day 14 with the amount of biofilm formed analyzed based on crystal violet staining.

At each time interval, the bacterial suspensions of the 24-well tissue culture plate were first aspirated out. An aliquot of 650μl of 0.1% crystal violet (Friendemann Schmidt, WA, Australia) was added to each well and the plate was gently swirled for 30 minutes to ensure uniform staining of the biofilm. The excess crystal violet stain was removed and the wells were washed thoroughly using 800μl sterile distilled water to remove any excess crystal violet stain. The crystal-violet stained biofilm was finally destained using 1ml of 95% ethanol (VWR Prolabo, Lutterworth)-5% acetic acid (Friendemann Schmidt, WA, Australia) which serve as destaining solution. The solution collected was measured at OD$_{600nm}$ on a SmartSpec Plus spectrophotometer (Bio-Rad, Hercules, CA, USA). The amount of crystal violet absorbed by the biofilm (CV$_{600nm}$) was determined by taking the mean absorbance value. All experiments were performed technical triplicates and in difference day of biological triplicates.

At the same time, bacterial turbidity value of each bacteria strains were taken for every time points. The bacterial turbidity value of each strain was taken in order to determine the growth rate of the strains. All experiments were performed in biological and technical triplicates and specific biofilm unit (CV$_{600nm}$/OD$_{600nm}$) value was calculated.

Illumina Library Preparation and Sequencing

DNA was extracted from 3-day old H. pylori using the Bacterial DNA mini kit (Stratec Molecular, Berlin, Germany) according to the manufacturer’s instructions. Illumina sequencing library was prepared and samples were sequenced on a MiSeq sequencer (Illumina, San Diego, CA, USA) as described in Perkins et al. [24]. Sequences used in this study have been deposited with GenBank under accession numbers listed in Table 1.

Identifying genes associated with biofilm formation

Biofilm forming ability of H. pylori standard strains J99 and 31 sequenced clinical strains were determined. Based on highest biofilm specific unit value, there strains were classified into 3
groups (good, moderate and poor biofilm formers. Cumulative frequency (%) was calculated using the formula:

\[
\text{cumulative frequency}\% = a + \left(\frac{x}{32} \times 100\right)
\]

Where,

\(a\) = previous cumulative frequency;
\(x\) = number of strains which showed the particular highest biofilm specific unit value

The rate of biofilm formation for each of the strains was determined by using the formula below:

\[
\text{Rate of biofilm formation} = \frac{\text{biofilm formation day } b - \text{biofilm formation day } a}{\text{day } b - \text{day } a}
\]

Table 1. Accession numbers and details of *H. pylori* genomic sequences used in this study.

