H.pylori virulence factors with gastric cancer

The Relationship between *cagH* , *cagL, cagG,* and *orf17* Genotypes of *Helicobacter pylori cag* Pathogenicity Island with Peptic Ulcerations and Gastric Cancer

- Negin Raei*, Saeid Latifi-Navid, Saber Zahri

Abstract

Helicobacter pylori (*H. pylori*) specific genotypes have been correlated with an increased risk of gastrointestinal disease in Iran. The aim of this study was finding the relevance of *H. pylori* virulence factors (*cagH* , *cagL* , *cagG* , and *orf17*) with gastric cancer (GC) and peptic ulcerations (PU) in Iran. The *H. pylori* strains were isolated from different ethnic and geographic origin within Iran and genotyped. Overall, the frequency of *cagH* , *cagL* , *orf17* , and *cagG* genotypes was 74. 1%, 98%, 79. 9%, and 70. 4% in patients with PU, respectively, while in non-atrophic gastritis (NAG) the frequency was 59. 6% for *cagH* , 82. 6% for *cagL* , 61% for *orf17* , and 74. 7% for *cagG* . The frequency of *cagH* , *cagL* , *cagG* , and *orf17* was 50%, 83. 3%, 61. 9%, and 54. 8% in GC group, respectively. No association was found between the mentioned genotypes and the risk of GC in Iran ($P = P$ -value > 0. 05); however, *cagL* and *orf17* genotypes were correlated with an increased risk of PU in Iran ($P = 0.021$ for *cagL* and $P = 0.015$ for *orf17*), *cagH* and *cagG* genes showed no consistent relationship with PU in Iran ($P > 0.05$).

Introduction

Gastric Cancer (GC) is the third common cancer that ends up death in the world,** the incidence rate of this cancer varies across different geographical areas.** The high incidence areas for gastric cancer has geographical distribution in Asia.** Ardabil province that is located in north west of Iran

and close to Caspian sea, is reported as area that has high incidence rate of gastric adenocarcinoma in Iran.** Gastric cancer has high rate in male and is the third cancer after breast and colorectal cancer in female in Iran.**The concrete reasons for prevalence of gastric cancer, which varies in different geographical areas, remain unknown; however data show that interaction between host and environmental factors, specially *Helicobacter pylori* (*H. pylori*) infection may play a remarkable role. This bacterium is found in more than 50 % of the world population and it is a well-known human pathogen. *H. pylori* is the cause of acute or chronic gastritis, peptic ulcerations (PU) disease, gastric adenocarcinoma and gastric mucosa-associated lymphoid tissue (MALT) lymphoma. In most of cases the infection remains asymptomatic during life, just 20% develops to peptic ulcer or gastric carcinoma.*** *H. pylori* results in acute or chronic infection, the acute infection decreases the acidity in the gastric, which promotes the bacterial colonization,* this hypochlorhydria stimulates gastritis that predisposes to atrophic gastritis, intestinal metaplasia, dysplasia and finally carcinoma.* *H. pylori* infection in the gastric triggers a mucosal inflammatory response and is the cause of different clinical outcomes in humans.* *H. pylori* harbors a pathogenicity island that is called *cag* pathogenicity island (*cag* PAI). The *cag* PAI is the main virulence factor that encodes type IV secretion system (T4SS), several genes within *cag* PAI are associated with an increased interlukin-8 (IL-8) production, which is produced by gastric epithelial cells.** Among the bacterial factors, the adherent ability to gastric epithelial cells is the most important factor in order to initiate the gastric inflammatory response and is crucial for IL-8 induction.**CagL is bacterial factor that is called a pilus protein (a VirB5 orthologe).* Integrin α5β1 is a receptor on

gastric epithelial cell that CagL binds to; by its arginine-glycine-aspartate (RGD) motif** , this interaction activates α5β1 receptor and facilitates and triggers** the delivery of bacterial CagA oncoprotein through T4SS into the epithelial cell.* CagL is associated with two other *cag* PAI proteins ; CagI and CagH.* All the three factors are highly needed for injection of CagA into the epithelial cell, and it is demonstrated that those factors are responsible for T4SS pili formation. *cagG* is another factor that is located upstream of *cagA, cagG* is the gene that shows some homology to adherence-related or motility-related genes of other bacteria in bioinformatics study or gene analysis.*** *orf17* is the gene that shows some homology to the gene of other bacteria in bioinformatics study, the study shows that in *Dickeya zeae* there is a gene which has homology to *orf17* in *helicobacter pylori* and 36% identity is reported in the pblast studies. *Dickeya zeae* , is the aerobic/anaerobic phytopathogene bacterium that is the cause of soft rot disease in a broad range plants species, specially many crops that are economically crucial.** this bacterium causes bacterial foot rot and it is reported in many asian countries particularly china which the disease mostly begins at the ligulus.**In this study we aimed to find out the distribution of four target genes (*cagH* , *cagL* , *cagG* , and *orf17*) in *H. pylori* genome and their relationship with gastrointestinal disease in Iran.

