STOMACH

Identification of a genetic marker of Helicobacter pylori strains involved in gastric extranodal marginal zone B cell lymphoma of the MALT-type

P Lehours, S Dupouy, B Bergey, A Ruskoné-Foumestraux, J C Delchier, R Rad, F Richy, J Tankovic, F Zerbib, F Mégraud, A Ménard


Background and aims: Gastric extranodal marginal zone B cell lymphoma of the mucosa associated lymphoid tissue (MALT)-type (MZBL) is a rare complication of Helicobacter pylori infection. Currently, no bacterial factor has been associated with the development of this disease. Our aim was to identify genes associated with lymphoma development.

Methods: We used subtractive hybridisation as a tool for comparative genomics between H pylori strains isolated from a patient with gastric MZBL and from a patient with gastritis only.

Results: When gastric MZBL strains were compared with gastritis strains, two open reading frames (ORFs) were significantly associated with gastric MZBL: JHP950 (74.4% v 48.7%, respectively; p = 0.023) and JHP1462 (25.6% v 2.6%, respectively; p = 0.004). The prevalence of JHP950 was 48.8% (p = 0.024) in duodenal ulcer strains and 39.3% (p = 0.006) in gastric adenocarcinoma strains, which makes this ORF a specific marker for gastric MZBL strains. In contrast, the prevalence of JHP1462 was 16% (p = 0.545) and 35.7% (p = 0.429) in duodenal ulcer and adenocarcinoma strains, respectively. These ORFs were present in reference strain J99 but not in reference strain 26695. JHP950 is located in the plasticity zone whereas the other, JHP1462, is located outside. Both encode for H pylori putative proteins with unknown functions.

Conclusion: Despite its low prevalence, the ORF JHP1462 can be considered a candidate marker for H pylori strains involved in severe gastroduodenal diseases. In contrast, the ORF JHP950 has a high prevalence, and is the first candidate marker for strains giving rise to an increased risk of gastric MZBL strains. Further confirmation in other studies is needed.

Helicobacter pylori is the main cause of gastritis and is involved in the pathogenesis of several gastrointestinal diseases—for example, duodenal ulcer and gastric cancer. H pylori is the first bacterium identified as being carcinogenic in humans. It has been related to two types of stomach neoplasia: gastric adenocarcinoma and mucosa associated lymphoid tissue (MALT) lymphoma. The relationship of H pylori infection and extranodal marginal zone B cell lymphoma of the MALT-type (MZBL) is now well established. Indeed, H pylori is the cause of approximately 80% of gastric MZBL. Gastric MZBL arises from a H pylori associated chronic gastritis which consists of polymorph and lymphocyte infiltration, and evolves from an initially polyclonal lymphoid proliferation towards a monoclonal lymphoid population. This evolution is accompanied by strain specific. We postulated that among these inconstant genes, operating alone or clustered together, some could be involved in gastric MZBL development.

The aim of our study was to identify genetic markers associated with H pylori strains isolated from patients with gastric MZBL in comparison with strains obtained from age matched dyspeptic patients with gastritis only. Using subtractive hybridisation, a technique which allows identification of genes or sequences present in a particular strain in comparison with a chosen control strain.

EXPERIMENTAL PROCEDURES

Bacterial strains

The H pylori strains used for subtractive hybridisation were: (1) a strain called “tester”, isolated from a 29 year old male patient with gastric MZBL; and (2) a strain chosen as a control, called “driver”, from a 74 year old male patient with gastritis but no lymphoid infiltration, atrophy, or intestinal metaplasia. These two strains presented a common profile for six known virulence genes: that is, cagA+, iceA+, vacA+, iceC+, hilD+, and sabA. Nevertheless, a gene cluster comprised of iceA1 allele, sabA “on” status, and hopZ “off” status was identified in strains for which the odds of having a gastric MZBL was 10 times higher than that of the gastritis strains.

Following sequencing of the whole genome of two H pylori strains (J99 and 26695), DNA arrays were used on various strains and the results showed that 22% of the genes were indeed strain specific. We postulated that among these inconstant genes, operating alone or clustered together, some could be involved in gastric MZBL development.

The aim of our study was to identify genetic markers associated with H pylori strains isolated from patients with gastric MZBL in comparison with strains obtained from age matched dyspeptic patients with gastritis only, using subtractive hybridisation, a technique which allows identification of genes or sequences present in a particular strain in comparison with a chosen control strain.