| Strain | Accession no. | Genome size (bp) | Genes | Proteins | %GC Contigs | Max length (bp) | N50 (bp) | Coverage | Ref |
|--------|---------------|------------------|-------|----------|-------------|----------------|----------|----------|-----|
| J99    | NC_000921     | 1,643,831        | 1,559 | 1,471    | 39.2        | 1               | 1,643,831 | -        | Alm et al, 1999 [25] |
| UM032  | NC_021215     | 1,593,537        | 1,549 | 1,438    | 38.8        | 1               | 1,593,537 | -        | Yalda et al., 2013 [26] |
| UM023  | AUSK000000000 | 1,624,154        | 1,562 | 1,518    | 38.7        | 34              | 485,260  | 183,178  | -   |
| UM037  | AUSI000000000 | 1,724,611        | 1,645 | 1,574    | 38.9        | 60              | 234,132  | 80,609   | -   |
| UM065  | AUSM000000000 | 1,587,249        | 1,500 | 1,461    | 38.9        | 39              | 334,064  | 163,534  | -   |
| UM066  | AUSJ000000000 | 1,694,309        | 1,590 | 1,562    | 38.6        | 34              | 319,894  | 146,858  | -   |
| UM077  | AUSQ000000000 | 1,620,877        | 1,565 | 1,527    | 38.8        | 53              | 328,671  | 187,040  | -   |
| UM085  | AUSP000000000 | 1,645,640        | 1,568 | 1,524    | 38.7        | 50              | 314,517  | 94,888   | -   |
| UM111  | AUSR000000000 | 1,663,383        | 1,581 | 1,536    | 38.7        | 38              | 245,602  | 110,134  | -   |
| UM045  | AONO000000000 | 1,602,114        | 1,559 | 1,439    | 39.0        | 20              | 110,455  | 146,660  | This study |
| UM054  | AONL000000000 | 1,594,474        | 1,570 | 1,422    | 39.1        | 81              | 103,788  | 39,593   | 148  |
| UM087  | LFDR000000000 | 1,657,861        | 1,610 | 1,456    | 38.9        | 17              | 260,866  | 153,641  | This study |
| UM119  | LFBY000000000 | 1,592,071        | 1,564 | 1,439    | 38.8        | 18              | 266,685  | 153,847  | 153  |
| UM122  | LFBX000000000 | 1,593,819        | 1,556 | 1,427    | 39.1        | 23              | 346,159  | 96,269   | 124  |
| UM131  | LFBZ000000000 | 1,598,266        | 1,551 | 1,422    | 39.2        | 16              | 375,408  | 158,380  | 151  |
| UM139  | LFCA000000000 | 1,617,602        | 1,574 | 1,454    | 38.9        | 20              | 231,133  | 90,197   | 180  |
| UM147  | LFLE000000000 | 1,633,271        | 1,567 | 1,442    | 39.5        | 58              | 272,592  | 145,059  | 153  |
| UM152  | LFIS000000000 | 1,633,599        | 1,568 | 1,445    | 39.0        | 16              | 434,701  | 188,275  | 169  |
| UM158  | LFCB000000000 | 1,634,841        | 1,588 | 1,448    | 38.9        | 27              | 240,532  | 185,427  | 153  |
| UM163  | LFJR000000000 | 1,656,598        | 1,601 | 1,475    | 38.5        | 16              | 314,921  | 140,690  | 175  |
| UM165  | LFIR000000000 | 1,651,447        | 1,611 | 1,471    | 38.9        | 20              | 414,265  | 125,433  | 116  |
| UM202  | LFKE000000000 | 1,672,999        | 1,664 | 1,494    | 38.9        | 62              | 117,987  | 40,452   | 214  |
| UM246  | LFKI000000000 | 1,644,161        | 1,592 | 1,466    | 38.6        | 16              | 277,351  | 174,853  | 200  |
| UM276  | LJXI000000000 | 1,666,681        | 1,689 | 1,512    | 38.8        | 61              | 330,937  | 158,795  | 157  |
| UM291  | LFJK000000000 | 1,569,053        | 1,547 | 1,418    | 38.8        | 28              | 219,046  | 123,332  | 135  |
| UM300  | LFIT000000000 | 1,591,764        | 1,535 | 1,410    | 39.2        | 14              | 259,105  | 193,002  | 142  |
| UM352  | LFKK000000000 | 1,623,273        | 1,573 | 1,443    | 38.7        | 19              | 330,865  | 156,967  | 202  |
| UM370  | LFKL000000000 | 1,582,384        | 1,552 | 1,422    | 38.7        | 23              | 200,250  | 93,564   | 174  |
| UM408  | LFIU000000000 | 1,576,513        | 1,524 | 1,391    | 39.1        | 26              | 265,219  | 111,980  | 253  |
| UM411  | LFKM000000000 | 1,631,715        | 1,582 | 1,440    | 39.0        | 17              | 271,501  | 151,419  | 201  |
| UM520  | LEOV000000000 | 1,617,859        | 1,605 | 1,444    | 38.7        | 51              | 177,886  | 52,801   | 160  |

doi:10.1371/journal.pone.0166835.t001
**H. pylori** genomic sequences were annotated and compared using RAST (Rapid Annotation using Subsystem Technology) (http://rast.nmpdr.org/) [28–30]. Sequence-based comparison among **H. pylori** strains was carried out using the SEED viewer of RAST with percentage similarity above 80% was used. Two-tailed Fisher’s exact test was carried out using SPSS (IBM statistic 20, Armonk, New York, USA). P-value <0.05 was considered statistically significant.

Next, the presence and absence of genes-of-interest identified by comparative genomic analysis were confirmed by conventional end-point PCR using primers and conditions listed in Table 2.

**Scanning electron microscopy (SEM)**

Scanning Electron Microscopy (SEM) was performed as described by Lemos et al. (2004) with modifications [31]. The biofilms were grown on the coverslips (Fig 1). The coverslips were washed 2 times in sterile phosphate buffered saline (PBS). The coverslips were then fixed with 2% (wt/vol) glutaraldehyde in 0.1 M PBS (pH 7.4) and post-fixed with 1% osmium tetroxide in 0.1M phosphate buffered saline. The fixed cells were washed with 0.1 M PBS and dehydrated with 50%, 75%, 95%, 100% ethanol. The samples were dried using critical point dryer (Bal-Tec CPD 0300) and examined using JSM-5600 scanning electron microscope (JEOL, Peabody, MA, USA) operating at 10kV.

