Genomic analysis of *Helicobacter pylori* from Andhra Pradesh, South India: Molecular evidence for three major genetic clusters

Niyaz Ahmed*, Aleem A. Khan†, Ayesha Alvi†, Santosh Tiwari†, C. S. Jyothirmayee*, Farhana Kauser*, Mahboob Ali† and Chittoor M. Habibullah†,§

*Pathogen Evolution Group, Centre for DNA Fingerprinting and Diagnostics, Nacharam, Hyderabad 500 076, India
†Centre for Liver Research and Diagnostics, Deccan College of Medical Sciences, Hyderabad 500 058, India

Earlier analyses of a very few selected gene loci of the gastric pathogen *Helicobacter pylori* indicated that different genotypes predominate in different human populations. This interpretation was tested and extended by whole genome fingerprinting and genotyping of fluorescent amplified-fragment length polymorphisms (FAFLP) of *H. pylori* isolates from Andhra Pradesh. Highly reproducible FAFLP profiles derived from *EcoRI*/*Mse*I restricted fragments revealed more or less similar genotype for 70 independent isolates with sufficient clonal groupings. These genotypes were however distinct from the European strains (*H. pylori* J99 and *H. pylori* 26695) and the reference strain NCTC 11637. We were able to genotype the isolates based on differential amplification of a total of 31 genomic loci representing important genes involved in molecular pathogenicity, adhesion/motility, toxicity/chemotaxis, DNA replication, translation and metabolism. Phylogenetic analyses revealed major affinities among the isolates and separated them as three different strain groups. This suggests colonization by at least three dominant strain types without much evidence for genetic recombination.

THE gastric pathogen, *H. pylori*, has been a major cause of peptic ulcer disease and is an early risk factor for gastric carcinoma. It is a known gastric pathogen that infects more than half of the world population[1,2]. Clinical isolates of *H. pylori* obtained from different individuals and ethnic groups in the world exhibit substantial genomic diversity than other bacterial species[3,4]. The *H. pylori* infection is very common in India where millions of adults and children are at the risk of developing gastric inflammation, ulcers and carcinoma[1,5]. To our knowledge, there has been no (or very little) effort made to date to analyse the genetic characteristics of the underlying strains representing Southern India, although studies related to genotypic analyses of isolates from Eastern part of India (Calcutta) have been published[1,3,5].

§For correspondence. (e-mail: cmhabib@rediffmail.com; niyaz@cdfd.org.in)

Much of our information on genetic structure of *H. pylori* is based on the studies conducted on European/American strains. Many of these studies describe *H. pylori* population as panmictic[6–8]. This could be due to synonymous base substitutions, gene deletions and insertion of mobile elements in conserved genes[1,3] that may be enhanced by possible inter-strain recombination in case of co-colonization[6–8]. In contrast, most other well-characterized bacterial pathogens are much more strongly clonal[9] such as the *Mycobacterium tuberculosis* complex[10]. Genomics and comparative genomics[11–13] have further proved plasticity of the *H. pylori* genome mainly due to sequence divergence (~ 3 to 5%) typically found in essential genes. On the other hand significantly clonal groupings have also been described to exist among *H. pylori*[6] populations in the world.

It has long been hypothesized that *H. pylori* diversity would be enhanced if humans differ in their food habits (and thereby in gastric environments) and in traits that are important to individual strains such as highly specific immune responses and or availability of receptors helpful in adhesion[14]. However, it depends on patterns of transmission in communities that are dictated by lifestyle and personal hygiene. This is especially true in the Indian context where in spite of a great ethnic diversity, there is a lot of similarity in food habits, lifestyle and practice of medicine among the communities. At the same time, many ethnic groups and tribes are culturally and linguistically distinct and do not intermarry with people of other communities and tribes. This is particularly true for the south Indians of Telugu linguistic group who are mainly Dravidians and married consanguinously for millennia. Their genetic separation from other Indian communities during much of the human history has already been discussed[15]. By analogy, isolates recovered from such population will indeed be distinct and possibly clonally evolved. Such distinctiveness needs to be dissected out at genome wide interface along with its significance in pathogen evolution and disease manifestation. These concerns and our continued interest in studying pathogen evolution through geno-

mic analyses of bacterial isolates[16–19], led to the present study towards describing *H. pylori* population structure in Andhra Pradesh, through a powerful, whole-genome-fingerprinting method of FAFLP[16–19] analysis.

Materials and methods

Bacterial isolates

Seventy independent isolates of *H. pylori* were cultured from gastric biopsies obtained after informed consents from patients of both the sexes with different family backgrounds who have undergone upper gastrointestinal endoscopy and were diagnosed for gastritis and duodenal ulcers. *H. pylori* was isolated by inoculating the minced biopsy specimen on Brucella agar (Becton-Dickinson, USA) supplemented with 7% sheep blood and antibiotics (Vancomycin, 6 mg/l, Amphotericin-B, 3 mg/l, Polymixin-B, 2500 IU/l) (Sigma, USA) and incubating it at 37°C under microaerophilic conditions for 3–7 days. *H. pylori* colonies were identified based on morphology, Gram staining, a positive urease test and subsequent gene-specific PCR tests for the presence of 16S ribosomal RNA gene fragments. Genomic DNA for FAFLP analysis was prepared by the cetyl-trimethyl-ammonium bromide (CTAB) method from all the isolates[1,3]. Genomic DNA of *H. pylori* reference strains 26695, J99 and NCTC 31167 were a gift from Douglas E. Berg, Washington University, St. Louis, USA.