Materials and methods:

Collection of tissue specimens

The tissue specimens were collected over 3 years from 2011 to 2014 in Iran; from patients with various gastric diseases.

H. pylori isolation and culture

Antrum and body biopsies were used for H. pylori culture. Biopsies were cultured on selective Brucella agar (Merck, Germany) involving 5% sheep blood, 10 mg/L of vancomycin (Zakaria, Iran), 5 mg/L trimethoprim (MP Biomedicals, France), 2. 5 IU/L polymyxin B (MP Biomedicals, France), and 8 mg/L amphotericin B (Bristol-Myers Squibb, USA). Cultures were incubated at 37 â-¦C under microaerophilic situation for 3-7 days. Bacterial isolates were identified as H. pylori on the basis of Gram-stained morphology and positive urease, catalase, and oxidase tests. Bacterial isolates were harvested in brain heart infusion broth (Merck) enriched with 20% glycerol and 10% inactivated horse serum and stored at −70 â-¦C.

Histological assessment and classification

All biopsies were taken from the gastric body (corpus) and antrum of patients with different gastrointestinal disease, and were used for histopathalogical examination. The biopsy specimens were initially formalin-fixed and paraffin-embedded. Sections of 4µm were obtained and stained with hematoxylin-eosin, Giemsa, and Alcian blue- periodic acid Shiff (pH 2. 5). By the use of Sydney system, GC was classified and graded. 4ms. abdi

DNA extraction

Using protocol DNGTM plus kit (Cinna Gene, Tehran, Iran), DNA was extracted from the urease, oxidase, and catalase positive gastric biopsy specimens, before using the kit, the tissue was completely broken down

using the scalpel. Extracted DNA was kept at -20ËšC until polymerase chain reaction (PCR) amplification was carried out.

PCR amplification

The first PCR was conducted in order to detect the *H. pylori* -specific 16S rDNA gene*. Using HP1-2 16S Rrna primer PCR amplification was performed (Optimized annealing temperature: 56 °C) in 30-µl reaction mixture and the PCR amplification products (1500 bp in size), which were 16S rDNA, were detected in electrophoreses that proved the specimens were infected by *H. pylori* (*H. P +*). The second PCR was done in order to detect the target genes which were *cagH* , *cagL* , *cagG* , and *orf17* (Optimized annealing temperature (°C): 52, 54, 50, and 50; respectively). The PCR amplification was carried out in 30-µl reaction mixture involving MgCl2 = 1 µl, PCR buffer = 3 µl, dNTPmix = 0. 5 µl, primers (reverse and forward mixture) = 1 µl, enzyme = 0. 2 µl, template DNA = 5 µl, and D. W = 19. 3 µl (total volume = 30 µl); and the PCR amplification products (*cagH* = 1113 bp *, cagL* = 263 bp, *cagG* = 398 bp, and *orf17* = 546 bp in size) were detected.

Statistical analysis

The relationship between the *cagH* , *cagL* , *cagG* and *orf17* genotypes of *H. pylori cag* pathogenicity island and clinical outcomes (including GC, PU, and non-atrophic gastritis or NAG) were analyzed. The SPSS statistical software package version 18. 0 was used for all statistical analyses in this study. With chi-square test and the fisher exact probability the analysis was done. A *P* -value of < 0. 05 indicated significance. A simple and multiple logistic regression was used to calculate the odd ratio (OR) and 95% confidence

interval (CI) of the clinical outcomes by including age, sex and *H. pylori* genotypes. The OR and CI were used to estimate the risk of PU and GC in the current study.

Results

A total of 242 strains isolated from *H. pylori* -positive Iranian patients (59. 5% male and 40. 5 female; 37. 6% age >= 55 and 62% age < 55), were examined in the present study. The strains involving 146 (60. 3%) with NAG, 54 (22. 3%) with PU, and 42 (17. 4%) with GC. The total frequency of *cagH* , *cagL* , *cagG* , and *orf17* genotypes was 61. 2%, 86. 4%, 71. 5%, and 64%, respectively.