**Results**

**Optimization of conditions for *in vitro* biofilm formation**

**H. pylori** biofilm formation was observable at the air-liquid interface of wells after three days of incubation in a humidified 10% CO2 incubator at 37˚C. Optimum conditions for biofilm formation were chosen based on maximum biofilm formation without compromising bacterial growth. Comparing between different concentrations of β-cyclodextrin (0.5% to 2.0%) and fetal bovine serum (5% to 15%) as supplement to BHI broth, 1% β-cyclodextrin was shown to be optimum for biofilm formation. Biofilm formation in BHI with 1% β-cyclodextrin broth showed at least one-fold higher than BHI supplemented with fetal bovine serum (5% to 15%) and β-cyclodextrin (0.5% and 2.0%). Biofilm formation in microaerophilic condition was shown to be 3-fold higher than ambient atmospheric condition. As osmolality of the medium

| Table 2. Primers used for amplification of **H. pylori** biofilm genes. |
|---|
| Gene | Primer sequence (5’- 3’) | Product size (bp) | PCR condition |
| hypothetical protein K747_10375 | Forward: CATCTCGGCTGTGAGGGGT | 426 | 95˚C, 5min; 35x (95˚C, 30s; 53˚C, 30s; 72˚C, 1min); 72˚C, 7min |
| | Reverse: TCTTCTTGTCTTTTGGCGAT | | |
| hypothetical protein K747_09130 | Forward: GAGTGGGATAGAGTTAGAAC | 777 | 95˚C, 5min; 35x (95˚C, 30s; 60˚C, 30s; 72˚C, 1min); 72˚C, 7min |
| | Reverse: GTATTAGCCGCTGCTTC | | |
| flagellar protein | Forward: GTGAGTTGTGGCAATAGGGGT | 388 | 95˚C, 5min; 35x (95˚C, 30s; 51˚C, 30s; 72˚C, 1min); 72˚C, 7min |
| | Reverse: AGGCCACTGAGTTTTTAGGT | | |
| alpha-(1,3)-fucosyltransferase | Forward: TCCACCTTACTACACCCGT | 1279 | 95˚C, 5min; 35x (95˚C, 30s; 57˚C, 30s; 72˚C, 1min); 72˚C, 7min |
| | Reverse: AGTCCTAAAAGAGGGTGAGC | | |
| hypothetical protein K747_06625 | Forward: GGCCTACCACACCATACGGGT | 1089 | 95˚C, 5min; 35x (95˚C, 30s; 57˚C, 30s; 72˚C, 1min); 72˚C, 7min |
| | Reverse: TGACCGGCTCTTTTGGTGCA | | |
| outer membrane protein (homD) | Forward: GACGCTCAAGGCAAGGTAGT | 1409 | 95˚C, 5min; 35x (95˚C, 30s; 57˚C, 30s; 72˚C, 1min); 72˚C, 7min |
| | Reverse: AACACATCGATCCTCCACCC | | |
| cag pathogenicity island protein | Forward: AACGCTCCATCAAGAGCCA | 1332 | 95˚C, 5min; 35x (95˚C, 30s; 57˚C, 30s; 72˚C, 1min); 72˚C, 7min |
| | Reverse: CCGCCTCTTTGCTCTTFACT | | |

doi:10.1371/journal.pone.0166835.t002
was increased (from original 0.5% to 1% and 2% sodium chloride), biofilm formation and specific biofilm unit were reduced. In addition, the optimum condition for both biofilm formation and growth occurred at pH 7.0. Thus, BHI broth supplemented with 0.4% yeast extract and 1% β-cyclodextrin at pH 7.0 and microaerophilic condition was determined to be optimum for the study of biofilm formation of *H. pylori* in this study.

**Identification of genes associated with biofilm formation**

The median biofilm specific unit value (cumulative frequency = 50%) was determined to be between specific biofilm unit of 1.0 and 1.1. Strains producing biofilm with specific biofilm unit < 1.0 were classified as poor biofilm formers while those > 1.1 were classified as good biofilm formers. Those specific biofilm units between 1.0 and 1.1 were classified as moderate biofilm formers. Twelve (37.5%) of the tested strains were classified as good biofilm formers, while another 12 (37.5%) strains were poor biofilm formers. The remaining 8 (25.0%) strains were moderate biofilm formers (Table 3). Highest specific biofilm unit was shown to be significantly different between poor, moderate and good biofilm formers (one-way ANOVA p-value = 0.001).

Based on rate of biofilm formation from day 0 to day 7, there was clear segregation between poor and good biofilm forming strains (Table 4). Student’s t-test revealed that there were
significant differences in rate of biofilm formation between poor, moderate and good biofilm forming strains across all time point (p-value < 0.05), except between poor and moderate groups at day 3 to day 5 (p-value = 0.260). While good and poor biofilm forming strains displayed relatively consistent rate of biofilm formation across 7 days, moderate biofilm forming strains varied in rate of biofilm formation across time course of study. The moderate biofilm forming strains showed either declined biofilm forming rate by day 3 or delayed optimum biofilm formation starting from day 3 onwards. The rate of biofilm formation profiles was consistent with the highest specific biofilm unit profiles across multiple *H. pylori* strains.