Computer methods

The genome sequence data of *H. pylori* 26695 (http://www.tigr.org)[11] were used to model the FAFLP profiling *in silico* (computationally). The genome data were queried by sequence-based searches using Lasergene™ package (DNAstar Inc. Madison, USA) for the presence of restriction endonuclease sites and a possible compatibility for various hexacutter and tetracutter enzymes to be chosen for restriction ligation reactions. Based on the theoretical possibilities of getting a large number of highly discriminative fragments with *Eco*RI/*Mse*I combination, these enzymes were selected for the study. The data concerning the size and number of fragments predicted following digestion of the genome were organized in the form of a spreadsheet. The *Mse*I/*Mse*I and *Eco*RI/*Eco*RI fragment data were deleted from the spreadsheet. The fragment size data were then filtered with the incorporation of selective nucleotides (A, G, C or T) at the 3′ end of the staggered overhangs produced by *Eco*RI digest. The fragment size data were then adjusted to allow for the addition of primers during PCR, and those fragments predicted to be amplified with each of the chosen selective primers were identified.

Design and synthesis of adapters and primers

All the oligonucleotides were synthesized on an ABI 392 (Perkin Elmer, FosterCity, USA) synthesizer. Fluorescent labelled oligos were specifically labelled on the synthesizer using fluorescent amidites. After desalting with Sephadex G25 (Amersham Pharmacia Biotech, USA), oligonucleotides were directly used as adapters and primers for FAFLP analysis. AFLP adapters consisted of a core sequence and an enzyme-specific sequence[16–19]. The non-selective forward primer for the *Mse*I adapter site was unlabelled. The reverse primer for the *Eco*RI adapter site, which contained a selective base A, G, C or T was labelled with a fluorophore (FAM, JOE, NED or TAMRA).

Restriction ligation reactions

Simultaneous restriction ligation reactions were carried out according to the method described earlier[16].

Preselective and selective amplification

For the purpose of initial low level selection of the FAFLP markers, preselective, exponential, amplification was carried out with primers anti-sense to the microbial adaptors on both the sides of the restricted fragments. This, essentially, helped in purifying the target away from sequences that amplify only linearly, i.e. those with one modified end. With the target concentration of 4.0 µl of diluted DNA prepared by the restriction ligation step and 0.5 µl of each of the preselective primers of 10 µM concentration, amplification was carried out in a 20 µl volume with the ready to use amplification cocktail (PE Biosystems, Foster City, USA) containing *Taq* polymerase, dNTPs and 10X PCR buffer. PCR cycling was carried out in a GeneAmp 9700 thermal cycler (PE Applied Biosystems, USA) for 18 cycles comprising three steps each of 94°C for 20 s, 55°C for 30 s and 72°C for 2 min. Preselective products were diluted 20 folds and stored at 4°C until further use. The *Eco*RI- and *Mse*I- modified fragments were re-amplified using fluorescent dye labelled primers ensuring selective amplification. Unlabelled and labelled primers were used at the concentration of 5 and 1 pmoles/µl respectively in a multiplex fashion in a 15 µl volume including 1.5 µl of the diluted preselective amplification product. Thermal cycling was carried out with a touch down PCR module. Initial denaturation was carried out at 94°C for 2 min followed by individual cycles of denaturation (94°C for 20 s), annealing (66–56°C for 30 s,), and extension (72°C for 2 min). The annealing temperature, beginning with 66°C, was decreased by one degree in each of the individual cycles. This was followed by 20 cycles of amplification with annealing temperature of 56°C. A hold of 30 min at 60°C was given finally before shifting the samples to 4°C for storage.

Gel electrophoresis and data collection

FAFLP products were loaded on to an ABI Prism 377 XL-96 DNA sequencer along with a red-coloured internal lane standard GS-500 Rox (Applied Biosystems, USA). Fragment separation was continued for 5 h on a 5% denaturing polyacrylamide gel. Fragments were detected and compiled by the ABI Data Collection™ (Applied Biosystems, USA) software. Gel images were generated and all the lanes were extracted for making individual electropherograms. Fragment analysis was performed with the Genescan Analysis™ 3.1 package (Applied Biosystems, USA). Individual sample files were then exported to the Genotyper™ 2.5 software for computer-assisted genotyping.