Abbreviations: RT-PCR, reverse transcription-polymerase chain reaction; MALT, mucosa associated lymphoid tissue; gastric MZBL, gastric extranodal marginal zone B cell lymphoma of the MALT-type; MCA, multiple correspondence analysis; ORF, open reading frame
Fe´de´ration Francophone de Cance´rologie Digestive and the d’Etude Franc¸ois des Lymphomes Digestifs (GELD) of the years were used. These strains were isolated from patients MZBL patients (27 men and 16 women, mean age 48.2 (13.4) H pylori ‘‘on’’. kanamycin and grown overnight at 37 ˚C under agitation. manufacturer’s recommendations. A total of 139 recombinants were cloned in the pTAdv plasmid (Clontech) according to the manufacturer’s recommendations. Twenty eight DNA samples extracted from H pylori strains isolated from German patients with gastric adenocarcinoma (12 men and 16 women; mean age 64.3 (9.4) years) were also included in this study. The two H pylori strains for which the genome has been sequenced, 26695 and J99, were used as reference strains.25 Genomic DNA was extracted from isolated colonies using a commercial kit (QiaGen SA, Courtaboeuf, France). The integrity of the genomic DNA template was controlled by polymerase chain reaction (PCR) with primers specific to a region of the 16S rDNA of H pylori. Twenty eight DNA samples extracted from H pylori strains isolated from German patients with gastric histological gastritis only (29 men and 10 women; mean age 48.3 (14.7) years) and 41 H pylori strains isolated from age matched duodenal ulcer patients (34 men and 7 women; mean age 49.7 (14.2) years) were also included. Isolation of H pylori from gastric biopsy specimens was carried out using standard culture methods, as previously described.25

Subtractive hybridisation

The PCR select bacterial genome subtraction kit (Clontech, Palo Alto, California, USA) was used with the ‘‘tester’’ (gastric MZBL) and ‘‘driver’’ (gastritis) strains. The manufacturer’s recommendations were followed with the exception of the choice of the restriction enzyme and the method of evaluation of the subtractive efficacy. The restriction enzyme AclI was chosen because of its ability to generate DNA fragments, from ‘‘tester’’ and ‘‘driver’’ strains, in the same molecular weight range (from 2.5 kb to 0.25 kb). The efficacy of subtractive hybridisation was evaluated using amplification of a fragment of H pylori 16S rDNA in a real time PCR experiment in a LightCycler thermocycler (Roche Diagnostics, Neuilly sur Seine, France). Results were plotted versus time, represented by cycle number, to produce a measure of PCR amplification at the end of each PCR cycle. A difference of 20 threshold cycles was observed between subtracted and unsubtracted products, indicating an important reduction in 16S rDNA abundance in the subtracted sample and unsubtracted products, indicating an important reduction 25

Screening of the subtractive library

To determine if the subtractive clones were really ‘‘tester’’ specific, primers for each clone were designed using the conserved sequence between the sequence identified and the corresponding gene of H pylori in database sequences (http://www.probes.toulouse.inra.fr/multialin/multialin.html).26 The specificity of our primers was tested on both H pylori reference DNAs. For the sequences with no homology found in the databases, the specificity of the amplified fragment was determined by direct sequencing and dot blot hybridisation between ‘‘tester’’ (gastric MZBL) and ‘‘driver’’ (gastritis) DNAs, as described below.

Dot blot hybridisation

Screening of the PCR sequences specific for the ‘‘tester’’ strain was performed using our collection of H pylori gastric MZBL and gastritis strains. Dot blot analyses were carried out using the same batch of membranes, prepared with the same dilution of DNA. Genomic DNA (100 ng) of each strain (MZBL, gastritis, J99, and 26695 reference strains) were spotted on a Hybond N+ membrane (Amersham Pharmacia Biotech) by means of a Bio-Rad dot blot 96 well filtration system apparatus (Bio-Rad, Ivry-Sur-Seine, France). Hybridisation was performed using digoxigenin labelled probes of each ORF identified, as described previously.27 The probes were prepared by PCR after primer design in a conserved area between the ‘‘tester’’ DNA sequence and the J99 reference strain or between both the J99, 26695, and ‘‘tester’’ strain when the DNA sequence was present in all.

Reverse transcription-PCR

DNA was extracted from the ‘‘tester’’ strain and the reference strain 399 during the exponential growth phase using the RNAeasy kit (Qiagen). RNA samples were treated with Rnase-Free Dnase (Promega, Madison, Wisconsin, USA) and quantified spectrometrically. Each reverse transcription PCR (RT-PCR) was carried out in a 10 µl volume containing 250 ng of RNA, 10 µl of random oligonucleotide primer, 0.125 µl of RNAsine (Promega), 2 µl of 5× Q buffer, and 2 µl of RT-Taq mix. Complementary DNA was synthesised at 45°C for 30 minutes followed by 95°C heat inactivation of the RT Taq polymerase for 15 minutes, and then by the amplification program used for PCR amplification. Absence of DNA in RNA preparations was verified by omitting the RT amplification step. After 40 PCR cycles, products were separated electrophoretically on agarose gels.
Hp ylori marker for gastric MALT lymphoma

were also not tested. The seven remaining sequences (clones H, O, P, and Q) were distributed throughout the Hp ylori genome and were included in our dot blot experiments.