Through the comparison of annotated genes present in *H. pylori* genomic sequences of poor, moderate and good biofilm forming strains, with 3 hypothetical genes (Hypothetical protein K747_10375, Hypothetical protein K747_09130 and Hypothetical protein K747_06625) and 4 functional genes (flagellar protein, *cag* pathogenicity island protein, outer membrane protein, etc).
(homD) and alpha-(1,3)- fucosyltransferase) were identified to be significantly correlating with capability of forming biofilm (p-value < 0.05) (Table 5). The presences and absences of these interested genes were verified by using PCR method (Table 5). One of the genes identified is
Table 5. Genes with significant association with biofilm formation (p-value <0.05) confirmed by PCR.

| Genes                        | RefSeq reference no. | Helicobacter pylori strain (N = 32) | Fisher’s exact test (p-value) |
|------------------------------|----------------------|------------------------------------|------------------------------|
|                              |                      | Poor biofilm former (N = 8) | Moderate biofilm former (N = 12) | Good biofilm former (N = 12) |
| hypothetical protein K747_10375 | WP_015645358.1       | 0                                 | 7 (58.3%)                     | 12 (100%)                    | <0.001*                       |
|                              |                      | UM065, UM276, UM023, UM520, UM246, UM370, UM291 | UM122, UM032, UM087, J99, UM163, UM119, UM085, UM077, UM066, UM111, UM352 |
| hypothetical protein K747_09130 | WP_015645548.1       | 1 (12.5%)                         | 7 (58.3%)                     | 11 (91.7%)                   | 0.001*                        |
|                              |                      | UM045                             | UM202, UM276, UM408, UM520, UM246, UM370, UM291 | UM122, UM032, UM087, J99, UM163, UM119, UM085, UM077, UM066, UM111 |
| hypothetical protein K747_06625 | WP_015644951.1       | 1 (12.5%)                         | 8 (66.7%)                     | 11 (91.7%)                   | <0.001*                       |
|                              |                      | UM045                             | UM202, UM158, UM408, UM246, UM023, UM520, UM139, UM291 | UM122, UM032, UM087, J99, UM163, UM119, UM085, UM077, UM066, UM352 |
| flagellar protein            | WP_000744159.1       | 1 (12.5%)                         | 8 (66.7%)                     | 11 (91.7%)                   | <0.001*                       |
|                              |                      | UM045                             | UM202, UM065, UM158, UM408, UM246, UM023, UM165, UM370, UM139 | UM122, UM032, UM087, J99, UM163, UM119, UM085, UM077, UM066, UM352 |
| cag pathogenicity island protein | WP_015644978.1       | 0                                 | 7 (58.3%)                     | 8 (66.7%)                    | 0.005*                        |
|                              |                      | UM065, UM276, UM023, UM520, UM246, UM370, UM291 | UM032, UM163, UM119, UM085, UM077, UM066, UM352 |
| outer membrane protein (homD) | WP_002205258.1       | 0                                 | 7 (58.3%)                     | 10 (83.3%)                   | 0.001*                        |
|                              |                      | UM065, UM276, UM408, UM520, UM370, UM291, UM139 | UM122, UM087, J99, UM163, UM119, UM085, UM077, UM066, UM352 |
| alpha-(1,3)-fucosyltransferase | WP_015646026.1       | 0                                 | 6 (50.0%)                     | 10 (83.3%)                   | 0.001*                        |

* p-value<0.05 is considered statistically significant

To confirm that cag pathogenicity island (cag-PAI) genes were involved in H. pylori biofilm formation, biofilm formation by wild-type and deletion mutants of cagA and cagPAI were compared. Bacteria turbidity and biofilm formation were increasing day 3 to day 14 (Fig 2). This was consistent with observation by SEM that spiral form of H. pylori could still found in the biofilm on day 14. Deletion of cagA and cagPAI were demonstrated to reduce biofilm formation (Figs 2 and 3). The microcolonies in the biofilm were denser and bigger in wild type as compared to ΔcagA and ΔcagPAI. cagPAI knockout mutants formed sparsely distributed microcolonies at the air-medium interface (Fig 3). At higher magnification, it was observed that the extracellular matrix formed by the wild-type strain was denser than that of ΔcagA and ΔcagPAI. In addition, extracellular matrix formed by ΔcagA was film-like in structure compared to the filament-like structure form by wild-type. Extracellular matrix was scanty in the biofilm of ΔcagPAI.
The identification of a set of genes that partition the good biofilm forming *H. pylori* strains from their poor biofilm forming counterparts demonstrate that biofilm formation phenotype is influenced by multiple genetic factors of the bacteria. Motility and adhesion capabilities apparently are critical for the initiation of biofilm formation (and probably the subsequent propagation of biofilm).