Gel analysis and genotyping

Based on the presence or absence of monomorphic and polymorphic bands/peaks, different FAFLP profiles were identified as 'amplitypes'. These amplitypes were colour-coded, and superimposed in Genescan™ to estimate the marker size, fluorescence intensity, data points on the gel and frequency of the monomorphic bands. The Genescan™ data of individual isolates was exported to Genotyper™ 2.5 (Applied Biosystems, USA), for genotyping. Bands were sized and genotyped for all the isolates within the user-defined categories of marker sizes (allele bins). Presence or absence of markers within the categories was scored by a Genotyper™ macro that generated final output in the form of a binary table for all the lanes. Phylogenetic trees were generated from the binary data to delineate relatedness among the amplitypes.

Confirmation and locus assignment of FAFLP markers

Polymorphic fragments flanked by *Eco*RI and *Mse*I sites on either sides and with selective nucleotides (A, G, C, or T) at 3′ end of the *Eco*RI sites were mapped on the *H. pylori* 26695 genome sequence with the help of Lasergene Navigator (DNAstar Inc., USA). Locus assignment was carried out after sequencing analysis and BLAST confirmation. Fragments were selectively amplified and separated on a PAGE gel and were eluted and purified by using GeneClean kit (Bio101, USA). Automated sequencing of FAFLP fragments was carried out by using Big Dye Terminator™ reactions (Applied Biosystems, USA).

Results

Computer modelling and experimental evaluation of FAFLP

The combination of enzymes and primers for analysis of *H. pylori* had been predicted by modelling FAFLP using computer analysis of the sequenced and annotated genome of strain *H. pylori* 26695. *In silico* digestion of the full genome sequence revealed 18,869 sites for *Mse*I and 184 sites for *Eco*RI. However, the four primer combination used (*Eco*RI + A and *Mse*I + 0, *Eco*RI + G and *Mse*I + 0, *Eco*RI + C and *Mse*I + 0, and *Eco*RI + T and *Mse*I + 0) generated a total of 76 dif ferently sized fragments ranging in size from 50 to 500 bp. *H. pylori J99* produced 95 fragments as against a total of 70 fragments generated with the *H. pylori NCTC 11637* genome. On the other hand, the Hyderabad strains generated ~ 145 fragments, almost double than the number of fragments produced by *H. pylori 26695*. For Hyderabad strains the A-selective primer combination (*Eco*RI + A and *Mse*I + 0) produced 37 of 145 fragments (25.5%) while the C-selective primer combination (*Eco*RI + C and *Mse*I + 0) produced 31 of 145 fragments (21.3%). The G-selective primer combination (*Eco*RI + G and *Mse*I + 0) produced 35 of 198 fragments (17.7%) and the T-selective primer combination (*Eco*RI + T and *Mse*I + 0), produced 42 of 198 fragments (28.9%). Table 1 shows the presence of these precisely sized discriminating fragments generated in experimental FAFLP using non-selective *Mse*I primer and each of the four *Eco*RI-selective primers containing A, C, T, or G at their 3′ terminus.

Automated genotyping and mapping of FAFLP markers

FAFLP gel data consisted of precisely amplified fragments in the size range of 50 to 500 bp. The two-dimensional gel images were transformed with Genescan™ 3.1 software (ABI) into electropherograms; these were colour-coded, overlaid, and visually inspected for polymorphisms. For every single lane representing the whole genome fingerprint of the corresponding isolate, FAFLP data were found to be highly digitized, precise and reproducible.

After sequence-confirmation and BLAST searches, some of the FAFLP loci represented the ORFs as important as those associated with pathogenicity (cag pathogenicity island protein/HP0524), adhesion/motility (polar-flagellin/HP0751) and toxicity/chemotaxis (vacuolating cytotoxin/HP0887). Also, some of the genomic loci represented ORFs involved in DNA replication and translation (DNA pol. III-*b*/HP0500 and Ribosomal protein S19/HP1315). Besides this, some FAFLP markers were corresponding to the genes important in biosyntheses and metabolism such as threonine synthase (HP0098), selenocysteine synthase (HP1513), adenine–cytosine DNA methyl transferase (HP0054), DAP-decarboxylase (HP0290), acetyl co-A synthase (HP1045) and beta-ketoacyl carrier protein synthase (HP0558). Many of the loci represented some of the ORFs coding for hypothetical proteins and conserved hypothetical proteins in

the *H. pylori* genome. Details of the loci taken into consideration for genotyping have been given in Table 1.

Genetic relationships among H. pylori clinical isolates

Chromosomal FAFLP markers were analysed by neighbour joining trees derived from tables generated by Genotyper 2.5 software. It is important that most of the isolates of *H. pylori* worldwide are not very similar to each other and the isolates often can be differentiated upon DNA fingerprinting. However in our case, an unusual affinity among the isolates was observed based on band sharing. The isolates were clustered in three separated strain groups (Figure 1). Thirty-two of the 70 isolates analysed show some similarity to the European strain *H. pylori 26695* and clustered accordingly. Maximum observed genetic distance between the three strain types was ~ 30% only. Most of the isolates clustered with an intra-strain distance of less than 10%. Three independent isolates namely DMC9, DMC 18 and DMC 24 were found to be 100% similar and clustered on the same data point (Figure 1). Another two pairs of unrelated isolates (DMC 36 & DMC 41; DMC 50 & DMC 62) were found to be clonal in origin. Similar topology of the neighbour joining trees were obtained with different primer pairs (data not shown).