Among the 10 sequences matching with only strain J99, six were localised in the plasticity zone (clones I–N) and one of these corresponded to a deletion of two ORFs (that is, JHP950 and JHP961). These two ORFs were therefore PCR amplified for JHP950, JHP1462, and vacA genes, respectively. A–G (table 1) were distributed throughout the Hp ylori genome and were included in our dot blot experiments.

Among the 10 sequences matching with only strain J99, six were localised in the plasticity zone (clones I–N) and one of these corresponded to a deletion of two ORFs (that is, JHP960 and JHP961). These two ORFs were therefore PCR amplified using J99 DNA and also tested by dot blot. The four remaining sequences (clones H, O, P, and Q) were therefore PCR amplified using J99 DNA and also tested by dot blot. The four remaining sequences (clones H, O, P, and Q) were therefore PCR amplified using J99 DNA and also tested by dot blot.

The seven remaining sequences (clones I–N) were localised in the plasticity zone (clones J–N) and one of these corresponded to a deletion of two ORFs (that is, JHP950 and JHP961). These two ORFs were therefore PCR amplified for JHP950, JHP1462, and vacA genes, respectively. A–G (table 1) were distributed throughout the Hp ylori genome and were included in our dot blot experiments.

Among the 10 sequences matching with only strain J99, six were localised in the plasticity zone (clones I–N) and one of these corresponded to a deletion of two ORFs (that is, JHP960 and JHP961). These two ORFs were therefore PCR amplified using J99 DNA and also tested by dot blot. The four remaining sequences (clones H, O, P, and Q) were therefore PCR amplified using J99 DNA and also tested by dot blot. The four remaining sequences (clones H, O, P, and Q) were therefore PCR amplified using J99 DNA and also tested by dot blot.

The odds of an association with MZBL and its 95% confidence interval (CI) were compared within groups of strains and between the different strata of the summary variables obtained by the MCA.

Table 1

<table>
<thead>
<tr>
<th>Clone</th>
<th>Strain J99</th>
<th>Strain 26695</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>JHP305</td>
<td>HP322</td>
<td>Putative poly E-rich protein†</td>
</tr>
<tr>
<td>B</td>
<td>JHP428</td>
<td>HP476</td>
<td>Predicted glutamyl-tRNA synthetase†</td>
</tr>
<tr>
<td>C</td>
<td>JHP1031</td>
<td>HP1105</td>
<td>Putative LPS biosynthesis protein</td>
</tr>
<tr>
<td>D</td>
<td>JHP1044</td>
<td>HP1116</td>
<td>Putative</td>
</tr>
<tr>
<td>E</td>
<td>JHP1285</td>
<td>HP1371</td>
<td>Putative</td>
</tr>
<tr>
<td>F</td>
<td>JHP1301</td>
<td>HP1409</td>
<td>Putative</td>
</tr>
<tr>
<td>G</td>
<td>JHP1424</td>
<td>HP1402</td>
<td>Type I RE (restriction subunit)</td>
</tr>
<tr>
<td>H</td>
<td>JHP820</td>
<td>None</td>
<td>Predicted 1,2-glycosyltransferase†</td>
</tr>
<tr>
<td>I</td>
<td>JHP921†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>J</td>
<td>JHP922†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>K</td>
<td>JHP946†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>L</td>
<td>JHP949†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>M</td>
<td>JHP950†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>N</td>
<td>JHP960†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>O</td>
<td>JHP1297</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>P</td>
<td>JHP1409</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>Q</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clone</th>
<th>Strain J99</th>
<th>Strain 26695</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>S</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>T</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>U</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>V</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>W</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>X</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>Y</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>Z</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clone</th>
<th>Strain J99</th>
<th>Strain 26695</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>JHP305</td>
<td>HP322</td>
<td>Putative poly E-rich protein†</td>
</tr>
<tr>
<td>B</td>
<td>JHP428</td>
<td>HP476</td>
<td>Predicted glutamyl-tRNA synthetase†</td>
</tr>
<tr>
<td>C</td>
<td>JHP1031</td>
<td>HP1105</td>
<td>Putative LPS biosynthesis protein</td>
</tr>
<tr>
<td>D</td>
<td>JHP1044</td>
<td>HP1116</td>
<td>Putative</td>
</tr>
<tr>
<td>E</td>
<td>JHP1285</td>
<td>HP1371</td>
<td>Putative</td>
</tr>
<tr>
<td>F</td>
<td>JHP1301</td>
<td>HP1409</td>
<td>Putative</td>
</tr>
<tr>
<td>G</td>
<td>JHP1424</td>
<td>HP1402</td>
<td>Type I RE (restriction subunit)</td>
</tr>
<tr>
<td>H</td>
<td>JHP820</td>
<td>None</td>
<td>Predicted 1,2-glycosyltransferase†</td>
</tr>
<tr>
<td>I</td>
<td>JHP921†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>J</td>
<td>JHP922†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>K</td>
<td>JHP946†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>L</td>
<td>JHP949†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>M</td>
<td>JHP950†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>N</td>
<td>JHP960†</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>O</td>
<td>JHP1297</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>P</td>
<td>JHP1409</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
<tr>
<td>Q</td>
<td>JHP1462</td>
<td>None</td>
<td>Predicted DNA TC ComB8 homolog†</td>
</tr>
</tbody>
</table>

p values were determined by Fisher’s exact test.