The movement ability by different species of bacteria is linked to their ability to colonize various ecological niches, and is frequently related to pathogenesis and biofilm formation. The gene encoding for hypothetical protein K747_06625 is predicted to contain ParB-like and HNH nuclease domains. In *Actinomyces oris* K20, genes containing ParB-like nuclease domain has been shown to be involved in the formation of meshwork-like structures, which are found in some biofilm-producing bacteria [32]. In *P. aeruginosa*, par mutants have been observed to be impaired in motility suggesting a direct or indirect role of Par proteins in regulation of these processes [33]. ParA/ParB family of proteins together with the centromere-like DNA sequence *parS* is involved in chromosome and plasmid partitioning during bacterial cell division. In the absence of ParB, ParA is most probably in its dimeric state bound with ATP proficient to bind non-specifically to DNA [34]. Furthermore, *par* mutation also induced genes involved in c-di-GMP turnover and signaling [33]. Cyclic-di-GMP (cyclic diguanylate) is an important messenger ubiquitous in bacterial cells controlling various processes, e.g. switch between the motile planktonic and biofilm lifestyles of bacteria, virulence of animal and plant pathogens, antibiotic production, progression through the cell cycle and other cellular functions [34].

Another gene-of-interest encodes for a flagellar protein, which is also involved in bacterial motility. It is believed that the presence of flagella provide motility that aids kinetics of biofilm formation [35]. Motility provides by flagella enhance the recruitment of planktonic cells to the biofilm [36]. Besides, the adhesive properties of flagella proteins might also promote biofilm formation in *H. pylori*. This was supported by previous studies that demonstrate that *Campylobacter jejuni* flagella are needed for the initiation of biofilm formation by mediating adhesion on the surface [37]. In one study, microcolonies were formed on a coverslip with flagella forming bridges between organisms [38]. During initial reversible attachment, flagella permit individual planktonic bacteria to swim toward an appropriate biotic or abiotic surface. Subsequently, during irreversible attachment, these flagella become lost and adhesive organelles elaborated [39].

Biofilm forms a protective layer for Gram-negative bacteria to prevent attack by host immune system against one of its cell wall component, lipopolysaccharide (LPS) [40]. *H. pylori* alpha-(1,3)-fucosyltransferase, one of the many genes that involved in *H. pylori* LPS synthesis [41], is also predicted to contain the biofilm formation protein (YliH/bssR) domain, which has been shown to be induced in *Escherichia coli* biofilms [42]. In addition, *H. pylori*, alpha-(1,3/4)-fucosyltransferase is required for fucosylation of both type I (Le^a^) and type II (Le^b^) Lewis antigens [43]. The adhesive properties of Lewis X antigen enable first colonist to adhere to the surface of the liquid through weak adhesion force in the formation of biofilm [44].

One of the proteins found to be prevalent with good biofilm forming strains encodes for a cag pathogenicity islands protein. CagA protein, encoded by cag pathogenicity islands, has...
been identified to be induced in *H. pylori* biofilms [45]. CagA, which serves as a virulent factor of *H. pylori*, is injected into the host through the type IV secretion system and affect the individual [46]. Type IV secretion system-mediated transfer is essential in direct cell-cell contact [47]. It is believed that this direct cell-cell contact can control the biofilm behavior in *H. pylori* [48]. Furthermore, CagE, another protein encoded by the *cag* pathogenicity islands, has also been reported to be involved in the *H. pylori* biofilm formation [15,48]. *cagE* encodes for a cytoplasmic ATPase to allow the translocation of CagA protein that induces the gastric epithelial cells to secrete interleukin-8 (IL-8) and activate many intracellular signaling pathway [25]. Thus, the *cag* pathogenicity island may play an important role in *H. pylori* biofilm formation. Besides its role in bacteria-host interaction, CagA and the *cag* pathogenicity island may also have a role to play in bacteria-bacteria interaction in *H. pylori* biofilm formation.

**Conclusion**

Using a comparative genomics approach, data from this study suggest that biofilm formation in *H. pylori* might be influenced by multiples genes. Good and poor biofilm forming strain may differ genetically in terms of motility, adhesion and bacteria-bacteria interactions, which are important during the initiation of biofilm formation.

**Acknowledgments**

This study was funded in full by the University of Malaya-Ministry of Education (UM-MoE) High Impact Research (HIR) Grant UM.C/625/1/HIR/MoE/CHAN/13/3 (Account No. H-50001-A000030) and the National Medical Research Council (NMRC) Grant No. R-182-000-180-213. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. BJM is co-Founder of Tri-Med Distributers PTY LTD and Founder and Executive Director of Ondek PTY LTD.