Discussion

Genetic diversity of the gastric pathogen, *H. pylori* has been well studied[1,3,5–7]. However, this diversity is mainly due to synonymous substitutions and insertion–deletions due to frequent genetic recombination. It is also evident that *H. pylori* transmission occurs strictly within families in humans[2]. It is not clear how frequently the recombinational events occur in such cases where *H. pylori* is transmitted exclusively within families where mixed infections are a rare phenomenon, although recombinant strains are known to emerge during human infection[7]. It will therefore, be interesting to know the genetic structure of such slowly evolving bacteria and their impact on gastroduodenal pathology. For this reason, computer–assisted FAFLP analysis was selected as a whole genome scanning tool. As a result, the many FAFLP fragments generated from our isolates reflect the availability of many more *Eco*RI sites in these isolates as against the computer analysis of *H. pylori 26695* genome where only 184 sites have been predicted.

Table 1. Details of polymorphic *Eco*RI/*Mse*I-FAFLP fragments in *H. pylori*. Marker sizes include adapter sequences on both the ends

FAFLP markers (bp)	Genomic location (*H. pylori* 26695)	Locus	No. of isolates analysed	No. positive
83	*Mse*I + 0 417573–417623 *Eco*RI + C	HP0405	59	44
88	*Mse*I + 0 43348–433403 *Eco*RI + G	HP1113	34	16
89	*Eco*RI + C 526903–526959 *Mse*I + 0	HP0500	59	45
93	*Eco*RI + C 988093–988153 *Mse*I + 0	HP0477	59	41
102	*Eco*RI + C 618747–618816 *Mse*I + 0	HP0586	59	33
102	*Eco*RI + G 1588384–1588453 *Mse*I + 0	HP1632	34	16
103	*Eco*RI + G 367334–367404 *Mse*I + 0	HP0357	34	15
106	*Eco*RI + G 508695–508768 *Mse*I + 0	HP0506	34	21
111	*Eco*RI + C 306899–306977 *Mse*I + 0	HP0290	59	43
114	*Mse*I + 0 645118–645199 *Eco*RI + C	HP0912	59	22
118	*Eco*RI + T 1361387–1361472 *Mse*I + 0	HP0269	59	17
129	*Eco*RI + T 1090831–1090927 *Mse*I + 0	HP1028	59	25
134	*Mse*I + 0 593264–593364 *Eco*RI + C	HP0558	59	33
136	*Mse*I + 0 213959–214062 *Eco*RI + G	HP0205	34	20
137	*Eco*RI + C 1098736–1098838 *Mse*I + 0	HP1038	59	37
153	*Eco*RI + C 1378250–1378370*Mse*I + 0	HP1315	59	03
159	*Mse*I + 0 896887–897013 *Eco*RI + C	HP0486	59	15
160	*Eco*RI + A 1107482–1107610 *Mse*I + 0	HP1045	59	07
162	*Mse*I + 0 959146–959275 *Eco*RI + C	HP0098	59	39
169	*Eco*RI + G 570060–570197 *Mse*I + 0	HP0564	34	20
183	*Mse*I + 0 203214–203364 *Eco*RI + G	HP0196	34	11
190	*Eco*RI + G 550814–550971 *Mse*I + 0	HP0551	34	15
197	*Mse*I + 0 1615269–1615433 *Eco*RI + C	HP1537	59	06
200	*Mse*I + 0 176686–176853 *Eco*RI + G	HP0170	34	22
203	*Eco*RI + C 806495–806665 *Mse*I + 0	HP0751	59	18
212	*Mse*I + 0 55663–55845 *Eco*RI + T	HP0054	59	12
223	*Eco*RI + G 892319–892509 *Mse*I + 0	HP0893	34	14
232	*Mse*I + 0 1404879–1405078 *Eco*RI + G	HP1456	34	21
237	*Eco*RI + C 214069–214273 *Mse*I + 0	HP0208	59	05
339	*Eco*RI + C 938694–938999 *Mse*I + 0	HP0887	59	25
442	*Eco*RI + C 796076–796484 *Mse*I + 0	HP0741	59	01

This implies for the prevalence of a genetically distinct strain type in this region compared with the *H. pylori J99* and *H. pylori 26695* and *H. pylori* NCTC11637 genotypes (Figure 2). The genomic fragments amplified by us were corresponding to different coding and non-coding regions throughout the genome (Table 1) including both conserved genes such as those coding for universal proteins and variable regions. This indicates the spectrum of microevolution of *H. pylori* strains *vis-à-vis* base substitutions at the *Eco*RI/*Mse*I cleavage sites.