*Open reading frame (ORF) localised in the plasticity zone.

†Revised annotation according to Boneca and colleagues.

**M2, extranodal marginal zone B cell lymphoma of the MALT-type; LPS, lipopolysaccharide; RE, restriction enzyme; ME, modification enzyme; TC, transformation competence.

Statistical analysis

Statistical analysis was performed using STATA 7.0 statistical software (Stata Corporation, College Station, Texas, USA). A first comparison of the characteristics between MZBL and gastritis strains was performed using Fisher’s exact test.

In order to create a variable describing how the different strain characteristics were associated, a multiple correspondence analysis (MCA) was performed followed by a cluster analysis using the average aggregation method and Euclidian distance. Both methods used the Statbox 2.5 program (Grimmer Logiciels, Paris, France).

The odds of an association with MZBL and its 95% confidence interval (CI) were compared within groups of strains and between the different strata of the summary variables obtained by the MCA.

RESULTS

PCR analysis of the subtractive clones

We sequenced the inserts of 132 clones on both strands, of which 81 were non-redundant. PCR screening between “tester” (gastric MZBL) and “driver” (gastritis) DNAs of these 81 clones showed that 39 (48.2%) contained sequences which were specific to the “tester” strain used in the experiment. Homologies of these clones with the Hp ylori databases (http://www.tigrblast.tigr.org/cmr-blast) and the NCBI blast program (http://www.ncbi.nlm.nih.gov/BLAST) were analysed. The following classification was made: 26 matched with sequences from the two reference strains (J99 and 26695), 10 matched with only J99 specific sequences, none matched with only strain 26695, and three presented no homology with Hp ylori reference strains.

For 10 of the 26 sequences which matched with the two reference strains, PCR products obtained from the “tester” DNA were different from the expected size when compared with the “driver” DNA and Hp ylori J99 and 26695 reference DNAs. Direct sequencing of these products showed significant variations, such as insertions or deletions, and therefore they were not included in the dot blot screening of our strain collections to avoid impairing the analysis. Nine other sequences where variation corresponded to intergenic regions were also not tested. The seven remaining sequences (clones A–G) (table 1) were distributed throughout the Hp ylori genome and were included in our dot blot experiments.

Among the 10 sequences matching with only strain J99, six were localised in the plasticity zone (clones I–N) and one of these corresponded to a deletion of two ORFs (that is, JHP960 and JHP961). These two ORFs were therefore PCR amplified using J99 DNA and also tested by dot blot. The four remaining sequences (clones H, O, P, and Q) were therefore PCR amplified using J99 DNA and also tested by dot blot. The four remaining sequences (clones H, O, P, and Q) were therefore PCR amplified using J99 DNA and also tested by dot blot.
Dot blot analysis of the selected clones (MALT v gastritis)
The distribution of the 24 selected sequences was tested by dot blot using our collection of *H pylori* gastric MZBL and gastritis strains. The specificity of the sequences found to be "tester" specific by PCR was verified by (1) the absence of hybridisation on the plot corresponding to the "driver" strain and (2) the absence of hybridisation for the original sequences on the plots corresponding to the "driver" DNA, strain J99, and strain 26695. Indeed, negative results by PCR may be due to either an absence of the tested gene (true negative) or a lack of primer annealing due to interstrain variation in the sequences targeted by the primers. Accordingly, three sequences, which matched with the two reference strains and were initially classified by PCR as "tester" specific, were removed from our panel because of a positive hybridisation with the "driver" plot DNA.

The 21 remaining sequences were then tested. The corresponding genes and results obtained are summarised in table 1. The proportion of each ORF was compared between strains from patients with gastric MZBL (table 3) and with gastritis only. Two ORFs were significantly associated with gastric MZBL strains: JHP950 (74.4% v 48.8%, respectively; p = 0.023) and JHP1462 (25.6% v 26.4%, respectively; p = 0.001). JHP950 was located in the plasticity region of strain J99 whereas JHP1462 was located outside. These two ORFs encoded for two *H pylori* specific proteins of unknown function.

