Draft Genome Sequences of Helicobacter pylori Isolates from Malaysia, Cultured from Patients with Functional Dyspepsia and Gastric Cancer


Updated information and services can be found at:
http://jb.asm.org/content/194/20/5695

These include:

**REFERENCES**
This article cites 8 articles, 3 of which can be accessed free at:
http://jb.asm.org/content/194/20/5695#ref-list-1

**CONTENT ALERTS**
Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to another ASM Journal go to: http://journals.asm.org/site/subscriptions/
Draft Genome Sequences of Helicobacter pylori Isolates from Malaysia, Cultured from Patients with Functional Dyspepsia and Gastric Cancer

Selva Perumal Gunatilchumy,a Xinsheng Teh,b Yelda Khosravi,a Nur Siti Khadijah Ramli,a Eng Guan Chua,a Thevakumar Kapitha,a Joanne N. Mason,b Huey Tyng Lee,b Halimah Alias,b Nur Zafrah Zaidan,b Norzawani Buang M. Yassin,b Liang Chung Tay,b Stephen Rudd,a Hazel M. Mitchell,a Nadeem O. Kaakoush,c Mun Fai Loke,a Khean Lee Goh,a and Jamuna Vadivelu,a

University of Malaya, Kuala Lumpur, Malaysiaa; Malaysian Genomics Resource Centre Berhad, Kuala Lumpur, Malaysia;b; and University of New South Wales, Sydney, Australiac

Malaysia is among countries with an intermediate gastric cancer incidence, demonstrating significant differences in the three major ethnic groups (Malay, Chinese, and Indian) in Helicobacter pylori prevalence and gastric cancer incidence (2). The draft genomes of 10 closely related H. pylori isolates from the multiracial Malaysian population will provide an insight into the genetic diversity of isolates in Southeast Asia. These isolates were cultured from gastric biopsy samples from patients with functional dyspepsia and gastric cancer. The availability of this genomic information will provide an opportunity for examining the evolution and population structure of H. pylori isolates from Southeast Asia, where the East meets the West.

TABLE 1 Sequencing statistics and genome information

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>No. of scaffolds</th>
<th>No. of bases (≥ 500 bp)</th>
<th>No. of bases</th>
<th>Maximum scaffold size</th>
<th>Mean scaffold size</th>
<th>N50</th>
<th>Avg sequencing coverage (times)</th>
<th>Genome size (bp)</th>
<th>GC content (%)</th>
<th>Predicted no. of protein-coding sequences</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>FD568</td>
<td>155</td>
<td>1,592,946</td>
<td>72</td>
<td>1,604,711</td>
<td>437,216</td>
<td>17,497</td>
<td>78,927</td>
<td>1,494,164</td>
<td>38.54</td>
<td>1,587</td>
<td>hspEAsia; FD</td>
</tr>
<tr>
<td>GC26</td>
<td>135</td>
<td>1,582,963</td>
<td>77</td>
<td>1,597,232</td>
<td>203,946</td>
<td>18,173</td>
<td>54,475</td>
<td>1,534,861</td>
<td>38.58</td>
<td>1,613</td>
<td>hspEAsia; GC</td>
</tr>
<tr>
<td>FD506</td>
<td>245</td>
<td>1,573,056</td>
<td>58</td>
<td>1,608,477</td>
<td>505,763</td>
<td>22,085</td>
<td>94,183</td>
<td>1,535,289</td>
<td>38.18</td>
<td>1,672</td>
<td>hspEAsia; FD</td>
</tr>
<tr>
<td>FD577</td>
<td>115</td>
<td>1,592,290</td>
<td>59</td>
<td>1,609,138</td>
<td>314,417</td>
<td>24,775</td>
<td>59,288</td>
<td>1,580,091</td>
<td>38.40</td>
<td>1,601</td>
<td>hspEAsia; FD</td>
</tr>
<tr>
<td>FD662</td>
<td>134</td>
<td>1,632,154</td>
<td>58</td>
<td>1,659,218</td>
<td>174,469</td>
<td>25,557</td>
<td>62,179</td>
<td>1,481,802</td>
<td>38.52</td>
<td>1,640</td>
<td>hspEItaly; FD</td>
</tr>
<tr>
<td>FD719</td>
<td>140</td>
<td>1,617,260</td>
<td>72</td>
<td>1,628,999</td>
<td>220,324</td>
<td>19,662</td>
<td>92,906</td>
<td>1,471,471</td>
<td>38.90</td>
<td>1,612</td>
<td>hspEItaly; FD</td>
</tr>
<tr>
<td>FD703</td>
<td>126</td>
<td>1,622,324</td>
<td>70</td>
<td>1,637,264</td>
<td>204,996</td>
<td>21,842</td>
<td>53,160</td>
<td>1,540,544</td>
<td>38.85</td>
<td>1,611</td>
<td>hspEItaly; FD</td>
</tr>
<tr>
<td>FD430</td>
<td>161</td>
<td>1,620,357</td>
<td>70</td>
<td>1,638,660</td>
<td>214,779</td>
<td>20,009</td>
<td>75,391</td>
<td>1,559,901</td>
<td>38.89</td>
<td>1,644</td>
<td>hspEItaly; FD</td>
</tr>
<tr>
<td>FD355</td>
<td>117</td>
<td>1,625,362</td>
<td>62</td>
<td>1,633,165</td>
<td>319,872</td>
<td>23,029</td>
<td>86,768</td>
<td>1,556,024</td>
<td>38.91</td>
<td>1,588</td>
<td>hspEItaly; FD</td>
</tr>
<tr>
<td>FD423</td>
<td>190</td>
<td>1,599,675</td>
<td>80</td>
<td>1,620,118</td>
<td>147,294</td>
<td>17,089</td>
<td>74,214</td>
<td>1,547,587</td>
<td>38.89</td>
<td>1,627</td>
<td>hspEItaly; FD</td>
</tr>
</tbody>
</table>