Phylogenetic analysis based on these fragments revealed three dominant clusters independent of presence or absence of the *cag* PAI and *vac* allele status. This suggests the possibility of colonization by three dominant strains. This may be in contrast with the observations related to a highly divergent population genetic structure of *H. pylori* world-

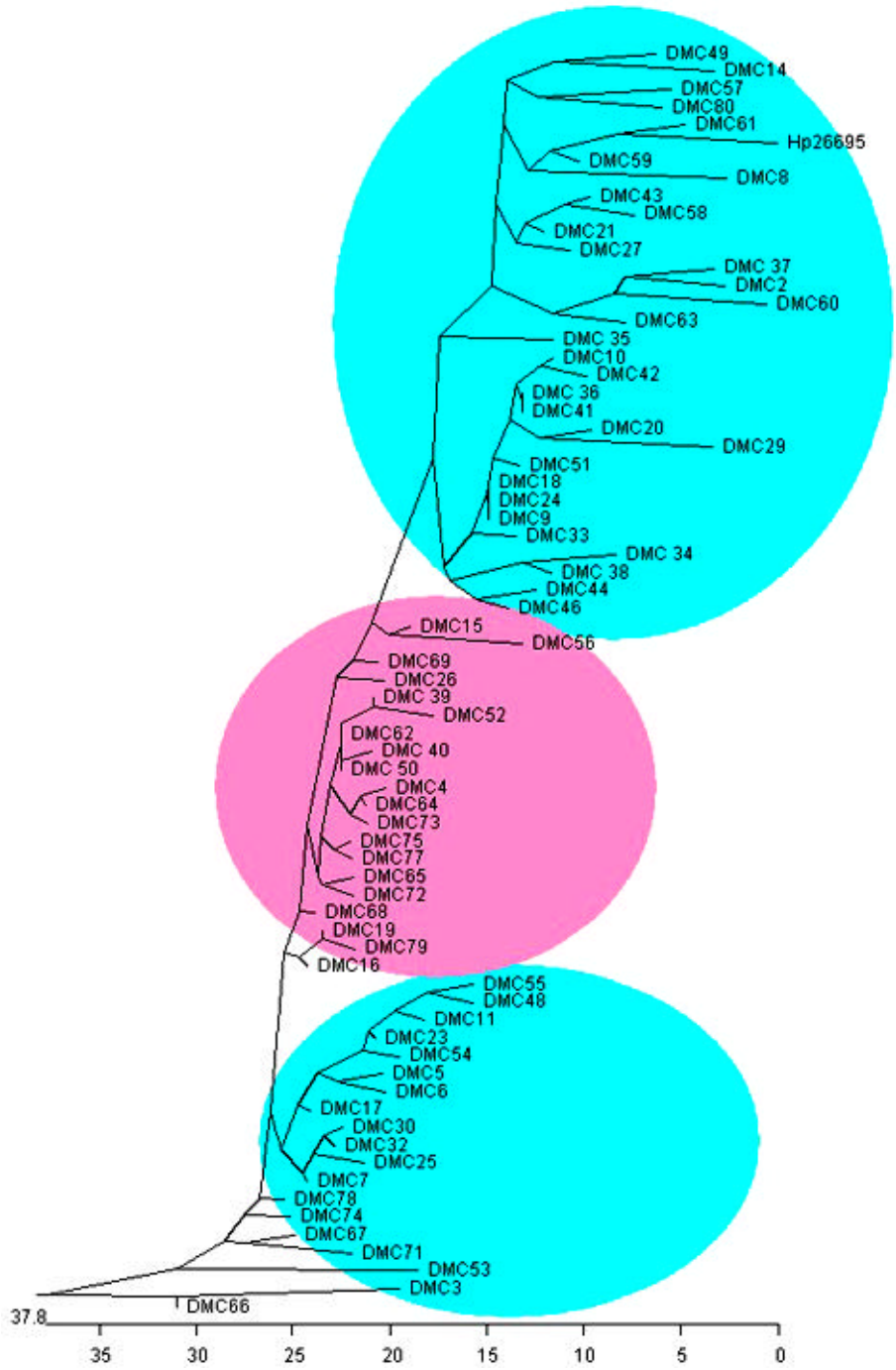

Figure 1. Neighbour joining tree generated using the binary data obtained from Genotyper™ macro. Isolates are divided into three distinct clusters based on the presence or absence of polymorphic bands within the Genotyper™ categories. The lower scale shows genetic dissimilarity among the isolates.

wide where no two independent isolates are described to be similar[1,3,4,6,20–22]. However, many of these observations are mainly based on insertion and deletion events within and in the vicinity of the cag PAI, allelic diversity of vacA sequences and synonymous nucleotide substitutions in structural genes. It is interesting that clonal groupings in our isolates were found to be independent of the disease outcome and the environmental factors such as dietary habits, smoking and alcoholism. Identical (and possibly clonal) genotypic structure of most of the independent isolates (Figure 3), therefore, is a novel finding as opposed to the studies conducted on the isolates from the West where the structure of *H. pylori* population is described as more diversified and panmictic[3,4,6–8]. Our observations appear to be more promising when we take into account the prevailing societal patterns (in the south Indian communities) that minimize the chances of horizontal transmission via oral contacts or sharing of food and utensils, etc. with unrelated individuals. Notably, the existing pattern of consanguinous marriages and closely maintained family relations and traditions might argue for a conserved, faithful, vertical transmission of a single strain over the years.