Study of the transcription of JHP950 and JHP1462 ORFs in gastric MZBL strain
RT-PCR was performed for JHP950 and JHP1462 using the primers designed for PCR screening of our subtractive library in order to verify if they were transcribed in our "tester" strain. The vacA gene was included as a positive control of RT-PCR using primers previously published.29 The results showed that both were expressed (fig 1).

Multiple correspondence analysis
Firstly, an MCA was performed for all variables obtained from dot blot. We were unable to identify gene clusters which were significantly associated with gastric MZBL or gastritis strains. Thus the analysis was focused on the two ORFs found associated with the gastric MZBL strains. Because the prevalence of JHP1462 in gastric MZBL and gastritis strains was low, it was impossible to keep it in our final analysis. In contrast, the ORF JHP950 was included in the data from our previous study in which we identified a gene cluster in gastric MZBL strains, which is absent in gastritis strains, and containing *iceA1*, *sabA* "on", and *bapZ* "off".19 Interestingly, the dendrogram obtained from the results of the cluster analysis of the *H pylori* gastric MZBL strain population showed that the ORF JHP950 was included in the formally identified gene cluster (data not shown). The possible association between the presence of gastric MZBL and the characteristics of strains cultured from these patients was tested. On analysis of the summary variables according to the MCA results, it was remarkable that the odds of having gastric MZBL among patients harbouring JHP950, *iceA1*, and *sabA* "on" strains were 10 times higher than for the other defined groups (odds ratio 10.3 (95% CI 1.2–86.0)) (table 2); indeed among a total of 11 strains having these three markers, 10 (90.9%) corresponded to gastric MZBL strains (p = 0.018) (fig 2).

Verification of the specificity of the two putative markers
The specificity of the association of the ORFs JHP950 and JHP1462 with gastric MZBL strains was evaluated by determining their prevalence in a collection of duodenal ulcer strains (n = 41) by dot blot analysis, as described above. The prevalence in duodenal ulcer strains of ORFs JHP950 and JHP1462 was 48.8% (p = 0.024) and 16% (p = 0.545), respectively.

The prevalence of ORFs JHP950 and JHP1462 on strains isolated from 28 gastric adenocarcinoma was also determined. As DNA obtained from these strains was not concentrated enough to perform dot blot analysis, PCR screening was performed with the primers used for dot blot probe synthesis. A prevalence of 39.3% (p = 0.006) for JHP950 and of 35.7% (p = 0.429) for JHP1462 ORFs was found (table 3).

These results confirm that JHP950 is a specific marker for *H pylori* strains isolated from gastric MZBL (table 3).

**DISCUSSION**
Following sequencing of the whole genome of two strains of *H pylori*, the possibility arose to study the genetic diversity of

---

**Table 2** Strength of the association of Helicobacter pylori genotypes and extranodal marginal zone B cell lymphoma of the MALT-type (gastric MZBL) status versus gastritis

<table>
<thead>
<tr>
<th>Total No of patients</th>
<th>Patients with MZBL</th>
<th>Patients with gastritis</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JHP950</td>
<td>82</td>
<td>43</td>
<td>39</td>
</tr>
<tr>
<td>Absent</td>
<td>31</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>Present</td>
<td>51</td>
<td>32</td>
<td>19</td>
</tr>
<tr>
<td>MCA summary variables</td>
<td>70</td>
<td>39</td>
<td>31</td>
</tr>
<tr>
<td>Others</td>
<td>59</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td><em>iceA1</em>, <em>sabA</em> &quot;on&quot;, and JHP950+</td>
<td>11</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>

OR, odds ratio, presented as univariate results; CI, exact binomial 95% confidence interval; MCA, multiple correspondence analysis.
Subtractive hybridisation was later validated by H pylori hybridisation, leading to the discovery of the cag pathogenicity island, a method similar to subtractive hybridisation and found that most were functional. We hypothesized that H pylori pathogenesis in gastric MZBL could be linked to currently unidentified bacterial factors, and therefore that subtractive hybridisation represented the most accurate approach.

Subtractive hybridisation is now considered an important tool for comparative prokaryotic genomics. Most of the studies performed previously compared bacterial strains with apparent enhanced capacity for human infection and/or interpatient transmission and colonisation. For example, genomic differences have been identified in a virulent strain of Klebsiella pneumoniae as well as in the uropathogenic Escherichia coli strain 536 compared with the non-pathogenic strain MG1655, and more recently a genomic locus was represented the most accurate approach.

The first category of sequences identified presented specific sequences which could provide a basis for developing bacteria from different species in order to identify nucleotide sequences that are characteristic of a strain or strain type. For example, H pylori strains to be compared is particularly important. Considering the fact that the natural evolution of H pylori infection begins with gastritis and moreover that host susceptibility factors of H pylori strains associated with severe gastroduodenal diseases.