**Author Contributions**

**Conceptualization:** MFL CGN BH KLG JV.

**Data curation:** MFL.

**Formal analysis:** EHJW CGN.

**Funding acquisition:** MFL.

**Investigation:** EHJW CGN.

**Methodology:** EHJW CGN EGC ACYT FP MFL BJM.

**Project administration:** JV.

**Resources:** KLG MFL BJM.

**Supervision:** KLG JV MFL.

**Writing – original draft:** EHJW CGN.

**Writing – review & editing:** BH MFL.
References

1. Kusters JG, van Vliet AHM, Kuipers EJ. Pathogenesis of *Helicobacter pylori* infection. Clin Microbiol Rev. 2006; 19(3):449–90. doi: 10.1128/CMR.00054-05 PMID: 16847081

2. IARC Working Group. Schistosomes, liver flukes and *Helicobacter pylori*. IARC Working Group on the evaluation of carcinogenic risks to humans. Lyon, 7–14 June 1994. IARC monographs on the evaluation of carcinogenic risks to humans / World Health Organization, International Agency for Research on Cancer. 1994. 61:1–241. PMID: 7715068

3. Fock KM, Katelaris P, Sugano K, Ang TL, Hunt R, Talley NJ, et al. Second Asia-Pacific consensus guidelines for *Helicobacter pylori* infection. J Gastroen Hepatol 2009; 24(10):1587–600. doi: 10.1111/j.1440-1746.2009.05982.x PMID: 19788600

4. Graham DY, Fischbach L. *Helicobacter pylori* treatment in the era of increasing antibiotic resistance. Gut. 2010; 59(8):1143–53. doi: 10.1136/gut.2009.192757 PMID: 20525969

5. Stark RM, Gerwig GJ, Pitman RS, Williams NA, Greeman J, et al. Biofilm formation by *Helicobacter pylori*. Lett Appl Microbiol. 1999; 28(2):121–6. doi: 10.1046/j.1365-2672.1999.00481.x PMID: 10063642

6. Percival SL, Suleman L. Biofilms and *Helicobacter pylori*: Dissemination and persistence within the environment and host. World J Gastroint est Pathophy sio. 2014 15; 5(3):122–32. doi: 10.4291/wj gp.v5. i3.122 PMID: 25133015

7. Carron MA, Tran VR, Sugawa C, Coticchia JM. Identification of *Helicobacter pylori* biofilms in human gastric mucosa. J Gastro intest Surg. 2006; 10(5):712–7. doi: 10.1016/j.gassur.2005.10.019 PMID: 16713544

8. Donlan RM, Costerton JW. Biofilms: survival mechanisms of clinically relevant microorganisms. Clin Microbiol Rev. 2002; 15(2):167–93. doi: 10.1128/CMR.15.2.167–193.2002 PMID: 11932229

9. Dunne WM Jr. Bacterial adhesion: seen any good biofilms lately? Clin Microbiol Rev. 2002; 15(2):155–66. doi: 10.1128 /CMR.15.2.155–166.2002 PMID: 11932228

10. Mohamed JA, Huang DB. Biofilm formation by enterococci. J Med Microbiol. 2007; 56(Pt 12):1581–8. doi: 10.1099/jmm.0.47331-0 PMID: 18033823

11. Harrison JJ, Turner RJ, Marques LR, Ceri H. A new understanding of these microbial communities is driving a revolution that may transform the science of Microbiology. Am Sci. 2002; 93:508. doi: 10.1511/2005.6.508

12. Ikeno T, Fukuda K, Ogawa M, Honda M, Tanabe T, Taniguchi H. Small and rough colony *Pseudomonas aeruginosa* with elevated biofilm formation ability isolated in hospitalized patients. Microbiol Immunol. 2007; 51(10):929–38. doi: 10.1111/j.1348-0421.2007.tb03989.x PMID: 17951982

13. Cammarota G, Branca G, Ardito F, Sanguinetti M, Ianigro G, et al. Biofilm demolition and antibiotic treatment to eradicate resistant *Helicobacter pylori*: A clinical trial. Clin Gastroenterol Hepatol. 2010; 8(9):817–820.e3. doi: 10.1016/j.cgh.2010.05.006 PMID: 20478402

14. Cammarota G, Sanguinetti M, Gallo A, Postera B. Biofilm formation by *Helicobacter pylori* as a target for eradication of resistant infection. Aliment Pharmacol Ther. 2012; 36(3):222–230. doi: 10.1111/j.1365-2036.2012.05165.x PMID: 22650647

15. Cole SP, Harwood J, Lee R, She R, Guiney DG. Characterization of monospecies biofilm formation by *Helicobacter pylori*. J Bacteriol. 2004; 186(10):3124–32. doi: 10.1128/JB.186.10.3124–3132.2004 PMID: 15126474