Interestingly, these findings are broadly congruent with the observations made by Achtman and colleagues[6] describing that clonal origins of *H. pylori* could still be discerned despite a rich history of inter-strain recombination and that this might reflect possible founder effects and geographical isolations. Based on this working hypothesis, we suggest that distinctiveness of *H. pylori* genotypes in South India could be due to strain variants of *H. pylori* that are particularly well adapted to these patients and evolved clonally in three lineages. These variants might have spread through the families mostly vertically and recombination has not been sufficiently frequent to break the evidence for clonal descent. In an other scenario, recombination seems to have occurred for most genes studied

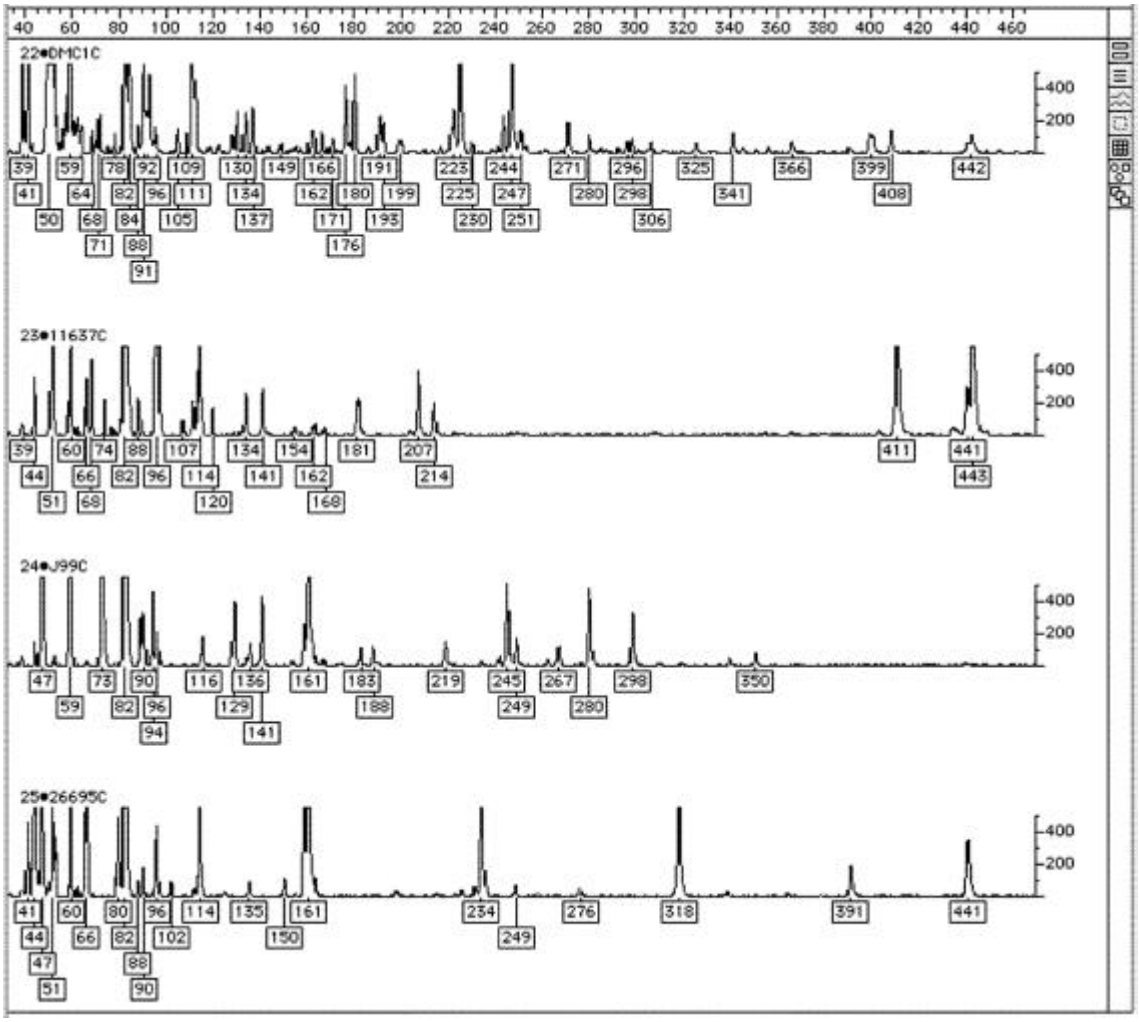

Figure 2. FAFLP genotypes of the four independent isolates [South Indian DMC1 (plot DMC1C), *H. pylori* NCTC11637 (plot 11637C), *H. pylori* J99 (plot J99C) and *H. pylori* 26695 (plot 26695C)] showing the individual *Mse*I + 0/*Eco*RI + C FAFLP patterns. The fragment sizes attributed to the peaks recognized by Genotyper™ software are shown. Traces are labelled with their sizes (rounded to integer) in base pairs. Scale on the right represents intensity of the FAFLP products on the gel (Genotyper peak height).