The prevalence of ORF JHP1462 (table 3) in MZBL strains was indeed low (25.6%) but when compared with gastritis strains the difference was not significant (p = 0.545). However, no significant difference was found between gastric MZBL strains and duodenal ulcer and gastric adenocarcinoma strains (table 3), indicating that JHP1462 is not a specific genetic marker for H pylori gastric MZBL strains but could be considered more generally as a candidate marker for H pylori strains associated with severe gastroduodenal diseases. It is located outside the plasticity zone, in common with two thirds of H pylori strain specific genes.

In a subtractive hybridisation experiment, the choice of strains to be compared is particularly important. Considering the fact that the natural evolution of H pylori infection begins with gastritis and moreover that host susceptibility factors of developing gastric MZBL are currently unknown, this choice is especially difficult. We based our choice on the pathology (gastric MZBL versus gastritis) and age of the patients from whom the strains were isolated (29 year old patient with gastric MZBL, 74 year old patient suffering from gastritis). The advanced age of the “driver” strain patient was chosen to minimise the possibility that gastritis would evolve to more severe gastrointestinal disease. As the mean age of patients suffering from gastric MZBL is close to 50 years, we could anticipate a particular pathogenic property in the tester strain issued from a much younger patient. Furthermore, the two strains were chosen because of their common profile for six known H pylori virulence genes—that is, cagA+, cagE+, vacA+1,11 babA2, iceA1, and sipA- “on”, and one adhesin sahA “on”.

MCA analysis, initially carried out on all of our dot blot data, failed to identify a gene cluster associated with gastric MZBL strains. However, despite the fact that our collection of gastric MZBL strains is the largest of those published, the number of strains included in this study may not be sufficient to efficiently perform MCA.

The first category of sequences identified presented significant homologies with genes present in strains 399 and 26695 but with important deletions or insertions. Such variations can modify the functional status of a gene or the functionality of the deduced protein. Moreover, it has been suggested that mixtures of matching and divergent sequences can facilitate quite dramatic changes in bacterial phenotype. We originally expected to find new H pylori genes and indeed we identified three sequences of unknown genomic localisation, corresponding to new putative coding regions. Nevertheless, none of these sequences was significantly associated with gastric MZBL. Such results could not have been found using a standard approach for detecting genetic diversity, and thus our findings provide more information for the future in investigating the particular properties of these strains. The results of the dot blot experiments (MALT versus gastritis) highlighted two ORFs (JHP950 and JHP1462). Globally, repartition of both ORFs (JHP950 and JHP1462) was statistically different (p = 0.016 and p = 0.002, respectively) when comparing the four H pylori strain collections studied.

The prevalence of ORF JHP1462 (table 3) in MZBL strains was indeed low (25.6%) but when compared with gastritis strains the difference was not significant (p = 0.545). However, no significant difference was found between gastric MZBL strains and duodenal ulcer and gastric adenocarcinoma strains (table 3), indicating that JHP1462 is not a specific genetic marker for H pylori gastric MZBL strains but could be considered more generally as a candidate marker for H pylori strains associated with severe gastroduodenal diseases. It is located outside the plasticity zone, in common with two thirds of H pylori strain specific genes.

In contrast, the prevalence of the JHP950 ORF in gastric MZBL strains was significantly higher than in strains isolated from duodenal ulcer and gastric adenocarcinoma. The difference in prevalence for this marker was in the same range as the difference between the prevalence of cagA in duodenal ulcer and gastric adenocarcinoma strains compared with gastritis strains. Indeed, in the past we looked for the cagA gene in relation to clinical presentation in a collection of 167 H pylori strains. The cagA gene was found in 82.6% of duodenal ulcer strains versus 63.3% of gastritis strains. These results were confirmed later in several European studies. The cagA gene was first considered as a marker of virulence for H pylori. Indeed, following the discovery of the cag pathogenicity island and of the interaction of the CagA protein with the epithelial cell, it is now considered to be a true virulence factor associated with severe clinical outcome (that is, duodenal ulceration and gastric adenocarcinoma). Similarly, the ORF JHP950 is part of the plasticity zone of strain J99. This region is approximately 45 kb long in strain J99 and 68 kb long in strain 26695, and contains one third of the strain specific genes. Occhialini et al studied several ORFs in the plasticity region of strain J99 by hybridisation and found that most were functional. We verified by RT-PCR that JHP950 and JHP1462 were transcribed. They were originally annotated as H pylori specific proteins with unknown functions (that is, they have no orthologs in other bacterial sequences). This classification is

### Table 3 Prevalence of JHP950 and JHP1462 open reading frames (ORFs) in H pylori strains isolated from patients with gastritis, duodenal ulcer, and gastric adenocarcinoma in comparison with strains isolated from patients with gastric MZBL

<table>
<thead>
<tr>
<th>Strain</th>
<th>Gastritis</th>
<th>Duodenal ulcer</th>
<th>Gastric adenocarcinoma</th>
<th>Gastric MZBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>JHP950</td>
<td>48.7 (p = 0.023)*</td>
<td>48.8 (p = 0.024)*</td>
<td>39.3 (p = 0.006)*</td>
<td>74.4</td>
</tr>
<tr>
<td>JHP1462</td>
<td>2.6 (p = 0.004)*</td>
<td>16.0 (p = 0.545)</td>
<td>35.7 (p = 0.429)</td>
<td>25.6</td>
</tr>
</tbody>
</table>

*Significant values versus gastritis strains.