All isolates were positive for the well-described housekeeping genes, which include atpA (a gene encoding the ATP synthase subunit A chain), glr (a glutamate racemase gene), ppa (an inorganic pyrophosphatase gene), efp (an elongation factor p gene), trpC (a bifunctional indole-3-glycerol phosphate synthase gene), fur (a ferric uptake regulation protein gene), and cysS (a cysteinyler-tRNA synthetase gene). In addition, all isolates were also positive for virulence genes: the cag pathogenicity island (PAI), vacA, and homAB.

It was predicted that the assembled genomes in this study contain approximately 1,620 genes (average), which is consistent with the H. pylori 26695 and J99 genomes, which contain 1,590 and 1,495 genes, respectively (1, 9). Based on the genomes of 26695 and J99, Salama et al. (5) and Gressmann et al. (3) attempted to provide an estimate of the number of genes belonging to the core

References

1. Rana, A. (2012) The draft genomes of 10 closely related H. pylori isolates from the multiracial Malaysian population will provide an insight into the genetic diversity of isolates in Southeast Asia. These isolates were cultured from gastric biopsy samples from patients with functional dyspepsia and gastric cancer. The availability of this genomic information will provide an opportunity for examining the evolution and population structure of H. pylori isolates from Southeast Asia, where the East meets the West. ACEJMB 194(20), p. 5695–5696.

2. Sudesh, A., Khoo, K. H., and Wong, R. (2012) All isolates were positive for the well-described housekeeping genes, which include atpA (a gene encoding the ATP synthase subunit A chain), glr (a glutamate racemase gene), ppa (an inorganic pyrophosphatase gene), efp (an elongation factor p gene), trpC (a bifunctional indole-3-glycerol phosphate synthase gene), fur (a ferric uptake regulation protein gene), and cysS (a cysteinyler-tRNA synthetase gene). In addition, all isolates were also positive for virulence genes: the cag pathogenicity island (PAI), vacA, and homAB.

3. Ang, T. S., and Natarajan, V. (2012) It was predicted that the assembled genomes in this study contain approximately 1,620 genes (average), which is consistent with the H. pylori 26695 and J99 genomes, which contain 1,590 and 1,495 genes, respectively (1, 9). Based on the genomes of 26695 and J99, Salama et al. (5) and Gressmann et al. (3) attempted to provide an estimate of the number of genes belonging to the core
genome of *H. pylori*, their estimates being 1,281 and 1,111 genes, respectively. In comparison, using the predicted genes from this study, which spans two subpopulations (hpAsia2/hspIndia and hpEastAsia/hspEAsia) and two disease groups, the core genome of *H. pylori* was extrapolated to contain no more than 760 genes. With less than 50 percent of its gene pool being well conserved across the entire *H. pylori* species, this study suggests that *H. pylori* may be genetically even more diverse than previously thought.

In conclusion, the availability of sequences of these closely related isolates will provide a platform for further analysis of genomic variability and plasticity, as well as bacterial evolution. Most importantly, data presented in this study have highlighted a need to take into consideration geographical and population variations in future genomic studies.

**Nucleotide sequence accession numbers.** The *H. pylori* draft genomes in this study have been deposited as a whole-genome shotgun project (BioProject ID no. PRJNA165757) at DDBJ/EMBL/GenBank under the accession numbers AKHM00000000 (*H. pylori* FD423), AKHN00000000 (FD430), AKHO00000000 (FD506), AKHP00000000 (FD535), AKHQ00000000 (FD568), AKHR00000000 (FD577), AKHS00000000 (FD703), AKHT00000000 (FD662), AKHU00000000 (FD719), and AKHV00000000 (GC26). The version described in this article is the first version, accession numbers AKHM01000000 to AKHV01000000.

**ACKNOWLEDGMENT**

We thankfully acknowledge support received from the University of Malaya-Ministry of Higher Education (UM-MOHE) High Impact Research (HIR) grant (reference UM.C/6251/HR/MOHE/CHAN-02; account no. A000002-50001, “Molecular Genetics”).

**REFERENCES**

4. Reference deleted.