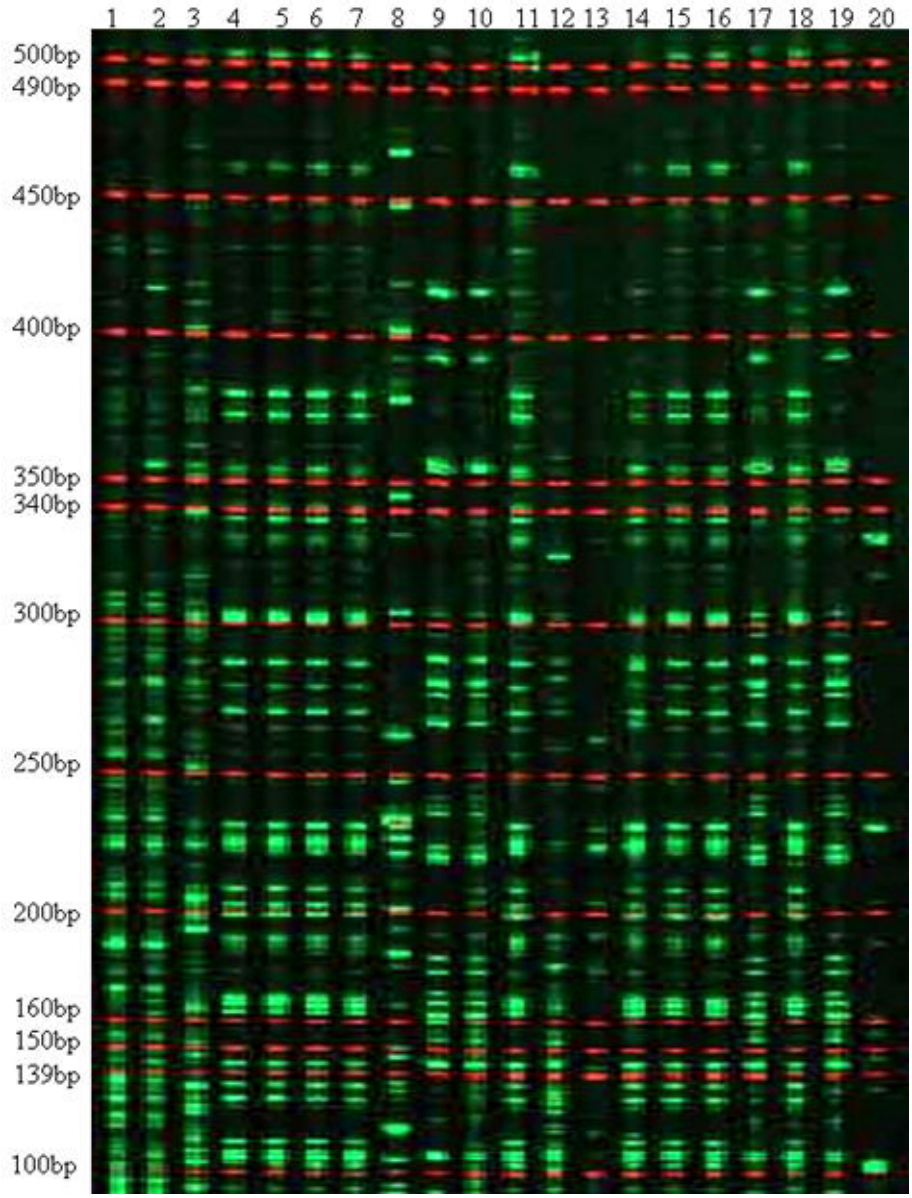

Figure 3. Two dimensional gel image of representative FAFLP patterns for *H. pylori* isolates from Andhra Pradesh. The red-coloured bands indicate internal lane standards (Genescan Rox 500).

(such as cag PAI, vac alleles, flagellins, outer membrane proteins, etc.)[1,3,4,6–8,21–23], but has not totally disrupted genetic affinities on a whole genome basis. This was possibly due to the fact that recombination, although possible, will be a rare event if the host is not colonized by multiple strains (horizontal infections).

FAFLP profiles generated in this study can be used for electronic archiving and retrieval for inter-laboratory comparison and are suitable for storage in epidemiological databases for comparative analyses. Given that the two *H. pylori* genomes sequenced to date are each from ethnic Europeans, genomic FAFLPs modelled on these data should facilitate identification of novel loci from additional representative strains, especially from understudied non-European populations. FAFLP-based identification and characterization of such loci which are abundant in the Asian than in the Western gene pool may lead to newer insights into the mechanisms of *H. pylori* colonization, microevolution carriage and virulence.

1. Mukhopadhyay, A. K. *et al*., Distinctiveness of genotypes of Helicobacter pylori in Calcutta, India. *J. Bacteriol*., 2000, **182**, 3219–3227.
2. Westblom, T. U., Czinn, S. J. and Nedrud, J. G., Gastroduodenal disease and Helicobacter pylori: pathophysiology, diagnosis and treatment. *Curr. Top. Microbiol. Immunol*., 1999, **241**, 1–313.
3. Kersulyte, D. *et al*., Differences in genotypes of Helicobacter pylori from different human populations. *J. Bacteriol*., 2000, **182**, 3210–3218.