The number of strains included in this study may not be sufficient to efficiently perform MCA.

The first category of sequences identified presented significant homologies with genes present in strains 399 and 26695 but with important deletions or insertions. Such variations can modify the functional status of a gene or the functionality of the deduced protein. Moreover, it has been suggested that mixtures of matching and divergent sequences can facilitate quite dramatic changes in bacterial phenotype. We originally expected to find new H pylori genes and indeed we identified three sequences of unknown genomic localisation, corresponding to new putative coding regions. Nevertheless, none of these sequences was significantly associated with gastric MZBL. Such results could not have been found using a standard approach for detecting genetic diversity, and thus our findings provide more information for the future in investigating the particular properties of these strains.

The results of the dot blot experiments (MALT versus gastritis) highlighted two ORFs (JHP950 and JHP1462). Globally, repartition of both ORFs (JHP950 and JHP1462) was statistically different (p = 0.016 and p = 0.002, respectively) when comparing the four H pylori strain collections studied.

The prevalence of ORF JHP1462 (table 3) in MZBL strains was indeed low (25.6%) but when compared with gastritis strains the difference was not significant (p = 0.545). However, no significant difference was found between gastric MZBL strains and duodenal ulcer and gastric adenocarcinoma strains (table 3), indicating that JHP1462 is not a specific genetic marker for H pylori gastric MZBL strains but could be considered more generally as a candidate marker for H pylori strains associated with severe gastroduodenal diseases. It is located outside the plasticity zone, in common with two thirds of H pylori strain specific genes.

In contrast, the prevalence of the JHP950 ORF in gastric MZBL strains was significantly higher than in strains isolated from duodenal ulcer and gastric adenocarcinoma. The difference in prevalence for this marker was in the same range as the difference between the prevalence of cagA in duodenal ulcer and gastric adenocarcinoma strains compared with gastritis strains. Indeed, in the past we looked for the cagA gene in relation to clinical presentation in a collection of 167 H pylori strains. The cagA gene was found in 82.6% of duodenal ulcer strains versus 63.3% of gastritis strains. These results were confirmed later in several European studies. The cagA gene was first considered as a marker of virulence for H pylori. Indeed, following the discovery of the cag pathogenicity island and of the interaction of the CagA protein with the epithelial cell, it is now considered to be a true virulence factor associated with severe clinical outcome (that is, duodenal ulceration and gastric adenocarcinoma).

Similarly, the ORF JHP950 is part of the plasticity zone of strain J99. This region is approximately 45 kb long in strain J99 and 68 kb long in strain 26695, and contains one third of the strain specific genes. Occhialini et al studied several ORFs in the plasticity region of strain J99 by hybridisation and found that most were functional. We verified by RT-PCR that JHP950 and JHP1462 were transcribed. They were originally annotated as H pylori specific proteins with unknown functions (that is, they have no orthologs in other bacterial sequences). This classification is...
still accurate after publication of a revised annotation and comparative analysis of \textit{H pylori} genomes.** Our results are in line with the study of Salama et al who found that the most important class of strain specific genes was more frequent among strain specific genes (58%) than in the core of the genome (21%) which contains genes encoding, for example, metabolic functions.15

A limit of our study was the use of PCR instead of hybridisation for detection of the ORFs JHP950 and JHP1462 in adenocarcinoma strains. Indeed, there are few collections of \textit{H pylori} strains obtained from patients with gastric adenocarcinoma in Europe. The DNA material obtained did not allow us to perform dot blot hybridisation but only detection by PCR. The latter technique can have false negative results because of a lack of primer binding due to limited differences in sequences. Therefore, we tested the accuracy of our method on our strain collections and found a sensitivity of 90% (JHP950) and 95% (JHP1462) versus dot blot hybridisation, which confirms the relevance of our approach.