Association between the virulence factors and PU in Iran: The frequency of the *cagH* , *cagL* , *orf17* , and *cagG* genotypes was 74. 1%, 98. 1%, 79. 6%, and 70. 4%in PU group (51. 8% for duodenal ulcer and 37. 03% for gastric ulcer) although in the control group (NAG) the following frequencies were reported: 59. 6% for *cagH* , 82. 9% for *cagL* , 61% for orf17, and 74. 7% for *cagG* . The simple logistic regression analysis illustrated that the *cagL* and *orf17* genotypes were remarkably associated with an increased risk of PU, the OR (95%CI) 10. 95 (1. 446-82. 935) and 2. 504 (1. 193-5. 235) respectively, ($P = 0.021$ for *cagL* and $P = 0.015$ for *orf17*). No significant relationship was found between the *cagH* and *cagG* genotypes and the risk of PU ($P > 0.05$). Analysis for combined genotypes showed that *cagH* / *cagL* and *cagH* / *orf17* were associated with an increased risk of PU, the OR (95%CI) was 9. 756 (1. 264-75. 303) for *cagH* / *cagL* and 2. 861 (1. 167-7. 013) for *cagH* / *orf17* ($P = 0.029$ and 0. 02, respectively). No significant correlation was found between *cagH* / *cagG* , *cagL* / *cagG* , *cagL* / *orf17* ,

cagG / *orf17* , *cagH* / *cagG* / *cagL* , *cagH* / *cagG* / *orf17* , *cagH* / *cagL* / *orf17* , and *cagL* / *cagG* / *orf17* combined genotypes and the risk of PU (*P* > 0. 05).

Multiple logistic regression analysis illustrated that *cagL* genotype was independently and significantly correlated with the age and sex-adjusted risk (Ø¨Ù‡ Ø¬Ø§ÛŒ Ø§ÛŒÙ† Ø¬Ù…Ù„ Ù‡ Ú†ÛŒ Ù…ÛŒ ØªÙˆÙ†Ù… Ø¨Ù†ÙˆÛŒØ³Ù…ØŸØŸ) for PU, the OR (95%CI) was 9. 557 (1. 219-7. 185) (*P* = 0. 032).

Association between the virulence factors and GC in Iran: Of the 42 GC patients, 42. 8% were with cardia cancer (CC), and 57. 1% with non-cardia cancer (NCC) and 38. 1% with diffuse-, 57. 1% with intestinal-, and 2. 3% with mucin producing-type adenocarcinomas, and 2. 3% with invasive squamous cell-type carcinoma. Chi-square test demonstrated no correlation between *cagH* , *cagL* , *cagG* , and *orf17* genotypes and risk of GC (*P* > 0. 05). statistical analysis for combined genotypes illustrated that *cagH* / *cagG* reduced the risk of GC, the OR (95%CI) was 0. 388 (0. 166-0. 908) (*P* = 0. 029). However, No remarkable association was found between *cagH/cagL* , *cagH* / *orf17* , *cagL* / *cagG* , *cagL* / *orf17* , *cagG* / *orf17* , *cagH* / *cagG* / *cagL* , *cagH* / *cagG* / *orf17* , *cagH* / *cagL* / *orf17* , and *cagL* / *cagG* / *orf17* combined genotypes and risk of GC (*P* > 0. 05).

Discussion

The presence of *H. pylori* in the gastric mucosa has been known as an essential risk factor of different gastrointestinal disease including; NAG, PU, and GC.** there are several virulence factors within *H. pylori* genome that might take part in mucosal damage.** This study investigated the relevance

of various virulence factors (*cagH* , *cagL* , *cagG* , and *orf17*) and their relationship with sever gastrointestinal disease. Studies showed that *cagG* mutant didn't result in inflammatory response or increase proliferation, and also demonstrated that lack of *cagG* gene leads in loss of CagA translocation/phosphorylation.**Recent reports showed that loss of *cagG* genotype have reduced adherence to epithelial cells.* An in vivo study in China have also illustrated no relationship between *cagG* and clinical outcomes. Most reports, suggested that lack of *cagG* genotype leads in complete removal of *H. pylori* IL-8 induction.**However, other studies suggested that copmlete deletion of *cagG* gene resulted in no reduction in IL-8 induction,** according to these data it is impossible to distinguish whether or not lack of inflammation with *cagG* mutant is associated with reduced colonization, loss of T4SS, or both.* A study on 145 isolates in china showed that *cagG* was detected in 91. 7% *H. pylori* isolates, 100% of isolates from patients with PU (Duodenal Ulcer and Gastric Ulcer) were *cagG* positive which was high , but not statistically significant ($P > 0.05$). *cagG* genotype was known as a conservative gene in chines population and there was no significant differences in the frequencies of *cagG* gene in isolates from patients with various digestive disease.** Hsu *et al* reported that of the 120 isolates from patients with different gastrointestinal disease in Korea 86. 7% (104/120) were *cagG* positive.** Mizushamia *et al* used PCR to investigate the distribution of *cagG* gene in 236 clinical *H. pylori* isolates in Japan, and found that 97% of isolates were *cagG* positive. For both isolates from Korea and Japan ($P > 0.05$), demonstrating no significant association with gastrointestinal disease.**In other study level of gastric mucosal inflammation was compared in the antrum and body (corpus) of both *cagG*