4. Logan, R. P. H. and Berg, D. E., Genetic diversity of Helicobacter pylori. *Lancet*, 1996, **178**, 3934–3938.
5. Datta, S. *et al*., Virulence genes and neutral DNA markers of Helicobacter pylori isolates from different ethnic communities of West Bengal, India. *J. Clin. Microbiol*., 2003, **41**, 3737–3743.
6. Achtman, M. *et al*., Recombination and clonal groupings within Helicobacter pylori from different geographical regions. *Mol. Microbiol*., 1999, **32**, 459–470.
7. Kersulyte, D., Chalkauskas, H. and Berg, D. E., Emergence of recombinant strains of Helicobacter pylori during human infection. *Mol. Microbiol*., 1999, **31**, 31–43.
8. Suerbaum, S. *et al*., Free recombination within Helicobacter pylori. *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 12619–12624.
9. Smith, J. M., Smith, N. H., O'Rourke, M. and Spratt, B. G., How clonal are bacteria? *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 4384–4388.
10. Wheeler, P. R. and Ratledge, C., Metabolism of *Mycobacterium tuberculosis*. In *Tuberculosis: Pathogenesis, Protection, and Control* (ed. Bloom, B. R.), American Society for Microbiology, Washington, DC, 1994, pp. 353–385.
11. Alm, R. A. *et al*., Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori. *Nature*, 1999, **397**, 176–180.
12. Tomb, J. F. *et al*., The complete genome sequence of the gastric pathogen Helicobacter pylori. *Nature*, 1997, **388**, 539–547.
13. Doig, P. *et al*., Helicobacter pylori physiology predicted from genomic comparison of two strains. *Microbiol. Mol. Biol. Rev*., 1999, **63**, 675–707.
14. Dubois, A. *et al*., Transient and persistent infection of non-human primates with Helicobacter pylori: implications for human disease. *Infect. Immunol*., 1996, **64**, 2885–2891.
15. Bamshad, M. *et al*., mtDNA variation in caste populations of Andhra Pradesh, India. *Human Biol*., 1996, **68**, 1–28.
16. Ahmed, N. *et al*., Whole genome fingerprinting and genotyping of multiple drug resistant (MDR) isolates of Pseudomonas aeruginosa from endophthalmitis patients in India. *Infection Genet. Evol*., 2002, **1**, 237–242.
17. Ahmed, N. *et al*., Distinctiveness of Mycobacterium tuberculosis genotypes from Human Immunodeficiency Virus Type 1-Seropositive and -Seronegative patients in Lima, Peru. *J. Clini. Microbiol*., 2003, **41**, 1712–1716.
18. Ahmed, N., Alam, M., Majeed, A. A., Rahman, S. A., Cataldi, A., Cousins, D. and Hasnain, S. E., Genome sequence based, comparative analysis of the fluorescent amplified fragment length polymorphisms (FAFLP) of tubercle bacilli from seals provides molecular evidence for a new species within the Mycobacterium tuberculosis complex. *Infect. Genet. Evol*., 2003, **2**, 193–199.
19. Carroll, I. M., Ahmed, N., Beesley, S. M., Khan, A. A., O Moráin, C. A. and Smyth, C. J., Fine Structure Molecular Typing of Irish Helicobacter pylori Isolates and their Genetic Relatedness to Strains from Four Different Continents. *J. Clin. Microbiol*., 2003, **41**, (in press).
20. Akopyanz, N., Bukonov, N. O., Westblom, T. U. and Berg, D. E., PCR based RFLP analysis of DNA sequence diversity in gastric pathogen Helicobacter pylori. *Nucleic Acids Res*., 1992, **20**, 6221–6225.
21. Bamford, K. B. *et al*., Helicobacter pylori: comparison of DNA fingerprints provides evidence for intrafamilial infection. *Gut*, 1993, **34**, 1348–1350.
22. Garner, J. A. and Cover, T. L., Analysis of genetic diversity in cytotoxin-producing and non-cytotoxin-producing Helicobacter pylori strains. *J. Infect. Dis*., 1995, **172**, 290–293.
23. Taylor, N. S. *et al*., Long term colonization with single and multiple strains of *H. pylori* assessed by DNA Fingerprinting. *J. Clin. Microbiol*., 1995, **33**, 918–923.
24. Dimola, S. and Caruso, M. L., Helicobacter pylori in animals affecting the human habitat through food chain. *Anticancer Res*., 1999, **19**, 3889–3894.
25. Dore, M. P., Sepulveda, A. R., Osato, M. S., Realdi, G. and Graham, D. Y., Helicobacter pylori in sheep milk. *Lancet*, 1999, **354**, 132.

ACKNOWLEDGEMENTS. This study was funded by a research grant from the Department of Biotechnology, New Delhi to NA and CMH. We thank Seyed E. Hasnain, Director, CDFD, Hyderabad for his encouragement and support.

Received 3 June 2003; revised accepted 16 September 2003