16. Merritt J, Qi F, Goodman SD, Anderson MH, Shi W. Mutation of luxS affects biofilm formation in *Streptococcus mutans*. Infect Immun. 2003; 71(4):1972–9. doi: 10.1128/IAI.71.4.1972–1979.2003 PMID: 12654815

17. Sela S, Frank S, Belausov E, Pinto R. A mutation in the luxS gene influences *Listeria monocytogenes* biofilm formation. Appl Environ Microb. 2006; 72(8):5653–8. doi: 10.1128/AEM.00048-06 PMID: 16885324

18. Shemesh M, Tam A, Steinberg D. Differential gene expression profiling of *Streptococcus mutans* cultured under biofilm and planktonic conditions. Microbiology (Reading, England). 2007; 153( Pt 5):1307–17. doi: 10.1099/mic.0.2006/002030-0 PMID: 17464045

19. Vega AE, Persia FA, Salinas-Ibanez GA, Alarcon T, Lucero Estrada SMC, Cortinas TI, et al. *Helicobacter pylori* biofilm formation and gene expression on abiotic surfaces using cyanobacterial extract. J Life Sci. 2012; 6:1317–27.

20. Yoshida A, Kuramitsu HK. Multiple *Streptococcus mutans* genes are involved in biofilm formation. Appl Environ Microb. 2002; 68(12):6283–91. doi: 10.1128/AEM.68.12.6283–6291.2002
21. Wu MC, Lin TL, Hsieh PF, Yang HC, Wang JT. Isolation of genes involved in biofilm formation of a Klebsiella pneumoniae strain causing pyogenic liver abscess. PloS one. 2011; 6(8):e23500. doi: 10.1371/journal.pone.0023500 PMID: 21858144

22. Chalker AF, Minehart HW, Hughes NJ, Koretki KK, Lomnet MA, Brinkman KK, et al. Systematic identification of selective essential genes in Helicobacter pylori by genome prioritization and allelic replacement mutagenesis. J Bacteriol. 2001; 183(4):1259–68. doi: 10.1128/JB.183.4.1259–1268.2001 PMID: 11157938

23. Tan S, Berg DE. Motility of urease-deficient derivatives of Helicobacter pylori. J Bacteriol. 2004; 186(3):885–8. doi: 10.1128/JB.186.3.885–888.2004 PMID: 14729719

24. Perkins TT, Tay CY, Thirriot F, Marshall B. Choosing a benchtop sequencing machine to characterise Helicobacter pylori genomes. PloS one. 2013; 8(6):e67539. doi: 10.1371/journal.pone.0067539 PMID: 23840736

25. Aim RA, Ling LS, Moir DT, King BL, Brown ED, Doig PC, et al. Genomic-sequencing on multiple related isolates of the human gastric pathogen Helicobacter pylori. Nature. 1999; 176–80. doi: 10.1038/16495 PMID: 9923682

26. Khosravi Y, Rehvathy V, Tan MH, Guntiethcium SP, Teh X, Wang S, Baybayan P, et al. Comparing the genomes of Helicobacter pylori clinical strain UM032 and mice-adapted derivatives. 2013; 5:25. doi: 10.1186/1757-4749-5-25 PMID: 23957912

27. Rehvathy V, Tan MH, Gunalechchumy SP, Teh X, Wang S, Baybayan P, et al. Multiple genome sequences of Helicobacter pylori strains of diverse disease and antibiotic resistance backgrounds from Malaysia. Genome Announc. 2013; 1(5). doi: 10.1128/genomeA.00687-73

28. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. BMC genomics. 2008; 9:75. doi: 10.1186/1471-2164-9-75 PMID: 18261238

29. Lemos JAC, Brown TA, Burne RA. Effects of RelA on key virulence properties of planktonic and biofilm populations of Streptococcus mutans. Infect Immun. 2004; 72(3):1431–40. doi: 10.1128/IAI.72.3.1431–1440.2004 PMID: 14977948

30. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, et al. RASTk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep. 2015; 5:8365. doi: 10.1038/srep08365 PMID: 25666858

31. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 2014; 42(Database issue):D206–14. doi: 10.1093/nar/gkt1226 PMID: 24293654

32. Lemos JAC, Brown TA, Burne RA. Effects of RelA on key virulence properties of planktonic and biofilm populations of Streptococcus mutans. Infect Immun. 2004; 72(3):1431–40. doi: 10.1128/IAI.72.3.1431–1440.2004 PMID: 14977948

33. Mashimo C, Kamitani H, Nambu T, Yamane K, Yamanaka T, Sugimori-Shinozuka C, et al. Identification of the genes involved in the biofilm-like structures on Actinomyces oris K20, a clinical isolate from an apical lesion. J Endodont. doi: 10.1016/j.joen.2012.08.009 PMID: 23228256