positive and negative group, it was illustrated that the level of inflammation was relatively higher in *cagG* positive than *cagG* negative group, both in antrum and body, ($P > 0.05$). resulting *cagG* genotype has no relationship with intensity of gastritis.**The prior results concordant with our results, we concluded that *cagG* genotype had no remarkable association with digestive disease in Iran ($P > 0.05$).

A recent study on *cagL* genotype in India, illustrated that 86. 6% of isolates were *cagL* positive.***other studies in Malaysia, Singapore, and Taiwan demonstrated > 85% of the isolates were *cagL* positive,** furthermore, the presence of RGD motif in the isolates suggested that the CagA oncoprotein translocation is mediated by an RGD dependent pathway, therefore it is crucial to find out whether or not the *cagL* amino acid sequence polymorphisms are correlated with clinical outcomes.??? A study on 61 isolates from patients with digestive disease in Iran, showed that 96. 7% were *cagL* positive. This report was concordant with the results from Taiwan (98. 6%)***most of the patients were *cagL* positive, but no remarkable association was detected between the cagL genotype and clinical outcomes ($P > 0.05$).** these results are in consistent with our results in GC group but not in PU group, demonstrating that *cagL* genotype is remarkably and independently associated with the risk of PU in Iran ($P = 0.021$), but no association was found with the risk of GC.

cagH and *orf17* have not been surveyed on genomics level; however, we found out the *cagH* genotype had no relationship with gastrointestinal disease in Iran ($P > 0.05$), and we concluded that *orf17* genotype had no association with GC group ($P > 0.05$) but there was a remarkable

relationship between *orf17* gene and an increased risk of PU in Iran (*P* = 0. 015).

Conclusion

We concluded that, if specific *H. pylori* virulent biomarkers can be known to predict PU or GC risk, patients with NAG can be selected for *H. pylori* eradication, and also several virulence factors must be studied simultaneously, in order to clarify the correlation between the virulence factors and gastrointestinal disease. It is proposed that the *orf17* and specially *cagL* genotypes of *H. pylori cag* PAI might be as beneficial factors for the risk prediction of PU, but not GC in Iran. This report is the first one regarding the relevance of *H. pylori orf17* genotype to PU in Iran(ØªÙˆ Ø¨Ø®Ø´ Ø¯ÛŒÚ¯Ù‡ Ø§ÛŒ Ø¨ÛŒØ§Ø±Ù… Ø§ÛŒÙ† Ø¬Ù…„ Ù‡ Ø±Ùˆ Ø¨Ù‡Øª Ù†ÛŒØ³ØªØŸØŸØŸ!!). No remarkable correlation was found between the *cagH* and *cagG* genotypes and the risk of both the PU and GC in Iran. Genotypes combination analysis showed that *cagH* / *cagG* decrease the risk of GC in Iran, but no consistent relationship was observed with PU, and *cagH* / *orf17* and *cagH* / *cagL* increased the risk of PU, not GC in Iran. (Ø§ÛŒÙ† Ù, Ø³Ù…Øª ØØ°Ù) Ø¨Ø´Ù‡ ÛŒØ§ Ù†Ù‡ØŸØŸ!! ØŸ)

Virulence factors	Disease type	*P*-value	OR	CI	Total strains
PU	NAG				
N (%)	N (%)				
cagH	40/54	87/146	0.	1.	0. 969- 127

	(74. 1)	(59. 6)	061	938	3. 873	
cagL	53/54 (98. 1)	121/146 (82. 9)	0. 021	10. 950	1. 446-82. 935	174
cagG	38/54 (70. 4)	109/146 (74. 7)	0. 542	0. 806	0. 403-1. 612	147
orf17	43/54 (79. 6)	89/146 (61)	0. 015	2. 504	1. 193-5. 253	132
cagH/ cagL	40/41 (97. 7)	82/102 80. 4)	0. 029	9. 756	1. 264-75. 303	122
cagH / cagG	32/40 (80)	70/90 (77. 8)	0. 776	1. 143	0. 455-2. 869	102
cagH/ orf17	34/41 (82. 9)	73/116 (62. 9)	0. 022	2. 861	1. 167-7. 013	107
cagL/ cagG	36/36 (100)	96/108 (88. 9)	—	—	—	132
cagL/ orf17	41/43 (95. 3)	81/97 (83. 5)	0. 071	4. 049	0. 888-18.	122