MCA analysis allowed us to show that the JHP950 marker is part of a gene cluster significantly associated with gastric MZBL strains. This gene cluster, which is comprised of JHP950, \textit{iceA}1 alleles, and a \textit{sabA} “on” status, partially validates our initial findings.16 The \textit{sabA} gene encodes for an adhesin. The association of JHP950 with \textit{iceA}1, a gene activated by contact with the gastric epithelium, and an \textit{adhesin} seems to be in line with the function of these genes.1718 This underlines the fact that JHP950 might have functional importance during interaction of the bacterium with cells. We can hypothesise that if JHP950 corresponds to a new \textit{H pylori} virulence factor, \textit{H pylori} strains harbouring this triple association of genes could have a pathogenic advantage. Indeed, the odds of having gastric MZBL among patients harbouring JHP950, \textit{iceA}1, and \textit{sabA} “on” strains were 10 times higher than for the other defined groups. Koehler et al recently reported that the vacA m2 allele was the predominant subtype in MALT lymphoma.19 We also previously reported a slight association between m2 genotype and MALT strains (when considering the m genotype only) but we found no association between \textit{vacA} s1m2 alleles and \textit{iceA}1 genotypes.20 However, we share the similar conclusion (that is, certain \textit{H pylori} subtype combinations have a predictive value for the development of gastric MZBL).

In conclusion, if JHP1462 cannot be considered a specific genetic marker of gastric MZBL strains but rather as a new candidate marker of \textit{H pylori} virulent strains in general, in contrast JHP950 can be considered as the first candidate marker of gastric MZBL. \textit{H pylori} strains. The association of JHP950 with gastric MZBL strains is similar to that between \textit{cytA} and duodenal ulcer strains. Moreover, JHP950 forms a gene cluster with \textit{iceA}1 and \textit{sabA} “on” which is associated with gastric MZBL strains. Future studies are needed to confirm this association and to understand the biological function of this ORF in order to gain insight into the pathogenesis of \textit{H pylori} and the development of gastric MZBL.

**ACKNOWLEDGEMENTS**

The authors thank the Groupe d’Etude des Lymphomes Digestifs (GELD) of the Fédération Française de Cancérologie Digestive, the Délégation de la Recherche Clinique de l’Assistance Publique-Hôpitaux de Paris (AP-HP), the Groupe d’Etude des Lymphomes de l’Adulte (GELLA) for supplying \textit{H pylori} strains, and the Conseil Régional d’Aquitaine and the Fonds de Recherche of the Société Nationale Française de Gastroentérologie for financially supporting the project.

---

**Authors’ affiliations**
P Lehours, S Dupouy, B Bergey, F Richy, F Zerbib, F Mégraud, A Ménard, Laboratoire de Bactériologie, Université Victor Segalen Bordeaux 2, Bordeaux, France

**REFERENCES**

A painful liver

Clinical presentation

A 62 year old woman was admitted to Elche University General Hospital for further evaluation of an intermittent right upper quadrant pain unrelated to meals or posture. She had previously been well with no history of liver disease, alcohol abuse, or abdominal trauma. On physical examination, there were no signs of chronic liver disease or portal hypertension. Results of laboratory studies, liver function tests, and ultrasound guided liver biopsy were also normal. Real time and colour Doppler ultrasonography showed an anechoic mass-like lesion at the porta hepatis with continuous non-pulsatile flow within the lesion. The gallbladder was normal with no gallstones visualized. Then, a helical computed tomography (CT) scan was performed.

Question

What does this investigation show (fig 1). See page 943 for answer.

This case is submitted by:

I Oliver, J Lacueva, R Calpena
Department of Pathology and Surgery, Elche University General Hospital, Elche (Alicante), Spain

E Girona
Department of Internal Medicine, Elche University General Hospital, Elche (Alicante), Spain

Correspondence to: Professor I Oliver, Camí de l’Almazara, 11, 03202, Elche (Alicante), Spain; israel.oliver@umh.es
Identification of a genetic marker of *Helicobacter pylori* strains involved in gastric extranodal marginal zone B cell lymphoma of the MALT-type

P Lehours, S Dupouy, B Bergey, A Ruskoné-Foumestraux, J C Delchier, R Rad, F Richy, J Tankovic, F Żerbib, F Mégraud and A Ménard

*Gut* 2004 53: 931-937
doi: 10.1136/gut.2003.028811

Updated information and services can be found at: [http://gut.bmj.com/content/53/7/931](http://gut.bmj.com/content/53/7/931)

This article cites 49 articles, 20 of which you can access for free at: [http://gut.bmj.com/content/53/7/931#BIBL](http://gut.bmj.com/content/53/7/931#BIBL)

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections

Articles on similar topics can be found in the following collections

- Stomach and duodenum (1689)
- Pancreatic cancer (660)
- Campylobacter, Salmonella, Shigella, Escherichia coli (242)
- *Helicobacter pylori* (218)

Notes

To request permissions go to: [http://group.bmj.com/group/rights-licensing/permissions](http://group.bmj.com/group/rights-licensing/permissions)

To order reprints go to: [http://journals.bmj.com/cgi/reprintform](http://journals.bmj.com/cgi/reprintform)

To subscribe to BMJ go to: [http://group.bmj.com/subscribe/](http://group.bmj.com/subscribe/)