34. Bartosik AA, Glabski K, Jecz P, Mikulska S, Fogtman A, Koblowska M, et al. Transcriptional profiling of ParA and ParB mutants in actively dividing cells of an opportunistic human pathogen Pseudomonas aeruginosa. PloS one. 2014; 9(1):e87276. doi: 10.1371/journal.pone.0087276 PMID: 24498062

35. Hengge R. Principles of c-di-GMP signalling in bacteria. Nat Rev Microbiol. 2009; 7(4):263–73. doi: 10.1038/nrmicro2109 PMID: 19287449

36. Svensson SL, Pryjma M, Gaynor EC. Flagella-mediated adhesion and extracellular DNA release contribute to biofilm formation and ssTress tolerance of Campylobacter jejuni. PloS one. 2014; 9(8):e106063. doi: 10.1371/journal.pone.0106063 PMID: 25166748

37. Lemon KP, Higgins DE, Kolter R. Flagellar motility is critical for Listeria monocytogenes biofilm formation. J Bacteriol. 2007; 189(12):4418–24. doi: 10.1128/JB.01967-06 PMID: 17416647

38. Moe KK, Mimura J, Ohnishi T, Wake T, Yamazaki W, Nakai M, et al. The mode of biofilm formation on smooth surfaces by Campylobacter jejuni. J Vet Med Sci/Jpn Soc Vet Sci. 2010; 72(4):411–6. doi: 10.1292/jvms.09-0339

39. Pru ß BM, Besemann C, Denton A, Wolfe AJ. A complex transcription network controls the early stages of biofilm development by Escherichia coli. J of Bacteriol. 2006; 188(11):3731–9. doi: 10.1128/JB.01780-05. PMID: 16707665

40. De Araujo C, Balestrino D, Roth L, Charbonnel N, Forester C. Quorum sensing affects biofilm formation through lipopolysaccharide synthesis in Klebsiella pneumoniae. Res Microbiol. 2010; 161(7):595–603. doi: 10.1016/j.resmic.2010.05.014 PMID: 20600864

41. Salaün L, Saunders NJ. Population-associated differences between the phase variable LPS biosynthetic genes of Helicobacter pylori. BMC Medical Genomics. 2006; 6:79. doi: 10.1186/1471-2180-6-79 PMID: 16981984
41. Ren D, Bedzyk LA, Thomas SM, Ye RW, Wood TK. Gene expression in *Escherichia coli* biofilms. *Appl Microbiol Biot.* 2004; 64(4):515–24. doi: 10.1007/s00253-003-1517-7

42. Rasko DA, Wang G, Palcic MM, Taylor DE. Cloning and characterization of the alpha(1,3/4) fucosyl-transferase of *Helicobacter pylori*. *J Biol Chem.* 2000; 275(7):4988–94. doi: 10.1007/s00253-003-1517-y PMID: 10671538

43. Mahdavi J, Borén T, Vandenbroucke-Grauls C, Appelmek BJ. Limited role of lipopolysaccharide Lewis antigens in adherence of *Helicobacter pylori* to the human gastric epithelium. *Infect Immun.* 2003; 71(5):2876–80. doi: 10.1128/IAI.71.5.2876–2880.2003 PMID: 12704161

44. Shao C, Sun Y, Wang N, Yu H, Zhou Y, Chen C, et al. Changes of proteome components of *Helicobacter pylori* biofilms induced by serum starvation. *Mol Med Reports* 8.6. 2013:1761–6. doi: 10.3892/mmr.2013.1712 PMID: 24100704

45. Pattis I, Weiss E, Laugks R, Haas R, Fischer W. The *Helicobacter pylori* CagF protein is a type IV secretion chaperone-like molecule that binds close to the C-terminal secretion signal of the CagA effector protein. *Microbiology.* 2007; 153(Pt 9):2896–909. doi: 10.1099/mic.0.2007/007385-0 PMID: 17768234

46. Zechner EL, Lang S, Schildbach JF. Assembly and mechanisms of bacterial type IV secretion machines. *Philosophical Transactions of the Royal Society of London B: Biological Sciences.* 2012; 367(1592):1073–87. doi: 10.1098/rstb.2011.0207 PMID: 22411979

47. Mattar R, Marques SB, Monteiro Mdo S, Dos Santos AF, Iriya K, Carrilho FJ. *Helicobacter pylori cag* pathogenicity island genes: clinical relevance for peptic ulcer disease development in Brazil. *J Med Microbiol.* 2007; 56(Pt 1):9–14. doi: 10.1099/jmm.0.46824–0 PMID: 17172510

48. Brókowska AMB, Duś I, Gościński G. The role of biofilm formation in pathogenesis of *Helicobacter pylori* infections. *Rev paper/Artykuł poglądowy.* 2013; 8:27–30. doi: 10.5114/pg.2013.34179