					463	
cagG/ orf17	31/35 (88. 6)	70/89 (78. 7)	0. 208	2. 104	0. 661- 6. 698	101
cagH/ cagG/ cagL	31/31 (100)	67/77 (87)	—	—	—	98
cagH/ cagG/ orf17	28/32 (87. 5)	61/75 (81. 3)	0. 438	1. 607	0. 485- 5. 323	89
cagH/ cagL/ orf17	34/35 (97. 1)	70/84 (83. 3)	0. 069	6. 800	0. 858- 53. 873	104
cagL/ cagG/ orf17	31/31 (100)	67/75 (89. 3)	—	—	—	98

Virulence factors	Disease type	*P* - value	OR*	CI‡	Total strains	
GC N (%)	NAG N (%)					
cagH	21/42	87/146	0.	0.	0. 340-	108

	(50)	(59. 6)	269	678	1. 351	
cagL	35/42 (83. 3)	121/146 (82. 9)	0. 945	1. 033	0. 412-2. 589	156
cagG	26/42 (61. 9)	109/146 (74. 7)	0. 108	0. 552	0267-1. 140	136
orf17	23/42 (54. 8)	89/146 (61)	0. 471	0. 775	0. 388-1. 550	112
cagH/ cagL	21/28 (75)	82/102 (80. 4)	0. 534	0. 732	0. 273-1. 960	103
cagH/ cagG	19/33 (57. 6)	70/90 (77. 8)	0. 029	0. 388	0. 166-0. 908	89
cagH/ orf17	20/38 (52. 3)	73/116 (62. 9)	0. 261	0. 654	0. 312-1. 372	93
cagL/ cagG	24/29 (82. 8)	96/108 (88. 9)	0. 378	0. 600	0. 193-1. 867	120
cagL/ orf17	22/28 (78. 6)	81/97 (83. 5)	0. 547	0. 724	0. 253-2. 070	103
cagG/	19/31	70/89	0.	0.	0. 178-	89

orf17	(61. 3)	(78. 7)	061	430	1. 039	
cagH/ cagG/ cagL	19/24 (79. 2)	67/77 (87)	0. 350	0. 567	0. 173-1. 861	86
cagH/ cagG/ orf17	19/30 (63. 3)	61/75 (81. 3)	0. 054	0. 396	0. 154-1. 018	80
cagH/ cagL/ orf17	20/26 (76. 9)	70/84 (83. 3)	0. 461	0. 667	0. 227-1. 959	90
cagL/ cagG/ orf17	19/23 (82. 6)	67/75 (89. 3)	0. 394	0. 567	0. 154-2. 089	86

Characteristics	Total frequency N (%)
Age	
<55	150/241 (62)
>= 55	91/241 (37. 6)
Sex	98/242 (40. 5)

Female = 0	
	144/242 (59. 5)
Male = 1	
Non-atrophic gastritis	146/242 (60. 3)
Peptic ulcer	54/242 (22. 3)
Duodenal ulcer	28/54 (51. 8)
Gastric ulcer	20/54 (37. 03)
Gastric cancer	
Cardia cancer	42/242 (17. 4)
Non-cardia cancer	18/42 (42. 85)
Intestinal-type	24/42 (57. 14)
adenocarcinoma	24/42 (57. 14)
Diffuse-type adenocarcinoma	16/42 (38. 09)
Mucin producing-type	1/42 (2. 38)
adenocarcinoma	1/42 (2. 38)
Invasive squamous cell-type	
carcinoma	

Gene	
and	Sequences (5á¿½- 3á¿½)
Primer	
cagH	5á¿½-ATGGCAGGTACACAAGCTAT-3á¿½

CagH-F	
CagH-R	5á¿½-TCACTTCACGATTATTTTAG-3á¿½
cagL	
CagL-15	5á¿½-AAAACACTCGTGAAAAATACCATATC-3á¿½
CagL-16	5á¿½-TCGCTTCAAAATTGGCTTTC-3á¿½
cagG	
CagG-F	5á¿½-TTATAAAATTAAATTACTATTTGC-3á¿½
CagG-R	5á¿½-GTGGTAAAAACGATGAATCTG-3á¿½
orf17	
Orf17-F	5á¿½-CTTGATTGATGAAAATTTGGTTG-3á¿½
Orf17-R	5á¿½-TTAGTGATATATTCATAATTTTCC-3á¿½